# Thèse de doctorat

Thèse présentée pour obtenir le grade universitaire de docteur

École doctorale 251 : Sciences de l'environnement spécialité « Océanographie »

Aix-Marseille Université

# Connectivité et structure génétiques des populations d'*Ericaria zosteroides* (Fucales) et *Laminaria rodriguezii* (Laminariales) des côtes françaises

Soutenue publiquement par

**Lauric REYNES**

Le 1er juin 2021



Sous la direction de :
**Thierry THIBAUT** et **Didier AURELLE**

Membres du jury :
**Frédérique VIARD** (rapporteur)
**Christophe DESTOMBE** (rapporteur)
**Myriam VALERO** (examinateur)
**Stéphane COUPÉ** (examinateur)
**Eric PANTE** (examinateur invité)

# Remerciements

Les mots me manquent, mais comment pourrais-je oublier celles et ceux qui ont contribués à l'élaboration de ma thèse ?

Premièrement, je remercie **Didier Aurelle** et **Thierry Thibaut** qui m'ont encadré et guidé tout au long de ce processus. Vos visions respectives de la science, et parfois vos désaccords, ont été riches d'enseignements.

Didier, si j'ai souhaité consacrer ma thèse à la génétique des populations, et aborder les problématiques environnementales actuelles sous cet angle c'est grâce à tes enseignements de Licence et Master. Tu m'as donné envie d'en savoir davantage sur cette discipline et tu m'as fait confiance en me donnant l'opportunité d'effectuer mon stage de Master 1 dans ton équipe. Tu as été disponible et j'ai particulièrement apprécié la facilité avec laquelle tu poses les bonnes questions sous la forme d'une hypothèse vérifiable.

Thierry, je pense à l'époque où j'étais encore en licence et que tu prenais un malin plaisir à critiquer mes dessins d'observations d'algues brunes. À vrai dire, malgré la bonne volonté dont je faisais preuve, cela ressemblait à tout sauf à une Fucales. Sachez par ailleurs que je ne suis pas l'auteur de celle (*Ericaria zosteroides*) figurant en page de couverture de ce mémoire (les honneurs reviennent à **Emma Rozis**). Quoi qu'il en soit Thierry, merci de m'avoir transmis ta passion pour ces écosystèmes que sont les forêts sous-marines. J'ai apprécié ta très grande connaissance du terrain et ta volonté de replacer ma curiosité scientifique, dans le contexte actuel d'érosion de la biodiversité.

Je tiens remercier la **Région Sud Provence-Alpes-Côte d'Azur** et mon partenaire socio-économique le **parc national Calanques** pour leur soutien financier à travers le programme "Emplois Jeunes Doctorants".

J'adresse tous mes remerciements au directeur de l'Institut Méditerranéen d'Océanologie (MIO), **Richard Sempéré** et à tous ceux et celles avec qui j'ai pu collaborer d'une quelconque manière, que ce soit pendant de courts échanges dans un couloir, une aide pendant une expérience ou bien une implication directe dans certains de mes travaux. Ce sont celles et ceux de mon équipe : **Aurélie Blanfuné**, **Charles-François Boudouresque**, **Marc Verlaque**, **Léo Berline**, **Sarah Bouchemousse**, **Thomas Changeux**, **Sandrine Ruitton**, **Delphine Thibault**, **Loic Guilloux**, **Jean-christophe Poggiale**, **Simone Richard** et **Nathalie Prévot**. Celles et ceux de l'équipe 3 - Microbiologie Environnementale Biotechnologie (MEB) : **Laurie Casalot**, **Sophie Guasco**, **Marc Garel**, **Magali Lescot**, **Valérie Michotey**, **Gwenola Simon**, **Léa Sylvi**, **Lisa Tanet**, **Cécile Militon**. Mes collèges de l'équipe 1 - Océanographie Physique, Littorale et Côtière (OPLC) avec qui j'ai défendu une approche interdisciplinaire de la connectivité des populations : **Cristele Chevalier** et **Christel Pinazo**. Je remercie le service d'administration et technique de l'OSU Pythéas et du MIO, je pense à **Dominique Estival**, **Aurélia Lozingot** et **Pascal Lesage**.

Je remercie aussi le service plongée de l'OSU Pythéas et les plongeurs de l'Ifremer et du GIS Posidonie pour leur soutien et leur implication dans ma thèse : **Dorian Guillemain**, **Stéphane Sartoretto**, **Bruno Belloni**, **Elodie Rouanet**, **Patrick Astruch**, **Adrien Goujard**.

Je remercie l'équipe d'Andromède Océanologie, notamment pour la prouesse technique d'avoir échantillonné l'une des espèces que j'étudie, jusqu'à 95 mètres de profondeur : **Florian Holon**, **Julie Deter** et **Laurent Ballesta**.

Merci aussi à l'unité de recherche Biologie évolutive et écologie des algues pour leur accueil durant les quelques semaines que j'ai passé à la Station Biologique de Roscoff.

Je tiens à remercier particulièrement **Myriam Valero** pour sa bienveillance, son franc-parler et sa forte implication dans ma thèse, **Stéphane Mauger** pour sa disponibilité et sa rigueur durant ma formation aux librairies de séquençages RAD. Je remercie également **Christophe Destombe**, **Louise Fouqueau**, **Lucie Jaugeon**, et **Claire Daguin** avec qui j'ai pu échanger lors de ma venue. Cette mission n'aurait pas été possible sans le soutien financier du **GDR Génomique Environnementale**.

Merci aux personnes extérieures aux institutions précédemment citées, je pense à **Stéphane Blanfuné** pour m'avoir permis d'utiliser un ROV, au service **Espaces Maritimes de la communauté de communes du Golfe de Saint-Tropez**, à **Laura Benestan** pour ces précieux conseils quand il s'agissait de mettre en relation les données de courantologies et de génétique, **Thierry Gosselin** et **Antoine Minne** pour nos échanges sur le traitement des données du séquençage RAD.

J'en viens maintenant à celles et ceux qui ont illuminé mes journées de thèse, je pense notamment à **Guillaume Marchessaux**, un collègue qui est devenu au fil du temps un ami très proche sur qui je peux compter, à **Quentin Langevin** pour l'aide apportée, j'ai beaucoup apprécié cette période où on se partageait une même paillasse. Je remercie aussi **Augustin Lafond**, la première personne qui est venu me parler au premier jour, il est devenu un très bon ami à l'image de **Thomas Lefort**, **Justine Legras**, **Amélie Vernale**, **Lucas Lhomond**, **Mélanie Ourgaud**, **Mathilde Mahe** et **Javier Tesán**. Je remercie aussi **Stéphine Botella** pour avoir consacré du temps à la relecture de ce manuscrit dans le but de débusquer les fautes d'orthographe ; elle est ma plus vieille amie.

Mes remerciements vont aussi à ma famille, mes parents pour leur soutien financier durant mes études, ma tante et ma grand-mère qui ont été là pour moi. Merci aussi à **Steeve Roussey** pour son soutien quotidien, la patience et la grande compréhension dont il a fait preuve. Sans nos innombrables moments d'évasion, bien que devenus rares par la rédaction de ce mémoire, cette thèse n'aurait jamais été ce qu'elle est.

Enfin, je tiens à remercier les rapporteurs et examinateurs de ce mémoire, pour l'honneur qu'ils me font d'être dans mon jury de thèse, **Frédérique Viard**, **Christophe Destombe**, **Myriam Valero**, **Stéphane Coupé** et **Eric Pante**.

*À ma grand-mère.*

# Sommaire

**1|Introduction générale ..... 7**

1.1 Contexte général de la thèse ..... 8

1.1.1 Le projet "Marine Forests" (MarFor) ..... 9

1.1.2 Écologie évolutive et changement global ..... 15

1.2 Représentation réduite du génome (génotypage RAD-Seq) ..... 19

1.2.1 Description de la méthode ..... 19

1.2.2 Applications du RAD-seq à différentes échelles d'étude ..... 22

1.3 Structure et diversité génétique des Laminariales et Fucales ..... 26

1.3.1 Echelles de différenciation génétique ..... 26

1.3.2 Reproduction végétative et implications évolutives ..... 34

1.3.3 Patrons d'adaptation locale chez les Laminariales et Fucales ..... 36

1.4 Les forêts marines méditerranéennes ..... 37

1.4.1 Les forêts à *Cystoseira sensu lato*, focus sur *Ericaria zosteroides* ..... 37

1.4.2 Les Laminariales des milieux profonds, focus sur *Laminaria rodriguezii* ..... 43

1.5 Objectifs de la thèse ..... 49

**2|La génomique des populations et la modélisation Lagrangienne soulignent le rôle des courants marins dans la dispersion des forêts marines à *Ericaria zosteroides* (Fucales) ..... 53**

2.1 Introduction ..... 57

2.2 Material and methods ..... 60

2.2.1 Sampling and demographic structure of populations ..... 60

2.2.2 DNA extraction and SNP discovery ..... 62

2.2.3 SNP filtering ..... 63

2.2.4 Species delimitation ..... 64

2.2.5 Outlier loci ..... 64

2.2.6 Population genetic variation ..... 65

2.2.7 Individual assignment analysis ..... 66

2.2.8 Hydrodynamic modeling ..... 66
2.2.9 Lagrangian modeling ..... 67
2.2.10 Environmental predictors of genetic structure ..... 68
2.2.11 Distance Based RDA: linking environment to genetic structure ..... 69
2.3 Results ..... 69
2.3.1 Demographic structure ..... 69
2.3.2 Bioinformatic and genotyping ..... 71
2.3.3 Species delimitation ..... 72
2.3.4 Genetic variation within populations ..... 72
2.3.5 Genetic structure and assignment tests ..... 73
2.3.6 Oceanographic connectivity ..... 75
2.3.7 Factors explaining the genetic structure at a local scale ..... 77
2.4 Discussion ..... 79
2.5 Acknowledgements ..... 84
2.6 Appendices ..... 84
2.7 Synthèse des résultats obtenus ..... 90

**3|Structure des populations et signatures génomiques de la clonalité chez *Laminaria rodriguezii* ..... 91**

3.1 Introduction ..... 96
3.2 Material and methods ..... 98
3.2.1 Sampling, DNA extraction ..... 98
3.2.2 ddRAD-sequencing ..... 99
3.2.3 Genotyping and SNP filtering ..... 100
3.2.4 Impact of SNP filtering on population genetic summary statistics ..... 101
3.2.5 Clonality in Laminaria rodriguezii ..... 102
3.2.6 Genetic diversity ..... 103
3.2.7 Population differentiation in *Laminaria rodriguezii* ..... 104
3.3 Results ..... 104
3.3.1 Sequencing and SNPs filtering ..... 104
3.3.2 Clonal structure of *Laminaria rodriguezii* ..... 105

3.3.3 Genetic diversity ..... 107
3.3.4 Population differentiation in *Laminaria rodriguezii* ..... 112
3.3.5 Negative $F_{IS}$ and high $F_{ST}$ in *L. rodriguezii* over SNP filtering methods ..... 112
3.4 Discussion ..... 114
3.4.1 RAD-seq as a powerful tool for detection of clonal lineages ..... 114
3.4.2 The genomic patterns of clonality ..... 115
3.4.3 Extreme levels of individual genomic heterozygosity ..... 117
3.4.4 Implications for conservation ..... 118
3.5 Conclusion ..... 119
3.6 Acknowledgements ..... 119
3.7 Appendices ..... 120
3.8 Structure génétique des populations de *L. rodriguezii* ..... 122
3.9 Synthèse des résultats obtenus ..... 125

**4|Discussion, conclusion et perspectives ..... 129**

4.1 L'adaptation locale chez *Ericaria zosteroides* ..... 130
4.1.1 Faux positifs ou signal d'adaptation locale ? ..... 130
4.1.2 Réponses physiologiques de *Cystoseira sensu lato* à un stress radiatif ..... 131
4.1.3 Evaluer la réponse physiologique des populations d'*E. zosteroides* ..... 132
4.2 Inférence des processus démographiques ..... 136
4.2.1 Estimation de tailles efficaces à la génération actuelle ..... 136
4.2.2 Vers la modélisation de l'histoire évolutive des populations ..... 144
4.3 Conclusion générale ..... 148

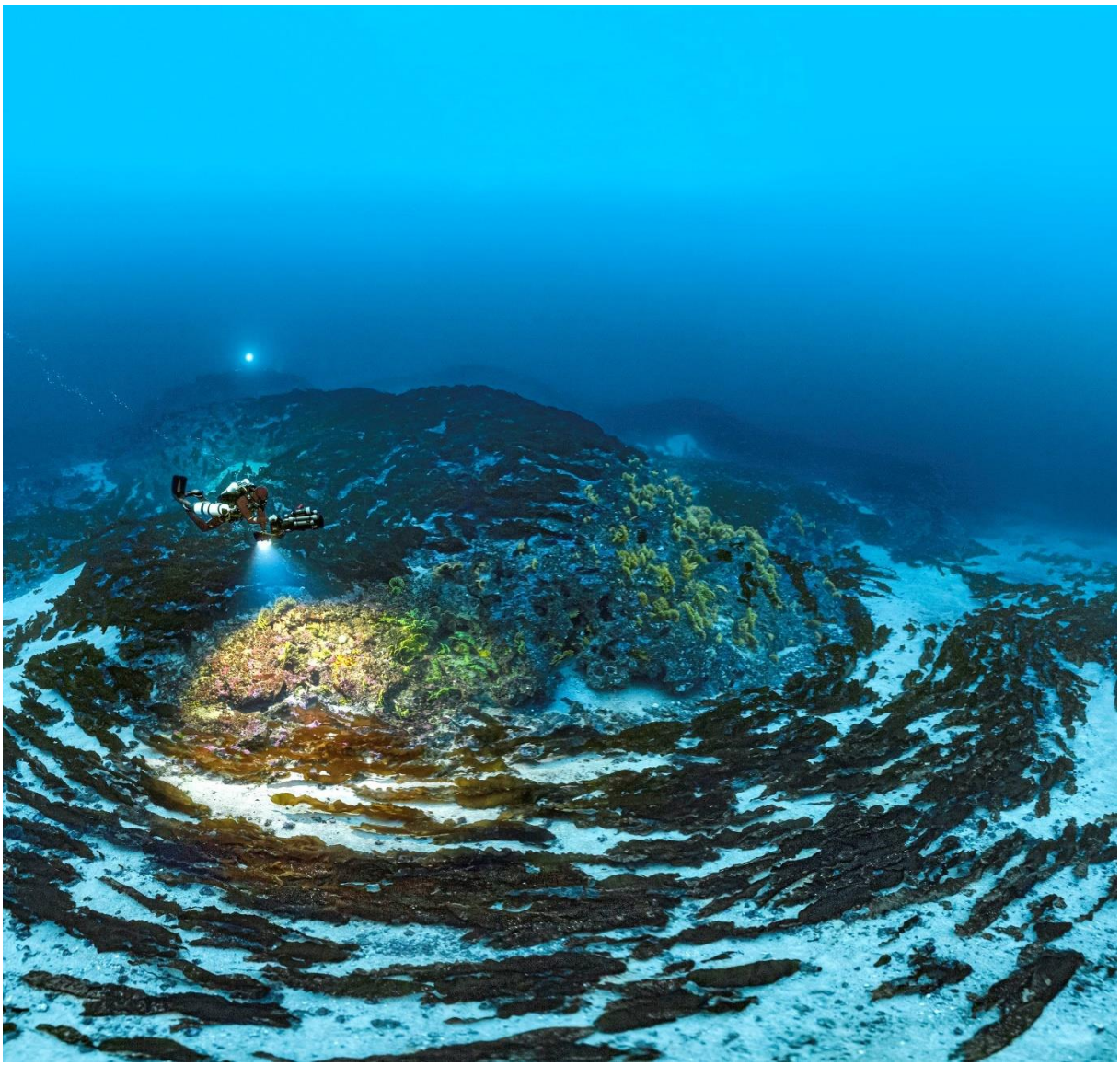
**Références ..... 151**

**Annexes ..... 178**

Annexe A | chapitre 1 ..... 178
Annexe B | chapitre 2 ..... 187
Annexe C | chapitre 3 ..... 195

# Chapitre 1

# Introduction générale



L'une des rares forêts marines à *Laminaria rodriguezii*, répertoriée à 76 mètres de profondeur au Banc du Magaud en Provence (© Laurent Ballesta).

## 1.1 Contexte général de la thèse

L'anthropocène désignant "l'âge de l'Homme", correspond à la proposition d'une nouvelle ère géologique qui est marquée par l'empreinte des activités humaines à une échelle planétaire (Houghton et al., 1990). Cette période correspond ainsi à un changement environnemental global, avec plusieurs composantes, dont le changement climatique.

Face à l'intensification de l'effet de serre qui entraine un réchauffement important de l'atmosphère, l'océan absorbe et stocke la majeure partie de l'énergie reçue. Le réchauffement des océans, la modification de la fréquence et de l'intensité des précipitations, le changement dans l'activité des tempêtes, la réduction de la masse de glace et de neige, l'élévation du niveau et l'acidification des océans sont autant de phénomènes liés au réchauffement climatique (GIEC, 2018). Tous ces phénomènes, dont les vagues de chaleur extrêmes ne sont pas sans impact sur les écosystèmes marins. Les populations d'organismes sessiles de la Grande Barrière de Corail et de la Méditerranée sont déjà impactées (Garrabou et al., 2009 ; Hughes et al., 2018). Le changement climatique n'est cependant pas la seule, ni dans beaucoup de cas la principale, pression anthropique qui affecte la biodiversité marine et le fonctionnement des écosystèmes. Les écosystèmes marins sont en effet impactés par la surexploitation, la pollution, la destruction, l'artificialisation des habitats et l'introduction d'espèces exotiques (Duarte, 2014 ; Rius et al., 2015 ; Alter et al., 2021).

Ces pressions anthropiques vont, seules ou en interaction, avoir des conséquences écologiques et démographiques majeures, par exemple en termes d'effectifs démographiques, d'aire de répartition ou de phénologie, avec des conséquences potentielles à plusieurs échelles écosystémiques (voir par exemple Walther et al. (2002) dans le cas du changement climatique). De tels changements vont aussi avoir des impacts sur les sociétés humaines, notamment en termes de niveau de la mer ou directement affectant la productivité des pêches (GIEC, 2018).

Dans ce contexte de changement global, une question majeure consiste à envisager quels sont les avenirs possibles de la biodiversité. La biodiversité, est une mesure complexe, intégrant plusieurs échelles : évolutive (la diversité des gènes, des espèces, des phylums etc.), fonctionnelle (diversité des fonctions : producteurs primaires, herbivores etc.), organisationnelle (diversité des habitats : fond meuble, herbier, forêt marines etc.), d'abondance

(dominance, rareté) et spatiale (diversité au sein d'un échantillon, d'une région, d'un continent etc.). C'est un processus, un état continuellement changeant, et la question est donc de savoir quelles seront les trajectoires prises par cette biodiversité dans un contexte de changement global. Ainsi dans le cas du changement climatique, plusieurs réponses possibles sont classiquement décrites à l'échelle des espèces ou des populations : extinction locale, changement d'aire de répartition, adaptation génétique, acclimatation, ou une combinaison de plusieurs de ces mécanismes (Parmesan, 2006). Le potentiel évolutif correspondra à la capacité de persistance, voire d'adaptation d'une population ou d'une espèce face à un changement environnemental. Afin d'étudier ces questions, il est nécessaire de combiner des approches écologiques et évolutives, notamment en ciblant des espèces écologiques clés, telles que les espèces ingénieurs d'écosystèmes ou clés de voûte (voir ci-dessous). Ce type d'approche permet de mieux étudier ces questions comme cela a été le cas dans le cadre du projet européen Biodiversà - MarFor auquel j'ai participé.

### 1.1.1 Le projet "Marine Forests" (MarFor)

***Les algues brunes***

Les algues brunes (Phaeophyceae) correspondent à une classe d'un peu plus de 2 000 espèces décrites (Guiry & Guiry, 2021), multicellulaires et photosynthétiques. L'origine des Phaeophyceae résulterait d'une radiation évolutive qui se serait produite au sein des Heterokonta (Straménopiles), il y a environ 310 millions d'années (Brown & Sorhannus, 2010). Les algues brunes constituent un groupe très diversifié avec peu de caractéristiques communes entre les principaux ordres (Fritsch, 1935). Elles possèdent, entre autres, une grande diversité d'organisations, de cycles de vie, de modes de reproduction, et d'exigences écologiques. Les approches phylogénétiques multilocus sont parvenues à un consensus selon lequel la diversification actuelle des algues brunes est associée à deux événements évolutifs majeurs (Bringloe et al., 2020). Le premier datant du milieu du Mésozoïque donna naissance au clade SSDO (Figure 1.1) pour Sphacelariales, Syringodermatales, Dictyotales, Onslowiales, celui-ci étant majoritairement représenté par l'ordre des Dictyotales. Le deuxième, plus tardif dans l'histoire des temps géologiques, correspond à la diversification au sein du clade nommé BACR pour "Brown Algal Crown Radiation" (Figure 1.1) (Silberfeld & Leigh, 2010 ; Bringloe et al., 2020).

Les relations phylogénétiques au sein de ce deuxième clade sont encore partiellement non résolues. C'est au sein du clade BACR que l'on retrouve les Laminariales et les Fucales, deux ordres qui constituent des écosystèmes de haute valeur écologique et socio-économique (Dayton, 1985 ; Steneck et al., 2002 ; Mineur et al., 2015 ; Wernberg et al., 2019).

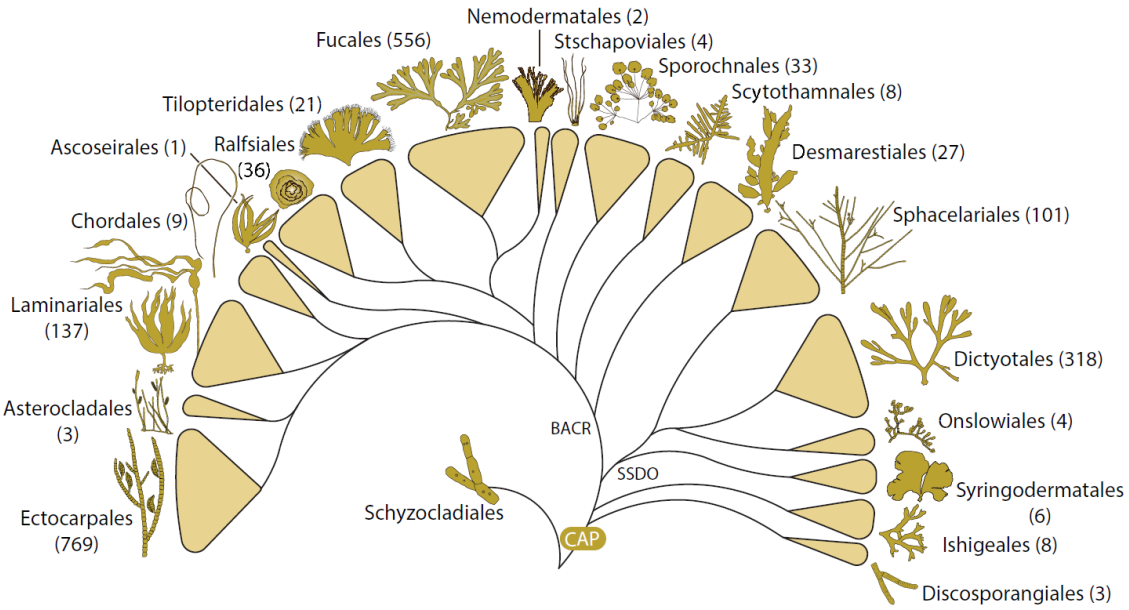



**Figure 1.1**. Classification phylogénétique des algues brunes (Phaeophyceae) : consensus obtenu à l'aide de 12 marqueurs génétiques (nucléaires, mitochondriaux et chloroplastiques). La présente figure est issue de Bringloe et al. (2020). Les parenthèses indiquent le nombre d'espèces décrites et acceptées pour chaque ordre. La ramification des lignées SSDO et BACR, et leur ancêtre commun (CAP) sont indiqués. Les deux espèces étudiées dans le cadre de cette thèse appartiennent à l'ordre des Fucales (*Ericaria zosteroides* (C.Agardh) Molinari & Guiry) et des Laminariales (*Laminaria rodriguezii* Bornet).

Le terme "forêts sous-marines" est ici employé pour désigner des écosystèmes marins dont la structure tridimensionnelle est générée par de grandes algues brunes (Wernberg & Filbee-Dexter, 2019), généralement de l'ordre des Laminariales et des Fucales (Figure 1.2). Ces deux ordres illustrent la notion d'espèces « clé de voûte » (« keystone species »), selon Paine (1969) et d'espèce ingénieur (« engineering species ») selon Jones et al. (1994). Cela tient à deux raisons principales : premièrement, parce que les populations qu'elles forment aux échelons inférieurs des réseaux trophiques sont la clé de voûte de la structure des communautés dont font partie bon nombre d'espèces (Mann, 1973 ; Steneck et al., 2002), deuxièmement, parce que leur production primaire et leur biomasse modifient les conditions abiotiques et

biotiques de l'environnement, créant ainsi de nouvelles conditions propices aux développements d'autres organismes (Carpenter & Lodge, 1986). Nous pouvons ainsi faire l'analogie avec les forêts tropicales terrestres : des écosystèmes parmi les plus productifs et diversifiés de notre planète abritant une multitude d'espèces.

Au cours du siècle écoulé, notamment depuis la fin de la Seconde Guerre Mondiale, les forêts sous-marines ont été particulièrement exposées aux impacts, potentiellement cumulatifs, du changement global. Les plus importantes pressions affectant les forêts marines sont la destruction de l'habitat par des engins de pêche, la pollution chimique, l'augmentation de la turbidité, le réchauffement des océans, le surpâturage par les herbivores (oursins et téléostéens) et la compétition avec des espèces invasives (Sivertsen, 1997 ; Thibaut et al., 2005 ; Norderhaug & Christie, 2009 ; Bartsch et al., 2016 ; Žuljević et al., 2016). Le constat qui en résulte est alarmant : le déclin des forêts sous-marines constitue une tendance mondiale (Steneck et al., 2002 ; Wernberg et al., 2019). Toutefois, la réponse de ces écosystèmes au changement global actuel est hautement variable selon les espèces et les populations considérées, mais également selon les causes du déclin (Araújo et al., 2016 ; Krumhansl et al., 2016). Par exemple, l'augmentation de la température de surface des océans serait l'un des principaux facteurs pour expliquer une baisse de l'abondance des populations de *Saccorhiza polyschides*, *Laminaria hyperborea* et *Laminaria ochroleuca* aux basses latitudes (sud-ouest de la péninsule Ibérique). Elle serait également l'une des causes à l'origine du déplacement progressif des populations intertidales vers des habitats subtidaux en ce qui concerne *Saccharina latissima* (Araújo et al., 2016). Néanmoins, la hausse des températures favoriserait la colonisation d'habitats en milieu boréal (Spitsbergen, Svalbard) qui jusqu'à présent étaient inhospitalier pour *Laminaria digitata* (Bartsch et al., 2016). Cette absence de réponse uniforme souligne, une seconde fois, la grande diversité des forêts marines et la difficulté à évaluer leurs états actuels et à prévoir leurs évolutions.

***Organisation et objectifs du projet MarFor***

Le projet européen Biodiversà Marine Forests (MarFor ; http://marfor.eu/) qui a débuté en 2017, a été développé dans ce contexte. L'essence même de ce projet a été de s'interroger sur l'évolution passée et la diversité actuelle des populations de kelps et de Fucales,

pour envisager les conséquences possibles du changement global actuel sur ces espèces à l'échelle des côtes européennes. Le terme anglophone kelp est ici employé pour désigner les grandes algues brunes de plusieurs mètres de long de l'ordre des Laminariales.

Le projet MarFor a été coordonné par le Centre des Sciences Marines (CCMAR, Faro, Portugal ; responsable du projet : Ester Serrão) avec une collaboration entre plusieurs pays européens, notamment la France (Station Biologique de Roscoff, le MIO), l'Italie (université de Bologne), l'Espagne (université de Malaga), l'Allemagne (Institut Alfred-Wegener, université de Cologne) ainsi que la Suède (université de Göteborg). En adoptant cette stratégie de partenariat, le projet MarFor a couvert une importante diversité de régions biogéographiques et les espèces caractéristiques de chacune d'entre elles. La finalité étant de porter notre attention sur des Laminariales et des Fucales, avec des différences importantes en termes d'aires de répartition, de niche écologique, de cycle de vie, de mode de reproduction, de potentiel de dispersion et d'adaptation.

Parmi les modèles d'étude sélectionnés, citons entre autres des Fucales, telles que *Ericaria zosteroides* (C.Agardh) Molinari & Guiry (anciennement connue sous *Cystoseira zosteroides*), *Ericaria amentacea* (C.Agardh) Molinari & Guiry (anciennement connue sous *Cystoseira amentacea*). Le projet a aussi considéré des macroalgues dont la rareté et la difficulté d'accès étaient jusqu'à présent un frein à leur étude, notamment des espèces profondes. Le kelp *Laminaria rodriguezii* Bornet, endémique de la mer Méditerranée (généralement au-delà de 70 m de profondeur) en est un bon exemple. Des espèces amphi-atlantiques, telles que *Laminaria digitata* (Hudson) J.V.Lamouroux et *Saccharina latissima* (Linnaeus) C.E.Lane, C.Mayes, Druehl & G.W.Saunders, ou d'autres avec une répartition boréale comme *Laminaria solidungula* J.Agardh, ont aussi pu être étudiées sur une large aire géographique grâce à ce projet. Les modèles d'étude propres à ma thèse seront présentés plus précisément dans la partie 1.4.

Le réseau de partenaires du projet MarFor a ainsi permis d'étudier une diversité d'espèces dans des contextes écologiques et phylogénétiques différents au sein des algues brunes. Ce projet a aussi permis de fédérer une communauté de chercheurs avec des approches à la fois différentes et complémentaires : écophysiologie des organismes (e.g. Diehl & Bischof, 2021), génétique des populations (e.g. Neiva et al., 2020) dont des approches en génomique (e.g. Guzinski et al., 2020 ; le présent travail) et en modélisation spatiale (e.g. Assis et al.,

2018a). Toutes ces disciplines ne sont pas hermétiques les unes aux autres, mais au contraire elles interagissent dans un cadre commun, par exemple entre génétique et physiologie des populations (e.g. Kinnby et al., 2020 ; Liesner et al., 2020), ce qui permet notamment d'étudier la réponse à des pressions de sélections divergentes (Kinnby et al., 2020). La combinaison d'approches de génétique des populations et de modélisation des courants marins (e.g. Buonomo et al., 2017 ; Nicastro et al. 2020), ou de modélisation de l'aire de distribution des espèces (e.g. Assis et al., 2018b) est utile pour envisager la dynamique spatiale et temporelle des populations.

***Le cas des forêts marines méditerranéennes***

Dans le cadre du projet MarFor, ma thèse vise à étudier deux espèces endémiques à la Méditerranée : la Fucales *Ericaria zosteroides* et le kelp *Laminaria rodriguezii*. L'accent mis sur la mer Méditerranée se justifie par l'impact du changement global sur ce bassin (Coll et al., 2012). La Méditerranée est l'une des régions du monde où les effets du changement climatique sont déjà visibles et alarmants (Garrabou et al., 2009 ; Rivetti et al., 2014). Par exemple, entre les années 1978 et 2003, la température de surface de la Méditerranée a augmenté de 1.4 °C (moyenne annuelle), ce qui est quatre fois supérieur à la moyenne mondiale (Belkin, 2009). Ces conditions sont propices à la tropicalisation progressive des communautés, favorisant notamment l'installation d'espèces lessepsiennes : de redoutables compétiteurs et prédateurs des forêts de Fucales (voir par exemple Lejeusne et al., 2010). La situation actuelle des forêts marines méditerranéennes est critique : les données disponibles témoignent d'une régression générale, voire d'extinctions locales pour ces trois espèces par exemple : *Laminaria rodriguezii* en mer Adriatique (Žuljević et al., 2016) et *Laminaria ochroleuca* et *Saccorhiza polyschides* dans le détroit de Messine (Araújo et al., 2016). Néanmoins, un manque important ou une absence de données scientifiques sur la distribution des Laminariales en Méditerranée (Bartsch et al., 2008 ; Araújo et al., 2016) limitent l'appréciation de leur état actuel de conservation. En revanche, la situation des Fucales est bien connue : les données actuelles font part d'une régression généralisée de la richesse spécifique et de l'abondance des populations sur le pourtour du bassin méditerranéen occidental (Thibaut et al., 2005 ; Thibaut et al., 2015 ; Blanfuné et al., 2016 ; Mancuso et al, 2018). Cependant, des signes de résiliences (Iveša et al., 2016 ; Blanfuné et al., 2019), voire de stabilité sont perçus chez certaines espèces lorsque les

mesures de conservation sont effectives dans les aires marines protégées (Thibaut et al., 2016a). Les causes de ce déclin, souvent multiformes et cumulatives, seront abordées dans la partie dédiée à la présentation des modèles d'études.




**Figure 1.2**. Quelques illustrations de forêts marines de Laminariales et Fucales des milieux infralittoraux, sublittoraux et profonds des côtes européennes. (A) une forêt de Fucales à *Ericaria zosteroides* (© Sandrine Ruitton), (B) de Laminariales à *Laminaria hyperborea* et *Undaria pinnatifida* (© Yann Fontana). *U. pinnatifida* est une espèce non-autochtone et invasive. (C) une forêt de Laminariales à *Laminaria digitata* de la zone infralittorale (© Christophe Destombe) (D) et de la zone circalittorale pour *Laminaria rodriguezii* (95 mètres de profondeur) (© Laurent Ballesta). (E) une forêt de Fucales à *Ericaria mediterranea* (© Lauric Reynes).

### 1.1.2 Écologie évolutive et changement global

L'étude des conséquences du changement global sur la biodiversité nécessite d'intégrer écologie et biologie de l'évolution. Cette intégration peut se décliner selon diverses approches et niveaux d'étude, avec un fort intérêt en gestion et conservation.

***Diversité génétique, extinction et adaptation :***

La diversité génétique est essentielle à l'adaptabilité des populations, et à la survie à long terme des espèces face à des environnements toujours fluctuants (Bijlsma & Loeschcke, 2012). D'une part, un niveau élevé de diversité génétique peut fournir une assurance évolutive à moyen ou long terme, par la présence d'allèles potentiellement avantageux, d'autre part des facteurs génétiques, incluant la perte de diversité, mais aussi la dépression de consanguinité, peuvent être source d'extinction (Frankham, 2005). Cependant le risque d'extinction dépend de caractéristiques qui sont propres aux espèces considérées, et la diversité génomique n'est pas forcément bien corrélée à ce risque (Teixeira & Huber, 2021). Par exemple, certaines espèces ont par nature un faible niveau de diversité génétique, celui-ci étant la conséquence de leur histoire démographique, notamment associée à un goulot d'étranglement passé (voir par exemple Zhou et al., 2013). Dans ces circonstances, la diversité génétique ne reflète guère le degré actuel des pressions anthropiques et le risque d'extinction des espèces (Teixeira & Huber, 2021). Par ailleurs, des populations dont la taille efficace est réduite ne seraient pas systématiquement exposées à la dépression de consanguinité, et inversement pour d'autres de grandes tailles (Teixeira & Huber, 2021). À plus court terme, le rôle de la diversité génétique a été suggéré dans la réponse immédiate (au sein d'une même génération) des populations face à une perturbation (Hughes & Stachowicz, 2004). Cependant les mécanismes sous-jacents et la généralité de ce type d'effet restent à préciser (voir par exemple Roger et al., 2012).

Malgré les incertitudes préalablement soulevées, et par une approche de précaution, la conservation et le suivi de la diversité génétique devraient être intégré aux politiques de gestion de la biodiversité (Laikre et al., 2020). Cependant, bien que les développements technologiques permettent des études à une échelle génomique sur des espèces non modèles, ce niveau de diversité n'est souvent qu'incomplètement pris en compte.

***Connectivité et flux de gènes :***

Les échanges entre populations (ou dispersion) jouent un rôle important dans les taux de croissance, les flux de gènes, et même la persistance des métapopulations voire des espèces (Lowe & Allendorf, 2010). On distingue classiquement la connectivité démographique pour laquelle les taux de croissance à l'échelle des populations sont affectés par les taux de migration entre populations voisines. La connectivité génétique correspond à l'impact des flux de gènes sur les processus évolutifs au sein des populations (Lowe & Allendord, 2010). Pour qu'il y ait flux de gènes il faut nécessairement une étape de dispersion entre populations (passage d'une larve, propagule ou adulte d'une population à une autre), ainsi qu'une contribution des migrants à la reproduction dans la population d'arrivée. Les inférences sur la connectivité entre populations vont aider à comprendre leur évolution (par exemple en termes de diversité), mais elles peuvent aussi être utilisées dans la gestion par exemple pour la définition d'aires marines protégées (Palumbi, 2003).

En milieu marin, la dispersion entre populations est difficile à analyser directement, du fait de l'ouverture du milieu, des distances en jeu et souvent de la toute petite taille des propagules, comme par exemple les larves et spores. L'utilisation des données de génétique des populations peut permettre d'estimer les flux de gènes entre populations par des approches directes et indirectes (Lowe & Allendorf, 2010). Les approches directes visent à assigner des individus à leur population d'origine à partir de leur génotype multilocus et des fréquences alléliques des populations testées. En ce sens, ces méthodes correspondent à des estimations de flux de gènes à court terme. Ces approches peuvent être aussi appliquées à des données de génomique des populations (Gosselin et al., 2016). Les approches indirectes utilisent les données de différenciation génétique entre populations pour estimer les flux de gènes en lien avec un modèle sous-jacent d'évolution, généralement d'interaction migration / dérive. Plusieurs méthodes existent maintenant pour de telles estimations utilisant différentes statistiques descriptives des données comme le spectre de fréquence allélique (voir par exemple : Gutenkunst et al., 2010 ; Prada & Hellberg, 2021). À l'inverse, l'analyse de la structure génétique peut être utilisée pour détecter des barrières aux flux de gènes (Cahill et al., 2017). En milieu marin, les inférences liées à la structure génétique ou aux flux de gènes peuvent être confrontées à la modélisation océanographique : ceci peut permettre de tester le rôle des courants dans la dispersion efficace (White et al., 2010 ; Alberto et al., 2011 ; Benestan et al.,

2016 ; Durrant et al., 2018 ; Nicastro et al., 2020). Les interprétations qui en découlent doivent cependant tenir compte de l'histoire des populations (Bierne et al., 2011).

***Modes de reproduction :***

Au même titre que les flux de gènes et la dérive génétique, le mode de reproduction a un impact important sur la diversité génétique des populations et indirectement sur le potentiel adaptatif (Ellegren & Galtier, 2016). Le mode de reproduction définit la manière dont sont transmis les gènes d'une génération à la suivante. Ceux-ci peuvent être transmis par le croisement entre des individus non-apparentés (« outcrossing »), des individus apparentés (« inbreeding »), par autofécondation (« selfing »). Par ailleurs, le matériel génétique peut être transmis sans la fécondation des gamètes : la reproduction clonale. Parmi les différents modes de reproduction des kelps et Fucales, l'un d'eux a particulièrement attiré notre attention : la clonalité partielle.

Une population partiellement clonale est une population dont le mode de reproduction fait intervenir une part de reproduction sexuée et une part de reproduction clonale. La reproduction clonale correspond à tout mode de reproduction qui aboutit à une copie génétique de l'organisme parental, aux erreurs de mutation près. La reproduction clonale se fait donc sans recombinaison, et elle peut se faire selon différentes modalités, dont la multiplication végétative et certains types de parthénogénèse (Jaron et al., 2021). En opposition, la reproduction sexuée implique donc un processus de recombinaison au cours du cycle de vie, lors de la méiose : les zygotes produits sont génétiquement uniques.

La clonalité partielle a des répercussions directes sur la diversité génétique des populations (Halkett et al., 2005 ; Duminil et al., 2007). En effet, en l'absence de méiose, et donc de recombinaison, certaines formes de clonalité comme la multiplication végétative favorisent un excès d'hétérozygotes par rapport à la proportion d'hétérozygotes attendue par reproduction panmictique (Birky 1996 ; Judson & Normark, 1996 ; Balloux et al., 2003 ; De Meeûs et al., 2007). Ceci pour deux raisons majeures : premièrement, l'hétérozygotie est conservée plutôt que réduite à chaque génération par l'absence de ségrégation des chromosomes (Judson & Normark, 1996), deuxièmement les mutations qui apparaissent au cours de l'évolution sont conservées à l'état hétérozygote, ce qui favorise la divergence des allèles au sein d'un même individu. Cela peut conduire à un effet Meselson (Balloux et al., 2003 ; De Meeûs et al.,

2007) bien que son importance soit discutée en fonction des organismes et de l'échelle de temps considérée (Jaron et al., 2021). Une des conséquences positive de la clonalité est de limiter la perte de diversité allélique par rapport à une situation de dérive génétique seule avec reproduction sexuée (Stoeckel & Masson, 2014). Toutefois, cette diversité allélique est issue de l'accumulation de mutations, dont certaines pourraient présenter un caractère délétère. En principe, la sélection purifiante tend à maintenir les mutations délétères en très faible fréquence ; or le rôle purifiant de la sélection pourrait être considérablement limité dans une population clonale ; l'effet du cliquet de Muller (Muller, 1964 ; Charlesworth, 2012). L'effet du cliquet de Muller est un argument en faveur de la théorie défendant que l'absence de sexe sur du long terme est une impasse pour l'évolution des espèces (Schwander, 2016).

S'intéresser aux conséquences de la clonalité partielle sur la diversité génétique nécessite en amont d'estimer, avec plus ou moins de certitude, le taux de clonalité, c'est-à-dire la fraction d'individus produit par voie clonale en comparaison à ceux résultant de la reproduction sexuée. Selon le taux de reproduction clonale, et les interactions avec d'autres forces évolutives telles que la dérive génétique et la mutation, la clonalité partielle va modifier la trajectoire évolutive des populations (Stoeckel & Masson, 2014 ; Reichel et al., 2016 ; Stoeckel et al., 2020 ; Krueger-Hadfield et al., 2021).

***Diversité adaptative :***

L'étude de la diversité adaptative, notamment de l'adaptation locale, peut être utile pour établir des projections de l'évolution des populations face au changement climatique (Peterson et al., 2019). Prendre en compte cette diversité est aussi utile en termes de conservation, par exemple en identifiant des cibles pour la protection (Xuereb et al., 2020), ou en orientant des projets de réintroduction. Le résultat de l'adaptation locale correspond à un patron dans lequel des génotypes originaires d'une population ont une valeur sélective plus élevée dans cet habitat que des génotypes venant d'autres habitats (Kawecki & Ebert, 2004). L'étude de l'adaptation locale peut se faire expérimentalement, par exemple par des transplantations croisées (Ledoux et al., 2015). Cependant ce type d'approche peut être difficile pour les organismes longévifs, ou difficile d'accès, comme les espèces profondes. Les outils de génomique des populations permettent de rechercher des signaux d'adaptation locale dans le génome, sans *a priori* (approches de type « bottom-up » ; voir par exemple Sork et al.,

2013). Ce type d'approche nécessite une couverture suffisante du génome (voir le débat sur le sujet : Lowry et al., 2016 ; Catchen et al., 2017 ; Lowry et al., 2017), par des méthodes relevant de la génomique « partielle », comme le séquençage RAD (Baird et al., 2008 ; voir partie dédiée à ces marqueurs).

Les différents points présentés ici, diversité, connectivité et adaptation, ne sont évidemment pas indépendants. La compréhension du niveau de diversité local nécessite de prendre en comptes les interactions flux de gènes / dérive, le mode de reproduction, et la sélection. L'évolution de l'adaptation locale dépend ainsi des flux de gènes entre habitats (Lenormand, 2012). Les analyses de génomique des populations, replacées dans un contexte écologique, permettent d'aborder ces différentes questions, en fonction des espèces étudiées et des échantillonnages disponibles.

## 1.2 Représentation réduite du génome (génotypage RAD-Seq)

### 1.2.1 Description de la méthode

Le principe du séquençage RAD (« Restriction-site Associated DNA sequencing », RAD-Seq), est de sous-échantillonner et de séquencer une partie du génome, en utilisant les sites de coupure d'enzymes de restriction comme « points de repères » pour ce séquençage. Chacun de ces sites, d'une taille de quelques paires de bases est spécifiquement reconnu par une enzyme de restriction (endonucléase), dont le rôle, est de cliver l'ADN en ce site spécifique. Cette méthode permettra de séquencer de courtes séquences adjacentes à ces sites de restriction. Dans la majorité des cas, toutefois voir Dorant et al. (2020) et Tigano (2020), le RAD-seq est employé pour identifier des SNPs ou « Single Nucleotide Polymorphism » ; il s'agit d'une variation d'un nucléotide dans une séquence d'ADN parmi les individus considérés.

Il existe de nombreux sites de restriction enzymatique, et leur fréquence de coupure est relativement variable selon les espèces considérées (Herrera et al., 2015). La bonne conservation de ces sites de coupure entre individus d'une même espèce permettra la comparaison des séquences à une échelle intraspécifique (Herrera et al., 2015). La possibilité de mutations au sein des sites de restriction peut cependant conduire à des pertes d'allèles (« allele drop-out ») qui peut avoir des conséquences sur les analyses de diversité génétique (Gautier et al., 2013). Lorsque la divergence génétique augmente, le nombre de sites partagés diminue,

mais le RAD-seq peut être utilisé dans certains cas, et à certaines échelles de temps, pour des approches de phylogénie (Cruaud et al., 2014, voir ci-dessous).

Le protocole initial du RAD-seq (Baird et al., 2008) a été adapté dans plusieurs variantes à l'image du 2b-RAD (Wang et al., 2012), ezRAD (Toonen et al., 2013) et ddRAD (Peterson et al., 2012). Néanmoins si ces différentes variantes permettent une plus grande souplesse en termes de coût et de temps passé en laboratoire (Davey et al., 2011), leur principe reste donc identique : réduire la part analysée des génomes eucaryotes pour s'affranchir de leur complexité structurelle. À titre d'exemple, la taille du génome nucléaire des algues brunes, pour des taxa où des données sont disponibles, serait comprise entre 166 Mb et 2005 Mb selon les espèces (Kapraun, 2005 ; Gregory et al., 2007). Pour *Laminaria digitata* la taille du génome serait proche de 600 Mb (communication personnelle, Mark Cock). Dans le cadre de cette thèse, j'ai opté pour un protocole ddRAD (Figure 1.3), ceci pour l'étude de la structure et la diversité génétique des populations d'*Ericaria zosteroides*, *Laminaria rodriguezii* et secondairement, *Laminaria digitata*. À la différence du protocole initial, la digestion du génome en ddRAD nécessite l'utilisation de deux endonucléases, une reconnaissant un site de restriction fréquent, l'autre un site de restriction plus rare. Spécifiquement, j'ai opté pour un couple d'endonucléases (*PstI* et *HhaI*) qui s'est avéré efficace antérieurement chez une algue brune, comme *Ectocarpus* sp. (Avia et al., 2017). Du fait de l'importante quantité de données obtenues par le séquençage à haut débit, l'analyse des séquences dérivées du RAD-seq nécessite un important travail bio-informatique. Ces étapes d'analyse comprennent notamment i) le contrôle qualité des séquences et la suppression des adaptateurs de séquençage, ii) le démultiplexage des librairies à l'aide de barcordes et d'index dont la complémentarité assure la correspondance à chaque échantillon iii) la vérification de l'architecture des libraires, notamment si les séquences R1 et R2 commencent par les motifs de restriction enzymatique attendue (e.g. CTGCA/G pour *PstI* et GCG/C pour *HhaI*), iv) l'assemblage des séquences en locus homologue par le biais de pipelines bio-informatiques, les plus employés étant Stacks (Catchen et al., 2011, 2013 ; Rochette et al., 2019) et PyRAD (Eaton, 2014), et finalement v) le génotypage des SNPs et la mise en forme des données pour des analyses ultérieures.

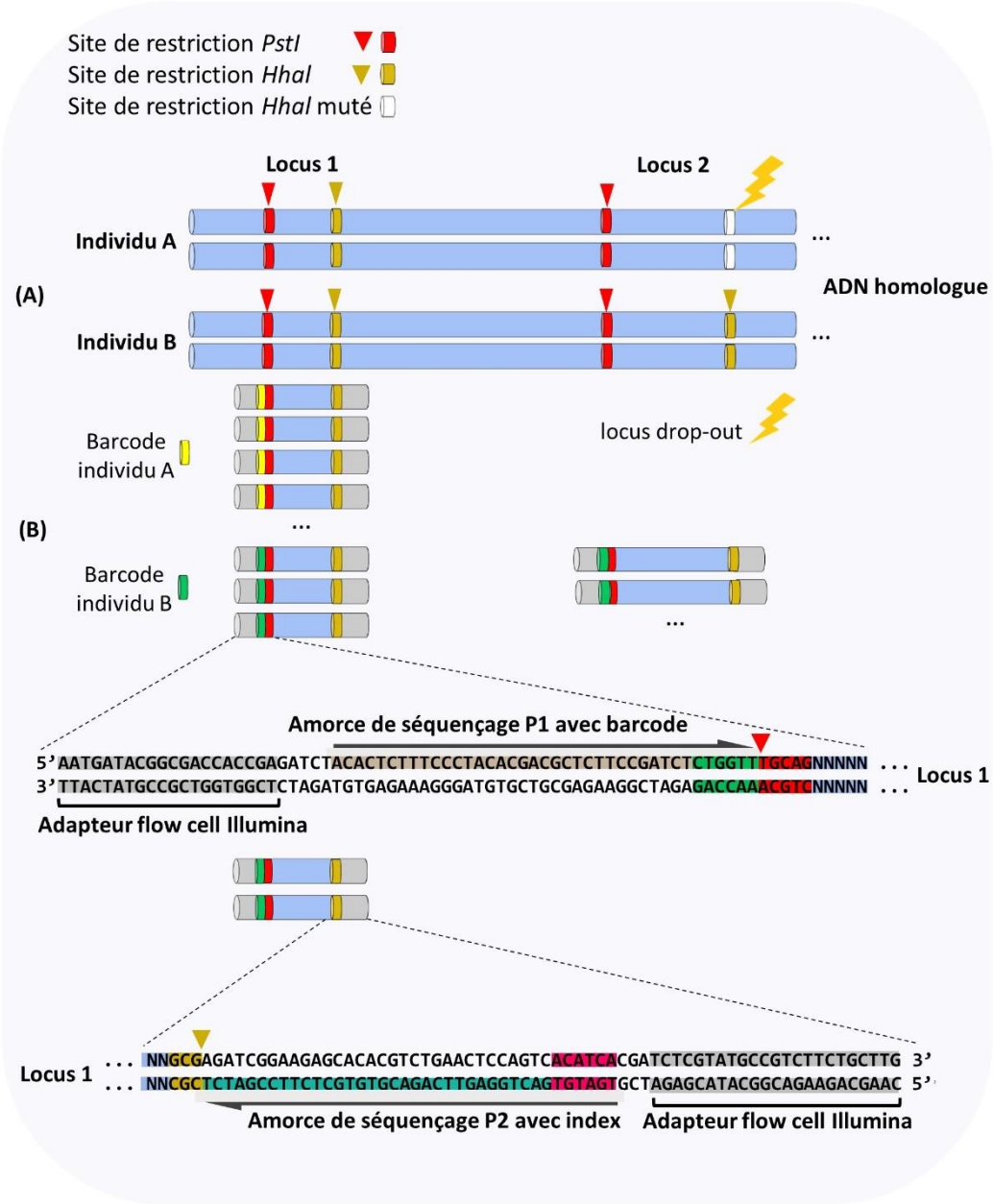



**Figure 1.3**. Représentation schématique du séquençage ddRAD-seq (double-digest) employé pour l'analyse de la diversité et la structure génétique des populations étudiées. (A) L'ADN génomique est digéré par le couple d'enzymes *PstI* et *HhaI*, deux locus sont concernés pour cet exemple. Le locus 1 est présent dans la librairie de séquençage pour les individus A et B alors que le locus 2 est absent chez l'individu A en raison d'une mutation dans le site de reconnaissance de *HhaI* ("locus drop-out") de celui-ci. A titre indicatif, il est davantage probable qu'une telle mutation survienne sur l'un des deux chromosomes homologues ("allele drop-out"). (B), la librairie de séquençage ddRAD-seq est construite de sorte à séquencer chaque extrémité d'un locus d'intérêt sur 150 pb à partir des amorces de séquençage Illumina P1 (extrémité *PstI*), incluant un barcorde individuel (6-12 nt) et P2 (extrémité *HhaI*), incluant un index (6 nt). Finalement, la librairie a été séquencé par un séquenceur Illumina HiSeq 4000. Pour davantage de précision sur le protocole utilisé se référer à l'Annexe A, Matériel A.1.

### 1.2.2 Applications du RAD-seq à différentes échelles d'étude

***Phylogénie et délimitation d'espèces***

À l'échelle interspécifique, le RAD-seq a été appliqué avec succès pour établir les relations phylogénétiques comme chez différentes familles de cétacés (14-19 Ma, Viricel et al., 2014), ou de coléoptères du genre *Carabus* (1,2-17 Ma, Cruaud et al., 2014). À titre d'information, Cruaud et al. (2014) générèrent par RAD-seq un volume d'information 400 fois supérieur à celui obtenu par le séquençage Sanger de six locus nucléaires et de trois locus mitochondriaux. Ceci leur a permis d'aboutir à une phylogénie robuste et résolue du genre *Carabus*. Néanmoins, Cruaud et al. (2014) démontrèrent que le nombre de locus homologues s'amenuise entre les taxa à mesure que leur distance évolutive augmente, illustrant ainsi la principale limite du RAD-seq pour résoudre des phylogénies profondes (Lemmon & Lemmon, 2013). En effet, plus les taxa divergent et moins ils sont susceptibles de conserver des sites de restriction enzymatiques en commun. La possibilité de réaliser des analyses phylogénétiques dépendra donc du taux de données manquantes autorisé dans le jeu de données final. Il est probable que cela dépendra aussi du groupe considéré et de son taux d'évolution moléculaire et des données disponibles. Par exemple, Cariou et al. (2013) ont démontré par l'utilisation de marqueurs simulés de type RAD-seq de 12 génomes d'espèces de drosophile, qu'il était possible de réaliser une phylogénie robuste de ces espèces dont la divergence évolutive est comprise entre approximativement 5 et 63 Ma. Cependant, un tri strict sur les données manquantes peut conduire à des jeux de données très, voire trop réduit. A ce sujet, Wagner et al. (2013) ont montré que parmi les 90 000 locus génotypés à partir de 156 individus de 16 espèces de cichlidés du lac Victoria, seuls 23 locus étaient partagés par l'ensemble des individus. Cependant, une réflexion sur les données manquantes, montre qu'accepter un certain niveau de données manquantes peut permettre de résoudre certaines relations phylogénétiques (Eaton et al., 2017).

L'étude de processus biologiques à une échelle intraspécifique, comme la connectivité entre populations, nécessite de procéder en amont une délimitation correcte des espèces (Pante et al., 2015a). Or la délimitation des espèces ne peut reposer que sur un seul locus, et le séquençage RAD fournit une approche puissante pour ces questionnements. Cela a par exemple été utilisé chez les octocoralliaires du genre *Chrysogorgia*, où des hypothèses d'espèces primaires formulées à partir d'haplotypes mitochondriaux du gène *mtMtuS*, ont été

réévaluées par des analyses phylogénétiques de marqueurs RAD-seq (Pante et al., 2015b). En ce qui concerne les kelps et Fucales, objets de cette thèse, les limites d'espèces ne sont pas toujours bien définies, et il est possible d'y découvrir des espèces cryptiques, c'est-à-dire des espèces pour lesquelles les critères morphologiques n'ont pas permis de détecter de différences suffisamment claires pour conclure à la présence d'espèces différentes. C'est notamment le cas de Fucales, comme par exemple *Durvillaea* sp. (Fraser et al., 2010, 2012), *Pelvetiopsis* sp. (Neiva et al., 2017) et de kelps, comme *Lessonia nigrescens* Bory de Saint-Vincent, 1826 (Tellier et al. 2009, 2011 ; González et al., 2012). Par ailleurs, l'adaptation locale à des habitats particuliers peut fortement restreindre les flux de gènes entre des espèces proches et donner lieu à la découverte d'une nouvelle entité évolutive. A ce sujet, les pressions de sélection le long d'un gradient bathymétrique seraient suffisamment fortes pour maintenir les limites d'espèces entre *Fucus vesiculosus* et le complexe *F. spiralis* (Billard et al., 2010 ; Zardi et al., 2011). Cette diversification a par ailleurs donné lieu à la description d'une nouvelle espèce au sein du complexe *F. spiralis*, distinguable sur la base de critères génétiques, physiologiques et morphologiques : *Fucus guiryi* G.I.Zardi, K.R.Nicastro, E.S.Serrão & G.A.Pearson (Zardi et al., 2011).

Dans une situation inverse, la plasticité phénotypique peut générer différents morphotypes qui peuvent conduire à une séparation erronée en différentes espèces. Par exemple, trois espèces morphologiques sont décrites au sein du complexe *Ericaria selaginoides* (Linnaeus) Molinari & Guiry (anciennement nommé *Cystoseira tamariscifolia*). Sur la base de microsatellites, Bermejo et al. (2018) constatèrent que ces espèces morphologiques ne correspondent pas à des groupes génétiques séparés, et ils ont conclu à la présence d'une seule espèce plastique. L'absence de distinction entre *Ericaria mediterranea* et *Ericaria amentacea* pour des marqueurs cytoplasmiques pourrait aller dans ce sens (Jódar-Pérez et al., 2020), mais pourrait aussi correspondre à une divergence récente et une évolution relativement lente des locus mitochondriaux et plastidiques chez ces espèces vicariantes. L'utilisation d'un plus grand nombre de marqueurs génétiques, notamment issu de séquençage RAD, est donc utile ici pour préciser les limites d'espèces.

***Polymorphisme neutre et structure génétique des populations***

À l'échelle intraspécifique, le RAD-seq à l'avantage d'offrir une image détaillée de la structure des populations à des locus neutres. Pour des groupes comme les kelps et les Fucales, en principe génétiquement différenciés à de courtes distances (voir la section 1.3), l'utilisation de microsatellites a donné des résultats concluants, pour par exemple, évaluer l'incidence de la discontinuité de l'habitat sur les flux de gènes (Billot et al., 2003 ; Robvieux, 2013), étudier le rôle des courants marins dans la connectivité génétique (Thibaut et al., 2016b ; Buonomo et al., 2017) ou encore suivre la dynamique d'une espèce invasive (Grulois et al., 2011 ; Salamon et al., 2020). Néanmoins, par le génotypage de centaines voire de milliers de locus, le RAD-seq accroît la précision des mesures de différenciation génétiques et la puissance des tests de différenciation (Waples, 1998). Cependant, ceci ne suffira probablement pas pour détecter de la structure génétique dans le cas d'espèces à forts flux de gènes et tailles efficaces. Dans ce dernier type de situation, l'utilisation de locus outliers (voir ci-dessous) peut permettre de révéler des barrières aux flux de gènes (Gagnaire et al., 2015 ; Robinet et al., 2020). En effet, Robinet et al. (2020) ont montré que l'effet indirect de la sélection aux locus outliers peut compenser le manque de signal à des locus neutres pour lesquels l'équilibre entre les flux de gènes et la dérive génique limite la différenciation génétique.

Pour les kelps et les Fucales, le génotypage de plus d'un millier de SNPs présente un avantage par rapport aux microsatellites, notamment en accentuant la résolution du partitionnement des individus sur la base de leurs génotypes multilocus (Guzinski et al., 2018 ; Guzinski et al., 2020 ; Le Cam et al., 2020). Par exemple, le génotypage de 4069 SNPs par RAD-seq a permis une séparation plus nette des populations du kelp *Saccharina latissima* (Figure 1.4A) en comparaison à celle de constatée par le biais de 18 microsatellites (Figure 1.4B) (Guzinski et al., 2020). Ceci est notamment visible pour la localité d'Helgoland dans le sud-est de la mer du Nord (HEL, Figure 1.4). Toutefois, la différenciation génétique des populations était significative pour les deux types de marqueurs. D'autres comparaisons entre microsatellites et RAD-seq, cette fois-ci conduites sur le kelp invasif *Undaria pinnatifida* ont rapporté que les patrons de diversité et de structure génétiques étaient similaires pour les deux types de marqueurs ; néanmoins, les microsatellites à l'inverse du RAD-seq étaient moins puissants pour détecter des différences génétiques à une échelle fine (Guzinski et al., 2018).

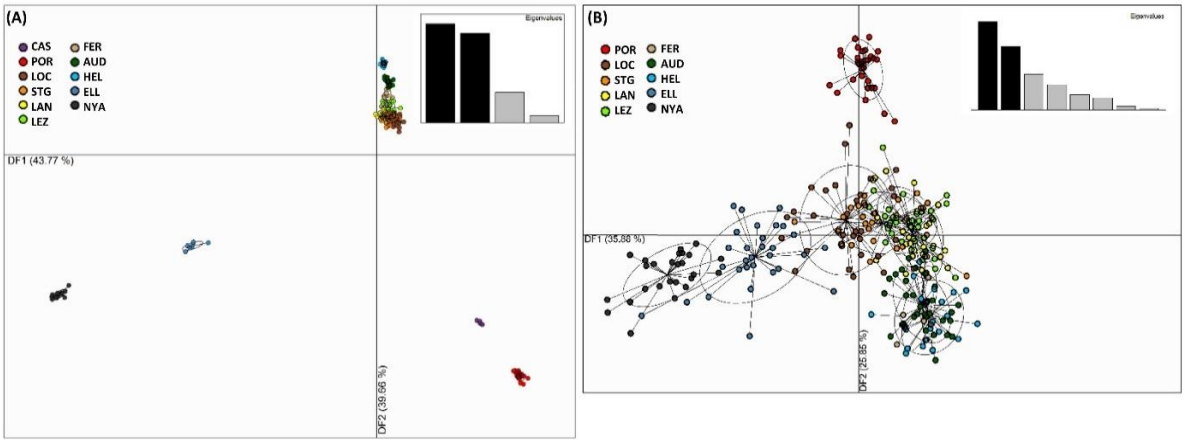



**Figure 1.4**. Axes 1 et 2 de l'Analyse Discriminante en Composantes Principales (Jombart et al., 2010) des populations de *Saccharina latissima*, (A) avec le génotypage de 4069 SNPs par RAD-seq parmi 11 localités, et (B) de 18 microsatellites parmi 10 localités (Guzinski et al., 2020). Chaque point représente un individu. Les individus sont regroupés par localité (voir légende dans l'angle supérieur gauche) mais ces dernières sont davantage séparées par RAD-seq.

***Détection des locus outliers et adaptation locale***

Le nombre de locus généré par RAD-seq est propice à la détection de régions génomiques impliquées dans l'adaptation locale (Catchen et al., 2017 ; McKinney et al., 2017a). Notamment, en multipliant le nombre de locus, on augmente la probabilité de détecter des régions influencées par la sélection. Selon le type de sélection considéré, ces régions du génome peuvent être caractérisées par un niveau de différenciation anormalement élevé (sélection directionnelle) ou faible (sélection balancée) par rapport à la distribution observée sur l'ensemble du génome. De tels locus sont qualifiés de locus déviants ("outlier loci"). Généralement, les locus outliers ne sont pas les cibles directes de la sélection, mais ils sont liés aux locus sous sélection ("hitch-hiking effect" ; Smith & Haigh, 1974). Néanmoins, les méthodes de recherche de locus sous sélection peuvent donner lieu à des biais, comme la détection de faux positifs (voir par exemple Narum & Hess, 2011 ; Bierne et al., 2013 ; Hoban et al., 2016 ; Pratlong et al., 2018 ; Booker et al., 2020). Enfin, la détection des locus outliers peut ne pas être directement liée à l'adaptation locale, mais au couplage d'incompatibilités génétiques avec des barrières environnementales (Bierne et al., 2011). Malgré ces limites qui, pour la plupart, ne sont pas intrinsèques à la méthode, le RAD-seq a permis d'ouvrir l'accès à l'étude de l'adaptation locale à une grande diversité d'organismes marins. À l'origine sur l'épinoche, une espèce de téléostéens sédentaires des milieux marins et lacustres (*Gasterosteus*

*aculeatus*, Hohenlohe et al., 2010), puis pour d'autres téléostéens comme la morue du pacifique (*Gadus macrocephalus*, Drinan et al., 2018), des décapodes comme le homard américain (*Homarus americanus,* Benestan et al., 2016), des octocoralliaires comme le corail rouge (*Corallium rubrum,* Pratlong et al., 2018), des bivalves comme les coquilles Saint-Jacques (*Pecten maximus* et *Pecten jacobeus*, Vendrami et al., 2019), la coque commune (*Cerastoderma edule*, Coscia et al., 2020) ou encore chez les algues brunes (*Ectocarpus* sp., Avia et al., 2017), dont des kelps (*Undaria pinnatifida*, Guzinski et al., 2018 ; *Saccharina latissima*, Guzinski et al., 2020) et des Fucales (*Fucus vesiculosus*, Kinnby et al., 2020).

## 1.3 Structure et diversité génétique des Laminariales et Fucales

### 1.3.1 Echelles de différenciation génétique

Les études antérieures ont montré que les populations de kelps et de Fucales présentent souvent une forte structure génétique. Une méta-analyse, incluant entre autres, 48 études conduites sur ces taxa a rapporté, un $F_{ST}$ moyen de 0,18 (± 0,19 SD par étude) pour une échelle spatiale de 390 km (± 602 SD par étude) (voir Annexe 2 dans Durrant et al., 2014). L'index $F_{ST}$ est une mesure la différenciation génétique des populations, il peut être calculé par paires de populations et sur l'ensemble.

En l'occurrence un $F_{ST}$ de 0,18 indique que les populations sont génétiquement différenciées. Néanmoins, cette valeur moyenne est à interpréter avec précaution car elle a été obtenue à partir d'études dont l'échelle spatiale considérée était très différente d'une espèce à une autre. Par exemple, pour le kelp *Postelsia palmaeformis*, le $F_{ST}$ a été considéré à l'échelle de quelques mètres (Kusumo et al., 2006) alors que pour la Fucales *Fucus serratus* celui-ci est apprécié entre des populations distantes de plus d'un millier de kilomètres (Coyer et al., 2006). Un tel niveau de contraste est imprécis pour établir des conclusions sur les échelles de différenciation génétique, notamment car la variance du $F_{ST}$ peut être dépendante de l'échelle d'étude. Cela est d'autant plus contraignant quand la structure génétique des espèces est décrite par un modèle d'isolement par la distance (IBD, Wright, 1943 ; Rousset, 1997). Le modèle d'IBD implique la différenciation génétique augmente avec les distances géographiques. Bien que ce patron soit fréquemment rapporté chez les kelps et Fucales (Billot et al., 2013 ; Tatarenkov et al., 2007 ; Durrant et al., 2014 ; Neiva et al., 2020) il n'est pas décrit de manière

systématique (voir par exemple Grulois et al., 2011 ; Thibaut et al., 2016b ; Guzinski et al., 2018).

Dans un scénario où la différenciation génétique augmente avec les distances géographiques, cela implique que les valeurs du $F_{ST}$ peuvent-être bornées par l'échelle d'études. Par exemple, quand la structure des populations du kelp *Laminaria hyperborea* a été étudiée par l'intermédiaire d'un échantillonnage hiérarchique, celui-ci incluant, entre autres, une distance de 0.3 km entre les sites (Robuchon et al., 2014), 66% des $F_{ST}$ étaient significatifs, contre 89% quand la même espèce a été étudiée à l'échelle d'une centaine de kilomètres (Evankow et al., 2019). Outre l'échelle d'étude considérée, plusieurs facteurs, non exclusifs peuvent être à l'origine de ces différences. Le premier connu comme « genetic sampling » résulte du fait que la dérive génétique tend à différencier les populations et que la distribution des fréquences alléliques qui en résulte est partiellement couverte par l'échantillonnage (Holsinger & Weir, 2009). Par conséquent, si deux échantillons sont tirés de cette distribution, leurs fréquences alléliques et le $F_{ST}$ associé peuvent être très différents. Par ailleurs, ce ne sont pas les mêmes microsatellites employés entre ces deux études. La distinction du $F_{ST}$ pourrait provenir du degré de polymorphisme des microsatellites, qui dans le second cas serait davantage propice pour détecter des différences génétiques significatives. Dans cet exemple, les populations sont dans un cas étudié dans le nord-ouest de la France (Rochuchon et al., 2014) et dans un autre le long la côte norvégienne (Evankow et al., 2019). La variation du $F_{ST}$ pourrait donc refléter une différence de diversité génétique à l'intérieur des populations, celle-ci étant façonnée, entre autres, par l'histoire biogéographique des régions étudiées (voir ci-dessous).

***Synthèse de la différenciation génétique***

Afin de compléter les synthèses existantes de la structure génétique des kelps (Valero et al., 2011) et plus généralement des macroalgues (Durrant et al., 2014), j'ai dressé un état des lieux de la différenciation génétique des populations de kelps et Fucales des côtes européennes. Trois échelles spatiales ont été considérées : une échelle très fine de l'ordre de quelques mètres (< 1km), une échelle locale (1-10 km) et régionale (≥ 10 km). La différenciation génétique a été appréciée par le $F_{ST}$ et quand l'information était disponible, j'ai établi à partir de cet index le pourcentage de comparaison significative par étude et pour chaque espèce. Cette analyse a été limitée à des marqueurs microsatellites et des SNPs, des marqueurs

comme les RFLP et les AFLP n'ont pas été considérés. De même, des espèces non-autochtones introduites au siècle actuel sur la côte européenne comme *Undaria pinnatifida* et *Sargassum muticum* ne sont pas considérées dans cette analyse car leurs différenciations génétiques sont étroitement associées à leurs dynamiques d'invasion (voir par exemple Guzinski et al., 2018).

Il ressort de cette analyse que la structure génétique des kelps autochtones est en principe peu étudiée à l'échelle locale (1-10 km) ; or quand celle-ci était évaluée, certaines populations de *L. digitata*, *L. hyperborea* et *S. latissima* étaient différenciées (Tableau 1.1). Dans ces circonstances, le pourcentage de $F_{ST}$ significatifs est compris entre 66% et 100%, ce qui témoigne pour ces espèces d'une structure génétique à courte distance. En ce qui concerne les Fucales, la très grande majorité des études ont montré qu'il existait des différences significatives à l'échelle locale (1-10 km) (Tableau 1.2) et cette différenciation est parfois même constatée à une distance de quelques mètres. C'est par exemple le cas des populations de *Fucus vesiculosus* (1 m ; Tatarenkov et al., 2007), *Ericaria amentacea* (10 m ; Buonomo et al., 2017) et *Fucus serratus* (~ 500 m ; Jueterbock et al., 2018) pour lesquelles 100% des $F_{ST}$ étaient significatifs.

***Potentiel de dispersion et différenciation génétique***

L'amplitude de la différenciation génétique résulte de la balance entre la dérive génétique qui tend à différencier les populations d'une part, et les flux de gènes qui tendent à homogénéiser la diversité génétique au sein d'une espèce d'autre part. Pour des espèces disposant d'un faible potentiel de dispersion, comme les kelps et les Fucales (Santelices, 1990 ; Kinlan & Gaines, 2003), les flux de gènes sont en principe restreints à courte distance. Par exemple, la distance de dispersion moyenne pour des Fucales comme *Cystoseira compressa* et *Ericaria amentacea* est de l'ordre de 20 à 40 cm (Mangialajo et al., 2012). Pour des kelps comme *Undaria pinnatifida*, la dispersion des spores serait limitée à approximativement 10 mètres (Grulois, 2010) et jusqu'à 200 m pour, par exemple, *Laminaria hyperborea* (Fredriksen et al., 1995). Il n'est donc pas étonnant de constater chez ces espèces une structure génétique à courte distance. Cette dispersion très localisée s'explique par la formation en quelques heures d'une membrane de fixation, d'une durée de vie très courte des propagules (spores ou zygotes) et par leur faible capacité à être maintenus en suspension dans la colonne d'eau. Ceci implique que le recrutement a préférentiellement lieu dans ou à quelques mètres de la

population source. Cela est par exemple constaté pour les kelps *Macrocystis pyrifera* et *Pterygophora californica*, deux espèces dont la viabilité des spores est de l'ordre de 32 heures (Reed et al., 1988). Par ailleurs des évènements occasionnels, tels que des tempêtes, peuvent participer à la remise en suspension des spores dans la colonne d'eau. Dans ces conditions, l'hydrodynamisme local est susceptible de faciliter la dispersion des spores à de plus longues distances (Reed et al., 1988). En effet, Reed et al. (1988) constatèrent chez *P. californica* qu'une dispersion asymétrique à longue distance (jusqu'à 4 km) était précédée d'une tempête de forte intensité. Néanmoins, concernant les Fucales, l'émission des gamètes survient lorsque les conditions hydrodynamiques sont calmes, celles-ci ont l'avantage de limiter la dilution des gamètes et de favoriser la syngamie (Serrão et al., 1996) mais aussi de limiter la dispersion des zygotes sur de plus longues distances. Cette stratégie de dispersion implique que la plupart des kelps et Fucales ont une dynamique basée sur un système d'auto-recrutement (voir par exemple Guzinski et al., 2018 ; Salamon et al., 2020), avec un degré de connectivité génétique modéré à très limité entre les populations.

Toutefois, les flux de gènes peuvent être soutenus par le décrochage et la dérive de structures fertiles, telles que les rameaux pour les Fucales et les sporophytes pour les kelps (Thiel & Gutow, 2005). La dérive de ces fragments peut ainsi favoriser une dispersion à plus longue distance que celle des spores et des zygotes. Les principaux facteurs à l'origine du décrochage des structures fertiles sont : le broutage par des herbivores, ceux-ci étant susceptibles de sectionner le crampon ou le stipe des kelps (Duggins et al., 2001), le décrochage naturel des rameaux secondaires durant le cycle de vie de certains Fucales (Arenas et al., 1995), l'hydrodynamisme local, notamment, la conséquence de tempêtes (Duggins et al., 2001). Des facteurs anthropiques, comme l'ancrage des bateaux, les activités de plaisances, telles que la pêche et la plongée sous-marine sont également soupçonnées d'en être la cause (voir par exemple Grulois et al., 2011). Une dispersion par dérive serait considérée comme effective et occasionnelle chez des kelps comme *Saccorhiza polyschides* (Assis et al., 2016), *Laminaria ochroleuca* (Assis et al., 2018b) et *Undaria pinnatifida* (Grulois et al., 2011) et des Fucales telles que *Ericaria amentacea* (Thibaut et al., 2016b ; Buonomo et al., 2017). Néanmoins, le niveau de flux gènes associé à ces évènements serait trop restreint pour contrebalancer la différenciation des populations par dérive génétique.

***Interactions avec le mode de reproduction, l'habitat et le contexte biogéographique***

Le degré de différenciation génétique des kelps et des Fucales est influencé par leur faible potentiel de dispersion mais également par des facteurs tels que le mode de reproduction, la zonation verticale (en fonction de la profondeur), plus généralement l'habitat, des barrières géographiques infranchissables (grande plages) ou encore le contexte biogéographique, notamment climatique et océanographique des populations.

Concernant le mode de reproduction, la différenciation génétique serait davantage limitée chez une espèce monoïque, comme *Fucus spiralis* ou *Fucus guiryi* que pour une espèce dioïque, telle que *Fucus vesiculosus* (Almeida et al., 2017). En effet chez une espèce monoïque (autocompatible), la perte de diversité génétique par autofécondation limiterait la différenciation génétique. Toutefois, la zonation verticale de ces trois espèces pourrait également en être la cause, *F. vesiculosus* étant situé en limite inférieure de la zone intertidale par rapport à *F. spiralis* et *F. guiryi.* En effet, la zonation verticale est un facteur essentiel pour expliquer la structure génétique des populations (Kelly & Palumbi, 2010). Les kelps intertidaux seraient davantage différenciés en comparaison à leurs homologues subtidaux (Valero et al., 2011). Par exemple, la structure des populations est davantage marquée pour le kelp intertidal *L. digitata* que pour le kelp subtidal *L. hyperborea* (Robuchon et al., 2014). L'une des hypothèses pouvant expliquer ces différences est que la dispersion d'une espèce intertidale est uniquement possible lorsque les individus sont immergés, ceci limitant la dispersion des spores.

Par ailleurs, la discontinuité de l'habitat est un facteur important pour expliquer la structure génétique des populations, notamment chez des espèces comme les kelps et les Fucales qui ont un faible potentiel de dispersion. A ce sujet, la différenciation génétique des peuplements continus du kelp *L. digitata* n'est pas significative à une distance inférieure à 10 km dans la partie nord et sud de la Bretagne (Billot et al., 2003). Néanmoins dans la baie de Saint-Malo (nord-ouest de la France) *L. digitata* a une distribution fragmentée : ces populations considérées comme marginales sont isolées d'autres peuplements continus par l'absence d'un habitat favorable. Il en résulte que pour cette baie, la différenciation génétique peut être effective à une centaine de mètres (Robuchon et al., 2014) avec un $F_{ST}$ très supérieur sur le plan spatial (Billot et al., 2003 ; Couceiro et al., 2013) et temporel (Valero et al., 2011) par comparaison à celui des peuplements continus. L'habitat joue également un rôle clé dans la dispersion des Fucales, comme cela a été rapporté en Méditerranée pour *Ericaria*

*amentacea*. En effet chez cette espèce, la continuité d'un habitat favorable est un prérequis à la dispersion ; cette dernière s'effectuant préférentiellement de proche en proche par les courants marins (Thibaut et al., 2016b ; Buonomo et al., 2017). A cet effet, les courants marins, et plus généralement le contexte océanographique actuel, sont des facteurs structurant les populations, et pouvant influencer la différenciation génétique (White et al., 2010), pour les Fucales (Thibaut et al., 2016b ; Buonomo et al., 2017 ; Nicastro et al., 2020), comme pour les kelps (Alberto et al., 2011 ; Durrant et al., 2018 ; Assis et al., 2018b).

Par exemple, pour les populations du kelp *Laminaria ochroleuca* Bachelot de la Pylaie étudiées du sud du Maroc au sud-ouest de l'Angleterre, la dispersion des spores et la dérive de sporophytes sont dépendantes des conditions océanographiques contemporaines (Assis et al., 2018b). Chez cette espèce, séparée historiquement en deux lignées lors du Quaternaire, le degré de différenciation génétique est inversement corrélé à la probabilité de dispersion par les courants marins (Figure 1.5a), ce qui indique que les flux de gènes s'effectuent majoritairement au sein d'une même région océanographique, et que ceux-ci sont très rares entre les différentes régions. Pour parvenir à ce résultat, les régions océanographiques (Figure 1.5b) ont été caractérisées par le couplage d'un modèle de courantologie à des simulations Lagrangiennes, ceci pour estimer la probabilité que la dispersion soit soutenue par les courants marins (connectivité océanographique).

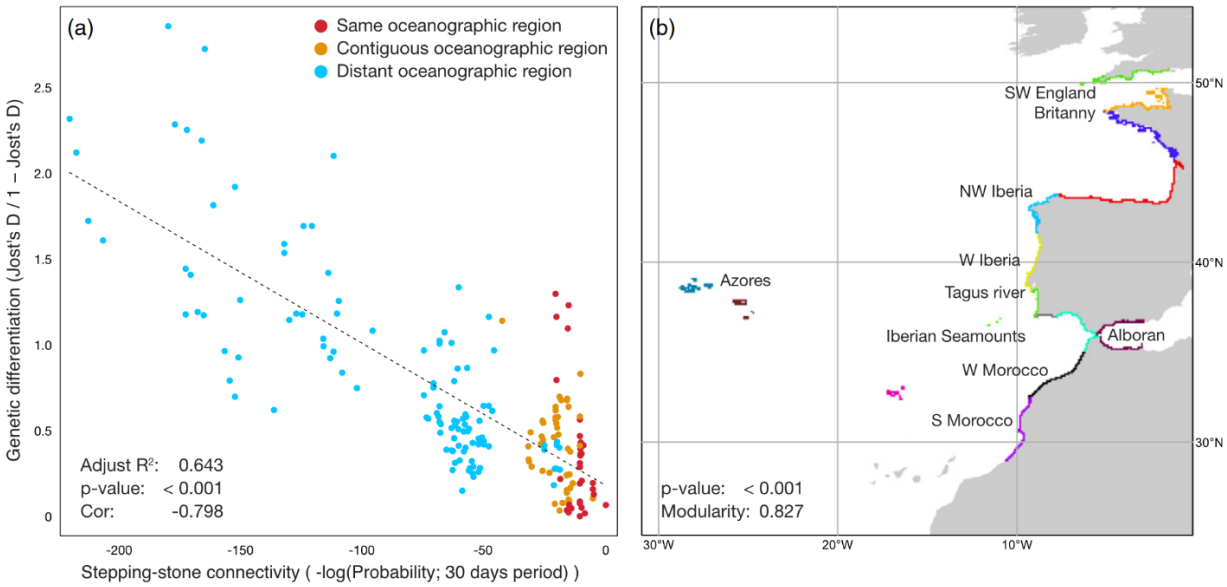



**Figure 1.5**. (a) Analyse de la corrélation entre la différenciation génétique des populations du kelp *Laminaria ochroleuca* et la probabilité de dispersion par les courants marins (b) fonction des régions océanographiques considérées (Assis et al., 2018b).

**Tableau 1.1**. Échelles de différenciation génétique des Laminariales (kelp) des côtes européennes. Chaque croix indique l'échelle à laquelle des différenciations génétiques sont constatées alors qu'un tiret matérialise le fait que cette dernière n'a pas été considérée compte tenu de la distance minimale d'échantillonnage (Dmin). $\%F_{ST}$ fait référence au pourcentage de $F_{ST}$ significatifs par échelle ou pour toutes distances confondues quand l'information n'était pas disponible. *No* indique que la différenciation n'est pas significative à l'échelle considérée.

| Laminariales (kelp) | Zone géographique | Dmin (km) | Différenciation génétique | | | $\%F_{ST}$ | Marqueurs | Références |
|---|---|---|---|---|---|---|---|---|
| | | | < 1 km | 1-10 km | ≥ 10 km | | | |
| *Laminaria digitata* | France (nord-ouest) | ≤ 2 | *No* | - | - | 0% | microsatellites (7) | Billot et al., 2003 ; Valero et al., 2011 ; Couceiro et al., 2013 |
| | | 10 | - | × | - | 100% | | |
| | | 50 | - | - | × | 100% | | |
| | France (nord-ouest) | 0,3 | × | - | - | 81% | microsatellites (12) | Robuchon et al., 2014 |
| | | 10 | - | × | - | | | |
| | | 60 | - | - | × | | | |
| | Atlantique (nord-est) | ≥ 100 | - | - | × | 100% | microsatellites (12) | Liesner et al., 2020 |
| | Atlantique (nord-est/ouest) | ≥ 50 | - | - | × | | microsatellites (12) | Neiva et al., 2020 |
| | Royaume-Uni, Norvège | 10 | - | × | × | 100% | microsatellites (12) | King et al., 2020 |
| *Laminaria hyperborea* | France (nord-ouest) | 0,3 | *No* | - | - | 66% | microsatellites (11) | Robuchon et al., 2014 |
| | | ≥ 10 | - | × | - | | | |
| | | ≥ 60 | - | - | × | | | |
| | Norvège | ≥ 100 | - | - | × | 89% | microsatellites (8) | Evankow et al., 2019 |
| | Irlande | 20 | - | - | × | | microsatellites (8) | Schoenrock et al., 2020 |
| *Laminaria ochroleuca* | Atlantique (nord-est), Alboran | ~10 | - | - | × | | microsatellites (15) | Assis et al., 2018b |
| *Saccorhiza polyschides* | | ~50 | - | - | × | 96% | microsatellites (7) | Assis et al., 2016 |
| *Saccharina latissima* | Atlantique (nord-est) | ~100 | - | - | × | 100% | microsatellites (25) | Guzinski et al., 2016 |
| | | ~100 | - | - | × | 100% | microsatellites (25) | Guzinski et al., 2020 |
| | | | - | - | × | 100% | 4069 SNPs | |
| | Atlantique (nord-est/ouest) | ~300 | - | - | × | 100% | microsatellites (10) | Luttikhuizen et al., 2018 |
| | Mer du Nord, Mer Baltique | ~50 | - | - | × | 93% | microsatellites (12) | Nielsen et al., 2016 |
| | Irlande (Strangford Lough) | ~2 | - | × | × | 86% | microsatellites (6) | Mooney et al., 2018 |
| | Norvège | ~100 | - | - | × | 97% | microsatellites (9) | Evankow et al., 2019 |
| | | 20 | - | - | × | 86% | microsatellites (11) | Næss, 2019 |

**Tableau 1.2**. Échelles de différenciation génétique des Fucales des côtes européennes. Chaque croix indique l'échelle à laquelle des différenciations génétiques sont constatées alors qu'un tiret matérialise le fait que cette dernière n'a pas été considérée compte tenu de la distance minimale d'échantillonnage (Dmin). $\%F_{ST}$ fait référence au pourcentage de $F_{ST}$ significatifs par échelle ou pour toutes distances confondues quand l'information n'était pas disponible. *No* indique que la différenciation n'est pas significative à l'échelle considérée.

| Fucales | Zone géographique | Dmin (km) | Différenciation génétique | | | $\%F_{ST}$ | Marqueurs | Références |
|---|---|---|---|---|---|---|---|---|
| | | | < 1 km | 1-10 km | ≥ 10 km | | | |
| *Fucus ceranoides* | Atlantique (nord-est) | 140 | - | - | × | 98% | microsatellites (9) | Neiva et al., 2012a |
| | Ibérique (nord-ouest) | ~15 | - | - | × | 99% | microsatellites (7) | Neiva et al., 2012b |
| *Fucus serratus* | Atlantique (nord-est) | ~0,5 | × | × | × | 100% | microsatellites (22) | Jueterbock et al., 2018 |
| | Atlantique (nord-est) | 0,2 | | × | × | 100% | microsatellites (7) | Coyer et al., 2003 |
| *Fucus vesiculosus* | Norvège, mer Baltique | 0,01 | × | × | × | 100% | microsatellites (5) | Tatarenkov et al., 2007 |
| | Atlantique (nord-est/ouest) | ~100 | - | - | × | 86% | microsatellites (9) | Assis et al., 2014 |
| | France (nord-ouest) | ~10 | - | × | × | 100% | microsatellites (7) | Almeida et al., 2017 |
| *Fucus spiralis* | France (nord-ouest) | ~10 | - | *No* | × | 50% | microsatellites (7) | Almeida et al., 2017 |
| *Fucus guiryi* | France (nord-ouest) | ~10 | - | × | × | 50% | microsatellites (7) | Almeida et al., 2017 |
| *Ericaria amentacea* | France (Provence, Corse) | 18 | - | × | × | 100% | microsatellites (6) | Robvieux, 2013 |
| | France (Baie de Marseille) | 3 | - | × | × | 100% | microsatellites (6) | Thibaut et al., 2016 |
| | Italie (sud-est), Sicile | 0,01 | × | × | × | 100% | microsatellites (10) | Buonomo et al., 2017 |
| *Ericaria selaginoides* | Méditerranée (Alboran, algéro-provencal, Ionienne), Altantique (nord-est) | ~50 | - | - | × | | microsatellites (6-10) | Bermejo et al., 2018, Buonomo et al. 2018 |
| *Ericaria mediterranea* | | | | | | | | |
| *Cystoseira compressa* | | | | | | | | |
| *Gongolaria elegans* | Méditerranée (Ibérique nord-est) | 9 | - | *No* | × | 66% | microsatellites (8) | Medrano et al., 2020 |

### 1.3.2 Reproduction végétative et implications évolutives

Parmi les différents mécanismes existants de reproduction asexuée (voir De Meeûs et al., 2007 ; Jaron et al., 2021), je me suis intéressé à l'une des rares formes de clonalité décrites dans le genre *Laminaria* : la multiplication végétative. Celle-ci prend place par l'élongation d'un stolon sur lequel se développent de nouvelles frondes (feuilles). Parmi les 29 espèces décrites du genre *Laminaria* (Guiry & Guiry, 2021), la multiplication végétative serait unique à deux d'entre elles : *Laminaria rodriguezii* (Figure 1.6A) (Huvé, 1955) et *Laminaria sinclairii* (Harvey ex J.D.Hooker & Harvey) Farlow, Anderson & Eaton (Demes & Graham, 2011) ; une espèce de l'Indo-Pacifique (Figure 1.6B).

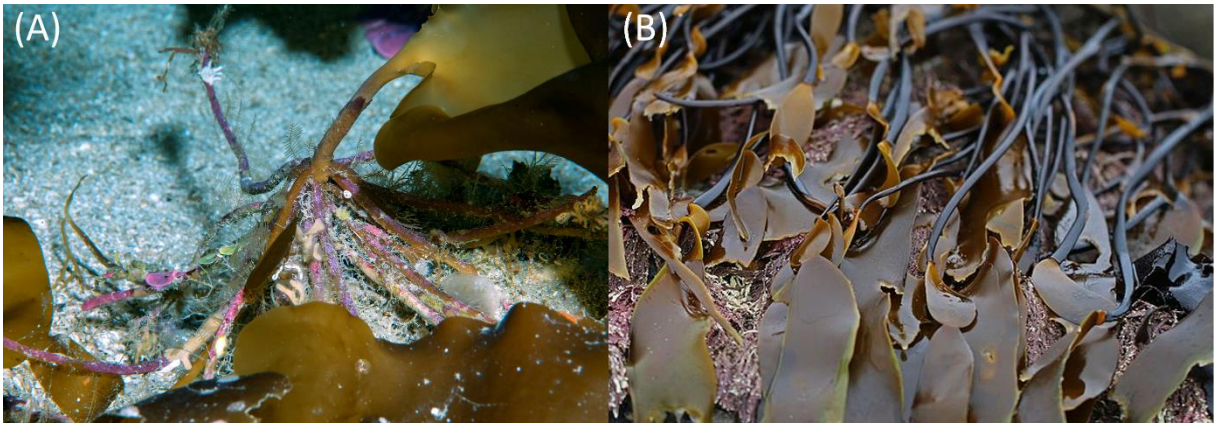



**Figure 1.6**. Les deux espèces du genre *Laminaria* pour qui la reproduction végétative opère par l'élongation d'un stolon, ici visible pour (A) le kelp des milieux profonds *Laminaria rodriguezii* (© Laurent Ballesta) et (B) le kelp de la zone intertidale de l'Indo-Pacifique *Laminaria sinclairii* (© Peter Pearsall).

La reproduction végétative est très rare chez les Fucales mais elle est néanmoins suspectée d'être le seul mode de reproduction des sargasses holopélagiques *Sargassum natans* et *Sargassum fluitans* vivant dans l'Atlantique Nord. En effet, aucune structure reproductrice n'a été identifié chez ces taxa en plus de 100 ans d'observations. Ce constat corrobore avec les patrons de la diversité génétique observés chez ces espèces à des locus microsatellites (ANR Origins, données non publiées). Les signatures de la clonalité sont marquées : un fort excès d'hétérozygotes, plusieurs lignées multilocus, dont certaines de taille relativement conséquente, jusqu'à 187 échantillons (Antiste, 2020). De manière surprenante, ces lignées multilocus sont pour la plupart partagées entre des échantillons distants de milliers de kilomètres. Celles-ci pourraient être le produit de la reproduction végétative : des lignées clonales, ou

alors refléter un manque de variabilité des marqueurs, qui en conséquence ne permettrait pas de discriminer des différences entre les individus (Arnaud-Haond et al., 2005). A titre d'exemple, pour une espèce qui se reproduit très majoritairement par autofécondation, comme la sargasse benthique invasive *Sargassum muticum*, la diversité génétique est si réduite que l'utilisation de microsatellites laissait penser que les individus étaient génétiquement identiques sur des milliers de kilomètres (Le Cam et al., 2020). Néanmoins, le génotypage d'un plus grand nombre de marqueurs (huit mille locus RAD-seq) a permis de discriminer plusieurs lignées multilocus (Le Cam et al., 2020).

Comme précédemment mentionné, la clonalité limite la diversité génotypique et peut favoriser une accumulation de mutations délétères (voir partie 1.1.2). Par ailleurs, la reproduction clonale favorise la persistance des génotypes au cours des générations là où la reproduction sexuée casse les liaisons existantes entre les allèles pour en créer de nouvelles. La persistance des génotypes clonaux interroge sur l'avantage évolutif de ce mode de reproduction. Un avantage clair de la clonalité est qu'un seul individu est suffisant à l'établissement d'une nouvelle population, ce qui n'est pas le cas d'une espèce non-hermaphrodite pour laquelle la reproduction mobilise les deux sexes séparées (Lynch, 1984). Ceci confère donc aux clones un haut potentiel de colonisation.

Une des hypothèses avancées pour expliquer le maintien de la reproduction asexuée à moyen terme serait celle du génotype « à usage général » (Baker, 1965 ; Lynch, 1984). Cette hypothèse propose que la sélection façonne au cours des générations un génotype généraliste qui possède une valeur sélective élevée pour une large gamme de paramètres environnementaux. Bien que cet effet soit difficilement mesurable, notamment chez les kelps et les Fucales pour qui l'âge des clones n'est pas connu, la persistance chez d'autres groupes de lignées clonales dont l'âge serait estimé à des milliers d'années, interroge.

La clonalité serait également favorisée dans des environnements marginaux dans lesquels l'absence de recombinaison est bénéfique pour conserver les génotypes localement adaptés à des conditions extrêmes alors que le flux de gènes serait susceptible de conduire à la mal-adaptation (Peck et al., 1998). L'un des exemples emblématiques à ce sujet concerne la Fucales *Fucus vesiculosus.* La très faible salinité de la partie nord de la mer Baltique (< 6 psu) fait de ce bassin un milieu inhospitalier pour les Fucales, notamment car le succès de la reproduction sexuée serait très limité en eau saumâtre. Or, *F. vesiculosus* est abondamment présent

dans cet environnement. Les analyses génétiques réalisées par Tatarenkov et al. (2005) ont ainsi révélé que les populations de *F. vesiculosus* en eau saumâtre étaient constituées en très grande proportion d'individus clonaux. L'hypothèse que la reproduction végétative permet le maintien de populations en milieux extrêmes, est intéressante à envisager dans le cas de *L. rodriguezii*. Il s'agit en effet de l'une des très rares espèces du genre *Laminaria* à être associée à des milieux profonds (voir la présentation de cette espèce ci-dessous).

### 1.3.3 Patrons d'adaptation locale chez les Laminariales et Fucales

Les patrons d'adaptation locale des kelps et des Fucales ont notamment été étudiés par l'intermédiaire d'expériences de cultures en milieu contrôlé. Celles-ci permettent d'apprécier le degré d'acclimatation et/ou d'adaptation des populations locales à des facteurs environnementaux, tels que la température ou la salinité. En revanche, la détection de locus outliers, nécessitant l'application d'outils génomiques, tels que le séquençage RAD est encore très limitée chez ces organismes.

L'une des espèces pour qui les patrons d'adaptation locale ont été étudiés est le kelp *Laminaria digitata*. L'utilisation de locus microsatellites, supposés neutres, a révélé chez cette espèce l'existence d'une structure génétique hiérarchique, marquée par la différenciation des lignées en limite sud de celles en limite nord de l'aire de répartition (Liesner et al., 2020). Par ailleurs, la réponse physiologique à la température des populations était hétérogène : les populations exposées à des périodes estivales les plus chaudes (Helgoland, nord-est de l'Allemagne et Quiberon, nord-ouest de la France) seraient davantage thermotolérantes que leurs homologues des zones boréales (Spitsbergen, océan Arctique) (Liesner et al., 2020). Cependant, les populations les plus thermotolérantes en limite sud, sont des populations marginales pour lesquelles la diversité génétique à des locus neutres est faible (Liesner et al., 2020 ; Neiva et al., 2020), ceci questionnant sur leur capacité d'adaptation au changement global et aux nombreux facteurs de stress qui y sont associés.

Quand l'information génomique était accessible, le couplage d'expériences en milieu contrôlé à des conditions stressantes et la détection de locus potentiellement impliqué dans l'adaptation locale a donné des résultats très encourageants pour apprécier le potentiel adaptatif de ces espèces. Ainsi, les travaux de Kinnby et al. (2020) ont rapporté en mer Baltique avec un séquençage de type RAD que les populations du Fucales *Fucus vesiculosus* sont

génétiquement différenciées à une échelle d'une dizaine de kilomètres, qu'il existerait des différences physiologiques entre les populations en réponse à un stress hypohalin et que ces différences émaneraient de locus impliqués dans des mécanismes d'adaptation à la salinité (de 10 à 18 psu selon les populations). En effet, les locus outliers identifiés dans les séquences RAD correspondent à des régions géniques dont la fonction serait en lien avec l'adaptation physiologique à des conditions hypohalines. En utilisant une approche similaire par RAD-seq, Guzinski et al. (2020) ont montré chez le kelp *Saccharina latissima* en Atlantique Nord qu'il existait une forte corrélation entre la fréquence de certains allèles à des locus outliers des populations et la température de surface de l'océan. Ce qui pourrait laisser entrevoir chez cette espèce un patron d'adaptation locale le long d'un gradient latitudinal de températures. Ainsi, les résultats obtenus sur *F. vesiculosus* et *S. latissima* indiquent que les populations de ces deux espèces pourraient réagir aux changements globaux de manière contrastée selon leurs capacités de thermotolérance (*S. latissima*) et leurs capacités d'adaptation à des conditions hypohalines (*F. vesiculosus*).

## 1.4 Les forêts marines méditerranéennes

### 1.4.1 Les forêts à *Cystoseira sensu lato*, focus sur *Ericaria zosteroides*

En Méditerranée, les forêts marines sont majoritairement représentées par le genre *Cystoseira sensu lato* (*Cystoseira* C. Agardh*, Ericaria* Stackhouse, *Gongolaria* Boehmer). Celui-ci regroupe un ensemble d'environs 46 taxa décrit, dont 44 sont présents en Méditerranée (Blanfuné, 2016 ; Guiry & Guiry, 2021). Chez ce groupe la délimitation des espèces n'est pas toujours claire (Bermejo et al., 2018) et la position phylogénétique des différents clades est encore source de débat (Draisma et al., 2010 ; Bruno de Sousa et al., 2019 ; Orellana et al., 2019 ; Molinari & Guiry, 2020). A titre d'exemple durant les trois années du projet MarFor, la taxonomie de *Cystoseira zosteroides* C.Agardh a été revue à deux reprises, passant de *Carpodesmia zosteroides* (C.Agardh) Greville à *Ericaria zosteroides* (C.Agardh) Molinari & Guiry. Par ailleurs, la présence d'espèce cryptiques et d'espèce synonymes peut être source de biais pour des études à une échelle intraspécifique, comme l'analyse de la connectivité entre populations (Pante et al., 2015a). Dans ce contexte, j'ai réalisé en amont de cette thèse, un travail visant à définir la position phylogénétique des échantillons de *E. zosteroides* au sein de la

classification de *Cystoseira sensu lato*. Cela a été fait en employant initialement des marqueurs mitochondriaux et chloroplastiques (par exemple, le locus *tRNALys-IGS*, Figure 1.7), puis des marqueurs RAD-seq (voir Materials B.1, partie 2.6). L'objectif était de tester si la diversité génétique a bien été étudiée au niveau intraspécifique La classification obtenue avec le locus mitochondrial *tRNALys-IGS* et l'homogénéité génétique des individus qui en ressort (Figure 1.7) est en accord avec l'expertise taxonomique de *E. zosteroides* faite sur la base de critères morphologiques. Néanmoins l'emploie d'un plus grand nombre de locus devrait permettre de confirmer ces résultats.

Les forêts marines à *Cystoseira sensu lato* constituent en zone littorale (de la surface jusqu'à 70-80 m de profondeur dans les eaux les plus claires) des habitats structurés et complexes des écosystèmes de substrat dur dominés par des producteurs primaires. Par leurs rôles de nurseries et de frayères, en particulier pour des espèces d'intérêt économique, par la diversité des habitats qu'elles constituent et par leurs interactions avec les autres écosystèmes littoraux (exportation de matière organique), la conservation des écosystèmes à *Cystoseira sensu lato*, est devenu une préoccupation centrale d'un point de vue écologique et socio-économique (Boudouresque, 1971 ; Verlaque, 1987 ; Ballesteros, 1990 ; Robvieux, 2013 ; Blanfuné, 2016 ; Capdevila, 2017).

En Méditerranée, le déclin des forêts à *Cystoseira sensu lato* est associé à l'impact de l'homme sur les écosystèmes littoraux des grandes métropoles. Les causes les plus fréquentes sont : la qualité des eaux (pollution chimiques, Sales et al., 2011, turbidité et eutrophisation, Thibaut et al., 2005 ; Iveša et al., 2016), la destruction de l'habitat (construction d'infrastructure, Blanfuné et al., 2016 et filets de pêche abandonnés, Thibaut et al., 2005, 2015), le surpâturage par les herbivores, tels que les oursins (Verlaque, 1987 ; Thibaut et al., 2005, 2015) et l'introduction d'espèces invasives comme la caulerpe *Caulerpa cylindracea* (Iveša et al., 2015).

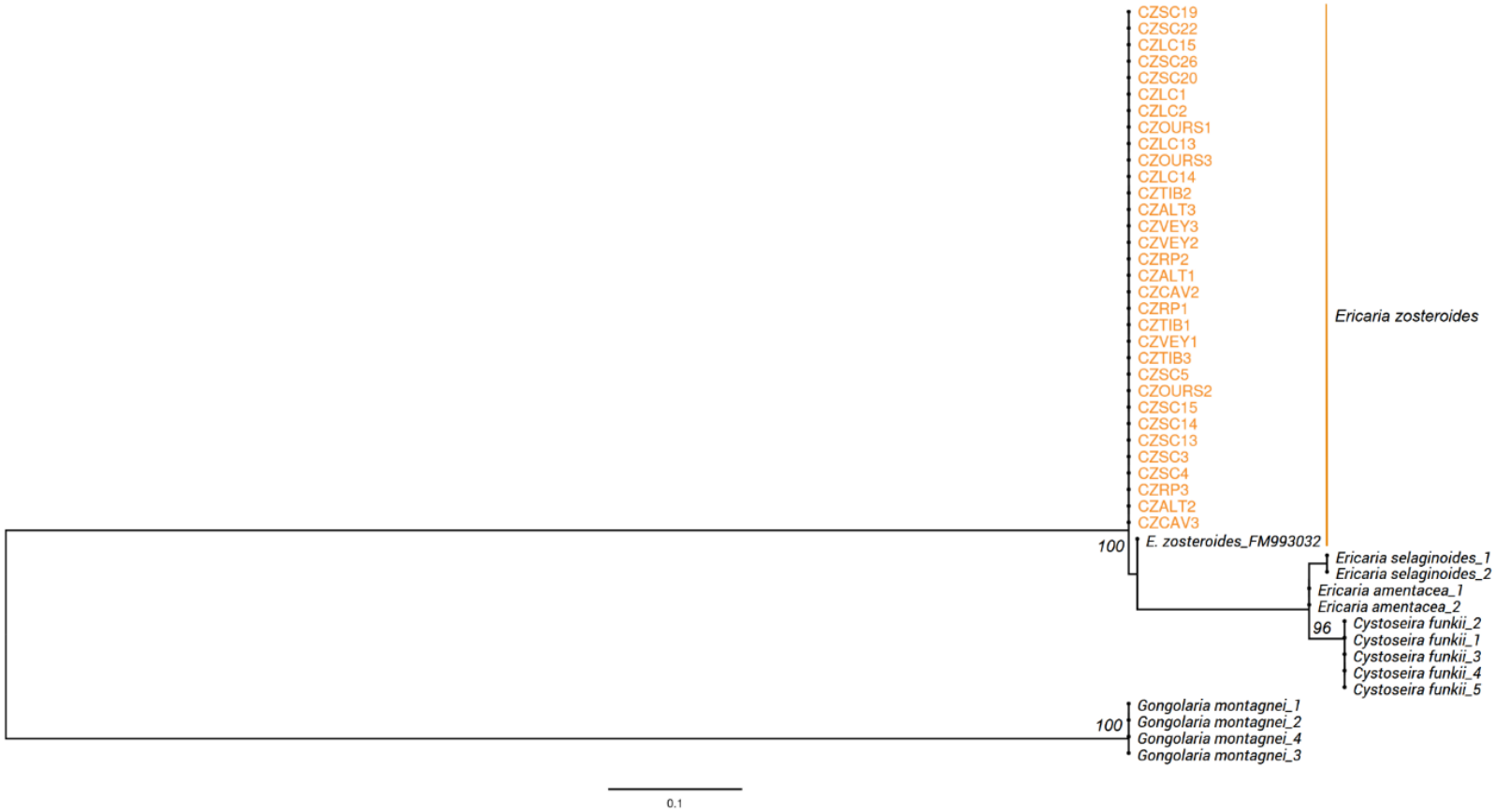



**Figure 1.7**. Arbre phylogénétique de *Cystoseira sensu lato*., raciné sur *Gongolaria montagnei* Indiquant la position phylogénétique d'une partie des échantillons de *Ericaria zosteroides* (Provence et Corse) employée dans le cadre de cette thèse (souligné en orange) par rapport à l'individu précédemment séquencé (séquence FM993032, Genbank) en Sicile par Draisma et al. (2010). Les différent taxa employés ont été échantillonnés durant ce projet. L'arbre a été reconstruit à partir du locus mitochondrial *tRNALys-IGS* (290 nt), les séquences des amorces, et les conditions de PCR sont identique au protocole de Draisma et al. (2010). L'arbre a été reconstruit avec IQ-TREE v.1.6.1 (Nguyen et al., 2015) en employant ModelFinder (Kalyaanamoorthy et al., 2017) et 1000 ultrafast bootstraps (Hoang et al., 2018). Le modèle d'évolution le plus parcimonieux retenu par ModelFinder est TN+F+G4. Seules les valeurs d'ultrafast bootstraps supérieure à 90% sont indiquées sur l'arbre.

La fragmentation des habitats littoraux et les extinctions locales qui résultent de ces nuisances sont un phénomène préoccupant au regard du très faible potentiel de dispersion des Fucales (voir la section « Structure et diversité génétique des forêts marines »). Les principaux risques à cet égard sont une diminution des flux de gènes et de la taille efficace des populations, avec des effets potentiellement délétères compte tenu de la réduction de la diversité génétique à envisager.

En réponse au changement climatique, le faible potentiel de dispersion des algues brunes pourrait limiter la réorganisation rapide de leurs aires de répartition (Assis et al., 2017). Ces conclusions reposent sur le constat d'un déplacement de leurs niches de quelques kilomètres par an par une dispersion des spores (un maximum de 5 km par an pour *S. polyschides*, Assis et al., 2017). Néanmoins, la capacité de colonisation de ces espèces par la dérive de fragments fertiles est un facteur à considérer pour évaluer le potentiel d'une réorganisation spatiale et compatible avec la vitesse du changement du climat.

La situation des Fucales en Méditerranée est quelque peu différente car la topologie de cette mer (un bassin semi-fermé) limite une réorganisation spatiale de ces espèces vers des habitats potentiellement favorables à de hautes latitudes. Cela a été constaté pour *Fucus virsoides* qui en quelques années est devenu proche de l'extinction en Adriatique (A. Falace cm. pers.). Les prédictions de Buonomo et al. (2018) rapportèrent que d'ici 2100, des Fucales comme *Ericaria amentacea* et *Cystoseira compressa* pourraient subir une réduction drastique de leurs aires de répartitions actuelles. Dans un scénario de changement climatique des plus optimistes (RCP 2.6, ~490 ppm $CO_2$, Moss et al., 2010), cette réduction s'élèverait à 52% et 48% pour ces deux espèces, respectivement (Buonomo et al., 2018). Néanmoins, l'acclimatation actuelle de *E. amentacea* et *C. compressa* à des environnements stressants, comme les cuvettes du littoral (Thibaut et al., 2014) dénotent que la réponse de ces espèces à la hausse des températures ne sera vraisemblablement pas homogène d'une population à une autre. En effet, en période estivale la température de ces cuvettes peut être supérieure à celle des prédictions les plus pessimistes du modèle de Buonomo et al. (2018). De plus, pour une espèce davantage associée à des conditions Atlantiques comme *Ericaria selaginoides,* l'extinction des populations en limite sud pourrait être en partie atténuée par une expansion vers le nord-est de l'Atlantique. Néanmoins, ce scénario de changement climatique impliquerait la disparition des populations pour lesquelles la diversité génétique est la plus élevée (sud-ouest de la

péninsule Ibérique, Maroc ; Buonomo et al., 2018 ; Bermejo et al., 2018), questionnant ainsi sur la capacité à venir des populations faiblement diversifiées à s'adapter à des situations naturelles changeantes.

*Ericaria zosteroides* est une espèce endémique à la Méditerranée, sa répartition s'étend à travers le bassin occidental à l'exception de la mer d'Alboran (Giaccone, 1973). Elle se rencontre fréquemment entre 15 et 50 mètres de profondeur, bien que cette amplitude soit conditionnée par la clarté des eaux (avec un record à 100 m aux îles Baléares et à Antibes au 19ème siècle). En effet, *E. zosteroides* nécessite de recevoir entre 1% et 0,3% de l'irradiance de surface pour être en mesure de se développer, préférentiellement sur des fonds rocheux balayés par des courants de fond (Giaccone, 1973). *Ericaria zosteroides* est une espèce longévive caractérisée par de faibles taux de renouvellement des populations (Capdevila, 2017) ; elle est en effet particulièrement sensible à certaines perturbations naturelles ou anthropiques (arrachage par les filets de pêche ou par des tempêtes) (Ballesteros et al., 2009 ; Capdevila et al., 2015 ; Thibaut et al., 2015 ; Capdevila, 2017) mais n'est pas soumises à l'influence des herbivores (saupes, oursins) ni à la destruction de l'habitat et est peu sensible au changement de turbidité (Thibaut et al., 2005). L'espèce est fréquente sur les côtes françaises avec de vastes populations en Corse, dans le Var et les Bouches-du-Rhône.

Il s'agit d'une espèce constituée d'un thalle dressé pérenne, solidement fixé au substrat par un disque basal, d'une taille maximale de 25 cm (axe primaire) sur lequel se développent annuellement des tophules (organes de réserves remplis de mannitol) d'où émergent les ramifications primaires puis secondaires, ramifiés en plusieurs branches (Figure 1.8).

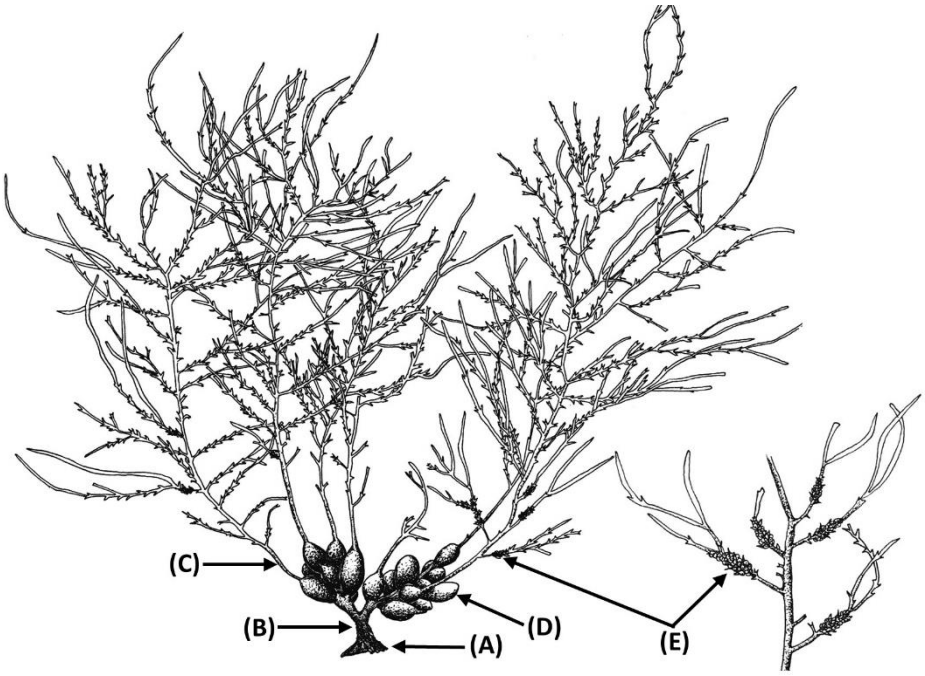



**Figure 1.8**. Morphologie d'*Ericaria zosteroides*, (A) disque basale, (B) axe primaire, celui-ci pouvant atteindre 25 cm de hauteur, (C) rameaux primaires, (D) tophules lisses, (E), réceptacles à la base des rameaux primaires. Les rameaux primaires et secondaires sont caducs, leur décrochage en période de fertilité (avril-septembre) est suspecté d'être un vecteur de flux de gènes (dessinatrice : Emma Rozis).

En présence de rameaux secondaires et tertiaires, l'espèce peut atteindre une hauteur de près d'un mètre (Ballesteros, 1990). Les structures reproductrices ou réceptacles se développent généralement à la base des rameaux primaires. Ces structures abritent les conceptacles, ceux-ci contiennent eux même les gamétocystes mâles (anthéridies) qui libéreront des spermatozoïdes biflagellés et femelles (oogones) qui libéreront huit oosphères (ovules), sans flagelle. Comme d'autres Fucales, *E. zosteroides* est une espèce hermaphrodite (les sexes ne sont pas séparés) qui dispose d'un cycle de reproduction monogénique diplontique (Figure 1.9). La reproduction est oogame et le zygote produit est entouré d'une membrane de fécondation qui va rapidement se fixer et développer des rhizoïdes pour fixer la plantule au substrat (Guern, 1962 ; Susini, 2006). On retrouve ce type de reproduction chez la plupart des Fucales. L'analyse de la diversité génétique des populations semble indiquer que l'autofécondation peut se produire chez ces espèces : un déficit en hétérozygotes est constaté dans la très grande majorité des populations étudiées (Robieux, 2013 ; Thibaut et al., 2016b ; Buonomo et al., 2017, 2018 ; Bermejo et al., 2018 ; Medrano et al., 2020). Par ailleurs, les zygotes ont une

flottabilité négative, leur dispersion serait donc limitée ; une distance de dispersion de l'ordre de 6,42 mètres en moyenne a été mesurée de manière expérimentale chez *E. zosteroides* (Capdevila et al., 2018).

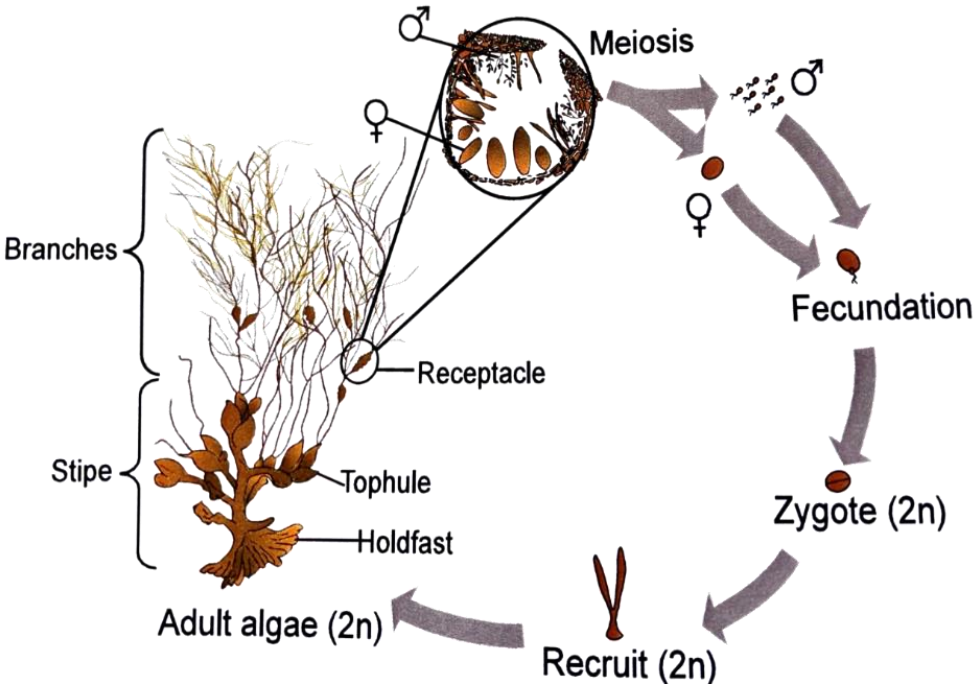



**Figure 1.9**. Cycle de vie d'*Ericaria zosteroides* pour laquelle la phase diploïde (2n) est dominante sur la phase haploïde (limitée à quelques cellules). Lors de la formation des spermatozoïdes flagellés et des oosphères par les gamétocystes mâles et femelles après méiose, les gamètes sont expulsés et la syngamie donnera un zygote, puis une plantule de quelques millimètres jusqu'à l'obtention d'un individu mature. La présente figure est issue de Capdevila (2017).

### 1.4.2 Les Laminariales des milieux profonds, focus sur *Laminaria rodriguezii*

La plupart des forêts marines à Laminariales sont présentes dans les milieux côtiers des régions tempérées et polaires de l'hémisphère nord (Kain, 1979 ; Steneck et al., 2002). À l'échelle des temps géologiques, les cycles de glaciation du Quaternaire ont modifié la distribution et l'abondance de ces espèces, avec des répercussions sur la diversité génétique actuelle des populations (Robuchon et al., 2014 ; Assis et al., 2016a ; Neiva et al., 2018 ; King et al., 2020 ; Neiva et al., 2020).

L'histoire biogéographique des kelps méditerranéo-atlantiques et méditerranéens, incluant *Saccorhiza polyschides* (Lightfoot) Batters, *Phyllariopsis brevipes* (C.Agardh) Henry & South, *Phyllariopsis purpurascens* (C.Agardh) Henry & South, *Laminaria ochroleuca* Bachelot de la Pylaie et *Laminaria rodriguezii* Bornet est en revanche très différente de celle des kelps de l'océan Atlantique de la zone intertidale ou subtidale. En effet, en période interglaciaire (période actuelle), les eaux de surface plus tempérées de la Méditerranée, et les fortes conditions d'irradiance qui y règnent contraignent ces espèces à affinité d'eaux froides vers des milieux profonds pour y trouver des conditions propices à leur survie : généralement, entre 20 et 50 mètres de profondeur pour *S. polyschides* et *Phyllariopsis* sp. (Flores-Moya, 2012), et jusqu'à 95 mètres pour *L. rodriguezii* en Corse (Fredj, 1972), et *L. ochroleuca* dans le détroit de Messine (Giaccone, 1972). La répartition verticale des kelps méditerranéens est associée, tout comme leurs homologues en régions tropicales et subtropicales (tels que, *Eisenia galapagensis* Taylor, *Laminaria brasiliensis* Joly & Oliveira et *Laminaria abyssalis* Joly & Oliveira), à la présence de remontées d'eaux froides chargées en nutriments (Druehl, 1981 ; Steneck et al., 2002). Ces structures océanographiques appelées « upwelling côtier » offrent des conditions propices au développement des kelps des milieux profonds. En effet, lorsque la couche de mélange se situe à une profondeur supérieure à celle du point de compensation de la photosynthèse, la probabilité d'y observer une de ces espèces augmente (Graham et al., 2007).

Les milieux profonds offrent des conditions environnementales stables, favorables à la persistance à long terme des kelps. Par analogie aux refuges climatiques des zones périglaciaires, les refuges en profondeur seraient propices au maintien de populations d'une grande diversité génétique et spécifique. Cette hypothèse a été avancé au regard des patrons de diversité génétiques des rares kelps qui ont été étudiés en milieux profonds (*S. polyschides*, Assis et al., 2016 et *L. ochroleuca*, Assis et al., 2018b). À l'échelle des bassins océaniques, la diversité génétique de ces deux espèces porte l'empreinte des changements climatiques du Quaternaire, alors qu'à une échelle locale, elle serait déterminée par la profondeur. En effet, les populations les plus profondes (*S. polyschides*, 20-35 mètres *; L. ochroleuca,* 20-50 mètres) sont génétiquement plus diversifiées que celles de surface (1-2 mètres, Assis et al., 2016 ; Assis et al., 2018b). Malheureusement, les contraintes d'accès et d'échantillonnages à de plus grandes profondeurs impliquent que la diversité génétique et le degré de connectivité des populations des kelps caractéristiques de ces milieux sont encore très mal connues. Le kelp *L. rodriguezii*

est l'une de ces espèces ; l'absence de données génétique est un point limitant pour évaluer le degré d'isolement des populations et leur état de conservation.

*Laminaria rodriguezii* se rencontre généralement entre 60 et 95 mètres de profondeur où par sa grande taille il peut former de vastes forêts marines balayées par d'importants courants de fond (Žuljević et al., 2016). Il s'agit d'une espèce endémique de Méditerranée, extrêmement rare, et qui se développe dans des conditions abiotiques très spécifiques, qui ne sont que partiellement connues. Son habitat est composé de coralligène de plateaux (Ballesteros, 2006) et est strictement associé à des eaux très transparentes et soumises à de très forts courants. *Laminaria rodriguezii* se développe à de très faibles intensités lumineuses et nécessite des températures basses et constantes (Fredj, 1972 ; Giaccone, 1973). Par exemple, pour les sites échantillonnés entre 71 et 76 mètres de profondeur au Banc du Magaud (est de la Provence), les individus étaient exposés à une température de 13.8°C, une salinité de 38.2 et des conditions d'irradiance de 2-4 µmol photons $m^{-2}$ $s^{-1}$ (données issues d'une sonde CTD ; profil effectué en juillet 2019, Deter et al., 2020). Annuellement, seule 0,1% de l'irradiance de surface est reçue et la température varie entre 12-14°C (D. Lefevre, Pôle MIO-obs.).

Concernant l'histoire évolutive de *L. rodriguezii*, les analyses phylogénétiques effectuées de manière préliminaire dans le cadre de cette thèse avec quelques marqueurs (locus mitochondrial *cox1*, Figure 1.10), corroborent avec l'arbre phylogénétique que nous avons reconstruit avec les partenaires du projet MarFor par le séquençage et l'assemblage complet du génome mitochondrial et chloroplastique (Rana et al., 2021). L'analyse préliminaire du gène *cox1* indique que les populations étudiées dans cette thèse présentent pour ce locus le même haplotype que celui précédemment obtenu pour un individu d'Adriatique par Žuljević et al. (2016). Par ailleurs, les arbres phylogénétiques de ces différents marqueurs indiquent que *L. rodriguezii* serait très proche des kelps de l'Atlantique *L. digitata* et *L. hyperborea*. La proximité de ces espèces interroge sur l'affinité de *L. rodriguezii* pour des milieux profonds, notamment car la température ne serait pas un facteur limitant à sa croissance. En effet, l'optimum de croissance de *L. rodriguezii* ne situerait entre 10-20°C avec une très forte décroisse à 23°C (Pageault, 2020). Néanmoins, l'irradiance de surface est vraisemblablement le facteur le plus important à considérer pour expliquer la limite supérieure de la distribution de cette espèce.

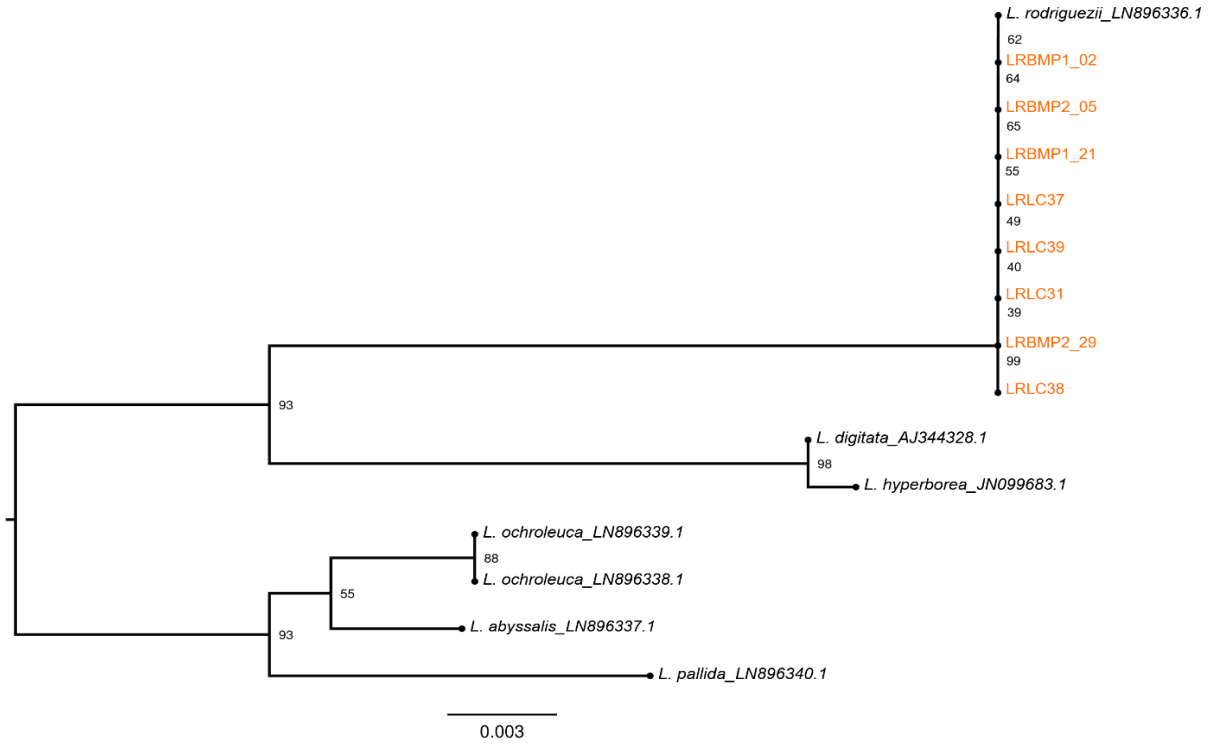



**Figure 1.10**. Arbre phylogénétique de *Laminaria* sp. (raciné sur le clade incluant *L. ochroleuca*) indiquant la position phylogénétique d'une partie des échantillons de *Laminaria rodriguezii* (est de la Provence) utilisés dans le chapitre 3 de ma thèse (en orange) par rapport à l'individu précédemment séquencé par Žuljević et al. (2016) en Croatie (LN896336.1). Les numéros d'accessions des séquences (Genbank) sont rapportés à la fin du nom des espèces. L'arbre a été reconstruit à partir du locus mitochondrial *cox1* (598 nt), les séquences des amorces, et les conditions de PCR sont identiques au protocole de Silberfeld et al. (2010). L'arbre a été reconstruit avec IQ-TREE v.1.6.1 (Nguyen et al., 2015) en employant ModelFinder (Kalyaanamoorthy et al., 2017) et 1000 ultrafast bootstraps (Hoang et al., 2018). Le modèle d'évolution le plus parcimonieux retenu par ModelFinder est K3Pu+F+I, les valeurs d'ultrafast bootstraps sont indiquées sur l'arbre.

*Laminaria rodriguezii* possède une distribution fragmentée à l'échelle du bassin occidental de la Méditerranée et de la mer Adriatique. Ces populations ont été répertoriées en Adriatique, principalement en Croatie, sur le pourtour des îles Baléares et de l'île d'Alborán, en Italie et Tunisie. Ces données sont pour la plupart issues de relevés datant des années 1960-1970, synthétisés par Boudouresque & Perret (1977), Giaccone et al. (2009), Žuljević et al. (2011) et Araújo et al. (2016) et d'une actualisation plus récente, témoignant d'une régression, voire d'extinction locale en Adriatique (Žuljević et al., 2016). En ce qui concerne les côtes françaises, l'espèce a été répertoriée dans l'est de la Provence (Banc du Magaud), en Corse, spécifiquement dans les parties nord-ouest (pointe de Revellata ; Fredj, 1972), nord-est (Parc

Naturel Marin du Cap Corse et de l'Agriate ; communication personnelle Andromède Océanologie) et sud (les Bouches de Bonifacio). Dans le cadre de cette thèse, une nouvelle population a été découverte à l'est de la Provence, faisant suite à des signalements par les pêcheurs de grandes algues brunes dans leurs filets. À l'aide d'un ROV, nous avons localisé cette dernière à 30 km du Banc du Magaud, au large du Cap Camarat entre 65 et 73 mètres sur un fond de coralligène (Figure 1.11). Bien que les causes à l'origine de la raréfaction de cette espèce soient encore méconnues, le chalutage, la remise en suspension du sédiment, et la diminution de la clarté des eaux sont les principaux facteurs susceptibles de conduire à l'extinction locale des populations (Žuljević et al., 2016).

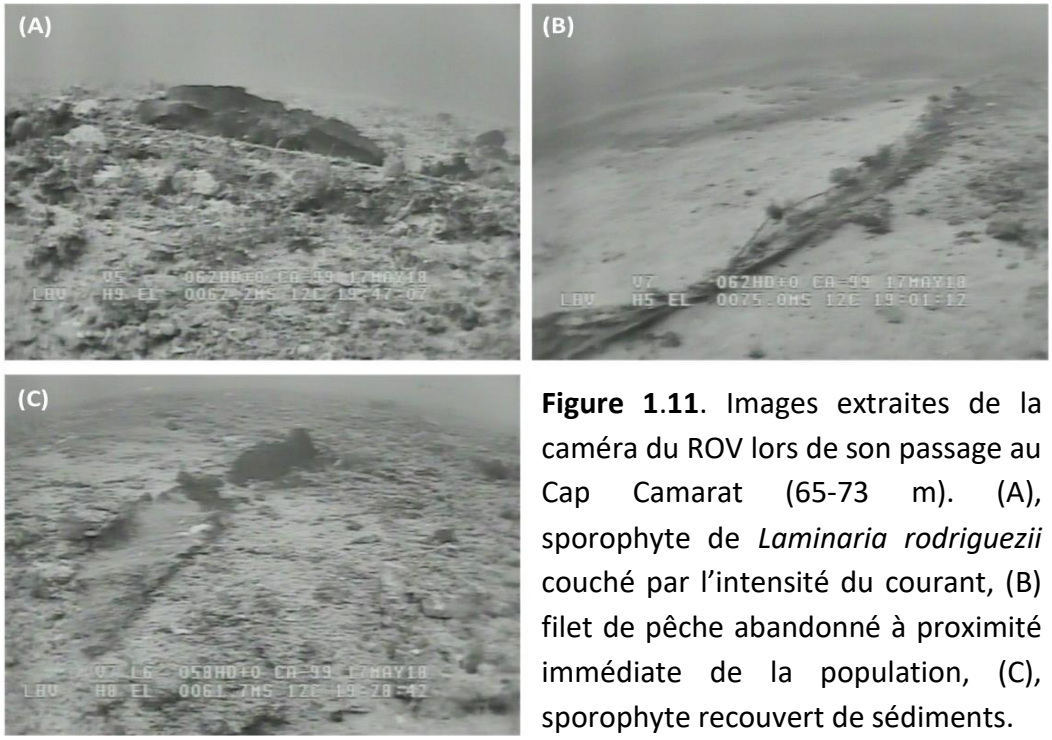



**Figure 1.11**. Images extraites de la caméra du ROV lors de son passage au Cap Camarat (65-73 m). (A), sporophyte de *Laminaria rodriguezii* couché par l'intensité du courant, (B) filet de pêche abandonné à proximité immédiate de la population, (C), sporophyte recouvert de sédiments.

*Laminaria rodriguezii* est constitué d'un stipe court qui porte une grande lame, la zone de croissance est située dans la zone méristématique et chaque année une nouvelle lame serait produite. Cet ensemble constitue le sporophyte qui est accroché par des stolons assez courts et très fragiles qui se détachent facilement et d'où émergent de nouvelles lames.

Le cycle de vie de *L. rodriguezii* est singulier en raison de la capacité des individus à se reproduire de manière sexuée et asexuée, par multiplication végétative. Le cycle de reproduction sexé est digénétique haplodiplontique hétéromorporphe avec dominance du sporophyte (jusqu'à 3 m de longueur), les gamétophytes étant microscopiques. Ce cycle vient très

récemment d'être observé par nos collègues allemands de l'Institut A. Wegener (Inka Bartsch com. pers.) avec qui nous collaborons. Sur les lames, des sores se forment, ceux-ci contiennent des sporocystes uniloculaires d'où émergeront des zoospores qui germeront en gamétophytes haploïdes dioïques mâles et femelles (Figure 1.12). Dans les gamétophytes mâles (à partir des anthéridies) les spermatozoïdes flagellés seront expulsés et dans les gamétophytes femelles à partir d'une oogone, des oosphères immobiles seront fécondées. Le jeune sporophyte se développera sur le gamétophyte femelle. Comme précédemment mentionné, la multiplication végétative prend place par l'élongation d'un stolon sur lequel se développent de nouvelles lames.

La fécondation entre les gamétophytes mâles et femelles d'un même sporophyte est connue chez les Laminaires et s'apparente à une forme d'autofécondation (« biparental inbreeding », Goecke et al., 2020). Les premières expériences en milieu contrôlé indiquent que ce mécanisme serait aussi possible chez *L. rodriguezii* (Inka Bartsch, communication personnelle). En revanche, nous ne savons pas si d'autres formes de reproduction asexuée ont lieu au niveau des gamétophytes, avec les conséquences évolutives que cela implique chez des espèces haplodiplobiontes (Krueger-Hadfield, 2021).

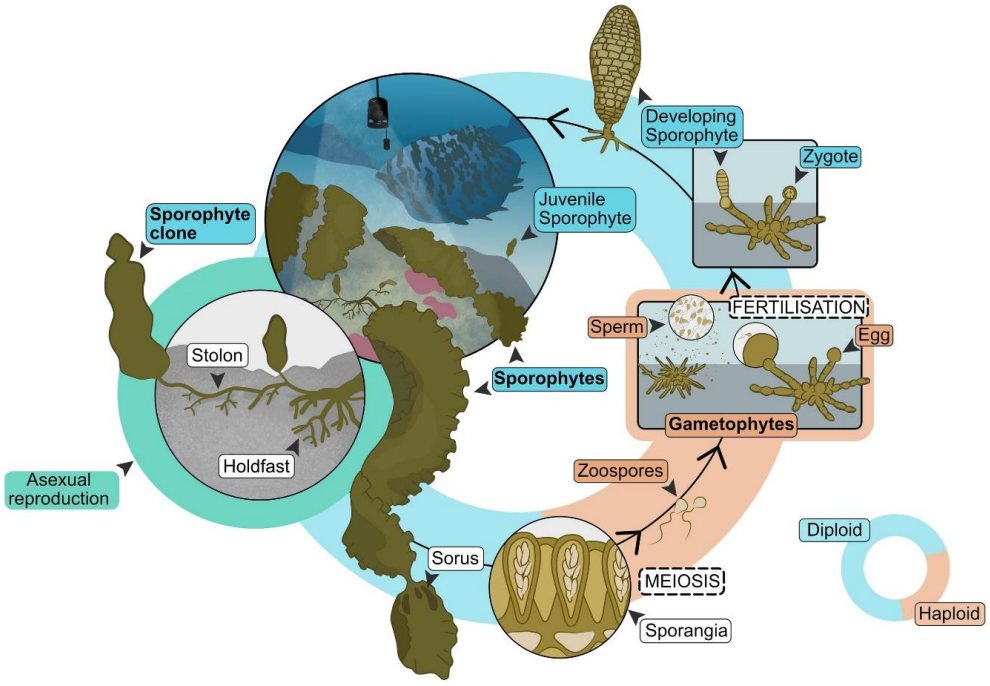



**Figure 1.12**. Cycle de vie de *Laminaria rodriguezii* caractérisé par l'alternance d'une phase sexuée, typique à la reproduction des espèces du genre *Laminaria* et d'une phase asexuée par multiplication végétative par l'élongation d'un stolon qui donnera lieu à un sporophyte clonal (© Antoine Minne).

## 1.5 Objectifs de la thèse

**Le premier objectif** de ma thèse a été de dresser un état initial de la diversité génétique et de la structure des populations du Fucales *Ericaria zosteroides* et du Kelp *Laminaria rodriguezii* le long des côtes françaises de la Méditerranée, deux espèces de hautes valeurs écologiques pour lesquelles ces informations sont inexistantes. Or, la diversité génétique renseigne, entre autres, sur la capacité de dispersion et le degré d'isolement des populations ; deux indicateurs essentiels à une prise de décision pour la conservation de ces espèces. Pour ce faire, nous avons opté pour un séquençage ddRAD (double-digest) de sorte à obtenir une résolution à l'échelle du génome des individus. Ceci pour évaluer le degré de différenciation génétique à des locus neutres au regard des flux de gènes et de la dérive génétique. Ce premier objectif est le socle commun aux chapitres 2 et 3 de ce manuscrit.

- Dans le chapitre 2, j'étudierai la différenciation génétique des populations de *E. zosteroides* à différentes échelles spatiales. Nous avons considéré : i) une échelle très fine, entre la partie supérieure et inférieure d'un tombant rocheux pour lequel les individus sont séparés par approximativement 20 mètres de profondeur, ii) une échelle locale par l'intermédiaire des populations de la baie de Marseille (ouest de la Provence) et celles présentes dans l'aire du Parc National des Calanques, iii) une échelle régionale entre l'est de la Provence et la Corse.

- Dans le chapitre 3, je présenterai la structure génétique des populations de *L. rodriguezii* à l'échelle de l'est de la Provence et du sud de la Corse. La structure génétique spatiale d'une des populations a aussi été considérée pour apprécier la différenciation à une distance de quelques mètres.

**Le deuxième objectif** a été d'étudier l'incidence des courants marins sur la connectivité génétique des populations de *E. zosteroides* et secondairement, le rôle de la sélection divergente dans un environnement hétérogène. Pour parvenir à cet objectif, nous nous sommes focalisés sur la baie de Marseille en procédant à une analyse par couplage d'un modèle de courantologie, de simulations de la dispersion des propagules par les courants marins et de la structure génétique des populations. Trois objectifs spécifiques ont été définis :

- Évaluer le rôle de la connectivité océanographique sur la structure génétique des populations à des locus neutres. Des paramètres démographiques, tels que la taille et la densité d'individus par site ont été mesurés pour considérer des hypothèses alternatives associées au recrutement.

- Rechercher des locus outliers afin d'identifier des régions du génome potentiellement impliquées dans l'adaptation locale des populations à la profondeur (irradiance, température).

**Le troisième objectif** est centré sur le mode de reproduction de *L. rodriguezii*, la clonalité partielle, afin d'en évaluer les conséquences sur la diversité génétique à l'échelle du génome des individus. Néanmoins, compte tenu des difficultés d'accès pour l'échantillonnage, la clonalité a été considérée uniquement au stade sporophytique de *L. rodriguezii*, les gamétophytes étant indétectables *in situ*. Ce troisième objectif est propre au chapitre 3 de ce manuscrit et il se décline en trois objectifs spécifiques :

- Produire des données empiriques à l'échelle du génome d'une espèce partiellement clonale avec comme modèle d'étude *L. rodriguezii*. La finalité étant d'évaluer si les signaux de la clonalité sont en accord avec les patrons de diversité génétique prédits en termes de distribution des valeurs de l'indice de fixation $F_{IS}$ (Stoeckel & Masson, 2014) et du déséquilibre de liaison (Stoeckel et al., 2020) pour différents taux de clonalité et de dérive génétique. L'objectif à terme est de pouvoir évaluer la part de reproduction sexuée et asexuée des populations de *L. rodriguezii* et les conséquences sur leur diversité génotypique et allélique.

- Evaluer la part respective de la clonalité partielle de celle de la dérive génétique sur la diversité génétique des populations. Cet objectif est ambitieux, néanmoins il a été considéré en optant pour une comparaison de la diversité génétique des populations de l'espèce partiellement clonale *L. rodriguezii* et l'espèce congénérique sexuée *Laminaria digitata*. Les populations de *L. digitata* ont été choisies en connaissance de leur degré de diversité génétique, précédemment analysé avec des microsatellites (Liesner et al., 2020).

- Délimiter les lignées clonales parmi les populations de *L. rodriguezii* en proposant une méthode qui considère les distances génétiques obtenu pour un même individu séquencé à deux reprises par ddRAD. L'objectif est d'arriver à une délimitation des lignées clonales qui tient compte des erreurs de génotypage dues à la méthode employée et aux mutations somatiques.

# Chapitre 2

# La génomique des populations et la modélisation Lagrangienne soulignent le rôle des courants marins dans la dispersion des forêts marines à *Ericaria zosteroides* (Fucales)



Une forêt à *Ericaria zosteroides*, tombant de l'Armoire (Baie de Carqueiranne) © Sandrine Ruitton

Ce chapitre correspond à un article publié dans la revue **Frontiers in Marine Science** (DOI: 10.3389/fmars.2021.683528, publié le 02 juin 2021).

***Contexte et objectifs***

Les précédents travaux en génétique des populations indiquent l'existence d'une forte structure génétique, à relativement courte distance, entre populations pour plusieurs espèces de *Cystoseira sensu lato*. Ces travaux ont concerné spécifiquement des espèces de l'étage de l'infralittoral supérieur comme *Ericaria amentacea*, *Ericaria selaginoides*, *Ericaria mediterranea*, *Gongolaria elegans* et *Cystoseira compressa* (Susini et al., 2007 ; Robvieux, 2013 ; Thibaut et al., 2016b ; Buonomo et al., 2017, 2018 ; Bermejo et al., 2018 ; Medrano et al., 2020). Par ailleurs, la différenciation génétique des populations de *E. amentacea* est corrélée aux courants marins de surface (Thibaut et al., 2016b), en combinaison avec le degré de fragmentation de l'habitat, notamment à la distance entre les tâches d'habitats (Buonomo et al., 2017). Ces précédents travaux soulignent l'importance d'une dispersion majoritairement localisée chez ces espèces. Cette faible distance de dispersion, en conjugaison avec une forte dérive locale, pourrait expliquer une telle structure génétique. Cependant, des épisodes rares de dispersion à longue distance doivent participer au fonctionnement de la métapopulation. Ces évènements sont soutenus par l'observation de nouvelles populations installées sur des substrats artificiels, tels que des digues et conduites sous-marines à une distance de plusieurs kilomètres des populations connues (Thibaut et al., 2014 ; Bonhomme et al., 2014). La colonisation de ces substrats artificiels par des populations installées récemment (événement datable) est une situation idéale pour étudier le potentiel de colonisation ou de recolonisation de ces espèces, potentiel qui n'est pas forcément reflété dans les flux de gènes observés.

Par ailleurs, la plupart des études précédemment mentionnées du genre *Cystoseira sensu lato* ont concerné des espèces peu profondes. Les études de connectivité associant génétique et courantologie à une échelle locale voire régionale (ici la baie de Marseille et la côte provençale), sont encore relativement limitées pour les espèces et populations plus profondes, au-delà de 40 m de profondeur notamment. Ainsi, la structure génétique des forêts marines d'algues brunes méditerranéennes de l'étage infralittoral inférieur et du circallitoral reste encore à explorer. Comme précédemment mentionné, l'association d'une

approche de génomique des populations, de démographie et de modèles de dispersion peut être utile pour estimer la contribution de variables environnementales (courant, isolement entre populations) sur la différenciation génétique observée. Il est ainsi possible de tester si l'observation d'une différenciation génétique à courte distance est en accord avec les capacités de dispersion modélisées à partir de la courantologie et de la biologie des stades de dispersion.

Dans ce contexte, j'ai étudié la connectivité génétique d'une espèce commune du circallitoral, la Fucales *Ericaria zosteroides*. Pour cette espèce endémique de Méditerranée, le faible potentiel de dispersion des zygotes (Capdevila et al., 2018), et la répartition fragmentée des populations sont deux éléments nous laissant à penser que la différenciation génétique peut être élevée à l'échelle de quelques centaines de mètres, voire quelques kilomètres. Par ailleurs, l'adaptation locale des populations du genre *Cystoseira sensu lato* est un domaine de recherche encore peu considéré. En Méditerranée la variation du régime thermique en fonction de la profondeur peut être un facteur sélectif important, favorable à l'évolution d'un patron d'adaptation locale (voir par exemple Pratlong et al., 2018). C'est notamment sur le gradient 10-50 m où des différences de régime thermique et de luminosité (quantité et composition) vont être très contrastées. Une telle diversité adaptative en lien avec le régime thermique devrait être prise en compte pour mieux comprendre la réponse de ces espèces au changement climatique. Il est donc important de tester cette hypothèse d'adaptation locale en fonction de la profondeur pour des Fucales disposant d'une distribution bathymétrique relativement large (15-50 mètres).

Le premier objectif a été de caractériser la structure démographique et génétique des populations d'*Ericaria zosteroides* à différentes échelles spatiales, puis dans un second temps d'évaluer la contribution des courants marins à la différenciation génétique des populations de la baie de Marseille (ouest de la Provence). C'est précisément la complémentarité d'une approche en génomique des populations par séquençage RAD à celle relevant de la modélisation Lagrangienne que nous avons considéré cet objectif. Ces deux approches ont par ailleurs été employées pour identifier l'origine la plus probable d'une population installée sur un substrat artificiel, afin d'estimer la distance de dispersion précédant cet événement de colonisation. Finalement, le nombre de marqueurs obtenus par l'approche RAD a permis de rechercher des signaux d'adaptation locale par une approche de type "locus outliers".

Frontiers in Marine Science (https://doi.org/10.3389/fmars.2021.683528)

## Population genomics and Lagrangian modeling shed light on dispersal events in the Mediterranean endemic *Ericaria zosteroides* (= *Cystoseira zosteroides*) (Fucales)

Lauric Reynes[1], Didier Aurelle[1,2], Cristele Chevalier[1], Christel Pinazo[1], Myriam Valero[3], Stéphane Mauger[3], Stéphane Sartoretto[4], Aurélie Blanfuné[1], Sandrine Ruitton[1], Charles-François Boudouresque[1], Marc Verlaque[1], Thierry Thibaut[1]

(1) Aix Marseille Université, Université de Toulon, CNRS, IRD, MIO, Marseille, France
(2) Institut de Systématique Évolution Biodiversité (ISYEB, UMR 7205), Muséum national d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, 57 rue Cuvier, CP 51, 75005 Paris, France
(3) IRL 3614, Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université, UC, UACH, CS 9 90074, Place Georges Teissier, 29688 Roscoff, France
(4) IFREMER, Zone Portuaire de Brégaillon, 83500 La Seyne-sur-mer, France

## Abstract

**Key words: Marine forests, population genomics, Lagrangian modeling, dispersal, gene flow**

Dispersal is a central process that affects population growth, gene flow, and ultimately species persistence. Here we investigate the extent to which gene flow occurs between fragmented populations of the deep-water brown algae *Ericaria zosteroides* (Turner) Greville (Sargassaceae, Fucales). These investigations were performed at different spatial scales from the bay of Marseille (western Provence) to Corsica. As dispersal of zygotes is shown to be limited over distances beyond a few meters, we used a multidisciplinary approach, based on Lagrangian modeling and population genomics to test the hypothesis that drifting of fertile parts of thallus (eggs on fertile branches), mediated by ocean currents, enable occasional gene flow between populations. Therefore we assessed the respective contribution of oceanographic connectivity, geographical isolation, and seawater temperatures to the genetic structure of this species. The genetic structure was assessed using 10 755 neutral SNPs and 12 outlier SNPs genotyped by dd-RAD sequencing in 261 individuals of *E. zosteroides*. We find that oceanographic connectivity is the best predictor of genetic structure, while differentiation in outlier SNPs can be explained by the depth of populations, as emphasized by the minimum seawater temperature predictor. However, further investigations will be necessary for clarifying how depth drives adaptive genetic differentiation in *E. zosteroides*. Our analyses revealed that local hydrodynamic conditions are correlated with the very high

divergence of one population in the Bay of Marseille. Overall, the levels of gene flow mediated by drifting were certainly not sufficient to counteract differentiation by local genetic drift, but enough to allow colonization several kilometers away. This study stresses the need to consider secondary dispersal mechanisms of presumed low dispersal marine species to improve inference of population connectivity.

## 2.1 Introduction

Genetic diversity is a prerequisite for the adaptive evolution of species. Accordingly, surveys of genetic diversity of rare and vulnerable species should be on the agenda of management and conservation frameworks (Hoban et al., 2020). This involves, among other requirements, the need to estimate the levels of genetic connectivity of species and their effect on the genetic structure of natural populations (Slatkin, 1985, 1987). Here, we will use the following definition of genetic connectivity: the degree to which the dispersal of individuals or propagules affect gene flow, with repercussions on evolutionary processes within populations, while demographic connectivity refers to the contribution of dispersal to the rate of population growth (Lowe and Allendorf, 2010). Genetic connectivity can counteract the potential loss of genetic diversity due to local inbreeding (i.e. inbreeding connectivity, Lowe and Allendorf, 2010), genetic drift, and the fixation of deleterious mutation (Keller and Waller, 2002). Demographic connectivity can maintain the stability of populations structured across space by increasing the range of distribution and enhancing the resilience of populations following disturbances (Gotelli, 1991; Lowe and Allendorf, 2010). Genetic connectivity can also counteract the evolution of local adaptation (Lenormand, 2002). The interplay between genetic connectivity, demography and local selection is then at the heart of the evolution of populations. In the marine environment, many organisms disperse through the release of many small propagules, which consequently require the use of indirect methods such as genetic inference and/or Lagrangian models for assessing dispersal. Combining these indirect methods in a multidisciplinary approach was seen as an effective way to create the most complete picture of population connectivity, notably in marine species with low dispersal potential for which connectivity may be driven by unsuspected occasional long-distance dispersal events (for review see Kinlan and Gaines, 2003).

Next-generation sequencing (NGS) has turned the study of genetic connectivity from the use of tens to thousands of markers, going up to whole individual genomes (Gagnaire, 2020). The use of this high number of markers allows more precise analysis of genetic differentiation, and increased power to detect genetic differences among populations (Waples, 1998; Gagnaire et al., 2015). Apart from allowing more powerful inferences, the genomic level allows taking into account selected or hitchhiker loci to detect barriers of dispersal, which is particularly useful in high-dispersal species (Gagnaire et al., 2015). Information about the spatial distribution of adaptive genetic diversity can also be considered in management and conservation strategies (Xuereb et al., 2020). Combining population-genomic inferences and Lagrangian dispersal models has helped to elucidate, with unprecedented clarity, how environmental factors and geography drove the genetic structure of marine species (e.g. Benestan et al., 2016; Dalongeville et al., 2017; Paterno et al., 2017; Xuereb et al., 2018a, 2018b; Bracco et al., 2019; Coscia et al., 2020). To our knowledge, such integrative frameworks at genome-scale have rarely been applied to brown algal species. The scale of the study is a critical point in such investigations of connectivity. We can define a fine scale (from a few centimeters and a dozen meters within a site) that can be used to study population structure and delimitation (demes) (e.g. Ledoux et al., 2010), a local scale (from hundreds of meters to a few kilometers) where scattered populations are shaped by local hydrodynamic currents (e.g. Palumbi, 2003), with a high variation mainly according to the wind direction, and a regional or continental scale where permanent oceanic currents are the main drivers of connectivity (e.g. White et al., 2010).

In the Mediterranean Sea, Fucales (*Cystoseira* C. Agardh*, Ericaria* Stackhouse*, Fucus* Linnaeus*, Gongolaria* Boehmer*,* and *Sargassum* C. Agardh) (see Molinari & Guiry, 2020 for an update of the taxonomy of the *Cystoseira sensu lato*) form marine forests and structure rocky reef assemblages. Most of the time, populations are patchy, from a few individuals to larger populations dwelling on rocks or boulders surrounded by detrital soft bottoms, coralligenous and *Posidonia oceanica* meadows (e.g. Ballesteros, 2006; Hereu et al., 2008; Boudouresque et al., 2017). Very few studies have focused on the connectivity of Mediterranean Fucales at different scales, and all concerned shallow waters species (i.e. Susini et al., 2007; Robvieux, 2013; Thibaut et al., 2016b, Buonomo et al., 2017; Medrano et al., 2020). Previous studies have reported that populations of *Ericaria amentacea* (C. Agardh) Molinari and Guiry

(synonym: *Cystoseira amentacea*) were genetically differentiated at local and regional spatial scales, from 100 meters to tens of kilometers (Susini et al., 2007; Thibaut et al., 2016b; Buonomo et al., 2017). Similar levels of genetic differentiation were also observed within the *Ericaria selaginoides* (Linnaeus) Molinari & Guiry complex (formerly called *Cystoseira tamariscifolia* complex) (Bermejo et al., 2018). Such high levels of genetic differentiation could have been expected given the low dispersal ability of these species, which is directly related to the short interval between external fertilization (i.e. mating process between motile male flagellate gametes and large female, non-motile gametes) and fixation of fertilized zygotes to the substrate. Large and non-buoyant zygotes of several species of Fucales (*Cystoseira* spp.*, Ericaria* spp. and *Gongolaria* spp.) sink and adhere to the substrate a few hours after fertilization (Guern, 1962, Falace et al., 2018). This mating pattern is expected to only allow restricted dispersal events (less than 40 cm to 10 m in *Ericaria* species; Mangialajo et al., 2012; Capdevila et al., 2018). Importantly, the local genetic structure of *Ericaria amentacea* in the Bay of Marseille was better explained by oceanographic distance rather than spatial distance (Thibaut et al., 2016b). Furthermore, and non-exclusively, the high population genetic differentiation in *Ericaria* and *Gongolaria* species is probably driven by the balance between restricted gene flow and strong genetic drift, as reported in some Mediterranean animal species (e.g. *Corallium rubrum*, Ledoux et al., 2010). Up to now, connectivity studies in *Ericaria* and *Gongolaria* have involved relatively small numbers of genetic markers (e.g. about ten microsatellites).

The connectivity of the deep-water species *Ericaria zosteroides* (C. Agardh) Molinari and Guiry (basionym: *Cystoseira zosteroides* C. Agardh) dwelling down in the lower zone of the circalittoral (down to 80 m depth) has never been investigated. *Ericaria zosteroides* is an endemic perennial and habitat-forming Mediterranean species and one of the most representative of the brown seaweeds of deep-water assemblages (Giaccone, 1973; Ballesteros, 1990; Ballesteros et al., 2009; Boudouresque et al., 2017). In addition to the uprooting of entire individuals by storms, the season's shedding of branches constitutes a vector for long-distance dispersal of eggs. During the last decades, anthropogenic disturbances (e.g. uprooting by fishing gear, increasing water turbidity, Thibaut et al., 2005; 2015), episodic climatic events and storms (Capdevila et al., 2015) have drastically fragmented the populations of this species, with possible repercussions on the connectivity among

populations. Furthermore, rising seawater temperatures due to climate change could have dramatic repercussions on the population recovery of this vulnerable species (Capdevila et al., 2019). As reported for other *Ericaria* species from the upper sublittoral zone, *E. zosteroides* is also able to colonize artificial habitats (Bonhomme et al., 2014, Thibaut pers. obs.) at distances greater than the dispersal scale of its zygotes (Capdevilla et al., 2018). This can indicate that occasional long-distance dispersal events have a great influence on the connectivity of *E. zosteroides*.

In the present study, we sampled ten sites situated less than 1 km to several hundreds of km apart to investigate the pattern of spatial structure at different scales. At the local scale (a few kilometers), we aimed at assessing the potential for distance dispersal connecting the different sites, perhaps driven by ocean currents. We therefore conducted inference of genetic structure with RAD sequencing (dd-RADseq; Peterson et al., 2012) in a complementary way with Lagrangian modeling to unravel how gene flow is mediated by drifting of fertile parts of the thallus. This multidisciplinary approach enabled us to answer the following questions: (i) Is *E. zosteroides* spatially structured at different spatial scales into genetically distinct populations? (ii) What is the respective influence of demography, oceanographic currents, and geographical distance on the genetic structure? (iii) Can we identify the sources of recruitment?

## 2.2 Material and methods

### 2.2.1 Sampling and demographic structure of populations

Two hundred and sixty-one individuals of *Ericaria zosteroides* were sampled at 10 sites in 2017 and 2018 by scuba diving, and a small fragment of the thallus (2-3 cm) was preserved in silica gel until DNA extraction. Populations were sampled in the north-western Mediterranean Sea from the Bay of Marseille (western Provence) to Corsica (Table 2.1, Figure 2.1).

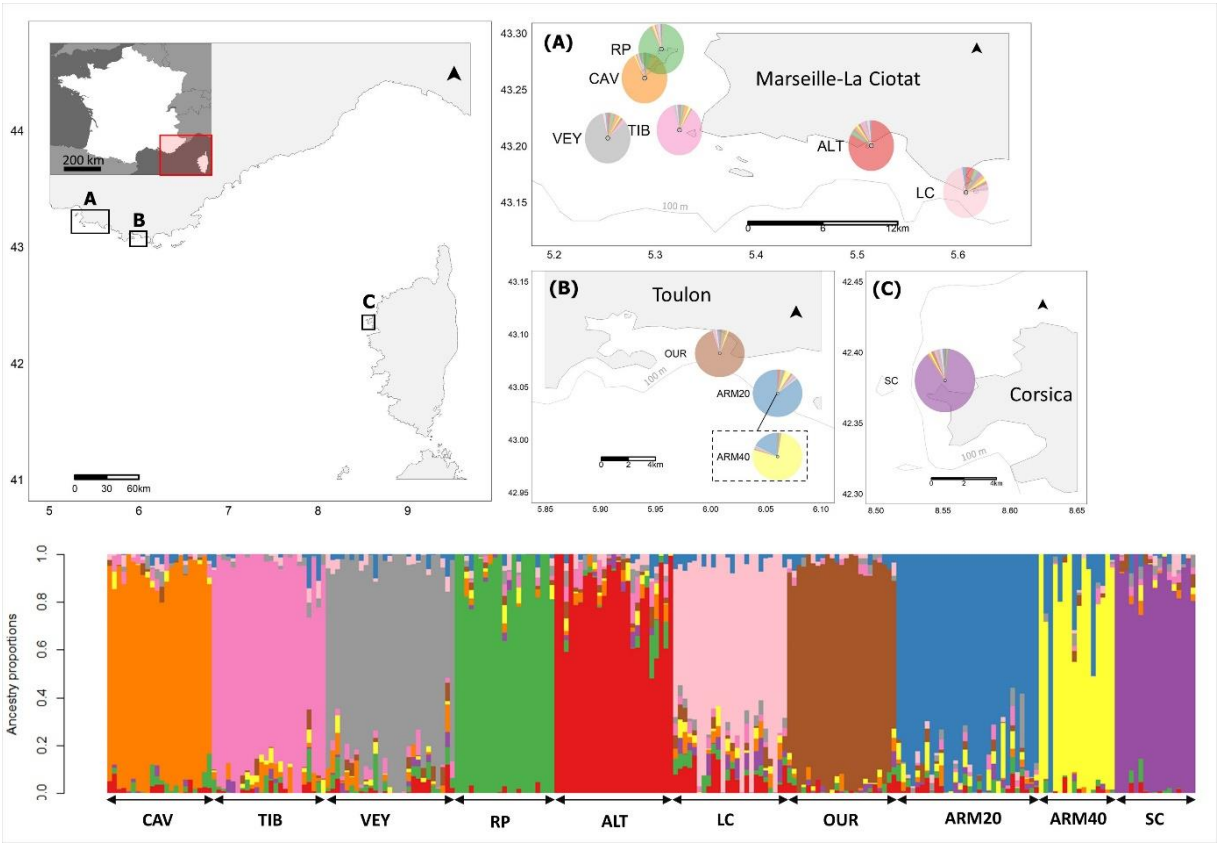



**Figure 2.1**. Sampling map of *Ericaria zosteroides* and genetic ancestry estimates derived from sNMF at K =10 clusters using the set of 10 755 putatively neutral SNPs. The pie plots indicate ancestry averaged over individuals per site, with different colors indicating admixture. The dotted box indicates that ARM20 and ARM40 populations have the same geographical position but differ by depth. Genetic ancestry is also reported at the individual level.

We included in the study a population settled on artificial substrates, the subsea-pipelines of the company Alteo (ALT) (Bonhomme et al., 2014). Furthermore, we focused on a site, L'Armoire, which is an isolated rocky bank from 22 to 40 m depth surrounded by sand where a continuous population of *E. zosteroides* is thriving over the whole of all the depth range of the bank (ARM). This rare spatial configuration offered a great opportunity to study vertical genetic connectivity. For this purpose, we separately sampled individuals at 22-24 m depth (ARM20) and 38-40 m depth (ARM40). The density of individuals of *E. zosteroides* was measured using random quadrats (1 $m^2$). A total of 214 quadrats were performed during the study. The length of the major axis of each individual is measured *in situ* with a plastic ruler and an accuracy of ± 5 mm. Individuals of less than 1 cm are considered as recruits of the year. From this length distribution, we defined cohorts according to discrete classes of 1 cm (see Hereu et al., 2008 for details); in total 289 individuals were measured. Because of the

difficulties of working at these depths, we could not measure the length of the axis at ARM40 and LC, and we could not measure the demographic parameters at SC. Densities were compared with a non-parametric Kruskal-Wallis test followed by a post-hoc pairwise Wilcoxon test. The distribution of the length was compared pairwise with a Kolmogorov-Smirnov test. The mean size of the main axis was compared using ANOVA followed by a pairwise Tuckey's test. In this study we define three scales: the fine scale for individuals within the same site placed up to 20 m apart, the local scale from Marseille to La Ciotat up to 29 km, and the regional scale from western Provence to Corsica, with a minimum distance by sea of 216 km between these two areas.

**Table 2.1**. Genetic diversity within populations of *Ericaria zosteroides* based on 10 755 neutral SNPs.

| Population | Code | Depth (m) | n | $H_o$ | $H_e$ (± SE) | $F_{IS}$ |
|---|---|---|---|---|---|---|
| **Western Provence (Marseille -La Ciotat)** | | | | | | |
| La Roche percée (Frioul Archipelago) | RP | 26-28 | 21 | 0.07 | 0.08 (0.001) | 0.04 ‡ |
| Cap Caveau (Frioul Archipelago) | CAV | 28-29 | 22 | 0.11 | 0.11 (0.001) | 0.02 ‡ |
| Tiboulen | TIB | 28-30 | 24 | 0.11 | 0.10 (0.001) | 0.00 |
| Le Veyron | VEY | 25-36 | 27 | 0.11 | 0.11 (0.001) | 0.01 ‡ |
| Conduite Alteo | ALT | 31 | 25 | 0.10 | 0.11 (0.001) | 0.04 ‡ |
| La Ciotat | LC | 45 | 24 | 0.12 | 0.12 (0.001) | 0.00 |
| **Eastern Provence (Toulon)** | | | | | | |
| Les Oursinières | OUR | 36 | 23 | 0.11 | 0.11 (0.001) | 0.01 ‡ |
| L'Armoire_20 | ARM20 | 22-24 | 30 | 0.09 | 0.09 (0.001) | -0.02 † |
| L'Armoire_40 | ARM40 | 38-40 | 16 | 0.09 | 0.09 (0.001) | -0.02 † |
| **Corsica** | | | | | | |
| Scandola | SC | 30-35 | 17 | 0.13 | 0.12 (0.001) | -0.01 |

n, number of individuals successfully genotyped; $H_o$, observed heterozygosity; $H_e$, expected heterozygosity and SE among loci; $F_{IS}$, average inbreeding coefficient. A significant deficit or excess in heterozygotes (P <0.001) is indicated as "‡" or "†", respectively.

### 2.2.2 DNA extraction and SNP discovery

Total genomic DNA was isolated with the Nucleospin® 96 plant kit II (Macherey-Nagel, Düren, Germany), according to the manufacturer's protocol. Then, SNP discovery was performed using double digest RAD sequencing libraries (dd-RADseq; Peterson et al., 2012). The protocol used to build ddRAD-seq libraries for sequencing is identical to that previously described in Reynes et al. (2021). Two hundred and seventy-one individuals of *Ericaria*

*zosteroides* were randomly distributed across two sequencing libraries, including ten replicates (one per sampling site), and 100 ng of genomic DNA per individual was double digested using the PstI and HhaI enzymes (NEB). Finally, dd-RADseq libraries were sequenced with the paired-end method (2×150 bp) on an Illumina Hiseq 4000 platform (Génome Québec Innovation Centre, McGill Univ., Montreal, Canada).

Paired-end raw reads were checked for control quality using FASTQC v.0.11.7 (Andrews, 2010) and trimmed to 137 bp after the removal of adapters sequencing by Trimmomatic (Bolger et al., 2014). High-quality reads were demultiplexed using Stacks's process_radtags (Catchen et al., 2011). Additionally, one individual genome of *E. zosteroides* was paired-end sequenced in short reads (2 x 150 bp) on an Illumina NovaSeq (Genewiz, Inc, South Plainfield, NJ). Then, *de novo* genome assembly was performed using SOAPdenovo v.2.41, with a minimum contig length of 1000 bp. The resulting draft assembly of the *E. zosteroides* nuclear genome was used as a reference genome to map paired-end reads of the RAD-sequencing experiment using the BWA-mem algorithm in BWA v.0.7.17 (Li and Durbin, 2009). Aligned reads were then filtered using Sambamba v.0.6.6 (Tarasov et al., 2015) to discard all possible multi-mapped reads to the genome. RAD loci were built from sets of alignments using the reference mode of the Stacks v2.4 pipeline (Rochette et al., 2019), then SNP discovery was done with the Bayesian genotype caller model (BGC; Maruki & Lynch, 2017) implemented in the Stacks v2.4. pipeline.

### 2.2.3 SNP filtering

Initially, we ran the population module of Stacks v2.4 requiring a locus to be genotyped at a frequency greater than 80% in at least 8 out of 10 sampling sites, while deleting SNPs with more than 50% of heterozygosity. Then, stringent quality filters were applied to the dataset (see Supporting Information Table B.1) using PlinkQC, the R version of PLINK v1.9 (Chang et al., 2015). This procedure consisted of i) discarding individuals outlying missing genotype frequency > 25% and those outlying three SD away from the mean rate of heterozygosity ii) dropping SNPs with missingness rate > 5% and SNPs with minor allele frequency (MAF) < 1%. Following this step, SNP error rate between replicates ranged from 0.009 (OURS) to 0.03 (SC) with an average rate of 0.01. Also, individual relatedness was checked using the KING-robust kinship estimator (Manichaikul et al., 2010), implemented in the Bigsnpr v1.3.0 R package

(Privé et al., 2018). We applied a KING-relatedness cutoff of 0.354 to discard duplicate samples and up to a third degree of relationship. Finally, SNP clumping was performed using plink to remove associated SNPs ($r^2 > 0.3$) using the MAF as a statistic of importance. At the end of this filtering process, the dataset comprised 11 067 SNPs among 229 individuals, with a genotyping rate of 98%.

### 2.2.4 Species delimitation

*Ericaria zosteroides* is present in sympatry with other species of *Cystoseira sensu lato*, mainly *Gongolaria montagnei* var. *compressa* and *E. funkii* (e.g. Hereu et al., 2008; Ballesteros et al., 2009). As outlined by Pante et al. (2015a), evolutionary units (e.g. species and populations) must be properly delimited to conduct a robust inference of population connectivity (see also Boero, 2010; Boudouresque et al., 2020). Hence, we tested for the existence of potential cryptic species in *E. zosteroides*, and subsequently validated our morphological inspection by performing an analysis of species delimitation (see Supporting Information Materials B.1).

### 2.2.5 Outlier loci

We searched for loci potentially influenced by selection, to build a neutral dataset used to infer population structure and connectivity. The putatively non-neutral (selected) loci were analyzed separately to test for adaptation to local conditions (see Supporting Information Materials B.2). We applied three tests to minimize the number of false positives loci assumed to be under selection. First, we used the $F_{ST}$ method described in Martins et al. (2016) and implemented in the LEA R package (Frichot et al., 2014, Frichot and François, 2015). This method starts by defining the optimum number of populations with ancestry estimates (Frichot et al., 2014). $F_{ST}$ values are then converted to z-scores and significant outliers are detected after controlling for a False Discovery Rate (FDR) of 0.05. Secondly, we used the Pcadapt v.4.3.1 R package (Luu et al., 2017). Pcadapt analyses the population structure by PCA, then it estimates to what extent each SNP is related to the first K principal components to identify outlier loci. Here, outlier tests were performed from the first six principal components, according to the percentage of explained variance for each PC. P-values were then computed by making a Gaussian approximation for each PC and by estimating the

standard deviation of the null distribution. Finally, P-values were corrected at 0.05 FDR. The third approach used here, implemented in Bayescan v. 2.1, estimates directly the probability that each locus is subject to selection using a Bayesian method (Foll and Gaggiotti, 2008). The test was run with default settings (5 000 iterations, 10 thinning intervals, 20 pilot runs - 5 000 iterations each, 10 prior odds). SNPs were identified as significant outliers after FDR correction of 0.10.

For the following analyses of genetic variation, we selected the set of putatively neutral SNPs by excluding those identified as an outlier by at least two out of the three methods. The set of putatively outlier SNPs used for the following analyses of local adaptation (see Supporting Information Materials B.2) was built by keeping all SNPs identified as an outlier by the three methods.

### 2.2.6 Population genetic variation

Spatial population structure was assessed with putatively neutral SNPs at fine (one site, two depths), local (six sites) as well as regional scales (ten sites). Two complementary approaches were used. First, a principal component analysis (PCA) was performed by singular value decomposition (SVD) of the centered genotype matrix, as implemented in Bigsnpr v. 1.3.0 and Bigstatsr v. 1.2.3 R packages (Privé et al., 2018). The missing genotypes were imputed according to the mean allele frequencies using the function *snp_fastImputeSimple* before performing the PCA. Second, the function sNMF of the LEA R package (Frichot et al., 2014, Frichot and François,2015) was used to estimate individual ancestry proportions. SNMF was run with 5 000 iterations and 10 repetitions for K value from 1 to 15. The best K value was defined using the cross-entropy criterion (see Supporting Information Figure B.1). Genetic differentiation was measured by calculating pairwise $F_{ST}$ values (Weir and Cockerham, 1984) between populations using the Genepop v. 4.7.5 R package (Rousset, 2008) and performing an exact test of genic differentiation (Raymond and Rousset, 1995) using the Markov chain method with default parameters, then a combination of all tests across loci (Fisher's method) was performed for each population pair.

Genetic diversity within populations was estimated using Genepop by estimating the average observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), and $F_{IS}$ (Weir and Cockerham, 1984). We tested for differences in $H_e$ between populations with a pairwise

Wilcoxon test adjusted for multiple comparisons (Bonferroni adjustment). Positive and negative $F_{IS}$ values were observed, so we tested departure from panmixia across all loci and populations using the Global Hardy-Weinberg test [Score (U) test] (Rousset and Raymond, 1995), implemented in Genepop. The test was performed twice: once by considering the alternative hypothesis of heterozygote excess and a second time with heterozygote deficiency. The P-value of each test was approximated by a Markov chain (MCMC) algorithm with the following settings: dememorization: 10 000, batch: 100, iterations per batch: 5 000. Then, a multiple score test (Rousset & Raymond, 1995) was performed to obtain a global P-value per population.

### 2.2.7 Individual assignment analysis

Individual assignment tests were used to test for contemporary gene flow and admixture between shallow (ARM20) and deep (ARM40) populations of Armoire, as well as to investigate the potential source population(s) for the colonization of ALT. Assignment tests were conducted using the Assigner v. 0.5.7 R package (Gosselin et al., 2016) and the implemented gsi_sim function (Anderson et al., 2008). We used the Training, Holdout & Leave-one-out (THL) method to avoid strong grading bias related to the fact that the selection of informative loci and population assignment is performed from the same set of individuals (Anderson, 2010). Specifically, individuals were distributed into two sets: a training set (including 50% of individuals per site), a holdout-set (including the other half of individuals per site). These analyses were standardized by considering the minimal sample size (i.e. 16 individuals). The selection of informative genetic markers was performed from the training dataset, thus identifying 14 sets of neutral SNPs ranked based on decreasing overall $F_{ST}$ values (50, 100, 200, 500, 1 000, 2 000, 3 000, 4 000, 5 000, 6 000, 7 000, 8 000, 9 000, and all SNPs). Finally, the individual assignment was performed using separately the 14 sets of SNPs for the hold-out set of individuals, and confidence intervals were obtained by bootstrapping individuals (i.e. 16 individuals per population were randomly chosen 10 times).

### 2.2.8 Hydrodynamic modeling

Oceanographic connectivity was assessed at local scale, from the Bay of Marseille to La Ciotat (western Provence), which is an area where the hydrodynamic processes showed

high spatio-temporal heterogeneity (Pradal & Millet, 2006). The MARS3D model (3D hydrodynamic Model for Applications at Regional Scale, IFREMER, Lazure & Dumas, 2008) in its RHOMA configuration (Pairaud et al., 2011; Fraysse et al., 2013) was used to generate three-dimensional oceanic circulation patterns. For our study, the model used 120 x 252 horizontal grid points at 400 m resolution and 30 sigma layers, with a corresponding time step of 30 s, took into account the forcing by the North-Western Mediterranean general circulation, using a nesting strategy with the large scale MARS3D-MENOR configuration (1.2 km resolution) (Nicolle et al., 2009). We expected that oceanographic connectivity was highly variable according to the year of simulation, as such Lagrangian modeling was performed over seven years from the ocean circulation patterns of the years 2008 to 2017. Hydrodynamic model were forced hourly by different atmospheric models (MM5, Arome, and WRF models). To ensure the robustness of our results, we focused on this long period considering a wide variety of typical atmospheric forcing and oceanic features of the region of Marseille in the calculation of the mean pattern of oceanographic connectivity. Notably, strong NW and SE winds (Pradal & Millet, 2006), the intrusion of the Northern Mediterranean Current (Barrier et al., 2016; Ross et al., 2016; Millet et al., 2018), and diluted waters from the Rhone River plume (Fraysse et al., 2014).

### 2.2.9 Lagrangian modeling

We first tested if dispersal of fragments of fertile thalli allows genetic connectivity between spatially distant populations, which requires that fertile fragments be transported to another population where they produce viable gametes which contribute to fertilization and reproduction in the target population. Detached secondary branches with receptacles are commonly observed around the *E. zosteroides* populations. Lagrangian simulations were performed with Ichthyop v. 3 targeting the breeding season of *E. zosteroides*, known to take place from May to September. Simulations were performed on the high-performance computing (HPC) cluster of the OSU Pythéas Institute (Marseille) and were repeated every day for the first 28 days of May, June, July, and August, amounting to a total of 784 release events per site over seven years. ICHTHYOP was parameterized to constraint dispersal of non-buoyant particles around the seafloor. Specifically, 1 000 particles were released each day from each of six locations within a radius of 400 m at the depth (± 2 m) of the study sites. The

computational time step (dt) was set to 400 s and particle position was recorded in the NetCDF output files every 1.33 hours. Pairwise oceanographic connectivity probabilities were defined as $P_{ij} = N_{i-j}/N_i$ where $N_i$ is the total number of released particles from site *i* and $N_{i-j}$ the total number of released particles from site *i* arriving in site *j* after 30 days of transport time; these probabilities were computed in MATLAB. To take into account the life span of the species, particles recorded at more than 100 m depth were not considered in the computation of $P_{ij}$. Finally, the probabilities of oceanographic connectivity between all pairs of sites were averaged over seven years of simulation to obtain the mean connectivity matrix and the mean time of transport.

### 2.2.10 Environmental predictors of genetic structure

To test the effects of environmental parameters on the genetic structure of *E. zosteroides,* we selected four sets of explanatory variables: oceanographic connectivity, geographical distance, population density and seawater temperature. Geographical distances were computed as distance-based Moran's eigenvector maps or dbMEM (Dray et al., 2006), which accounts for the effect of spatial variation among sites. To do this, pairwise Euclidean distances between sites were calculated using the codep v. 0.9.1 R package, then the distance matrix was used as input data for computing dbMEM eigenvectors with the adespatial v.0.3.8 R package. We kept the default settings for truncations. Oceanographic connectivity was computed as Asymmetric Eigenvector Maps or AEM (Blanchet et al., 2008), which turned out to be extremely useful for modeling asymmetric processes, such as ocean currents, and testing their effects on genetic structure (e.g. Benestan et al., 2016; Xuereb et al., 2018a; Coscia et al., 2020). For this purpose, we transformed the mean oceanographic connectivity matrix (Figure 2.7) from Lagrangian simulations into a nodes-to-edges matrix, which records the presence/absence of connectivity link between nodes. Specifically, fifteen connectivity links (i.e. edges) were reported among the six study sites considered (i.e. nodes). Edges were weighted by the probabilities of dispersal, and this weighted matrix was then used as input data to compute AEM predictors with the adespatial v.0.3.8 R package. Seawater temperature predictors were computed by extracting the temperature variables of the MARS3D model. Specifically, we considered minimum (Min_T) and maximum (Max_T) seawater temperatures, the annual mean (Mean_T), and variance (Var_T). These parameters were extracted from the

site's coordinates accounting for the depth of populations and were averaged over seven years of simulations. For the density predictor, we used the mean value of populations reported in our study (Table 2.2).

### 2.2.11 Distance Based RDA: linking environment to genetic structure

At a local scale (Marseille-La Ciotat), partial and global distance based RDA (Legendre and Anderson, 1999) were used to investigate, respectively, the separate and joint effects of oceanographic connectivity (AEM1; AEM-2; AEM-3; AEM-4; AEM-5), geographical distance (dbMEM-1), population demography (density), and seawater temperature (Min-T, Max-T, Mean-T, Var-T) on the response variable: individual coordinates of principal coordinate analysis (PCoA) based on the pairwise Euclidean genetic distance matrix. These analyses were carried out using the Vegan v. 2.5.6 R package (Oksanen et al., 2017) and repeated twice with the set of 10 755 neutral and 12 outlier SNPs. We considered all independent predictors, with Variance Inflation Factor (VIF) below 10 to avoid multicollinearity effects. Following the VIF threshold, Min-T and Var-T predictors were excluded in partial and global distance based RDA while the density predictor was only excluded in the whole model. Then, we applied stepwise forward selection to select the best predictors using the *ordiR2step* R function. The *ordiR2step* performs model choice solely on the adjusted coefficient of determination ($R^2_{adj}$) and selects the model with the highest one. The statistical significance of partial and global Distance Based RDA and their respective predictors were determined by analysis of variance (ANOVA) with 1000 permutations. Finally, $R^2_{adj}$ and the p-value for the F-tests were used to evaluate the validity of the models.

## 2.3 Results

### 2.3.1 Demographic structure

Populations of *Ericaria zosteroides* are spatially isolated in the studied zone, without a continuum between them. The density of populations was highly variable and significantly different between sites (KW test, H=125.25, df (7), p<0.001) (Figure 2.2). The mean population density ranged from 1.23 ± 0.25 ind.m$^{-2}$ recorded in the Bay of Marseille at RP, to 11 ± 0.62 ind.m$^{-2}$ at ARM20 (Table 2.2). The size distribution of individuals was statistically only similar

in 8 out of the 21 pairwise comparisons (Kolmogorov-Smirnov pairwise test, p>0.05) (see Supporting Information Table B.2). Only one cohort was observed in five populations (TIB, VEY, OUR, ALT, ARM20), two cohorts in the RP population, and three cohorts in CAV (see Supporting Information Figure B.2). The mean size of the main axis ranged from 3.5 ± 0.2 cm to 7.4 ± 0.5 cm in ALT and OUR, respectively. The mean size differed among sites ($F_{6, 282}$ = 11.91, p<0.001). Individuals at ALT were significantly smaller than all the other populations studied, except for the comparison with TIB (Tukey test) (Figure 2.2). According to the estimated growth rate, which ranged from 0.5 to 0.9 cm.year$^{-1}$ (Ballesteros et al., 2009; Navarro et al., 2011), the maximum size of the main axis seems to indicate that the oldest individuals reach 8 to 14 years on the pipeline (ALT), against 11 to 31 years in other populations (Table 2.2).

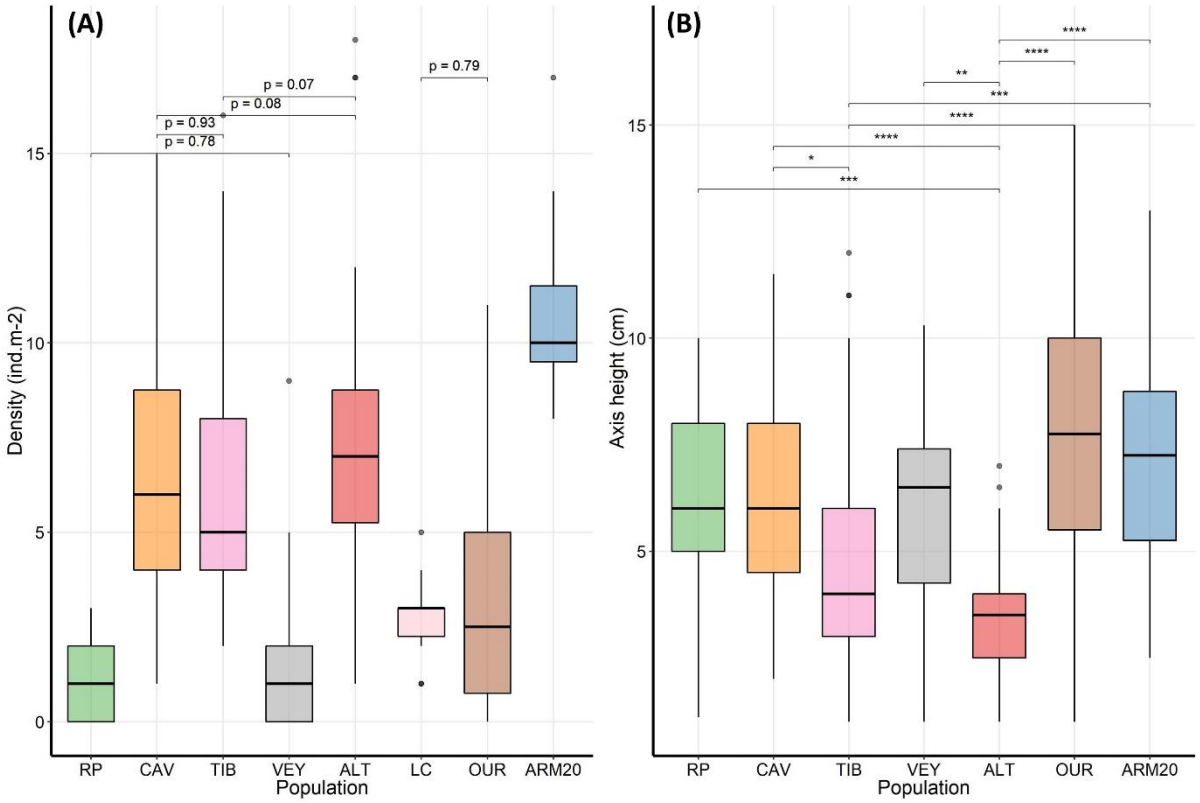



**Figure 2.2**. Demographic structure of populations of *Ericaria zosteroides*. (A) spatial distribution of the population density (ind.m$^{-2}$), all pairwise comparisons are significant (pairwise Wilcoxon test, p<0.001), except for those indicated by p>0.05. (B) Distribution of the height of the main axis of individuals, the significant difference between populations (Tukey test, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001) are reported.

**Table 2.2**. Demographic structure of populations of *Ericaria zosteroides*.

| Density (ind.m$^{-2}$) | | | | Axis height (cm) | | | |
|---|---|---|---|---|---|---|---|
| Population | Code | n | Average (± SE) | n | Average (± SE) | Max | Age (years) |
| **Marseille -La Ciotat** | | | | | | | |
| La Roche percée | RP | 22 | 1.23 (0.25) | 22 | 6.37 (0.49) | 10 | 10.9-20 |
| Cap Caveau | CAV | 42 | 6.14 (0.52) | 63 | 6.37 (0.37) | 11.5 | 12.5-23.5 |
| Tiboulen | TIB | 35 | 6.14 (0.57) | 58 | 4.82 (0.46) | 12 | 13-24.5 |
| Le Veyron | VEY | 40 | 1.32 (0.27) | 35 | 5.92 (0.37) | 10.3 | 11.2-21 |
| Alteo | ALT | 30 | 7.73 (0.73) | 41 | 3.55 (0.25) | 7 | 7.6-14.3 |
| La Ciotat | LC | 10 | 2.80 (0.39) | *NA* | *NA* | *NA* | *NA* |
| **Toulon** | | | | | | | |
| Les Oursinières | OUR | 20 | 3.15 (0.69) | 40 | 7.38 (0.75) | 15 | 16.3-30.6 |
| L'Armoire_20 | ARM20 | 15 | 11.0 (0.62) | 30 | 7.23 (0.67) | 13 | 14.1-26.5 |

The density of populations (n, number of quadrats per site; Average, average density, and SE), the size of individuals (n, number of individuals; Average, average with SE; and Max, maximum size of the main axis), and Age: estimate range of the age of individuals according to a growth rate estimation for the axis of 0.5 to 0.9 cm.year$^{-1}$. NA indicates not available.

### 2.3.2 Bioinformatic and genotyping

The draft genome of *Ericaria zosteroides* was sequenced (51.98 Gb of raw data), and *de novo* assembled into 123 092 contigs of a mean length of 2 872 bp. The ddRAD-sequencing experiment enabled us to generate and filter 974 562 924 paired-end high-quality reads among the 261 individuals of *E. zosteroides*, consisting of an average of 3 596 173 (± 1 356 820 SE) high-quality reads per individual. On average, 75.50 % (min: 27.98 %; max: 91.34 % per individual) of ddRADseq reads were successfully mapped to the *E. zosteroides* genome. Following the SNP calling process by Stacks v.2.4, mapped reads were assembled into 314 538 loci in which 721 374 SNPs were identified. The SNP filtering procedure (see Supporting Information, Table B.1) reduced the initial set of 721 374 SNPs to 21 606 SNPs, while also reducing the number of individuals from 271 to 229. Finally, clumping on MAF ($r^2 > 0.3$) allowed retention of 11 067 SNPs.

From the set of 11 067 SNPs, BayeScan detected 39 outlier SNPs, while pcadapt and sNMF identified 895 and 482 outliers, respectively. Twelve outliers were shared between the three outlier tests (Figure 2.3). Furthermore, 312 SNPs detected by at least two of the three methods were excluded from the set of 11 067 SNPs and those remaining (10 755 SNPs) constitute the neutral dataset.




**Figure 2.3**. Venn diagram denoting overlap between SNPs identified as an outlier with sNMF, Pcadapt and Bayscan. The set of putative selected SNPs was built by considering the 12 SNPs shared among these three methods.

### 2.3.3 Species delimitation

The results of the analysis of species delimitation are presented in Supplementary Materials B.1. All samples identified as *E. zosteroides* included in this analysis were grouped in a clade. The divergence among samples from the different populations analyzed here was much lower than the divergence among samples from different species.

### 2.3.4 Genetic variation within populations

The genetic diversity ($H_e$) of most of the populations differed among sites (Pairwise Wilcoxon Test, Bonferroni adjusted: $p < 0.001$), except for 8 of the 45 pairwise comparisons (see Supporting Information, Table B.3). The lowest levels of observed ($H_o$) and expected ($H_e$) heterozygosity were recorded in the population of RP ($H_o$ = 0.07, $H_e$ = 0.08) in the Bay of Marseille, while the highest ones were reported in Corsica at SC ($H_o$ = 0.13, $H_e$ = 0.12). The level of genetic diversity ($H_e$) on the pipelines (ALT) was similar to those reported in neighbouring populations, such as CAV (pairwise Wilcoxon test, Bonferroni adjusted: $p = 0.86$). On the isolated Armoire reef bank, the shallowest (ARM20) and deepest (ARM40) sites did not show any significant difference in genetic diversity (see Supporting Information, Table B.3), but the average values of $H_e$ and $H_o$ were low in this locality compared to other sites (Table 2.2.1). We observed a significant excess in heterozygotes at ARM20 and ARM40, while five of the ten populations were characterized by significant deficits in heterozygotes (Table 2.2.1).

The mean $F_{IS}$ ranged from slightly negative values at ARM20 ($F_{IS}$ = -0.02) and ARM40 ($F_{IS}$ = -0.02) to positive values, notably at RP ($F_{IS}$ = 0.04) and ALT ($F_{IS}$ = 0.04).

### 2.3.5 Genetic structure and assignment tests

The overall estimated $F_{ST}$ was 0.09, and it was slightly lower on a local scale within the Bay of Marseille ($F_{ST}$ = 0.08). Populations were significantly genetically differentiated from each other (exact G test for genic differentiation, P < 0.001), and pairwise $F_{ST}$ ranged from 0.05 to 0.14. The highest $F_{ST}$ values were observed for the comparison of the RP population with other ones (Figure 2.4). The lowest genetic differentiation ($F_{ST}$ = 0.05) was observed between CAV and TIB populations, as well as between shallow (ARM20) and deep (ARM40) populations at the Armoire site (Figure 2.4).

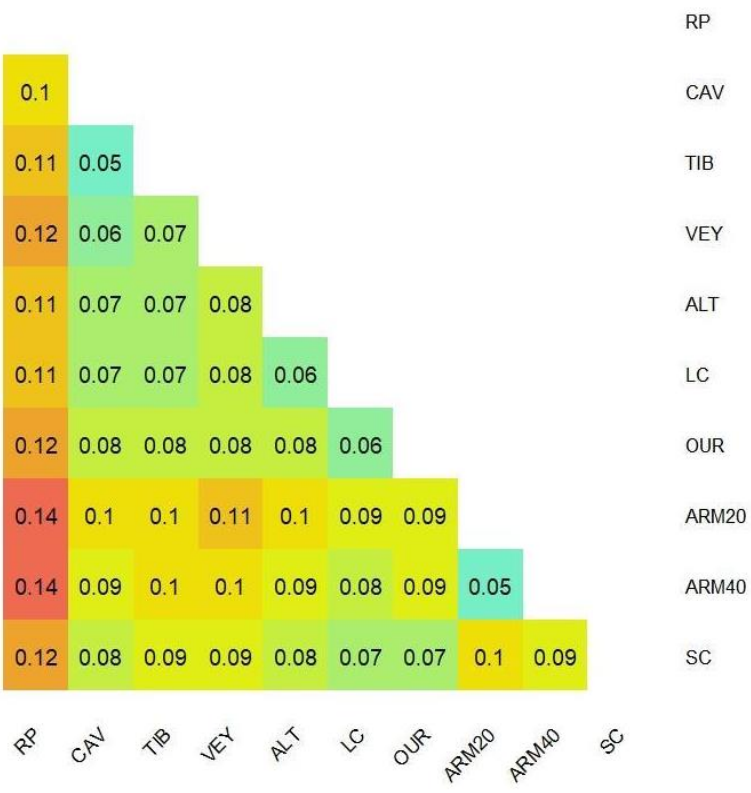



**Figure 2.4**. Heatmap of pairwise $F_{ST}$ between populations of *Ericaria zosteroides* using the set of 10 755 SNPs putatively neutral SNPs. Pairwise genetic differentiation was reported as significant among populations (exact G test for genic differentiation, p<0.001).

The principal component analysis (PCA), and sNMF clustering gave similar results and were congruent with the $F_{ST}$ values. The first axis of PCA (PC1, 19% of the variance) allowed a distinction between ARM20 / ARM40 individuals and other ones, while the second axis (PC2, 13% of the variance) discriminated the population of RP from other ones (Figure 2.5A). Additional PC axes showed the separation of the other populations (see Supporting

Information, Figure B.3). By performing the PCA only on individuals from the Bay of Marseille, ALT, and LC (Figure 2.5B), the first axis (PC1, 18% of the variance) differentiated individuals from RP from other sites, while the second axis (PC2, 17% of the variance) showed differentiation between three groups: one with individuals from ALT and LC, the second with individuals of CAV and TIB, the third with individuals of VEY. However, one individual of VEY was grouped with those of CAV and TIB.

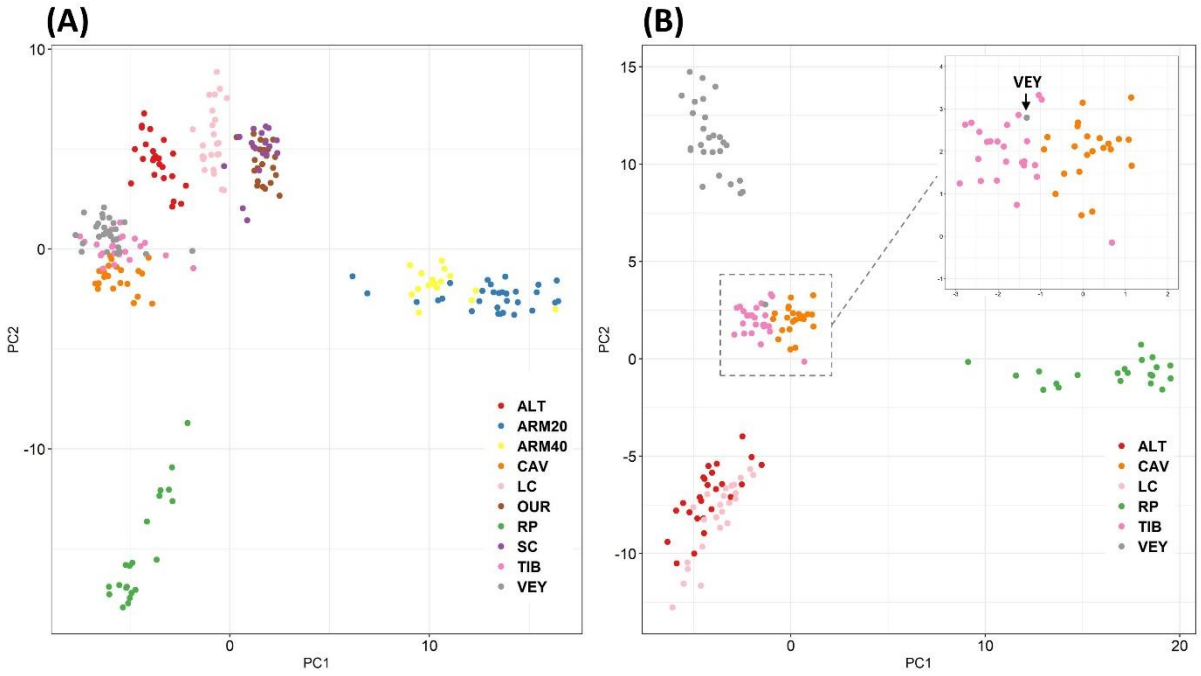



**Figure 2.5**. Principal component analysis (Axes 1 and 2) computed using the set of 10 755 putative neutral SNPs (A) with all individuals (n = 229) and (B) restricted to individuals (n = 143) sampled from the Bay of Marseille, Cassis to La Ciotat.

According to the minimum cross-entropy criterion, K= 10 seemed to be the most informative number of clusters in sNMF (see Supporting Information, Figure B.1). All individuals were clustered according to their populations of origin, except at ARM40 for which 17% of admixture was detected with ARM20. The admixture rate of ALT with LC was relatively low (6%) but higher than that observed with other populations (admixture range: 2-3%) (Figure 2.1).

The proportion of individuals assigned to their populations of origin reached 100% using the set of 10 755 neutral SNPs, except for ARM40 for which one individual was assigned to ARM20 (see Supporting Information, Figure B.4). In contrast, no individual of ARM20 was assigned to ARM40. The success of the assignment increased with the number of SNPs

considered (see Supporting Information, Figure B.5). The overall assignment success ranged from 73 % with the top-ranked 50 SNPs to 99 % with the top-ranked 500 SNPs, and reached a plateau at 100% by using more than 500 SNPs.

### 2.3.6 Oceanographic connectivity

According to the results of Lagrangian simulations, dispersal events among populations of *E. zosteroides* were scarce. The mean probability of oceanographic connectivity (P) was below 0.01 for most of the pairwise comparisons (Figure 2.6A). Across the Bay of Marseille, the simulations indicated that the population of RP lacked oceanographic connectivity with other populations. The highest probabilities of connection, still very low, were observed among the CAV, TIB, and VEY populations (P ranging from 0.04 to 0.06). Regarding the connectivity with the ALT population on pipelines, there was only a unidirectional connection from LC to ALT (Figure 2.7), with P = 0.02, and the corresponding mean time of transport (around ten days; Figure 2.6B) was the lowest for the comparisons tested here.

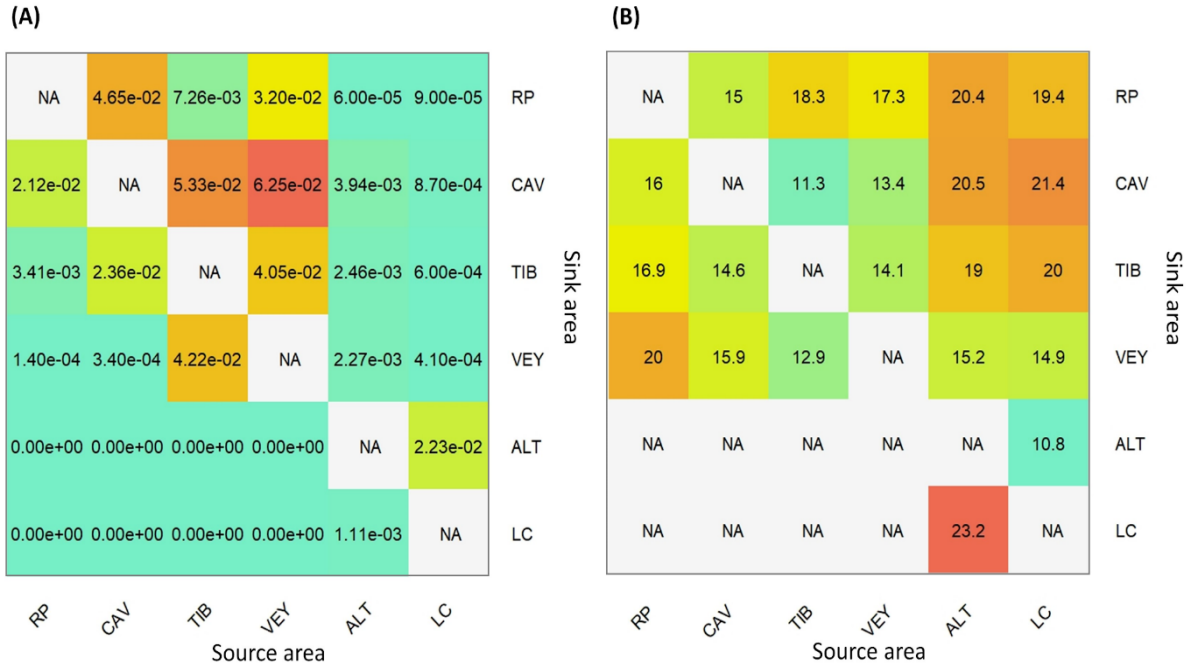



**Figure 2.6**. Pairwise oceanographic connectivity from the source area to sink area. (A) The average probability of dispersal between populations with (B) the associated mean time of particle transport (in day). NA value of the diagonal matrices denotes that self-recruitment was not considered; another NA indicates the absence of dispersal between populations to compute transport time.

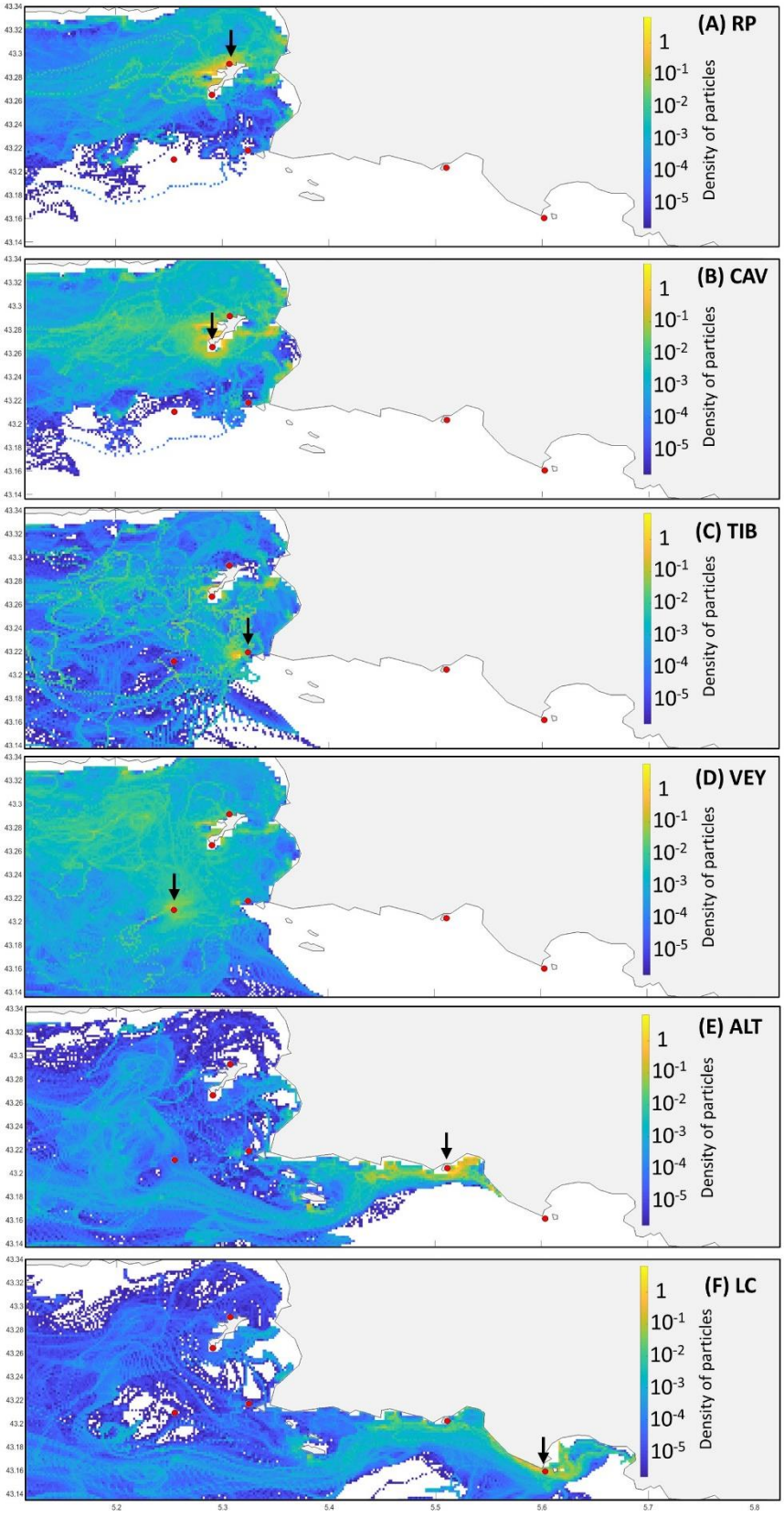



**Figure 2.7**. An example of a probability density map of the distribution of particles recorded at the local scale (Marseille - La Ciotat) over Lagrangian simulations of the year 2014. Probabilities are color-coded in each mesh cell (resolution of 200 meters), indicating no particle record (white) from low (blue) to high probability (yellow). For each panel (A-F), the black arrow indicates the source area considered among the six sites, materialized as a red dot.

### 2.3.7 Factors explaining the genetic structure at a local scale

For the neutral genetic structure of *E. zosteroides*, oceanographic connectivity relating to AEM-1, AEM-2, and AEM-4 predictors, as well as geographical distance relating to the db-MEM-1 predictor, were selected by the stepwise forward selection (Table 2.3), and included in the global distance based RDA. The global distance based RDA was significant ($P<0.001$) with an adjusted coefficient of determination ($R^2_{adj}$) of 0.11. When partitioning the respective effects of oceanographic connectivity, geographical distance, density and temperature in partial distance based RDA, all of them were significant predictors of neutral genetic structure, but oceanographic connectivity explained the greatest amount of variance with $R^2_{adj} = 0.11$ against $R^2_{adj} = 0.03$ for both geographical distance and density predictors, and $R^2_{adj} = 0.05$ for temperature predictors (Max-T and Mean-T). However, the temperature was not reported as a significant predictor of neutral genetic structure in the global Distance Based RDA (Table 2.3). The first axis of the global distance based RDA explained 30.3 % of the total variance and it was mainly explained by oceanographic connectivity with the AEM-1 predictor (Figure 2.8). This axis separated the individuals of RP from other ones. The second significant axis (accounting for 27.5% of the total variance) was predominantly defined by geographical distance with db-MEM1, secondly by oceanographic connectivity with AEM-2 and AEM-4 predictors. The geographical distance enabled the distinction between ALT-LC populations and the other ones of the Bay of Marseille, while oceanographic connectivity drove the genetic differentiation of the CAV, TIB, and VEY populations.

For the adaptive genetic structure (i.e. 12 potentially selected loci), seawater temperature (Min-T) was selected by the stepwise forward selection and included in the global Distance Based RDA. The Min-T predictor contributed to the differentiation between the deepest population (LC, depth 45 m) and the other ones (see Supporting Information Materials B.2).

**Table 2.3**. Global and partial distances based RDA (db-RDA). Environmental predictors of the neutral genetic structure selected by the stepwise forward selection (*ordiR2step*) as highlighted in bold are included in the db-RDA framework.

| | Oceanographic connectivity | Geographical distance | Demography | Seawater temperature | P model | $R^2_{adj}$ |
|---|---|---|---|---|---|---|
| Global db-RDA | **AEM-1**<br>**AEM-2**<br>AEM-3<br>**AEM-4**<br>AEM-5 | **dbMEM-1** | | Max-T<br>Mean-T | <0.001 | 0.11 |
| Partial db-RDA | **AEM-1**<br>**AEM-2**<br>AEM-3<br>**AEM-4**<br>**AEM-5** | | | | <0.001 | 0.11 |
| | | **dbMEM-1** | | | <0.001 | 0.03 |
| | | | **Density** | | <0.001 | 0.03 |
| | | | | **Max-T**<br>**Mean-T** | <0.001 | 0.05 |

All predictors retained in the models are significant at P <0.001 using ANOVA. The adjusted coefficient of determination ($R^2_{adj}$) and the P-value of the model are reported. Seawater temperature was not selected as a significant predictor in the global db-RDA.

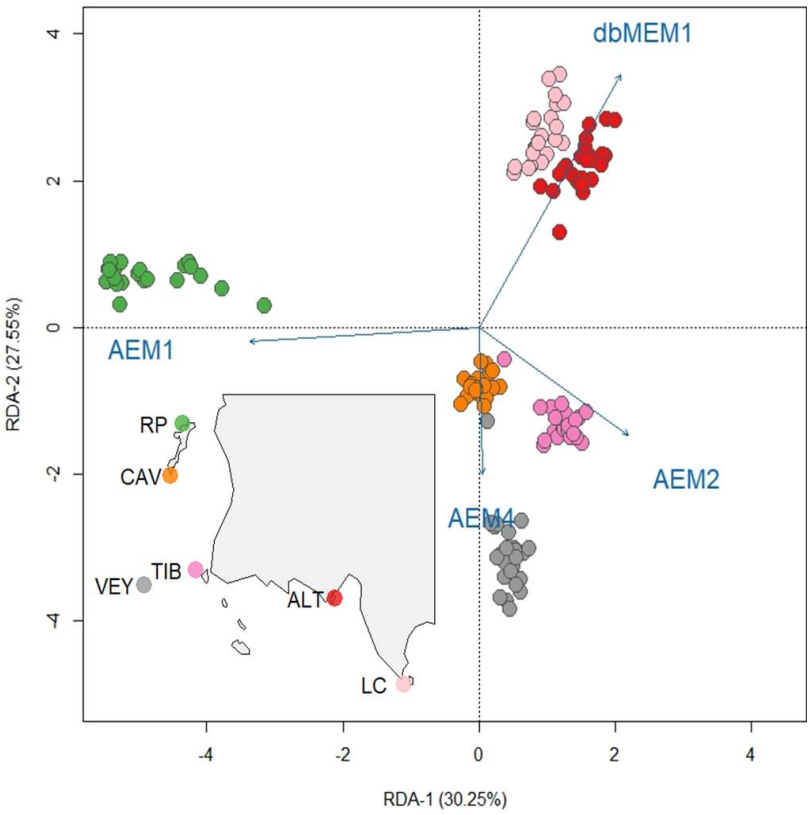



**Figure 2.8**. Distance-based RDA (db-RDA) using the set of 10 755 putatively neutral SNPs at the Marseille-La Ciotat site. The significant predictors are indicated by the directional vectors in the biplot (RDA axes 1 and 2) and individuals are represented as dots and are colored according to their population reported on the sampling map (bottom left corner).

## 2.4 Discussion

Populations of *Ericaria zosteroides* are genetically differentiated from each other, including at short distances, as reported in the Mediterranean related species *Ericaria amentacea* and *Gongolaria elegans* (synonym: *Cystoseira elegans*) (Susini et al., 2007; Thibaut et al., 2016b; Buonomo et al., 2017; Medrano et al., 2020). Such local genetic differentiation has already been observed in other brown algae, such as in the Mediterranean Kelp *Laminaria rodriguezii* (Reynes et al., 2021), the Atlantic Kelp *Laminaria digitata* (Billot et al., 2003; Robuchon et al., 2014), *Saccharina latissima* (Mao et al., 2020), *Undaria pinnatifida* (Grulois et al., 2011; Guzinski et al., 2018; Salamon et al., 2020), as well as in the Fucoids *Hormosira banksii* (Coleman et al., 2019) and *Fucus ceranoides* (Neiva et al., 2012b). These studies, from low to high-throughput genotyping approaches, point to the important genetic structuring generally observed in these species. Unexpectedly, the highest level of differentiation observed here did not correspond to the sites separated by the largest geographical distance (i.e. the Corsica sample from all other sites (~216 km)), but to the RP site from all others of the Bay of Marseille, sometimes separated by a spatial distance of only a few kilometers (~3 km). Such unexpected patterns of genetic structure can be the result of incorrect species identification (Pante et al., 2015a). For example, the differentiation of populations of the *Lithophyllum* calcareous red algae in the same area is shaped by a combination of cryptic species and intra-specific differentiation (de Jode et al., 2019). Within the *Ericaria* genus, the use of microsatellite loci challenged the traditional species delimitation in the *E. selaginoides* species complex (Bermejo et al., 2018, as *E. tamariscifolia*). However, morphological characters of *E. zosteroides* are clear-cut, and cryptic species within our sampling are unlikely as indicated by the molecular comparison with other congeneric species. All our analysis of connectivity then corresponds to intra-specific comparisons.

We clearly showed that bottom currents can play a pivotal role in the connectivity of the populations via the dispersal of fragments. Thus, the genetic structure of *E. zosteroides* was better explained by oceanographic connectivity (adjusted $R^2_{adj}$ = 11%) than geographical distance (adjusted $R^2_{adj}$ = 3%) using a variation partitioning method, a commonly used method in this kind of study (e.g. Benestan et al., 2016; Dalongeville et al., 2017; Xuereb et al., 2018a, 2018b; Coscia et al., 2020). Fucoids and kelps are usually known to disperse via sea-surface

rafting of fragments of the thallus (e.g. Coleman et al., 2011; Hawes et al., 2017; Durrant et al., 2018). Unlike the very shallow species *E. amentacea*, for which dispersal at the sea-surface has been highlighted directly with zygotes on a limited time dispersal (12 h) or with fragments entangled in floating debris or via seagull (Thibaut et al., 2016b; Buonomo et al., 2017), the dispersal of *E. zosteroides* is most probably done *via* fertile fragments of thallus rolling on the sea-floor. Dispersal of zygotes of *E. zosteroides* seems to be limited to 10 m (Capdevila et al., 2018), and vegetative reproduction does not occur in *Ericaria* spp. This mode of dispersal leads to episodic gene flow among spatially distant populations. Dispersal events above a few kilometers distance appear to be rare and require on average more than ten days of transport time. In addition to the randomness of the occurrence of currents, topography and unfavorable habitats to the development of *E. zosteroides* (*Posidonia oceanica* beds, sand, coralligenous) can become traps for these fragments and therefore considerably limit gene flow between populations, below what is suggested by our modeling approach.

The highly divergent population of RP in the Bay of Marseille is isolated by oceanographic connectivity and characterized by a low level of genetic variation and significant heterozygote deficiency. This population has the lowest individual density. All of these elements support the fact that the genetic diversity in RP is driven by strong local genetic drift and probably high self-recruitment. A meta-analysis of the genetic structure of populations in this area has shown for several species a barrier to gene flow near the Frioul Archipelago, where RP is located (Cahill et al 2017), but differences in sampling locations limit a more precise comparison with these results. The evolutionary history of the populations should be considered as well. The RP population could belong to a different genetic cluster that was previously differentiated, and the contact zone of this cluster with other populations could be trapped in this area of reduced oceanographic connectivity (Bierne et al., 2011).

More generally, the strong genetic structure of populations of *E. zosteroides* could be explained by another factor: the priority effect (De Meester et al., 2002), as suggested to explain high levels of genetic differentiation in Fucoids (e.g. Neiva et al., 2012b; Thibaut et al., 2016b). This hypothesis states that gene flow can be restricted not by dispersal but by the density and the size of individuals, which can limit new recruitments. Within *E. zosteroides* populations the recruitment is highly regulated by density-dependent effects (e.g. Ballesteros et al., 2009; Capdevila et al., 2015). If the mean density of our populations is within the range

of other *E. zosteroides* populations measured in other North-Western Mediterranean populations (Thibaut et al., 2005; Ballesteros et al., 2009; Hereu et al., 2008; Capdevila et al., 2015), the size range of the main axis is much smaller than those recorded in the MPA of Scandola in Corsica (Ballesteros et al., 2009) and Port-Cros (Hereu et al., 2008), and similar to those observed at the Balearic and Catalonia sites (Capdevila et al., 2015). The estimated age of the individuals studied here is lower than those recorded in the MPA of Scandola (Corsica) (Ballesteros et al., 2009). When individuals of *E. zosteroides* are extirpated during episodic disturbance events, the loss of individuals and free available space can promote recruitment (e.g. Navarro et al., 2011; Capdevila et al., 2015; Capdevila Lanzaco, 2017). The population studied on the isolated Armoire bank (ARM) had the highest density of individuals recorded in the study. Additionally, no recruits were observed here. The high genetic differentiation observed for this location (the second highest) may be driven by the combined effects of spatial isolation and priority effect, the latter reducing the success of zygote installation. In the case of RP, despite low density, oceanographic connectivity probably remains too low to allow recruitment. To test the effect of demography on the genetic structure of this species, a dedicated study should be performed with the analysis of more populations with different densities, if possible not correlated with the oceanographic connectivity.

The ARM site was also characterized by an excess of heterozygotes compared to panmixia, which is usually not observed in Fucoids (Coleman and Brawley, 2005; Thibaut et al., 2016b; Buonomo et al., 2017; Medrano et al., 2020). The mating system of Fucoids should rather create homozygous zygotes through selfing. The excess of heterozygotes observed here, but at this site only, could be explained by several non-mutually exclusive hypotheses. If few breeders contribute to the next generation, allelic frequencies may differ between individuals contributing to the male and female gamete pools because of binomial sampling error: this can create heterozygosity excess (Pudovkin et al., 1996). This effect can be more important in the case of self-incompatibility (Stoeckel et al., 2006). The selective advantage of heterozygotes compared to homozygotes (i.e. heterosis) could explain the negative $F_{IS}$ as has been suggested for the red algae *Gracilaria chilensis* (Guillemin et al., 2008). To test this last hypothesis, it would be interesting to study genotypic change between recruits and the oldest individuals. Heterozygote excess can also be observed under partial clonal reproduction (e.g. Stoeckel and Masson, 2014), but this has never been observed in *E. zosteroides*.

The observation of a significant genetic differentiation at ARM among quite close depths (20 vs 40 m) indicates a structuring along an environmental gradient which corresponds to differences in light intensity and quality, and in thermal conditions (Pratlong et al., 2018). The observation of a vertical genetic structure at that scale has already been reported in this area in three octocoral species based on microsatellite data (Ledoux et al., 2010; Mokhtar-Jamai et al., 2011; Cánovas-Molina et al., 2018), and more recently with RAD sequencing in *Corallium rubrum* (Pratlong et al., 2018). The level of genetic differentiation reported between shallow and deep individuals (ARM) are of the same order of magnitude as populations at the same depth (between CAV and TIB) but 5.8 km apart. This differentiation according to depth can be driven by a limited dispersal of sinking non-buoyant zygotes and fragments (Guern, 1962; Capdevila et al., 2018), by an effect of local adaptation for some loci, or by intrinsic genetic incompatibilities coupled with environmental barriers along this depth gradient (Bierne et al., 2011). Interestingly, our results suggest a possible effect of temperature on the genetic structure at outlier loci: nevertheless, this subject would require the analysis of more comparisons among different depths and thermal conditions.

Several studies had previously reported the colonization of new habitats by marine species previously thought to be low dispersal, such as *Ericaria* species, which can colonize habitats several kilometers away (Thibaut et al., 2014). The individuals of ALT (estimation from 8 to 14 years) were younger than those from other populations studied here. For this long-living species, the rate of natural mortality is extremely low in absence of disturbance (lower than 2%, Ballesteros et al., 2009). These results seem to indicate that ALT is one of the youngest populations of the study, which corroborates well with the first report of the species in 2014 (Bonhomme et al., 2014) since the installation of the pipelines in 1967. The high density of this younger population does not preclude the possibility of founder effects during colonization, with the reduction in genetic variation. However, the genetic variation ($H_e$) was not reduced in the founder population in comparison to the levels reported in other ones in the Bay of Marseille. This finding was similar to those in the recently founded populations of the Fucoid *Gongolaria elegans*, which showed no signs of reduction in genetic diversity (Medrano et al., 2020). According to pairwise oceanographic connectivity and genetic structure, the population of La Ciotat (LC) 9 km away could be the donor population.

The application of genomics in conservation strategy will be essential for monitoring species and taking decisive actions by prioritizing the evolutionary potential of populations (von der Heyden, 2017; Xuereb et al., 2020). The use of genomic information coupled with *in situ* measurements of the demography and modeling of dispersal is seen to be highly effective to unravel occasional connectivity events in supposed low dispersal marine species. Based on our findings, we propose that such an integrative framework must be the proper way to unravel the relative influence of environmental factors on the genetic structure of marine forests. For *E. zosteroide*s in some cases, the species can disperse over long distances and this new information implies another strategy for it is conservation. Thus, apart from including species in the list of protection (all the Mediterranean Fucales have been listed under the Barcelona convention as strictly protected), none of these lists of species have been transcribed into national law and the difficulty of identifying species makes this measure ineffective (Verlaque et al., 2019). Currently, the marine forests are surveyed in MPAs but there is no conservation plan to effectively protect these populations (lack of management of the impact of herbivores and fishing activities on marine forests). *Ericaria zosteroides* is one of the most long-living erect seaweeds in the Mediterranean, and although rarely impacted by severe natural disturbances (Capdevila et al., 2015, 2019) most of the time, the species is impacted by cumulative direct anthropogenic stressors such as the decline in water transparency (Thibaut et al., 2005) and fishing activities (Hereu et al., 2008).

Our results show that some populations of *E. zosteroides* are spatially isolated and genetically differentiated but some of them remain connected through occasional dispersal. Conservation should be focused at the population level and with a priority of protection for the most isolated populations. The most effective measure is to ban fishing activities and then to set up surveillance of the increase in turbidity mainly due to untreated sewage outfall. None of the studied populations are within a No Take Zone, except VEY. Even though the other populations near Marseille - TIB, VEY, ALT, LC - are in the core area of the Parc National des Calanques, and RP, CAV are within its “adjacent area“, all the populations are subject to anthropic disturbances. Our results stress that conservation strategies should be focused on the implementation of the proposed measures, allowing the preservation in priority of the evolutionary potential and the ecological function of these isolated populations. Monitoring the temporal connectivity patterns of these populations will be necessary to determine

whether isolated and vulnerable populations need assisted gene flow from a restoration perspective.

## 2.5 Acknowledgements

This project was funded by the MARFOR Biodiversa/0004/2015 project (http://marfor.eu/), and by the Agence Nationale de la Recherche, Grant/ Award Number: ANR-18-CE32-0001 (Clonix2D). Lauric Reynes also received support from the GDR Génomique Environnementale (GDR3692). The project leading to this publication has received funding from the European FEDER Fund under project 1166-39417. This work was also funded by a Ph.D. Fellowship from Région PACA and the Calanques National Park. The authors would like to thank the diving unit of OSU Institut Pythéas, the divers of the National Institute for Ocean Science (IFREMER), and the divers of GIS Posidonie for sampling *Ericaria zosteroides* populations; the Biogenouest genomics core facility (Genomer Plateforme Génomique at the Station Biologique de Roscoff) for their technical support, the staff of the "Cluster de calcul intensif HPC'' Platform of OSU Institut Pythéas (Aix-Marseille University, INSU-CNRS) for providing the computing facilities, and finally Laura Benestan for her invaluable support in producing Asymmetric Eigenvector Maps (AEM). Finally, we would like to thank Michael Paul for English language proofreading.

## 2.6 Appendices

**Materials B.1**. Species delimitation analyses

We built an additional dataset consisting of twenty-three individuals of *E. zosteroides* from several geographical locations and eight individuals belonging to other species of *Cystoseira sensu lato* including *C. spinosa var. compressa*, *C. funkii*, *C. amentacea*, *C. tamariscifolia*, *C. mediterranea*, *C. baccata,* and *C. brachycarpa.* The DNA extraction, dd-RAD-seq library preparation, sequencing and sequence quality filtering were executed in an identical way to *E. zosteroides* (see main text for details). The reads were mapped to the *E. zosteroides* genome with the BWA-mem algorithm in BWA v.0.7.17 (Li and Durbin, 2009), and the catalog of loci was built with the reference mode of the Stacks v2.4 pipeline (Rochette et al., 2019). Using the population module of Stacks, we selected loci exclusively shared by all

samples (--min-samples-overall = 1). Following this step, 11 010 SNPs (25 587 loci) were filtered from the raw catalog of 569 242 SNPs (166 493 loci). Finally, each variable site was exported in PHYLIP format for downstream analyses. Phylogenetic maximum-likelihood (ML) analyses were performed with IQ-TREE v.1.6.1 (Nguyen et al., 2015) using ModelFinder (Kalyaanamoorthy et al., 2017) and 1000 ultrafast bootstraps (Hoang et al., 2018). The best-fit model of nucleotide substitution according to Bayesian information criterion (BIC) was TVM+F.

The percentage of average mapped reads per individual ranged from 50.10% in *C. brachycarpa* to 89.67% in *C. mediterranea*. We reported that individuals of *E. zosteroides* belonged to the same clade of the phylogenetic tree, whatever the initial geographical origin, while the eight other species of *Cystoseira sensu lato* belonged to evolutionally distinct clades (Figure B.1.1). All of these clades were supported by high ultrafast bootstrap values (99 or 100%). The genome-wide nuclear distance, as well as the mitochondrial genetic distance (e.g. Draisma et al., 2010; Bruno de Sousa et al., 2019; Pérez et al., 2020), were large enough among *E. zosteroides*, *C. spinosa var. compressa, and C. funkii* to discriminate species.

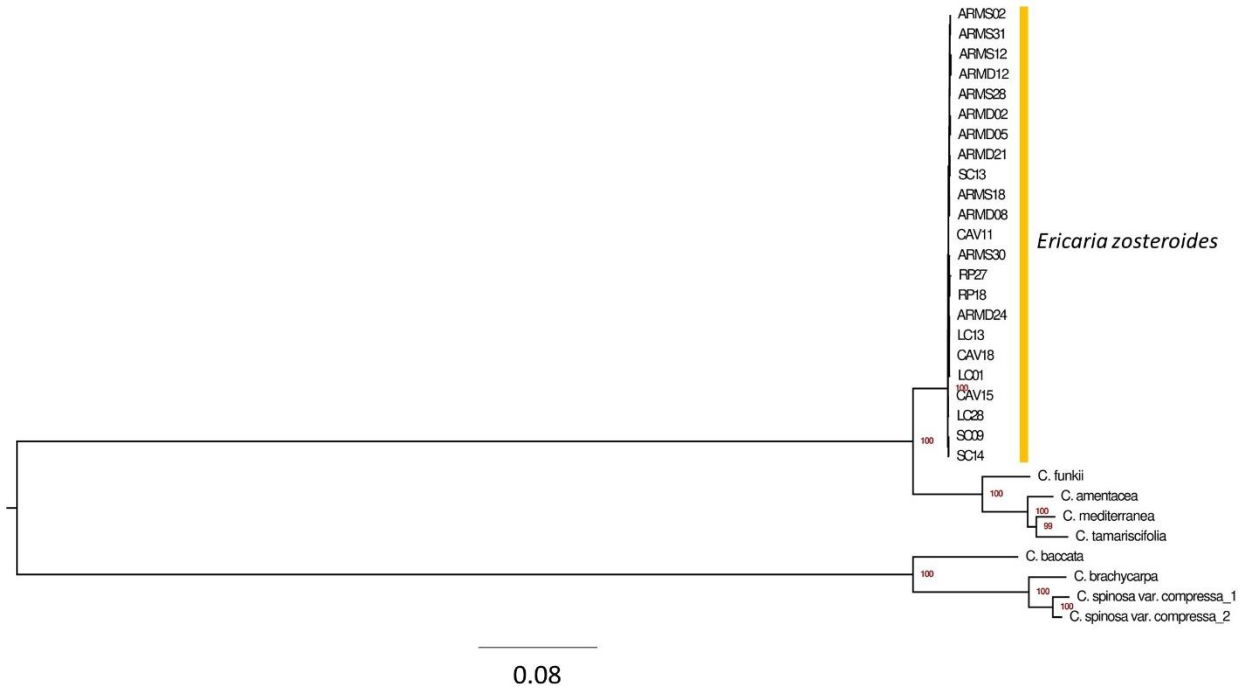


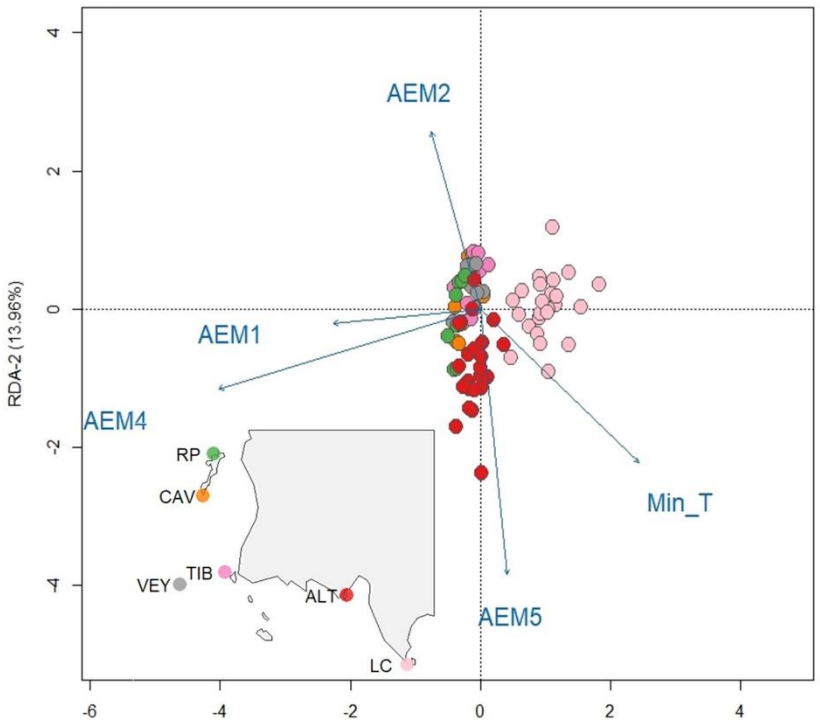

**Figure B.1.1**. ML phylogeny inferred using 11 010 SNPs in IQ-TREE. The tree was rooted using *Cystoseira baccata sensu lato* as an outgroup. The values in red indicate the percentages of ultrafast bootstrap resampling (n = 1000); the values below 90% are not shown.

**Materials B.2**. Genetic structure on the basis of outlier SNPs

This appendix presents details of the response of outlier SNPs to geographical and environmental predictors of genetic structure. We present the results of the distance-based RDA (db-RDA) according to the set of the 12 outlier SNPs reported in the main text, then we describe the spatial distribution of allelic frequencies for these outlier SNPs and we discuss our findings. According to the forward selection procedure, the minimum seawater temperature predictor (Min_T) explained 10% of the variance ($R^2_{adj}$ = 0.10, P < 0.001) in genetic structure while oceanographic connectivity with AEM-4, AEM-5, AEM-2, AEM-1 predictors explained 39% ($R^2_{adj}$ = 0.39, P < 0.001) of the variance (Figure B.2.1). Individuals seemed to be mainly differentiated along the first significant axis of the db-RDA (74.85% of the total variance), enabling the distinction of the deepest population (LC, depth = 45 m) from other ones that correspond to depths ranging from 25 to 36 m. The positive coordinate of individuals on this axis was driven by the minimum seawater temperature and was mainly related to the population of LC. Conversely, the negative coordinate of the axis was driven by oceanographic connectivity predictors such as AEM1 and AEM4, without specific differentiation of a population.

The allelic frequencies of the 12 putatively outlier SNPs were different between the deepest water population of LC and the other ones in the Bay of Marseille (Figure B.2.2), which is congruent with the results of the db-RDA. Among these outlier SNPs, five of them (i.e. 112036_64, 136434_44, 1521_81, 211417_31, and 57060_94) were particularly interesting since one of the two alleles tended to be absent in the Bay of Marseille and overrepresented in the population of LC. These allele frequencies ranged from 0.34 to 0.72 in the population of LC and from 0.03 to 0.06 within the Bay of Marseille (Figure B.2.2). The alleles of the population of LC are not specific to depth, in that they are observed at high frequencies in populations of Eastern-Provence and Corsica, whatever their depth. However, three outlier loci attracted much attention (i.e. 1521_81, 57060_94, and 211417_31) since major alleles reported at high frequency in the deepest water population were relatively less frequent in shallow populations. Also, although these alleles are too infrequent at the isolated sites of Armoire, they appeared to be more preserved in the deepest ARM40 than the shallowest ARM20 populations (Figure B.2.2).

(A)




(B)

| Population | Depth (m) | Min_T (°C) | Max_T (°C) | Mean_T (°C) | Var_T (°C) |
|---|---|---|---|---|---|
| RP | 26-28 | 11.73 | 18.17 | 14.03 | 1.16 |
| CAV | 28-29 | 11.73 | 17.66 | 14.01 | 1.1 |
| TIB | 28-30 | 11.69 | 19.34 | 14.66 | 3.0 |
| VEY | 25-36 | 12.15 | 16.69 | 13.98 | 0.71 |
| ALT | 31 | 12.29 | 16.5 | 13.96 | 0.71 |
| LC | 45 | 12.22 | 15.36 | 13.75 | 0.42 |

(C)

| Oceanographic connectivity | Seawater temperature | P model | $R^2_{adj}$ |
|---|---|---|---|
| | Min_T *** | <0.001 | 0.10 |
| AEM-4 ***<br>AEM-5 ***<br>AEM-2 ***<br>AEM-1 *** | | <0.001 | 0.39 |
| AEM-4 ***<br>AEM-5 ***<br>AEM-2 ***<br>AEM-1 * | Min_T *** | <0.001 | 0.43 |

**Figure B.2.1**. Distance-based RDA (db-RDA) using the set of 12 putatively outlier SNPs. (A) db-RDA biplot showing the significant predictors. The dots represent individuals and are colored according to their populations. (B) seawater temperature predictors averaged over seven years of simulations. (C) significant predictors retained following the forward selection procedure using both oceanographic connectivity and seawater temperature predictors simultaneously and independently from each other. Geographical distance was not retained as a significant predictor.

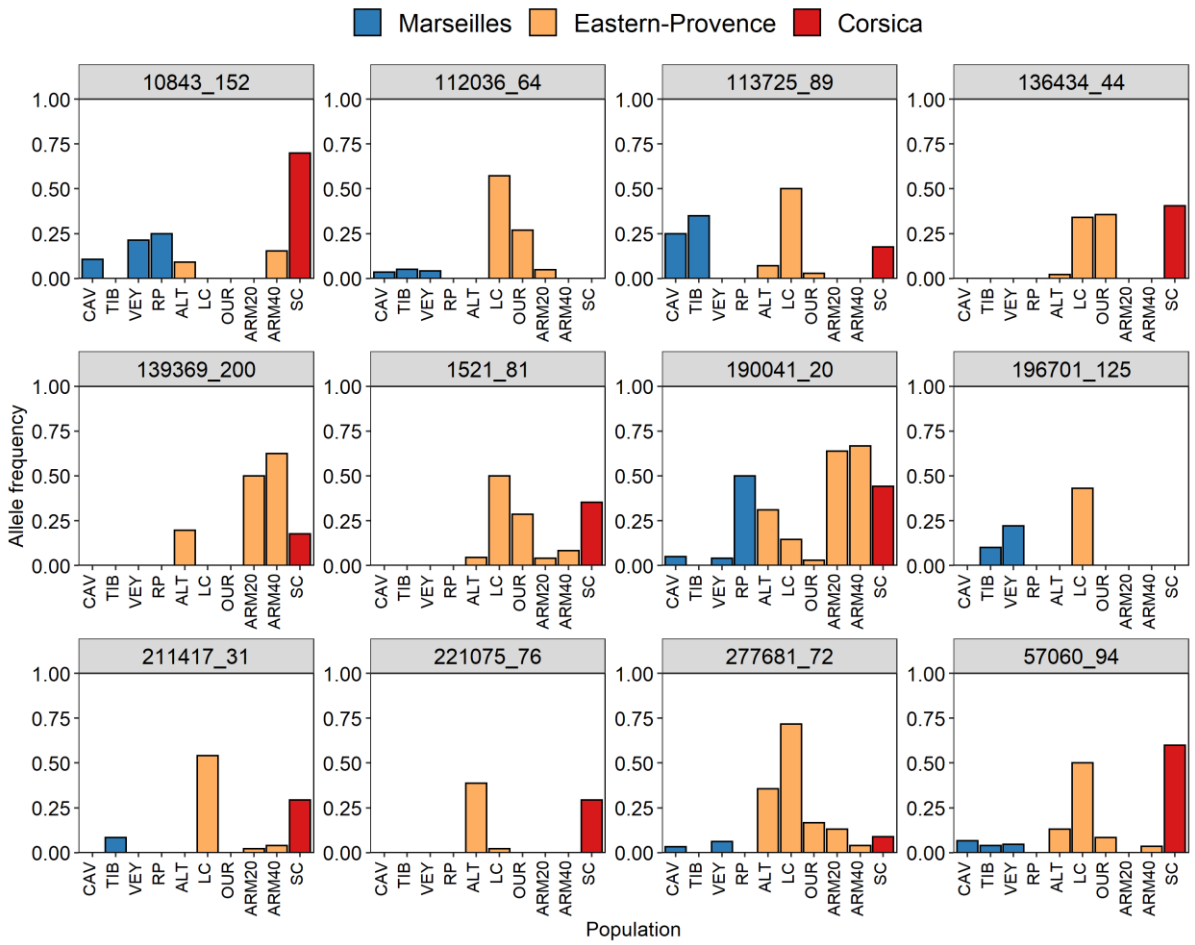



**Figure B.2.2**. Spatial distribution of the minor allele frequencies (MAF) of the 12 putatively outlier SNPs at the scales of the bay of Marseille (blue), Eastern-Provence (yellow), and Corsica (red).

We reported from the neutral set of SNPs that the isolated bank populations of Armoire had one of the lowest levels of genetic variation reported in our study. This is discussed in the main manuscript as the effects of restricted gene flow from other regions and strong intensity of genetic drift, which potentially limit the strength of divergent selection with depth. Furthermore, our results indicated that directional gene flow from shallow to deep populations occurred at Armoire. Gene frequency change by natural selection can be slowed by homogenizing gene flow, imposing a limit of local adaptation (Lenormand, 2002). Hence, the lack of shift in allele frequency from shallow to deep individuals at Armoire for the outlier loci could be the consequence of gene flow, which would potentially limit local adaptation along this depth gradient. Further studies will be necessary to conclude whether the pattern of differentiation by depth reported in the population of LC can be interpreted as a result of

local adaptation. Nevertheless, the fitness of *Ericaria zosteroides* was significantly impacted by increasing seawater temperature from 16°C to 24°C (Capdevila et al., 2019), suggesting that seawater temperature may involve changes in metabolic traits that may be under selection depending on the local thermal regime.

**Additional Supporting Information may be found in the Annexe B.**

**Figure B.1**. Variation in cross-entropy according to the number K of ancestral populations in sNMF.

**Figure B.2**. Distribution of the size of the main axis (cm) in the sampled populations; the sizes were distributed into fourteen discrete classes.

**Figure B.3**. Additional axis from PC3 to PC8 (PC3: 12%; PC4: 11%; PC5: 10%; PC6: 9%; PC7: 9%; PC8: 7%) of the PCA conducted on the neutral set of 10 755 SNPs and 229 individuals.

**Figure B.4**. Assignment tests over the different scales of the study. For each population (Sampled population, y-axis), the percentage of individuals assigned (Inferred population, x-axis) to their respective sources (green circle) or other putative sources (red circle). The size of circles is proportional to the percentage of assignment, (A) assignment tests restricted to the top 100 ranked SNPs (decreasing overall $F_{ST}$), (B) assignment tests for all neutral markers (10 755 SNPs).

**Figure B.5**. Assignment success (%) at the population level and over populations (Overall) as a function of the number of SNPs ranked based on decreasing $F_{ST}$ values.

**Table B.1**. Quality criteria applied to build the neutral and outlier dataset of *Ericaria zosteroides*. For each filtering step, the number of SNPs passing the filter (SNPs count), as well as the number of individuals (individuals count), are reported. The catalog of 314 538 loci was initially built using the Stacks v.2.4 bioinformatics pipeline, with the draft genome of *E. zosteroides* used as a reference.

**Table B.2**. Comparisons in the distribution of the size of the main axis between populations using the Kolmorov-Sirmov pairwise test. Non-significant P-values are in bold type ($p>0.05$).

**Table B.3**. Comparisons in genetic diversity (expected heterozygosity among SNPs) among populations using the Pairwise Wilcoxon Test. P-values were adjusted for multiple testing with the Bonferroni method. Non-significant P-values are in bold type ($p>0.001$).

## 2.7 Synthèse des résultats obtenus

Les résultats de ce chapitre ont mis en évidence que la différenciation génétique des populations de *E. zosteroides* est élevée et significative pour les sites considérés. Ces différences opèrent à courte distance (sur le tombant de l'Armoire), à une échelle locale (baie de Marseille) et à longue distance (entre l'est de la Provence et la Corse). Cette forte structure génétique est comparable à celle constatée pour *E. amentacea* dans la baie de Marseille (Thibaut et al., 2016b) et plus globalement à l'échelle de la Provence (Susini et al., 2007 ; Robvieux, 2013). Compte tenu du faible potentiel de dispersion de ces espèces et de la fragmentation du milieu côtier, nos résultats suggèrent que le degré de flux de gènes est trop limité pour contrebalancer la différenciation par dérive génétique. Par ailleurs, l'approche par modélisation Lagrangienne a montré qu'à l'échelle de la baie de Marseille, la différenciation des populations est davantage expliquée par les courants de fond ($R^2_{adj}$ = 0,11) que par la distance spatiale entre les taches d'habitats ($R^2_{adj}$ = 0,03). Ces résultats indiquent que la dérive de fragments fertiles à proximité du fond pourrait jouer un rôle important dans la connectivité génétique. Ce mécanisme de dispersion pourrait aussi expliquer la colonisation d'une conduite sous-marine à une distance de 9 km de la plus proche population connue. D'après les analyses de connectivité et d'assignations individuelles, cette population pourrait être la population source à l'origine de la colonisation. Cependant la nette différenciation observée entre cette dernière sur substrat naturel et la population nouvellement installée, laisse la porte ouverte à d'autres interprétations de ce processus de colonisation. Enfin, de manière surprenante, une population de la baie de Marseille est apparue très différenciée des autres, y compris celles distantes de seulement quelques kilomètres, une différenciation même plus forte que celle observée pour la population de Corse. Cette distinction est cohérente avec la connectivité océanographique réduite et la faible densité d'individus observée sur ce site. Néanmoins, en l'état actuel des hypothèses alternatives de couplages entre une barrière océanographique et une barrière endogène aux flux de gènes ne peuvent être exclues (Bierne et al., 2011). Par ailleurs, nous identifions 12 locus outliers susceptibles d'être impliqués dans une réponse adaptative des populations à la température. Les résultats de cette étude ont été finalement discutés, notamment à la lumière des mesures de conservation d'*Ericaria zosteroides* dans le Parc National des Calanques.

# Chapitre 3

# Structure des populations et signatures génomiques de la clonalité chez *Laminaria rodriguezii*



L'une des rares forêts marines à *Laminaria rodriguezii*, échantillonnée à 71-76 mètres de profondeur au Banc du Magaud (est de la Provence) durant l'expédition Gombessa 5 (© Laurent Ballesta).

Ce chapitre correspond à un article publié dans **Molecular Ecology** (*doi.org/10.1111/mec.15860,* publié le 24 février 2021) et il contient une analyse complémentaire de la structure génétique des populations de *Laminaria rodriguezii.*

***Contexte et objectifs***

J'ai étudié ici les signatures génomiques de la clonalité et la différenciation génétique des populations partiellement clonales du kelp *Laminaria rodriguezii*. Le taux de clonalité, et son interaction avec la dérive génétique et la mutation vont affecter diverses statistiques descriptives de la diversité et de la structure génétique des populations, rendant ainsi possible la détection d'un mode de reproduction clonal ou partiellement clonal par rapport à un mode de reproduction uniquement sexué (Balloux et al., 2003 ; Stoeckel & Masson, 2014 ; Stoeckel et al., 2020). Si des effets tels que la diminution de la diversité génotypique ou la présence d'excès en hétérozygotes sont déjà observables avec un nombre réduit de marqueurs génétiques, les attendus théoriques sur la distribution génomique des effets de la clonalité ouvrent de nouvelles perspectives pour l'étude de ce mécanisme. Ainsi, la variance du $F_{IS}$ augmente avec le taux de clonalité, de manière visible pour les plus forts taux de clonalité (taux de clonalité > 0,9) (Stoeckel et Masson, 2014). Un taux de clonalité élevé devrait aussi entraîner un accroissement du déséquilibre de liaison (LD) (Stoeckel et al., 2020). Ceci a pu être vérifié empiriquement pour des macroalgues dont le cycle de reproduction est digénétique haplo-diplontique (Krueger-Hadfield et al., 2021). Cette étude indique que le déséquilibre de liaison augmente à mesure que la richesse génotypique diminue et cela pour plusieurs taxa clonaux. Néanmoins, le déséquilibre de liaison peut être aussi élevé dans une population majoritairement sexuée que dans une population hautement clonale (taux de clonalité > 0,99). Ce constat dénote que des facteurs comme la taille efficace, la reproduction entre apparentés et l'autofécondation peuvent être une des causes à l'origine de l'augmentation du déséquilibre de liaison (Krueger-Hadfield et al., 2021).

L'utilisation d'approches de génomique des populations, comme le séquençage RAD, permet maintenant de tester ces attendus théoriques sur des populations naturelles d'espèces non modèles. Dans cette partie de ma thèse, j'ai donc appliqué le séquençage RAD (ddRAD-seq), à l'analyse de la distribution de statistiques de génomique des populations à l'échelle du génome des populations de *L. rodriguezii*. Afin d'interpréter ces résultats, j'ai effectué une comparaison des résultats avec ceux obtenus chez une espèce congénérique,

*Laminaria digitata* avec seulement la reproduction sexuée (même cycle de reproduction sexué que *L. rodriguezii*). Les populations de *L. digitata* ont été choisies pour leurs niveau de diversité génétique qui est variable d'un site à un autre (Liesner et al., 2020). La prise en compte de plusieurs populations avec des histoires, des niveaux de dérive génétique et de connectivité contrastés, a permis de s'interroger sur les potentiels effets confondants dans l'étude de l'impact génomique de la clonalité, par exemple sur les analyses de déséquilibre de liaison. Par ailleurs, l'analyse de données génomiques chez une espèce partiellement clonale a nécessité de s'interroger sur une méthode permettant de tenir compte des erreurs de séquençage et des mutations somatiques dans la délimitation des lignées clonales. La prise en compte de l'impact de la reproduction clonale m'a ensuite permis d'étudier la structure génétique des populations de *L. rodriguezii* afin d'évaluer leur degré d'isolement et les implications que cela peut avoir pour la conservation de cette espèce menacée.

Molecular Ecology (https://doi.org/10.1111/mec.15860)

## Genomic signatures of clonality in the deep water kelp *Laminaria rodriguezii*

Lauric Reynes[1], Thierry Thibaut[1], Stéphane Mauger[2], Aurélie Blanfuné[1], Florian Holon[3], Corinne Cruaud[4], Arnaud Couloux[5], Myriam Valero[2], Didier Aurelle[1,6]

*lauric.reynes@mio.osupytheas.fr

1 Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, Marseille, France
2 IRL 3614, Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université, UC, UACH, CS 90074, Place Georges Teissier, 29688 Roscoff, France
3 Andromède Océanologie, Montpellier, France
4 Genoscope, Institut de Biologie François-Jacob, Commissariat à l'Energie Atomique (CEA), Université Paris-Saclay, F-91057 Evry, France.
5 Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France
6 Institut de Systématique Évolution Biodiversité (ISYEB, UMR 7205), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, 57 rue Cuvier, CP 51, 75005 Paris, France

# Abstract

**Key words: Partial clonality, mode of reproduction, heterozygote excess, population genomics, kelp**

The development of population genomic approaches in non-model species allows for renewed studies of the impact of reproductive systems and genetic drift on population diversity. Here, we investigate the genomic signatures of partial clonality in the deep water kelp *Laminaria rodriguezii*, known to reproduce by both sexual and asexual means. We compared these results with the species *Laminaria digitata*, a closely related species that differs by different traits, in particular its reproductive mode (no clonal reproduction). We analysed genome-wide variation with dd-RAD sequencing using 4077 SNPs in *L. rodriguezii* and 7364 SNPs in *L. digitata*. As predicted for partially clonal populations, we show that the distribution of $F_{IS}$ within populations of *L. rodriguezii* is shifted toward negative values, with a high number of loci showing heterozygote excess. This finding is the opposite of what we observed within sexual populations of *L. digitata*, characterized by a generalized deficit in heterozygotes. Furthermore, we observed distinct distributions of $F_{IS}$ among populations of *L. rodriguezii*, which is congruent with the predictions of theoretical models for different levels of clonality and genetic drift. These findings highlight that the empirical distribution of $F_{IS}$ is a promising feature for the genomic study of asexuality in natural populations. Our results also show that the populations of *L. rodriguezii* analysed here are genetically differentiated and probably isolated. Our study provides a conceptual framework to investigate partial clonality on the

basis of RAD-sequencing SNPs. These results could be obtained without any reference genome, and are therefore of interest for various non-model species.

## 3.1 Introduction

Partial clonality, defined as a mode of reproduction where species reproduce both sexually and asexually, occurs in a wide variety of organisms across the tree of life, and is particularly widespread in fungi, plants, rotifers, cladocerans and insects (Halkett et al., 2005). Clonal reproduction can occur through different mechanisms such as fragmentation, vegetative reproduction, mitotic parthenogenesis, or some forms of meiotic parthenogenesis (De Meeûs et al., 2007; Jaron et al., 2020). Partial clonality has attracted considerable attention among evolutionary biologists due to the potential effects of asexuality on the evolutionary trajectory of species since reproductive modes have a major impact on the level of genome-wide genetic diversity and its distribution within and among populations (Halkett et al., 2005; Duminil et al., 2007). Moreover, reproductive modes profoundly affect other biological traits that play important roles in colonization processes and the spread of populations, such as population growth and dispersal (Kettenring & Mock, 2012; González de León et al., 2016), and in buffering and resilience after catastrophic events (Becheler et al., 2020).

Theoretical models predict that when clonality predominates, it will increase heterozygosity (Marshall & Weir, 1979; Balloux et al., 2003; De Meeûs et al., 2006), which may ultimately lead to higher levels of observed heterozygosity ($H_o$) than those expected ($H_e$) under panmixia. This deviation has been observed empirically through negative values of fixation index $F_{IS}$, reported in various partially clonal populations (in animals Adjeroud et al., 2014; Halkett et al., 2005; angiosperms Alberto et al., 2005; Arnaud-Haond et al., 2007a, and red and brown macroalgae Guillemin et al., 2008; Ardehed et al., 2015; Coleman & Wernberg, 2018, Pardo et al., 2019). Without sex, alleles do not segregate independently, which implies that heterozygosity can be preserved rather than reduced over generations (Judson & Normark, 1996). Increasing clonal reproduction can increase effective population size at the allelic scale but it decreases the genotypic effective population size: as a consequence asexual populations can maintain higher allele diversity but lower genotype diversity compared to sexual ones (Balloux et al., 2003). Asexual reproduction also leads to specific patterns of the distribution of $F_{IS}$ over the genome, characterized by a high occurrence of extreme, mainly negative,

values (Stoeckel & Masson, 2014; Reichel et al., 2016). These last theoretical studies reported, among other points, that high rates of asexuality increased the variance of the distribution of $F_{IS}$ compared to that obtained from fully sexual populations (Stoeckel & Masson, 2014) but also as compared to fully clonal populations (De Meeûs et al. 2006). Some variation in $F_{IS}$ among loci can also indicate the occurrence of rare recombination events as in the automictic parthenogenetic species *Artemia parthenogenetica* (Nougué et al., 2015). In addition, if asexuality is stable over generations, the accumulation of mutations in different alleles can increase sequence differences among alleles, which corresponds to the "Meselson effect" (Birky 1996; Judson & Normark, 1996; Balloux et al., 2003; De Meeûs et al., 2007). This effect has been empirically demonstrated in ancient asexual *Timema* stick-insects (Schwander et al., 2011) and in the human pathogen *Trypanosoma brucei gambiense* (Koffi et al., 2009; Weir et al., 2016) and *Candida albicans* (Nébavi et al., 2006). Conversely the long term persistence without sex of Bdelloid rotifers, referred as asexual scandals, has been challenged by the suggestion of non-canonical sex (Schwander, 2016).

In this paper, we investigated how population genomic studies of partially clonal species allow analysis of both the rate of clonality and its consequences for the distribution of $F_{IS}$. With the development of thousands of Single Nucleotide Polymorphism (SNP), we will test whether the observed empirical distribution of $F_{IS}$ follows the aforementioned predicted variations in $F_{IS}$ among loci. The development of reduced representation sequencing methods (see for review Davey et al., 2011) now allows such investigations in non-model species. These approaches are also useful to more precisely delineate clonal lineages, where microsatellite loci have failed to detect any genetic variation, as in the case of the selfing species *Sargassum muticum* in the introduction range (Le Cam et al., 2019). Among the diversity of reduced representation sequencing methods, Restriction Site Associated DNA sequencing (RAD sequencing; Baird et al., 2008) is useful to efficiently develop enough SNPs to study the genomic distribution of population genetic statistics.

We used the kelp *Laminaria rodriguezii* as a model species for the study of partial clonality. *Laminaria rodriguezii* is a deep water kelp endemic to the Mediterranean Sea, mainly reported from depths below 70 m, rarely found in shallower waters on seamounts or in upwelling systems (Ballesteros, 2006), and with a maximum depth recorded at 260 m in the Adriatic Sea (Ercegović, 1960). The rare populations of this species are exposed to various levels of disturbances, such as direct destruction by fishing gear and a decline in water

transparency, which has led to *L. rodriguezii* being classified as endangered (Zuljevic et al., 2016). Besides sexual reproduction, *L. rodriguezii* can reproduce asexually by stoloniferous or rhizoidal growth. This mode of asexual reproduction involves somatic cells and should lead to clonal lineages. Vegetative reproduction is only known in two *Laminaria* species: *L. rodriguezii* (Huvé, 1955) and the northeast Pacific *L. sinclairii* (Demes & Graham, 2011). *Laminaria rodriguezii* can also reproduce sexually with a haplo-diplobiontic life cycle involving the alternation of microscopic dioicous unisexual haploid gametophytes with large diploid sporophytes typically found in kelps (Bringloe et al., 2020). Even if we do not know the importance of self incompatibility in this species, self fertilisation (i.e. fertilisation between male and female gametophytes originating from a single sporophyte parent), was reported to be possible experimentally in most studied species of Laminariales (see the recent review of Goecke et al., 2020). *Laminaria rodriguezii* then presents a mixed reproductive system, for which the genetic consequences have not been studied. This deep water species is also characterized by small and highly fragmented populations (Araújo et al., 2016). Therefore, the populations of *L. rodriguezii* should be impacted by a reduction in genetic polymorphism through drift, a reduction that can be counterbalanced by the aforementioned effects of asexual reproduction on allele diversity (Balloux et al., 2003).

We specifically studied here whether the levels of asexual reproduction have discernable effects on the empirical genomic distribution of diversity and $F_{IS}$ among loci. We included populations of the fully sexual congeneric species, *L. digitata,* with different levels of microsatellite genetic diversity (Liesner et al., 2020) to take into account the consequence of the mode reproduction with the confounding effect of variation in the strength of genetic drift in our interpretations. *Laminaria digitata* is analysed only for comparative purposes, and we did not develop the study of the genetic structure of this species in this article. We also interpret our results in the light of previous theoretical studies, especially considering the full distribution of $F_{IS}$ among loci (Stoeckel & Masson, 2014). Among other factors, we tested whether the clonal rates in populations of *L. rodriguezii* lead to a shift toward highly negative $F_{IS}$ values among loci, as expected for the highest levels of clonality (Stoeckel & Masson, 2014). We used a large set of SNPs in order to access the most complete distribution of $F_{IS}$ values over populations of genomes. We also evaluated the robustness of our results by considering different cutoffs in the filtering of SNPs. The implications of the inferred genetic diversity and genetic

structure for the management and conservation of this rare deep water kelp *L. rodriguezii* are discussed.

## 3.2 Material and methods

### 3.2.1 Sampling, DNA extraction

Individuals of *L. rodriguezii* were sampled from four sites in the Mediterranean, three in Eastern Provence (Banc Magaud_1, Banc Magaud_2, Cap Camarat) and one in Southern Corsica (Bonifacio), between 65 to 76 meters depth (Figure 3.1). Sampling was done by scuba divers between June and August 2018. Sampling for DNA collection was performed by collecting a small piece of tissue from the blade of sporophytes. More precisely, a total of 47 samples were collected from sporophytes not connected to each other by a stolon, and with a minimum distance of two meters between sampled blades. This strategy aimed to promote the sampling of distinct genets. For one of the two Banc du Magaud sites, namely Banc Magaud_2, we used a different sampling design: the spatial distance between the sampled sporophytes was recorded along a regular transect of 20 meters to study the spatial genetic structure inside the population. Regarding *L. digitata*, we used samples from five northeastern Atlantic populations of *L. digitata* collected in 2018 (Table C.1, Supporting Information). Genomic DNA was extracted using the Nucleospin® 96 plant kit (Macherey-Nagel, Düren, Germany), according to the manufacturer's protocol.

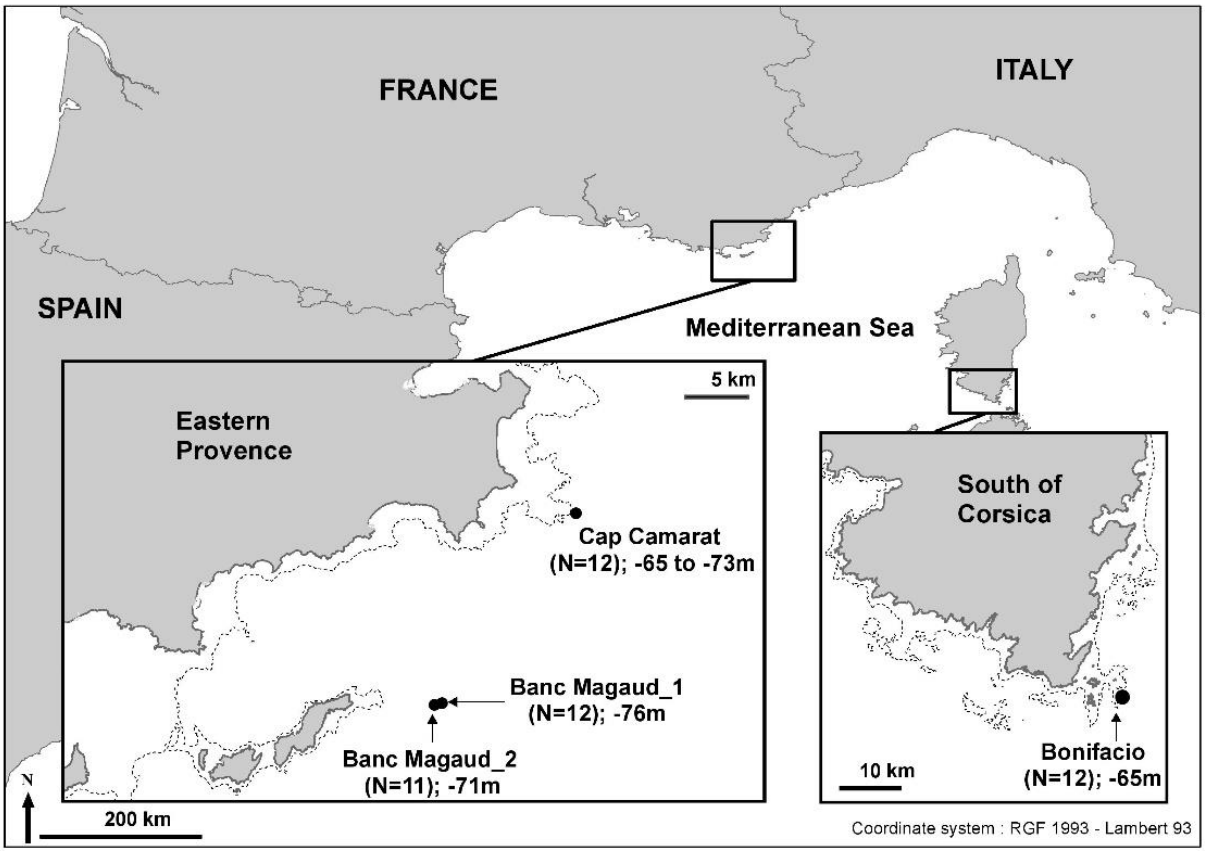



**Figure 3.1**. Sampling map of *Laminaria rodriguezii* in Eastern Provence and Southern Corsica, reporting the depth in meters of sampling sites and the number of individuals sampled per site (N).

### 3.2.2 ddRAD-sequencing

Two double-digest RAD-sequencing libraries (ddRAD-seq) with 47 individuals of *L. rodriguezii* and 125 individuals of *L. digitata* were prepared according to Peterson et al., (2012). To minimize bias in library preparation, all DNA samples were randomized across libraries, and two and five replicates of *L. rodriguezii* and *L. digitata*, respectively, were included. DNA concentrations were estimated with PicoGreen (Invitrogen, Carlsbad, CA, USA), and 100 ng of genomic DNA was double digested using PstI and HhaI (NEB) for 12 hours at 37°C. The purification of digested DNA was conducted by the NucleoMag NGS clean-up and size select kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's protocol. When these steps were completed, the ligation of sequencing adapters to restriction digested DNA was performed as follows: P1 adapters (overhang PstI), including unique barcodes (6-12 bp) and P2 universal adaptor (overhang HhaI) were ligated to DNA using 0.5 µl (400 U/µL) of T4 DNA ligase (NEB), 6 µl of 10X T4 ligase buffer (NEB) and incubated for 12 hours at room temperature. Before PCR amplification, an additional purification step was done with the

NucleoMag NGS clean-up and size select kit. Amplification by PCR was restricted to 15 cycles and performed with the Q5 hot Start High-Fidelity DNA polymerase kit (NEB). Then, samples were multiplexed, and automated size selected for a range between 400 and 900 bp using a Pippin Prep (Sage Science, Beverly, MA, USA). Finally, the libraries were sequenced with the paired-end method (2×150 bp) on an Illumina Hiseq 4000 platform (Génome Québec Innovation Centre, McGill Univ., Montreal, Canada).

### 3.2.3 Genotyping and SNP filtering

Raw quality reads were checked using FastQC v0.11.7 (Andrews, 2010) and trimmed to 137 bp after the removal of adapters sequencing by Trimmomatic (Bolger et al., 2014). Then, paired-reads were demultiplexed by both index and barcode using Stacks's process_radtags (Catchen et al., 2011). Thanks to the phylogenetic proximity between *L. rodriguezii* and *L. digitata* (Žuljević et al., 2016), we were able to use the draft genome of *L. digitata* (unpublished data, Mark Cock) to align our reads with the BWA-mem algorithm in BWA v0.7.17 (Li & Durbin, 2009). Aligned reads were then assembled into loci and finally, we called single nucleotide polymorphisms (SNPs) with the Stacks v2.4 pipeline (Rochette et al., 2019). Aligned reads of mapping quality below 30 (which means a probability of incorrect alignment below 0.001) were excluded to minimize errors resulting from the erroneous assembly of paralogous loci as a single locus. To avoid a strong linkage between SNPs, we kept only the first SNP of each RAD locus. Further filtering was performed on the basis of minor allele frequency, missing data rates for both individuals (i.e the percentage of missing genotypes per individual) and loci (the percentage of non-genotyped individuals per locus), using VCFtools v0.1.17 (Danecek et al., 2011) (Table 3.1).

**Table 3.1**. Quality filters used to build the final RAD-seq datasets of *Laminaria rodriguezii* and *Laminaria digitata*.

| Filtering thresholds | *Laminaria rodriguezii* | | | *Laminaria digitata* | | |
|---|---|---|---|---|---|---|
| | SNP counts | missingness | coverage | SNP counts | missingness | coverage |
| **Locus genotyped**<br>> 75% of the samples<br>> 75% of the populations<br>**SNP kept**<br>obs. heterozygosity < 80% | 70 026 | 13.97 % | 11.29 | 609 699 | 12.26 % | 14.11 |
| **Linkage disequilibrium**<br>single SNP per locus | 27 769 | 13.18 % | 11.88 | 58 830 | 13.45 % | 14.36 |
| **MAF**<br>MAF > 2% | 16 865 | 13.51 % | 12.18 | 21 090 | 15.01 % | 14.43 |
| **Missingness**<br>per individual < 30%<br>per SNP < 20% | 13 235 | 4.46 % | 13.91 | 15 150 | 4.73 % | 17.24 |
| **Coverage**<br>mean read depth > 15 X<br>mean read depth < 100 X | 4 366 | 2.27 % | 22.17 | 7 538 | 3.40 % | 21.63 |
| **Outliers detection**<br>putatively neutral | 4 077 | 2.29 % | 22.04 | 7364 | 3.39 % | 21.64 |

For each filtering step, the number of SNP passing the filter (SNPs count), as well as the mean missing data rate per individual (missingness) and the mean sequencing depth per individual (coverage) are reported. Loci were initially built in both species from the draft genome of *L. digitata* using the Stacks v.2.4 bioinformatics pipeline. The threshold of missing genotypes per individual < 30% has deleted four and nine individuals in *L. rodriguezii* and *L. digitata* datasets respectively. See Methods for details.

We aimed to identify the signatures of clonality from putatively neutral loci. To this end, we excluded outlier loci identified as involved under divergent selection with *pcadapt* v3.5 (Luu et al., 2017). Pcadapt starts by analysing population structure with a PCA. Then candidate outliers are identified as those which are excessively correlated with population structure (Luu et al., 2017). The outlier identified with pcadapt should then correspond to inflated signals of genetic differentiation.

### 3.2.4 Impact of SNP filtering on population genetic summary statistics

RAD-seq is subject to different sources of genotyping errors (Mastretta-Yanes et al., 2015) that may generate spurious inferences in surveys of genetic variation. However, the noise of artifactual SNPs can be efficiently overcome by applying a rigorous filtering strategy of raw RAD-seq datasets (e.g. McKinney et al., 2017b; O'Leary et al., 2018) and measuring its

effects on population genetic summary statistics (Shafer et al., 2017). Care was taken in this study to minimize the impact of artifactual SNPs on the true biological signal of the datasets (see Supporting Information "Materials C.1"). To achieve this, we studied the impact of different filtering cutoffs on several parameters of genetic diversity and departure from panmixia: i) the mean $F_{IS}$ and standard deviation (SD) among loci, ii) the overall $F_{ST}$ and SD among loci, iii) the SNP error rate, defined as the proportion of SNP mismatches between replicate pairs (Mastretta-Yanes et al., 2015). We have specifically explored the effect of the maximum rate of missing data allowed per SNP (max-missing), the minor allele frequency (MAF), and the mean read depth per locus across all individuals (meanDP), or by considering only loci for which genotypes were strictly identical between replicates. Further details are given in Supporting Information "Materials C.1".

### 3.2.5 Clonality in *Laminaria rodriguezii*

Because of somatic mutations and sequencing errors, we do not expect to get two strictly identical genotypes over thousands of loci even if they belong to the same clonal lineage. Consequently, we considered that genotypes with very low divergence corresponded to the same clonal lineage. We chose the maximum distance between replicates (d = 0.022 for *L. rodriguezii*, d = 0.020 for *L. digitata*, see Results) as a threshold to identify repeated genotypes. Distances were computed as the ratio of the number of allelic differences between two individuals (the Hamming distance, implemented in the function *diss.dist* of the Poppr v2.8.2 R package). Samples of *L. rodriguezii* separated by a distance lower than the aforementioned threshold were clustered into clonal lineages (referred here as MLLs Multi-Locus Lineages, see Arnaud-Haond et al., 2007b) using the UPGMA clustering algorithm (implemented in Poppr v2.8.2 R packages: Kamvar et al., 2014; Kamvar et al., 2015). The matrix of genetic distances was additionally used for the construction of a network with the neighbor-net method implemented in SplitsTree, which was used to visualize genetic proximity of samples from the same clonal lineages, and also the differences among populations (Huson, 1998). We computed the genotypic richness (R) index within populations (Dorken & Eckert, 2001), as follows: $R = \frac{G-1}{N-1}$, where G is the number of distinct genotypes detected in the population and N the total number of individuals sampled. We thus considered each cluster of clonal individuals as a single genotype to calculate G. The distribution of clonal membership

was analysed with the Pareto β index as described in Arnaud-Haond et al. (2007b) using the Poppr v2.8.2 R package. Finally, the average values of genetic diversity and population differentiation of *L. rodriguezii* were estimated in two complementary ways: firstly, including all genotyped individuals, secondly by keeping only one representative of each clonal lineage.

### 3.2.6 Genetic diversity

The within-population genetic diversity was assessed with the datasets of *L. rodriguezii* and *L. digitata*, after the removal of potentially selected loci. Given that the percentage of polymorphic loci (%P) may be affected by the variations in sample sizes, we used a random sampling approach implemented in a custom bash script (available at https://github.com/rebecca-cj/revegetation). The percentage of polymorphic loci was calculated from 100 random resamplings (with replacement between samples) for each sample size at the population level. In addition, estimates of observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), and $F_{IS}$ (Weir & Cockerham, 1984) were calculated with the GENEPOP v1.0.5 R package (Rousset, 2008). The same data sets were also analyzed with the HierFstat R package (Goudet, 2005) to compute the distribution of $F_{IS}$ values across loci. We used a two-sample Kolmogorov-Smirnov (KS) test to test for differences in the distribution of $F_{IS}$ between populations. We used the KS D statistic as the maximum distance between cumulative distributions to assess the magnitude of these differences. The P-value of the D statistic was computed using a bootstrapping method ($10^6$ bootstraps), with the Matching v4.9.7 R package (Sekhon, 2008). The distributions were divided into ten discrete classes of $F_{IS}$ for which the frequency of occurrence in populations was shown as histograms. The Global Hardy-Weinberg test [Score (U) test] (Raymond & Rousset, 1995), implemented in GENEPOP, was used to test departure from panmixia across all loci for the different populations. The test was performed twice: firstly by considering the alternative hypothesis of heterozygote excess and secondly by considering heterozygote deficiency. The P-value of each test was approximated by Markov chain (MCMC) algorithm with the following settings; dememorization: 10 000, batch: 100, iterations per batch: 10 000. Then, a multisample score test (Raymond & Rousset, 1995) was performed to obtain a global P-value per population. Finally, the two-sided P-value of the Global Hardy-Weinberg test [Score (U) test] was obtained by computing the following formula: Pmin + (1 - Pmax) where Pmin and Pmax are the smaller and the higher probabilities of unilateral tests, respectively (De Meeûs, 2012). Linkage

disequilibrium (LD) among loci was analyzed with the standardized index of multilocus association $\overline{r_d}$ (Agapow & Burt, 2001) in Poppr v2.8.2. The $\overline{r_d}$ index is a summary statistic of linkage disequilibrium based on the variance of pairwise distances among individuals (see details in Agapow & Burt, 2001). This index is particularly useful for datasets with a high number of loci for which pairwise approaches are difficult to apply. To avoid the consequences of a Wahlund effect, we computed the $\overline{r_d}$ index separately for each sampling site. Genotypes were permuted 1000 times among individuals inside populations to test for linkage disequilibrium.

### 3.2.7 Population differentiation in *Laminaria rodriguezii*

The overall level of genetic differentiation (overall $F_{ST}$) was estimated in Genepop v4.2.2. We performed a G-based test of genotypic differentiation across loci (Goudet et al., 1996) with the same software. The P-value for each locus was computed using the Markov chain method with default parameters (500000 iterations). The significance of the test was evaluated by the method of the harmonic mean P-values (HMP) (Wilson, 2019), implemented in the harmonicmeanp v3 R package. The test compares $P_R$, the HMP of a set of P-values $R$, to $\alpha_L$, the significance threshold for a false positive rate weighted by $\omega_R$, the total number of individual P-values (see Table 1 in Wilson, 2019). The HMP P-value is significant when $P_R \leq \alpha_L \omega_R$. We used a false positive rate of 0.05 and an equal weight for each P-value, considered as similar to Bonferroni correction.

## 3.3 Results

### 3.3.1 Sequencing and SNPs filtering

Approximately 1.2 billion (1 154 828 457) raw reads were produced from 49 samples of *L. rodriguezii* (i.e. 47 individuals and two replicates) and 130 samples of *L. digitata* (i.e. 125 individuals and five replicates). After quality filtering and read trimming, an average of 2.7 million (± 1.5 million SE) and 2.3 million (± 2.1 million SE) high-quality reads per individual were retained for *L. rodriguezii* and *L. digitata*, respectively. The percentage of reads mapped to the *L. digitata* genome, ranged from an average of 85.4% (min: 53.9%; max: 92.2%) for *L. rodriguezii*, to 94.2% (min: 67.9%; max: 97.6%) for *L. digitata*. The initial filtering of loci, which retained those shared by a high proportion of individuals, resulted in catalogs of 44 716 loci

and 60 942 loci for *L. rodriguezii* and *L. digitata*, respectively. At this stage, the number of SNPs was more than eight times higher for *L. digitata* than for *L. rodriguezii* (Table 3.1). Additional quality filtering steps followed by outlier exclusion resulted in the selection of 4077 putatively neutral SNPs among 43 individuals of *L. rodriguezii* and 7364 putatively neutral SNPs among 116 individuals of *L. digitata* (Table 3.1). Specifically, 289 SNPs (6.62%) and 174 SNPs (2.31%) were identified as outliers for *L. rodriguezii* and *L. digitata*, respectively. The rate of missing data was slightly higher in *L. digitata* (i.e. 3.36%) than in *L. rodriguezii* (i.e. 2.24%). The mean SNP error rate across replicates was estimated to be 1.06% in *L. digitata* and 3.89% in *L. rodriguezii* datasets.

### 3.3.2 Clonal structure of *Laminaria rodriguezii*

In *L. rodriguezii*, the distribution of pairwise genetic distances was discontinuous with a gap from the threshold d = 0.022 to the lowest distance reported above this threshold (d = 0.042) (Figure 3.2A). This threshold corresponded to the maximum distance observed among replicates. The use of this threshold allowed us to identify four different clonal lineages with several members: namely MLL-A, MLL-B, MLL-C and MLL-D, composed of three, three, seven, two samples respectively (see the network, Figure 3.3). By selecting only one sample for each clonal lineage, we finally retained 32 different samples in our data analyses from the 43 samples initially present in the whole dataset. At Bonifacio (Southern Corsica), we reported a relatively strong occurrence of clonal lineages (R = 0.20), with the dominance of a few large clones (β = 0.56), which contrasted with the low occurrence of clonal lineages (R = 0.90) of small size (β = 2.66) in Eastern Provence (Table C.2, Supporting Information). For instance, the maximum clonal size reported was seven at Bonifacio against three in Eastern Provence. The analysis of the spatial genetic structure along the transect at Banc Magaud_2 (Figure C.1, Supporting Information) revealed that: i) individuals belonging to the same clonal lineage were relatively close to one another (less than one meter) and ii) the genetic distance between individuals increased with spatial distance (Mantel test, R = 0.40, $p < 0.001$). For the other sites, the sampling protocol (with a minimum spacing of two meters), did not allow a similar investigation of the spatial genetic structure. Contrary to *L. rodriguezii*, no pairwise distance among distinct individuals of *L. digitata* was below the distance threshold defined by replicates (Figure 3.2B), indicating that there was no cluster of genotypes grouped in clonal

lineages. The lowest divergence among individuals in *L. digitata* was around d = 0.035, and corresponded to intra-population differences within the Helgoland population.

Besides the detection of clonal lineages, we observed in *L. rodriguezii* three individuals, with distinct long branches in the phylogenetic networks (indicated as outliers in Figure 3.3). These three outlier individuals shared a common characteristic: the highest proportion of heterozygous genotypes among loci per individual, with an average of 33% against 16% in other individuals (Figure C.2, Supporting Information). These outlier individuals also had the highest percentage of private alleles among all individuals (average percentage of private alleles at the individual level (Pr) = 20.3% for outlier individuals and Pr = 1.0% among the other 40 genotyped individuals, see Figure C.2, Supporting Information). By definition, these private alleles were different among the three individuals.

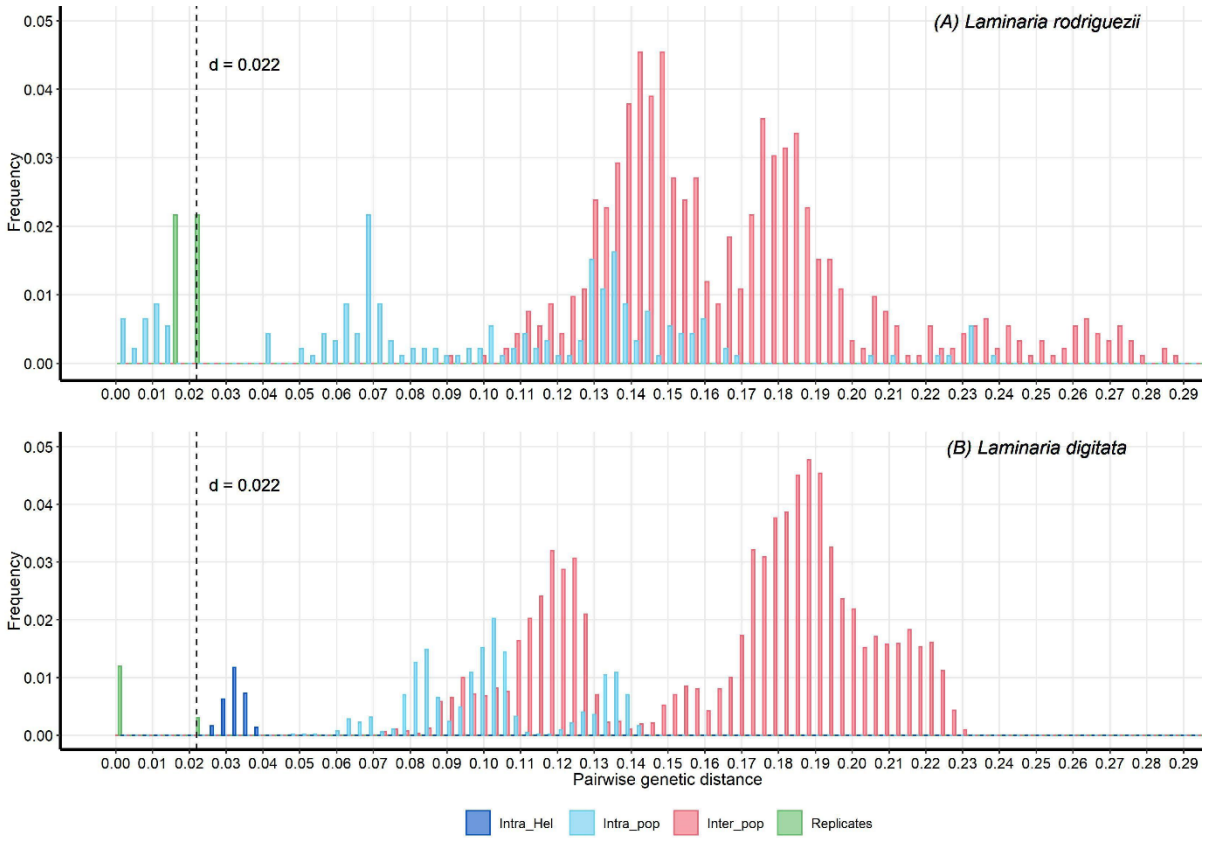



**Figure 3.2**. Frequency of pairwise genetic distances between individuals of (A) *Laminaria rodriguezii* and (B) *Laminaria digitata*. Intra-populational (Intra-pop), inter-populational (Inter-pop), within Helgoland (Intra-Hel) and replicates (Replicates) comparisons are colored according to the legend. The frequency of pairwise genetic distances across replicates was multiplied by 40 to create an optimum bin size. The vertical dotted line shows the distance threshold (d = 0.022) as reported as the maximum value of replicate pairs.

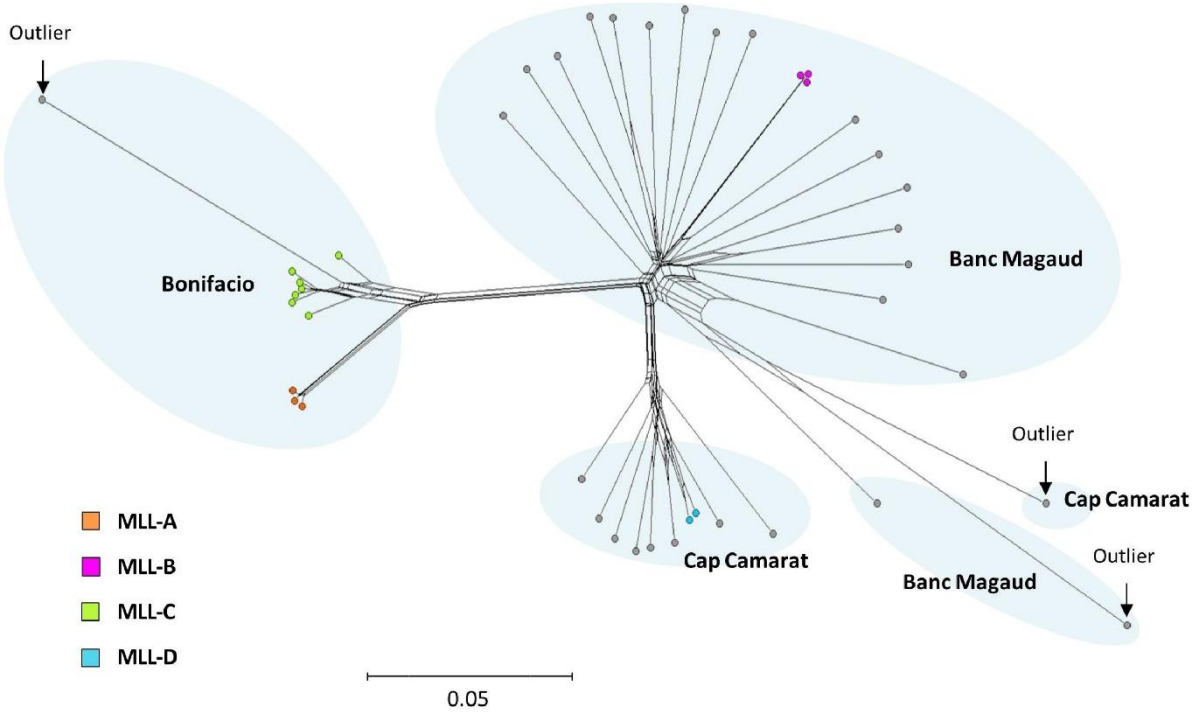



**Figure 3.3**. Phylogenetic networks of *Laminaria rodriguezii* represented with the neighbor-net method. Each dot represents an individual, Individuals belonging to the same clone (MLLs, Multi-Locus Lineages) are colored according to the legend. Individuals characterized by the highest proportion of heterozygous genotypes and private alleles are indicated as outliers.

### 3.3.3 Genetic diversity

*Laminaria rodriguezii* and *Laminaria digitata* displayed similar expected heterozygosity ($H_e$), with an average $H_e$ within populations of 0.14 and 0.15, respectively (Table 3.2). The proportion of polymorphic loci (P) within populations (estimated with the minimum sampling size n = 10 and 100 random resamples) was higher in *L. rodriguezii* (average P = 54.42%) than in *L. digitata* (average P = 39.39%) (see Table 3.2, Figure C.3, Supporting Information). We observed a departure from panmixia in both species, with an excess of heterozygotes in *L. rodriguezii*, and a deficit of heterozygotes in *L. digitata* (see Table 3.3 for *L. rodriguezii* and Table C.1, Supporting Information for *L. digitata*). The heterozygotes excess was statistically significant (Score [U] test, $P < 0.001$) within populations for *L. rodriguezii*, except for Banc Magaud_1. The heterozygotes deficit was significant (Score [U] test, $P < 0.001$) within populations of *L. digitata*.

**Table 3.2**. Comparison of population genetic summary statistics (mean and range of values) in *Laminaria rodriguezii* and *Laminaria digitata*.

| | %P | | | $H_o$ | | | $H_e$ | | | $F_{IS}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***L. rodriguezii*** | | | | | | | | | | | | |
| Average | | 54.42 | | | 0.16 | | | 0.14 | | | -0.14 | |
| Range | 40.96 | – | 72.14 | 0.11 | – | 0.19 | 0.11 | – | 0.18 | -0.47 | – | -0.02 |
| ***L. digitata*** | | | | | | | | | | | | |
| Average | | 39.39 | | | 0.11 | | | 0.15 | | | 0.26 | |
| Range | 21.44 | – | 54.89 | 0.09 | – | 0.14 | 0.12 | – | 0.19 | 0.12 | – | 0.43 |

%P, average percentage of polymorphic loci within populations rarified at n = 10 by random resampling (see methods for details), $H_o$, observed heterozygosity, $H_e$, expected heterozygosity, $F_{IS}$, fixation index.

**Table 3.3**. Genetic variation within populations of *Laminaria rodriguezii* (A) before and (B) after applying the clonal correction.

| **(A)** | | | | | | |
|---|---|---|---|---|---|---|
| **Population** | **Label** | **N** | **$H_o$** | **$H_e$** | **$F_{IS}$** | **$\overline{r_d}$** |
| Banc Magaud_1 | BM_1 | 10 | 0.18 | 0.17 | -0.02 | **0.01** |
| Banc Magaud _2 | BM_2 | 11 | 0.19 | 0.18 | **-0.04** | **0.08** |
| Cap Camarat | CAM | 11 | 0.11 | 0.11 | **-0.05** | **0.10** |
| Bonifacio | BON | 11 | 0.16 | 0.11 | **-0.47** | **0.27** |
| **(B)** | | | | | | |
| **Population** | **Label** | **N** | **$H_o$** | **$H_e$** | **$F_{IS}$** | **$\overline{r_d}$** |
| Banc Magaud_1 | BM_1 | 10 | 0.18 | 0.17 | -0.02 | **0.01** |
| Banc Magaud _2 | BM_2 | 9 | 0.19 | 0.19 | -0.01 | **0.05** |
| Cap Camarat | CAM | 10 | 0.11 | 0.11 | -0.02 | **0.10** |
| Bonifacio | BON | 3 | 0.20 | 0.16 | **-0.26** | **0.20** |

Clonal correction consists of keeping only one representative of clonal lineages. N, number of individuals successfully genotyped, $H_o$, observed heterozygosity, $H_e$, expected heterozygosity, $F_{IS}$, fixation index, $\overline{r_d}$, average multilocus linkage disequilibrium. $F_{IS}$ and $\overline{r_d}$ values highlighted in bold indicate significant heterozygote excess within populations (P-val < 0.001) and significant linkage disequilibrium (P-val < 0.001), respectively.

The distributions of $F_{IS}$ were significantly different for all intra and inter-specific comparisons of *L. rodriguezii* and *L. digitata* populations (KS tests, $p < 0.001$, D statistic: 0.07-0.53). At the intraspecific level, the most striking difference in the distributions of $F_{IS}$ (KS tests, D statistic: 0.26-0.27) was detected for *L. rodriguezii* between Eastern Provence (Banc Magaud_1, Banc_Magaud_2 and Cap Camarat) and Bonifacio (Figure 3.4A). The distributions of $F_{IS}$ in Eastern Provence were moderately skewed to negative values with a long tail to

positive values (skewness = 0.69; kurtosis = 4.07). The distribution of $F_{IS}$ in Bonifacio was highly skewed toward negative values (skewness = -0.95; kurtosis = 3.14) and showed a much higher proportion of extreme negative values as above to -0.8 (13.6%) in comparison to those measured in Eastern Provence (average, 0.6%). Similarly, the variance in $F_{IS}$ among loci was higher at Bonifacio (var = 0.13) than in Eastern Provence (var = 0.08, 0.07 and 0.06 for Banc Magaud_1, Banc Magaud_2, and Cap Camarat, respectively).

In contrast to these results, $F_{IS}$ values were mainly positive in *L. digitata*, whatever the population considered even if the distribution mode of $F_{IS}$ is always observed between -0.1 and 0 (Figure 3.4B). Extreme negative $F_{IS}$ (above -0.8) and even moderate to high negative values (those between -0.5 and -0.3) were almost absent in *L. digitata* with frequencies of 0.03% and 0.06%, respectively. The distribution mode of $F_{IS}$ was observed for slightly negative values between -0.1 and 0 whatever the population considered (Figure 3.4B). The asymmetry of the distributions, as revealed by the positive skewness (skewness = 0.98; kurtosis = 3.32) is indicative of a long tail toward high and extreme positive values (0.5 to 1). This class of extreme positive $F_{IS}$ corresponded to 28% of all loci in Helgoland. At the intra-specific level, the second major difference in the shape of the distribution of $F_{IS}$ was reported between Helgoland and all other populations of *L. digitata* (KS tests, D statistic 0.11-0.23).

The average multilocus estimates of $F_{IS}$ within populations ranged from 0.12 to 0.43 for *L. digitata* (Table C.1, Supporting Information) whereas mean values were mainly negative within populations in *L. rodriguezii*, with a range from -0.47 in Bonifacio to -0.02 at Banc Magaud_1 (Table 3.3A). After applying the correction for repeated genotypes, the average multilocus estimates of $F_{IS}$ increased to a range from $F_{IS}$ = -0.26 at Bonifacio to $F_{IS}$ = -0.01 at Banc Magaud_2 (Table 3.3B). After the correction for repeated genotypes, the heterozygosity excess only remained significant at Bonifacio (Score [U] test, $P < 0.001$).

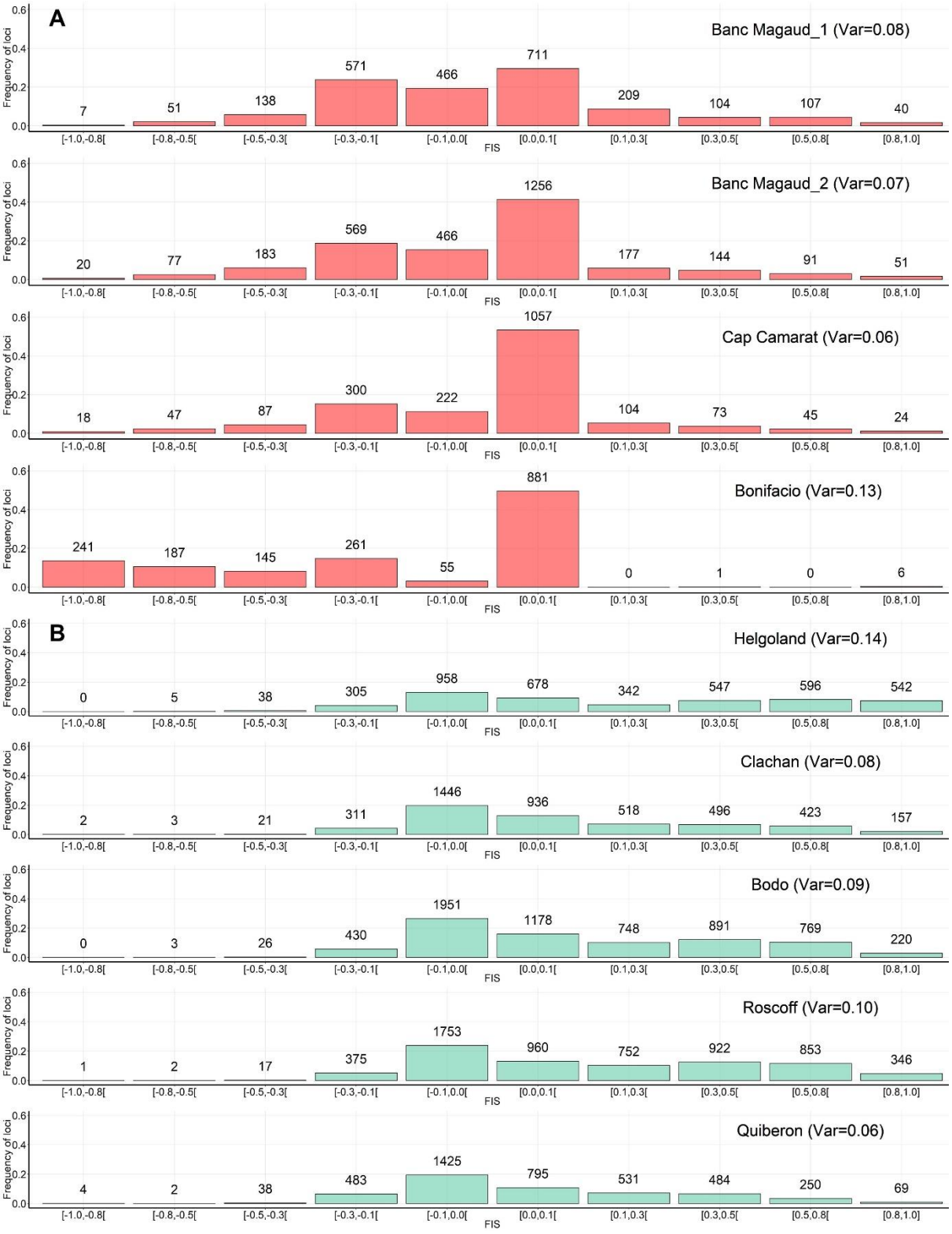



**Figure 3.4**. Distributions of the fixation index $F_{IS}$ among loci within populations of (A) *Laminaria rodriguezii* and (B) *Laminaria digitata* computed across 4 077 and 7364 putatively neutral SNPs, respectively. The $F_{IS}$ values were distributed into 10 discrete classes and the number of occurrences for each class is indicated on top of histogram bars. The variance of $F_{IS}$ among loci (Var) is indicated in brackets next to population labels.

The distribution of the index of multilocus association $\overline{r_d}$ (Figure 3.5) was nearly identical whatever the number of SNPs subsampled (i.e. 1000; 2000 and 3000 random SNPs). The mean $\overline{r_d}$ was statistically significant ($p < 0.001$) within all populations of both species. For *L. rodriguezii*, the $\overline{r_d}$ index reached high levels within all populations (mean $\overline{r_d}$ = 0.08 to 0.27), except at Banc Magaud_1 (mean $\overline{r_d}$ = 0.01) for which there was no repeated sampling of a clonal lineage (Table 3.3). The values of $\overline{r_d}$ decreased slightly by keeping only one representative of each clonal lineage (Table 3.3, Figure 3.5). For *L. digitata*, the mean $\overline{r_d}$ ranged between 0.01 to 0.03 within all populations (Table C.1, Figure 3.5), except for Helgoland that displayed a remarkably high level of linkage disequilibrium similar to values observed in *L. rodriguezii* (mean $\overline{r_d}$ = 0.20).

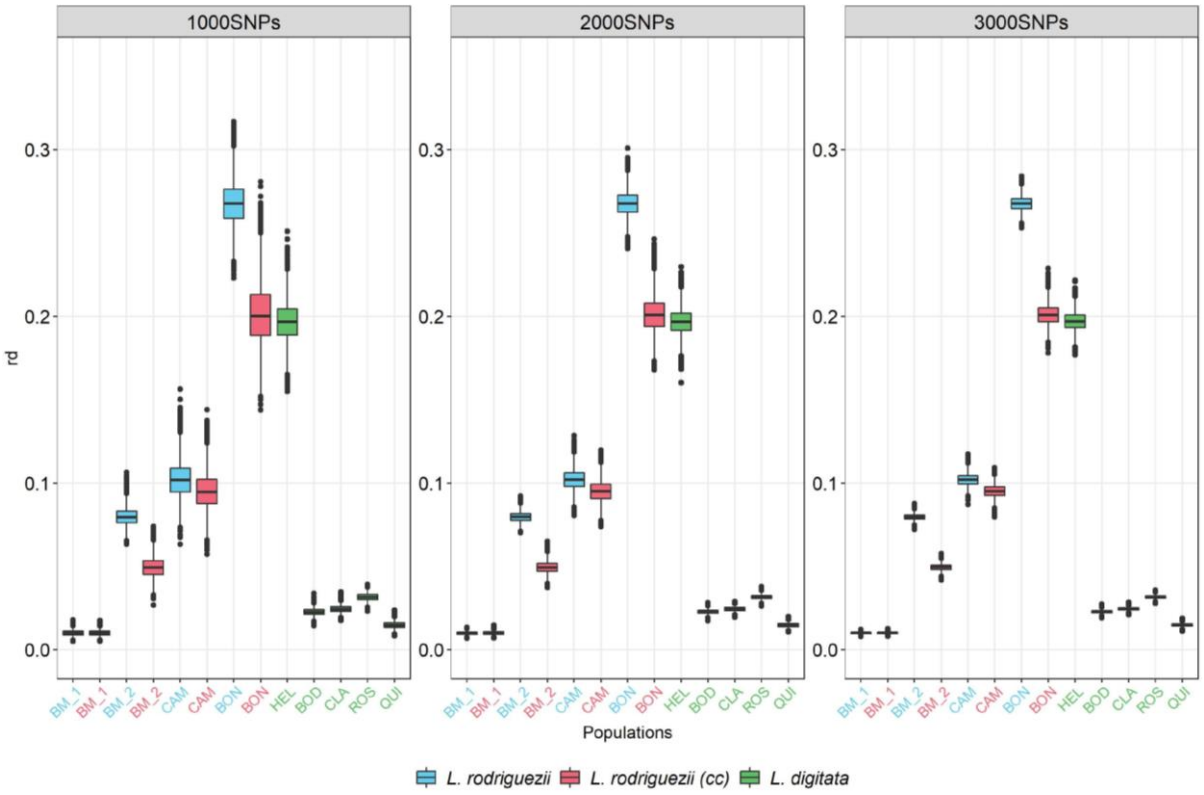



**Figure 3.5**. Distributions of the standardized index of multilocus association $\overline{r_d}$ within populations of *Laminaria rodriguezii* before (highlighted in blue) and after (highlighted in red) clonal correction (cc). Comparison of the index value was done with five populations of *Laminaria digitata* (highlighted in green). The $\overline{r_d}$ values were estimated using 10 000 iterations by random subsampling of 1000; 2000 and 3000 SNPs. The abbreviations used for populations are those reported in Table 3.3 for *L. rodriguezii* and Table C.1 for *L. digitata*.

### 3.3.4 Population differentiation in *Laminaria rodriguezii*

*Laminaria rodriguezii* exhibited high levels of population structure with a global $F_{ST}$ of 0.28 when considering all individuals (n = 43). After applying the clone correction (i.e. one sample per clonal lineage), the estimate of genetic differentiation decreased (global $F_{ST}$ = 0.18). After correcting for multiple testing, genotypic differentiation among populations of *L. rodriguezii* was significant (HMP significance threshold, $\alpha_L \omega_R = 0.032$; HMP P-value, $P_R = 0$), with and without the clone correction.

### 3.3.5 Negative $F_{IS}$ and high $F_{ST}$ in *L. rodriguezii* over SNP filtering methods

For *L. rodriguezii*, the mean $F_{IS}$ and overall $F_{ST}$ varied between the reference dataset and the filtered datasets (Table 3.4). However, the mean $F_{IS}$ was always negative, with high variance among loci, whatever the levels of stringency. In the same way, the overall $F_{ST}$ stayed high and it was never below 0.27. The maximum rate of missing data allowed per SNP (max-missing) had a very limited effect on mean $F_{IS}$ and overall $F_{ST}$, which ranged from -0.06 to -0.07 and from 0.27 to 0.30, respectively. The highest variance was associated with MAF cutoff as the filter excluded the lowest polymorphic loci, and subsequently increased the level of genetic variation in the dataset. Therefore, we choose to keep the MAF cutoff at 2% so as not to inflate the levels of genetic variation in *L. rodriguezii*. The SNP error rate decreased from 5.22% in the reference dataset to 2.74% in the dataset which was filtered by excluding loci with low coverage (i.e. meanDP > 15X). Genotyping errors were probably non-randomly distributed and occurred mostly in loci with low coverage, so we excluded them by applying meanDP > 15X.

**Table 3.4**. The effect of SNP filtering cutoffs on the mean (± SD) $F_{IS}$ over loci, overall (± SD) $F_{ST}$ over loci and the SNP error rate in *Laminaria rodriguezii*.

| Quality filter | cutoff | SNPs count | missingness (%) | coverage | mean (± SD) $F_{IS}$ | overall (± SD) $F_{ST}$ | SNP error rate (%) |
|---|---|---|---|---|---|---|---|
| Reference dataset | | 21 357 | 4.11 | 13.00 | -0.07(0.20) | 0.27(0.18) | 5.22 |
| max-missing (%) | < 15 | 21 114 | 3.97 | 13.08 | -0.07(0.20) | 0.27(0.18) | 5.17 |
| | < 10 | 19 468 | 3.33 | 13.52 | -0.07(0.20) | 0.27(0.18) | 4.94 |
| | < 5 | 14 480 | 1.95 | 14.88 | -0.07(0.19) | 0.28(0.18) | 4.30 |
| | 0 | 5 479 | 0.00 | 18.12 | -0.06(0.18) | 0.30(0.19) | 2.90 |
| MAF (%) | > 2 | 12 735 | 4.28 | 13.31 | -0.08(0.26) | 0.30(0.22) | 7.07 |
| | > 5 | 5 094 | 4.54 | 16.18 | -0.12(0.34) | 0.37(0.26) | 4.92 |
| | > 10 | 3 428 | 4.56 | 17.29 | -0.14(0.37) | 0.42(0.26) | 3.86 |
| meanDP (X) | > 5 | 19 245 | 3.59 | 13.71 | -0.08(0.20) | 0.28(0.18) | 4.88 |
| | > 10 | 11 773 | 2.52 | 17.41 | -0.11(0.22) | 0.31(0.21) | 3.92 |
| | > 15 | 6 163 | 1.90 | 21.85 | -0.13(0.23) | 0.33(0.22) | 2.74 |
| max-missing cutoff for fixed MAF (2%) and meanDP (>15) | < 20 | 3 911 | 2.14 | 21.83 | -0.15(0.29) | 0.34(0.25) | 3.71 |
| | < 10 | 3 850 | 1.99 | 21.81 | -0.14(0.28) | 0.35(0.25) | 3.70 |
| | < 5 | 3 464 | 1.41 | 21.90 | -0.12(0.28) | 0.35(0.25) | 3.53 |
| | 0 | 1 813 | 0.00 | 22.71 | -0.10(0.27) | 0.37(0.26) | 2.68 |
| matching genotypes | | 3 638 | 2.14 | 22.07 | -0.15(0.29) | 0.35(0.25) | 0.00 |

As thresholds, we considered the maximum rate of missing data allowed per SNP (max-missing), the Minor Allele Frequency (MAF), the mean read depth per locus across all individuals (meanDP). For each cutoff value we reported the number of SNP passing the filter (SNPs count), the mean missing data rate per individual (missingness), and the mean read depth per individual (coverage). We considered an additional set of SNPs (matching genotypes) for which replicates have exactly the same multi-locus genotypes (SNP error rate = 0).

## 3.4 Discussion

We used RAD-sequencing to analyze the genome-wide diversity of the rare *Laminaria rodriguezii*. We found that intra-individual genomic diversity is a powerful tool to identify clonal lineages. We also proposed that the observed $F_{IS}$ distribution and linkage disequilibrium at the genome scale can be compared with theoretical models to search for a discernable signal of clonality, even in this partially clonal reproducing species. Finally, we will discuss here the findings of high heterozygosity within some individuals that raise questions about the mechanisms allowing the persistence of such genomic features.

### 3.4.1 RAD-seq as a powerful tool for detection of clonal lineages

The identification of clonal lineages in natural populations was addressed by assessing the probability that identical or slightly distinct genotypes belong to the same clonal lineage. That can be achieved through the development of standardized methods allowing for investigation of the influence of somatic mutations and scoring errors (Arnaud-Haond et al., 2007). However, while this methodological framework is particularly robust for a small number of loci, RAD-seq and the amount of data generated by Next Generation Sequencing (NGS) brings new challenges. The main concern is that increasing the number of loci is equivalent to increasing the number of missing data and genotyping errors that can subsequently affect the detection of clonal lineages. This is especially the case for reduced representation sequencing (e.g. RAD-seq) sensitive to missing data and genotyping error, thus necessitating appropriate filtering strategy (e.g. Mastretta-Yanes et al., 2015; Shafer et al., 2017; Andrews et al., 2018; Boscari et al., 2019). How then can we find an adequate distance threshold for collapsing slightly distinct genotypes into clonal lineages? For that purpose, we can define a threshold using the pairwise genetic distances, a threshold below which genotypes are assumed to belong to the same clonal lineage. Some studies have fixed a cut-off based on the gap in pairwise genetic distances (e.g. Shrestha et al., 2014), while others have defined an arbitrary 95% genotype similarity threshold (e.g. Locatelli & Drew, 2019). Nevertheless, such approaches can be inaccurate, particularly when different sources of noise (i.e. genotyping error, missing data, somatic mutation) can accentuate differences between clonal genotypes above the detection threshold. This also depends on the genetic diversity of the populations: clonal lineages are more easily identified in populations where the diversity

among sexually produced individuals is high, i.e. in partially clonal species. In this study, we proposed defining a threshold on the basis of the pairwise genetic distance across replicates. While this approach seems to be useful in evaluating the genetic distance expected as a result of genotyping errors and somatic mutations, it has rarely been reported, except in *Leptopsammia pruvoti* (Boscari et al., 2019). For this organism genotyped with 1386 SNPs (with the 2b-RAD protocol), slightly distinct genotypes below a distance threshold set at d = 0.043 were collapsed into clonal lineages. For *L. rodriguezii*, this cut-off was lower (d = 0.022), even if of a similar order as in *L. pruvoti*. This difference highlights the importance of not considering a common threshold but rather of adjusting it using genotype mismatch across replicates. Note that using the same method on both species, the threshold was very similar between *L. rodriguezii* and *L. digitata.*

### 3.4.2 The genomic patterns of clonality

Our results show how the distribution of $F_{IS}$ among loci can be used to study species and populations with various signals of clonality including partial clonality. Signatures of clonality in natural populations are usually inferred from repeated genotypes, along with mean negative values of $F_{IS}$, and linkage disequilibrium across loci (Balloux et al., 2003; Halkett et al., 2005). Nevertheless, recent studies reported that one of these signatures, genotypic richness, had poor relevance for measuring the relative importance of sexual versus clonal reproduction (Arnaud-Haond et al., 2020), even using relatively large sample sizes (from 100 to 500 individuals) (Stoeckel et al., 2020). This is because the estimates of genotypic richness are strongly dependent on sampling strategy and density, basically decreasing with increasing sampling effort (Arnaud-Haond et al., 2007b; Gorospe et al., 2015; Becheler et al., 2017). In our study, by looking at the $F_{IS}$ distribution in the genome, we propose a new analysis which seems useful to detect partial clonal reproduction. Our approach confirmed the prediction that inter-locus variance of $F_{IS}$ (rather than an average value) increased as clonal rates increased (Balloux et al., 2003; Halkett et al., 2005; Stoeckel & Masson 2014; Reichel et al., 2016; Stoeckel et al., 2020). We also reported (1) contrasted distribution of $F_{IS}$ between species, with $F_{IS}$ shifted towards negative values in the partially clonal species *L. rodriguezii*, and towards positive values in the fully sexually reproducing species *L. digitata*; (2) striking differences in distributions of $F_{IS}$ among populations of *L. rodriguezii*, with higher inter-locus variance in $F_{IS}$ at Bonifacio (Corsica) than in the three localities of Eastern Provence.

Remarkably, the empirical distributions of $F_{IS}$ in Eastern Provence matched well with the simulation results of Stoeckel & Masson (2014) which predicted that most $F_{IS}$ values are concentrated between -0.3 and 0.1 below intermediate levels of clonality.

Genetic drift is another possible explanation for negative $F_{IS}$ values. In a finite population with random mating, genetic drift due to sampling effect indeed tends to create a slightly negative mean $F_{IS}$, but the effect is small, of order 1 / (2N) where N is the size of the population (see Figures 3 and 4 in Stoeckel & Masson, 2014). For the fully sexual dioicous species *L. digitata*, this drift effect could explain why the distribution mode of $F_{IS}$ is between -0.1 and 0 whatever the population considered. However, the probability of highly negative values increases when genetic drift is combined with intermediate rates of clonality (Stoeckel & Masson, 2014). Nevertheless, Stoeckel & Masson (2014) showed that values below -0.3 remain rare if the clonality rate is below 0.9. This is indeed what is observed for all *L. rodriguezii* populations, suggesting that their clonality rates are intermediate, except at Bonifacio, the Corsican population for which we observed the highest proportion of $F_{IS}$ below -0.3 and even below -0.8. In addition, the estimates of genotypic richness (R), Pareto (β), and linkage disequilibrium ($\overline{r_d}$) emphasized the degree of distinction between the Eastern Provence and Bonifacio populations, thus strengthening our hypothesis of a higher level of clonality in Corsica than in Eastern Provence. A more in-depth investigation of the combined effects of drift and clonal reproduction on the distribution of $F_{IS}$ would be useful in this context. Parameters other than reproductive mode differ between the two species compared here, such as the levels of genetic drift, the investigated geographical range, and potentially their demographic history. The population of Helgoland for the fully sexually reproducing species, *L. digitata*, reached a high level of linkage disequilibrium comparable to the Bonifacio population of *L. rodriguezii* for which the signature of clonality was clearly discernible. Nevertheless, Helgoland had a distribution of $F_{IS}$ markedly different from *L. rodriguezii*, with a majority of positive values. Helgoland is a rocky island surrounded by several hundreds of kilometers of sandy seafloor. This population shows the highest levels of positive $F_{IS}$, the lowest values of pairwise genetic distances between individuals, and the highest value of linkage genetic disequilibrium observed for *L. digitata*. This is most probably explained by the effect of low effective size and maybe of a bottleneck. This population also showed the lowest level of genetic diversity observed for *L. digitata* when genotyped with the microsatellite locus in the study of Liesner et al. (2020). There is a lack of case studies to conclude more precisely

on the role of drift and of demographic history in the distribution of inter locus $F_{IS}$ variance. However, the theoretical results of Stoeckel & Masson (2014) and Stoeckel et al. (2020) agree well with our experimental results. This paves the way toward a better assessment of the importance of clonality through genomic studies, but it remains the case that the estimation of clonality rates is still only possible using temporal sampling (Ali et al., 2016; Becheler et al., 2017).

### 3.4.3 Extreme levels of individual genomic heterozygosity

We reported an intriguing signal of genome-wide heterozygosity in three of the 43 individuals in *L. rodriguezii*. These outlier individuals were characterized by an exceptionally high proportion of heterozygous genotypes compared to other individuals. Given the relatively low sample sizes, we do not really know how these individuals depart from the population distribution in individual heterozygosity, but we can discuss several hypotheses regarding their origin. High heterozygosity at particular loci can correspond to paralogs assembled in the same RAD locus depending on assembly parameters (Verdu et al., 2016; Mckinney et al., 2017b; Nadukkalam Ravindran et al., 2018). However, for the following reasons, we would argue that polymorphisms detected within these outlier individuals do not correspond to artefactual loci: firstly, sequencing reads were mapped to the reference genome of a closely related species, followed by stringent filter criteria, including for maximum heterozygosity. Secondly, paired-end sequencing protocols provide additional information on correct read placement since RAD loci are retained when both reads are properly mapped (according to the insert length between forward and reverse reads). Thirdly, if paralogs were responsible for high levels of heterozygosity, we would not expect to observe this in only three individuals. On the other hand, recent studies have revealed that both cross-contamination among samples and environmental contamination, especially during library preparation and/or sequencing, can lead to artifactual results, with the sequencing of loci of different origins in the same sample (Laurin-Lemay et al., 2012; Ballenghien et al., 2017). Here, the three outlier individuals were also characterised by the highest proportion of private alleles at the individual level, and they were not grouped on the network. Such private alleles can then not be explained by cross-contamination among samples of the same species. Rather they could be explained by the accumulation of mutations along with clonal reproduction, thus increasing the divergence among loci in the same individuals, a potential signal of the

Meselson effect (Birky, 1996; Welsh & Meselson, 2000; De Meeûs et al., 2007). These outlier individuals could then belong to a relatively ancient asexual lineage. However, we lack temporal sampling to get more precise estimates of the rates of clonal reproduction, and therefore, the hypothesis of clonality to explain these outlier individuals remains highly speculative. A high frequency of heterozygous loci could also be observed following hybridization among well differentiated genetic lineages. Hybridization has indeed been proposed as the main driver of high heterozygosity in parthenogenetic animals (Jaron et al., 2020). Here, the hybrid hypothesis should involve several hybridizations from three different sources to explain the observed pattern of private alleles, as well as the genetic distance among the three outliers. It would therefore be interesting to analyze the genomic diversity of other *L. rodriguezii* populations in the Mediterranean in order to go further on this topic. Nevertheless, the populations of *L. rodriguezii* are scarce, and there are no other congeneric species in this area.

### 3.4.4 Implications for conservation

Our results emphasise that *L. rodriguezii* is spatially structured into genetically distinct populations. At Banc Magaud_2 we also evidenced a significant spatial genetic structure: regarding the low occurrence of clonality in this population, this structure is better explained by reduced dispersal through haploid spores. Although high levels of population genetic structure are not uncommon in kelps (e.g. Billot et al., 2003; Coleman et al., 2009; Durrant et al., 2018), our results contrast with the lack of genetic differentiation among populations of *L. digitata* spatially distant by up to 10 km (Billot et al., 2003; Robuchon et al., 2014). In this respect, the spatial scale of dispersal of *L. rodriguezii* is presumably smaller than that reported in other species of the genus *Laminaria* (Santelices, 1990). The patchy distribution and small population sizes of *L. rodriguezii* can probably explain this pattern of high and local genetic differentiation through strong local genetic drift and limited gene flow among populations. The identification of such a local genetic structure has important implications for management and conservation. The high genetic drift and limited gene flow can explain the lower genetic diversity expected in *L. rodriguezii* compared to *L. digitata*, a species with a wider geographical distribution, exhibiting semi-continuous and dense populations (see Hamrick & Godt, 1990; Frankham, 1996; Levy et al., 2016). Furthermore, the range of *L. rodriguezii* has been drastically reduced during the last half-century, especially in the Adriatic Sea, where the

species has suffered losses of 85% from its historical range (Žuljević et al., 2016). However, the observed heterozygosity and the percentage of polymorphic loci were higher in *L. rodriguezii* when compared to *L. digitata*. This can be explained by the conservation of heterozygous genotypes with low recombination rates (Judson & Normark, 1996; Balloux et al., 2003; see Meloni et al., 2013 for an empirical example), as genetic drift under clonal reproduction acts on genotype frequencies rather than allele frequencies (Stoeckel & Masson, 2014; Reichel et al., 2016). Further studies would be necessary to test with genomic data whether the current fragmented range and small populations of *L. rodriguezii* correspond to a demographic reduction from a larger ancestral population. The extensive cartography and monitoring of the biocenoses down to 100 m along French coasts do point to the rarity of this species. The conservation of these last isolated populations should become a high priority for stakeholders.

## 3.5 Conclusion

Our study shows that despite limitations on the detection of clonal lineages, SNPs derived from RAD sequencing are interesting markers for population genomic studies of partially clonal species. The interaction between demographic history and reproductive mode shaped the distribution of $F_{IS}$ and linkage disequilibrium in *L. rodriguezii*. Interestingly, even if our loci were mapped to a reference genome for assembly, we did not use the physical location of loci in our analyses, which indicates that similar approaches could be envisioned with *de novo* assembly. In order to decipher the respective impact of demography and reproduction in our results, one next step could be to use summary statistics linked with these distributions (such as higher moments) for evolutionary inferences, using for example Approximate Bayesian Computation (Csilléry et al., 2010).

## 3.6 Acknowledgements

This project was funded, by the MARFOR Biodiversa/0004/2015 project (http://marfor.eu/), and by the Agence Nationale de la Recherche, Grant/ Award Number: ANR-18-CE32-0001 (Clonix2D). Lauric Reynes also received support from the GDR Génomique Environnementale (GDR3692). The project leading to this publication has received funding from the European FEDER Fund under project 1166-39417. This work was also funded by a Ph.D. Fellowship from Région PACA and the Calanques National Park. The authors would like to thank Julie Deter, in charge of the scientific coordination during sampling, the team of research divers of Andromède Océanologie for the prowess in sampling

*Laminaria rodriguezii* at 76 meters depth, the Biogenouest genomics core facility (Genomer Plateforme génomique at the Station Biologique de Roscoff) for their technical support, Claire Daguin-Thiébaut and Quentin Langevin for useful discussions regarding the optimization processes in purpose-extracted high-quality DNA in brown macroalgae, Marck Cock providing the draft genome of *Laminaria digitata*, Thierry Gosselin for very useful discussions on the quality control of RADseq data, and finally Solenn Stoeckel for his comments on the manuscript. We deeply thank Thierry De Meeûs, three anonymous reviewers and the corresponding editor, Tatiana Giraud, for their comments that enabled us to greatly improve the current version of this manuscript. Finally, we would like to thank Michael Paul for English language proofreading.

## 3.7 Appendices

**Materials C.1**.The effect of SNP filtering on population genetic summary statistics.

To address the effect of SNP filtering we started from the output of the catalog of 867 513 loci of *L. rodriguezii*, in which 283 864 SNPs were initially reported using the Stacks v2.4 bioinformatics pipeline (Rochette et al., 2019). Then, we selected a subset of informative SNPs by requiring a locus to be present at a frequency of greater than 80% at each of the four study sites. Additionally, we excluded SNPs with more than 80% heterozygosity and selected the first SNP of each RAD locus. As reported in the present study, we eliminated four individuals (with from 35% to 63% of missing genotypes) for which RAD-sequencing has been unsuccessful. At the end of this step, 21 357 informative SNPs were genotyped among the 43 individuals of the present study, thus constituting the reference dataset for the following analyses. We applied many filtering cutoffs (see Table 3.4) to the reference dataset, using VCFtools v0.1.17 (Danecek et al., 2011), thus producing 15 VCF files. Using the same software, we computed the number of SNP passing the filter (SNPs count), the mean missing data rate per individual (missingness), and the mean read depth per individual (coverage) for each dataset. Finally, we computed both $F_{IS}$ and $F_{ST}$ values (Weir & Cockerham, 1984) using the HierFstat R package (Goudet, 2005) and the mean SNP error rate between replicates with the R function developed by Mastretta-Yanes et al., (2015). We also considered a dataset including SNPs for which genotypes are strictly identical between replicates (matching genotypes) to consider the effect of potentially problematic SNPs on population genetic summary statistics. The results of these analyses are presented in the main manuscript (see Table 3.4).

**Additional Supporting Information may be found in the Annexe C.**

**Figure C.1**. Spatial analyses at Banc Magaud_2.

**Figure C.2**. Genomic patterns of the three outliers as comparison to other individuals of *Laminaria rodriguezii*.

**Figure C.3**. Average percentage of polymorphic loci (%P) within populations of (A) *Laminaria rodriguezii* and (B) *Laminaria digitata*.

**Table C.1**. Genetic variation within populations of *Laminaria digitata*.

**Table C.2**. Genotypic richness (R) and Pareto (β) indexes estimated for populations of *Laminaria rodriguezii*.

## 3.8 Structure génétique des populations de *L. rodriguezii*

En complément des analyses de diversité génétique qui nous ont permis d'apprécier les signatures de clonalité dans les populations de *Laminaria rodriguezii*, j'ai effectué des analyses plus poussées de la structure génétique des populations avec plusieurs objectifs. Il s'agissait d'étudier les limites aux flux de gènes et le niveau de proximité relative entre les populations étudiées d'une part, et d'analyser l'origine des individus très hétérozygotes d'autre part. Ces analyses ont été réalisées à partir du jeu de données de l'article publié dans Molecular Ecology, soit 43 individus génotypés par ddRAD-seq à 4077 SNPs.

***Matériels et méthodes***

Plusieurs méthodes ont été utilisées pour étudier la structure génétique des populations. Les valeurs de $F_{ST}$ par paire (Weir & Cockerham, 1984) ont été estimées avec le package HierFstat v.0.04-22 R (Goudet, 2005), et des tests exacts de différenciation génotypique, implémentés dans Genepop v4.2.2 ont été réalisés par paire de populations à chaque locus (Rousset, 2008). Des estimations non biaisées des probabilités associées à ces tests ont été calculées à l'aide d'un algorithme de chaîne de Markov (MCMC) en gardant les paramètres par défaut de Genepop. Les probabilités des tests multiples (un test par locus) ont été combinées par la méthode de Fisher de sorte à obtenir une probabilité multi-locus. Nous avons également employé une analyse de factorisation matricielle non négative (sNMF) (Frichot et al., 2014) implémentée dans le package LEA v.2.4.0 R (Frichot & François, 2015). L'algorithme sNMF ne suppose pas la panmixie, ce qui est intéressant pour les espèces à systèmes de reproduction mixtes. L'algorithme sNMF a été exécuté avec 100 réplicas pour chaque valeur du nombre de groupe testé (K) et ceci pour K compris entre un et dix, les valeurs des autres paramètres correspondant aux valeurs par défaut. De plus, la structure de la population a été évaluée à l'aide d'une analyse en composantes principales (ACP) implémentée dans le package adegenet R (Jombart et Ahmed, 2011). Pour l'ACP, l'ensemble de données était centré et les valeurs manquantes ont été remplacées par les fréquences moyennes des allèles de la population.

***Une structure génétique à courte distance***

Comme indiqué précédemment, *L. rodriguezii* possède des niveaux élevés de structure génétique (voir la partie 3.3.4 « Population differentiation in *Laminaria rodriguezii* »). Les $F_{ST}$

par paire sont compris entre 0,01 pour la comparaison entre les deux sites du Banc du Magaud et 0,44 pour la comparaison entre les populations du Cap Camarat (est de la Provence) et celle de Bonifacio (sud de la Corse). Ces valeurs ont été obtenues en gardant qu'un seul clone par lignée clonale (n = 32). Par ailleurs, en conservant l'ensemble des individus génotypés (n = 43), donc en incluant les génotypes répétés des lignées clonales, les $F_{ST}$ par paire sont plus élevés et compris entre 0,02 et 0,48 (Table 3.5). De même, tous les tests de différenciation génotypiques sont significatifs à P <0,001, à l'exception de ceux entre les deux sites du Banc du Magaud pour lesquels les probabilités des tests de différenciation étaient de 1 par la méthode de Fisher, avant et après la correction des clones.

**Table 3.5**. $F_{ST}$ par paire, des populations de *Laminaria rodriguezii* génotypés à 4077 SNPs.

| **$F_{ST}$** | Banc Magaud_1 | Banc Magaud_2 | Bonifacio | Cap Camarat |
|---|---|---|---|---|
| Banc Magaud_1 | - | 0,02 | 0,35* | 0,20* |
| Banc Magaud_2 | 0,01 | - | 0,34* | 0,19* |
| Bonifacio | 0,29* | 0,25* | - | 0,48* |
| Cap Camarat | 0,19* | 0,17* | 0,44* | - |

Au-dessus de la diagonale, les valeurs ont été calculées en incluant tous les individus génotypés (n = 43), en dessous de la diagonale en incluant un représentant par lignée clonale (n = 32). * Tests de différenciation génotypique avec une probabilité inférieure à 0,001.

L'analyse sNMF révèle que le nombre de clusters le plus probable selon le critère d'information bayésien (BIC) est K = 6 (Figure C.4, Annexe C). Dans cette configuration, les individus sont regroupés par localité et ceux du Banc du Magaud appartiennent à un même groupe génétique (Figure 3.6A). Les entités supplémentaires correspondent à une partie des lignées clonales (Figure 3.6A ; MLL-A en bleu et MLL-B en orange, voir également Figure 3.3) et aux individus qualifiés d'outliers (Figure 3.6A ; en gris) en raison de leurs très fortes proportions d'allèles privés et de génotypes hétérozygotes par comparaison à d'autres individus (Figure C.2, Annexe C). Par ailleurs, l'analyse par ACP est en accord avec les regroupements obtenus par sNMF et le degré de différenciation génétique mesuré par le $F_{ST}$. La première composante principale, représentant 27% de la variance génétique, sépare le sud de la Corse de l'est de la Provence (Figure 3.6B). La deuxième composante principale, représentant 11% de la variance, opère la distinction entre le Cap Camarat et les sites du Banc du Magaud. Par ailleurs, les localités du Banc du Magaud ne sont pas discriminées par l'ACP.

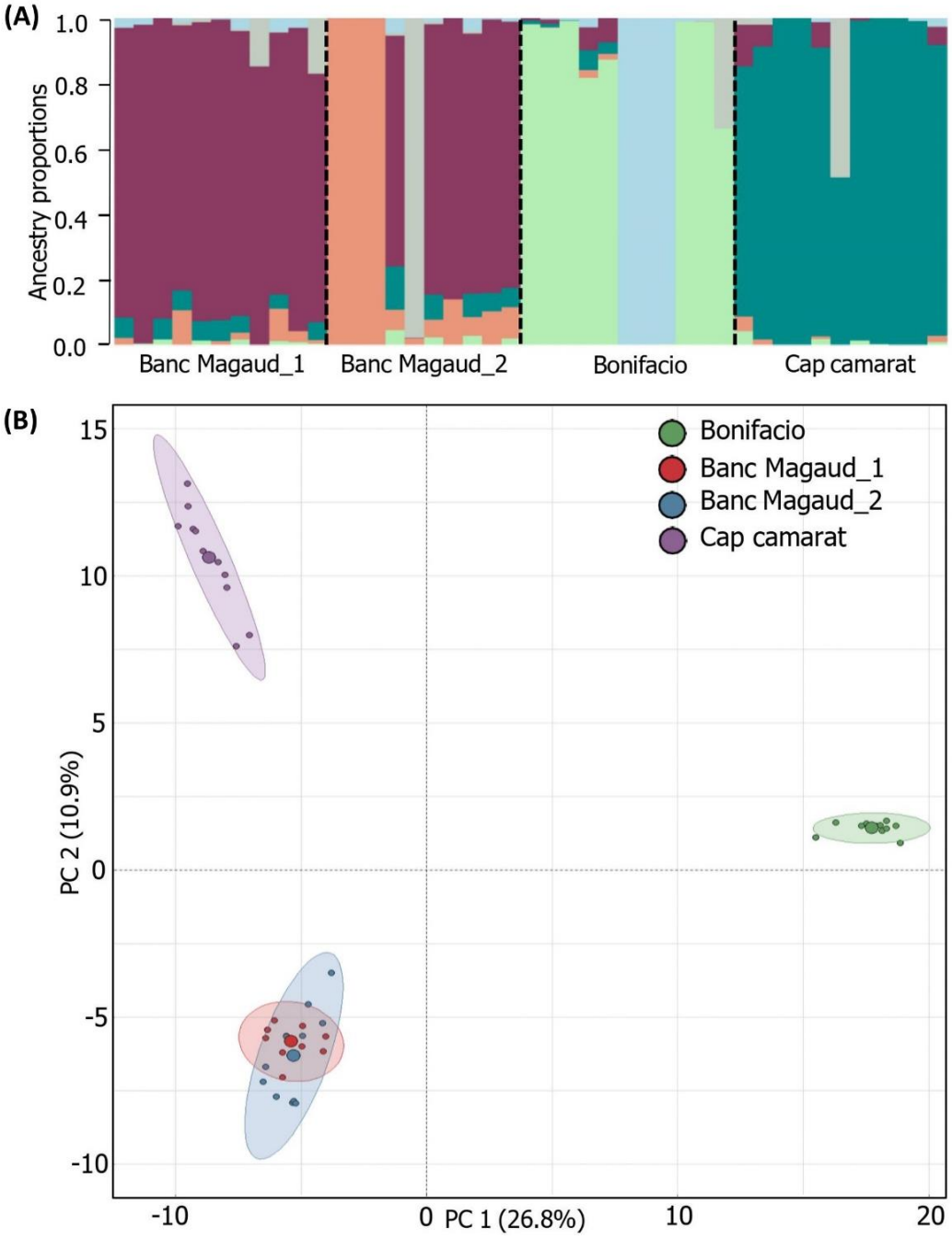



**Figure 3.6**. Structure génétique des populations de *Laminaria rodriguezii* en considérant 43 individus génotypés à 4077 SNPs. (A), l'analyse sNMF où chaque individu est représenté par une fine barre verticale partitionnée en K = 6 segments correspondant aux coefficients d'appartenance associées à une couleur. (B), nuage de points de l'ACP des deux premières composantes principales, chaque point indique un individu et les ellipses, l'inertie de chaque population.

## 3.9 Synthèse des résultats obtenus

L'analyse des données de séquençage RAD a mis en évidence un déplacement de la distribution du $F_{IS}$ vers des valeurs négatives au sein des populations de *L. rodriguezii*, avec un nombre élevé de locus présentant un excès d'hétérozygotes. Ces résultats sont à l'opposé de ce qui est obtenu dans les populations de *L. digitata* caractérisées par un déplacement de la distribution du $F_{IS}$ vers des valeurs positives et une forte proportion de locus avec un déficit en hétérozygotes. De plus nous constatons différents types de distribution du $F_{IS}$ parmi les populations de *L. rodriguezii*, ce qui suggère en accord avec les autres indices de diversité et les prédictions des modèles théoriques, que la population de Bonifacio (sud de la Corse) présente un taux de reproduction clonale plus élevée que les populations de l'est de la Provence. Néanmoins, nous constatons que la population de Bonifacio a un haut niveau de déséquilibre de liaison, d'une amplitude comparable à celui observé pour une des populations de *L. digitata*, spécifiquement la population marginale d'Helgoland en mer du Nord. Cette dernière observation confirme qu'un haut niveau de déséquilibre de liaison peut être la conséquence de facteurs autres que la clonalité. La distribution du $F_{IS}$ et sa variance à l'échelle du génome nous parait également être un trait important et particulièrement révélateur du taux de clonalité des populations. Par ailleurs, en plus des lignées clonales que nous avons pu identifier à partir d'un seuil de distance génétique défini à l'aide des réplicas de séquençage, nous avons observé la présence d'une minorité d'individus déviants qui sont caractérisés par une forte proportion de génotypes hétérozygotes (33% des locus) et d'allèles privés (20.3% des locus). Nous avançons plusieurs hypothèses pour expliquer la présence de ces individus déviants, l'une relève de la conséquence de la divergence à long terme des allèles d'un même locus, et ainsi être le signal potentiel d'un effet Meselson. Enfin, de tels génotypes très hétérozygotes pourraient être la conséquence d'hybridation entre espèces ou lignées divergentes : ce point nécessiterait un échantillonnage plus étendu pour être étudié.

En ce qui concerne la structure génétique de *L. rodriguezii* nous avons mis en évidence une forte différenciation génétique à l'échelle de l'est de la Provence mais aussi entre cette localité et la Corse. À titre de comparaison, une forte différenciation génétique a également été observée entre les populations corses et continentales chez le corail rouge *Corallium rubrum* par le génotypage de SNPs par séquençage RAD. En revanche à cette échelle géographique les valeurs de $F_{ST}$ sont bien plus faibles pour C. *rubrum* ($F_{ST}$ = 0,24 ; Pratlong et

al., 2018) que pour *L. rodriguezii*. Cela confirme la différenciation attendue entre les populations insulaires et continentales pour ces espèces peu dispersives. A l'échelle de l'est de la Provence, nos résultats indiquent un très faible degré de connectivité, suggérant que le potentiel de dispersion des spores de *L. rodriguezii* est faible et comparable à celui des kelps de l'infralittoral (Santelices, 1990). Par exemple pour le kelp *Laminaria hyperborea*, la dispersion des spores est majoritaire entre 50 et 200 mètres (Fredriksen et al., 1995). Cette distance semblerait compatible avec le faible degré de différenciation constaté entre les individus du Banc du Magaud, ceux-ci étant dans un continuum et seulement séparés par une distance de 200 mètres. La répartition très fragmentée de *L. rodriguezii,* et la rareté des populations sont deux facteurs susceptibles de limiter une dispersion de proche en proche. Pour des kelps de l'infralittoral supérieur comme *Laminaria digitata*, la fragmentation de l'habitat structure les populations, comme le montre la comparaison du degré de différenciation génétique entre des peuplements continues et d'autres isolés (Billot et al., 2003 ; Robuchon et al., 2014 ; Couceiro et al., 2013).

L'identification d'une telle structure génétique locale et la présence de dèmes très peu connectés ont des implications importantes pour la conservation de *L. rodriguezii*. En effet, les gestionnaires doivent considérer ces populations rares et isolées comme des unités uniques à préserver. Le manque de flottabilité de cette espèce, suggère que la dérive des fragments arrachés (pouvant faire plusieurs mètres de long) est localisée à quelques mètres bien que l'hydrodynamisme soit constant et très puissant. Cette hypothèse n'a pas été évaluée compte tenu de la taille trop restreinte de l'échantillon, néanmoins aucune des lignées locales est partagée entre plusieurs sites. Dans l'hypothèse où la dérive des fragments est un vecteur de flux de gènes, la disponibilité d'un substrat favorable, et la profondeur (irradiance) sont des facteurs qui conditionnent l'installation et le développement des gamétophytes. Or, l'espèce se développe sur des remontées rocheuses isolées au milieu du détritique côtier et entourées de profondeur qui ne sont pas compatibles avec ses exigences physiologiques, ce qui rend son expansion impossible. On trouve parfois des sporophytes arrachés en épave à de grande profondeur, ils alimentent en matière organique les écosystèmes profonds. Nous pouvons raisonnablement supposer que *L. rodriguezii* joue un rôle important dans les flux de carbone entre les habitats. Par ailleurs, cette laminaire crée un habitat structuré abritant des espèces commerciales de poissons et de crustacés (langoustes) ; ces zones sont connues des pêcheurs et affectées par l'activité de pêche. En effet, toutes les populations que nous avons

échantillonnées sont perturbées par des arrachages du aux filets de pêches. Néanmoins, la capacité étonnante de cette espèce à se fragmenter et à se reproduire par clonalité pourrait expliquer la persistance de populations déjà observée dans les années 1960 (Banc du Magaud ; lors de l'exploration de la soucoupe plongeante du navire océanographique la Calypso). Cependant, une attention particulière est à mettre en place pour les sites du Banc du Magaud et du Cap Camarat car ceux-ci sont dans la zone d'adhésion du Parc National de Port-Cros. Il en est de même pour la population de Bonifacio dans la Réserve Naturelle des Bouches de Bonifacio. *A minima* une information aux pêcheurs professionnels est nécessaire ainsi qu'un suivi de l'état de la structure des populations et de la diversité génétique/génotypique.

# Chapitre 4

# Discussion, conclusion et perspectives



Installation d'un support servant à transplanter les individus de *Ericaria zosteroides* pour l'étude de l'adaptation locale à la profondeur (© Sandrine Ruitton).

## 4.1 L'adaptation locale chez *Ericaria zosteroides*

Bien que prometteurs, les résultats portants sur la différenciation génétique des populations d'*Ericaria zosteroides* à des locus outliers restent préliminaires (Materials B.2, chapitre 2). D'autres investigations, comme l'annotation de génome, des expériences de réponses physiologiques à un stress thermique et radiatif seront nécessaires pour considérer le potentiel adaptatif des populations.

### 4.1.1 Faux positifs ou signal d'adaptation locale ?

L'analyse basée sur les distances génétiques et les prédicteurs environnementaux (db-RDA, Figure B.2.1, chapitre 2) a mis en évidence que le signal de différenciation porté par les 12 SNPs outliers est corrélé au minimum de température et correspond à une nette distinction entre le site le plus profond analysé ici (La Ciotat, 45 mètres) et les autres sites moins profonds (baie de Marseille, 25-36 mètres). Ces résultats supportent l'hypothèse d'une réponse adaptative des individus à la profondeur, en l'occurrence la température ou à un autre facteur corrélé à cette variable, telle que la luminosité. Ces résultats sont préliminaires pour plusieurs raisons. D'une part dans un contexte de forte structure génétique, les méthodes de recherches de locus outliers sont susceptibles de donner lieu à des faux positifs (Bierne et al., 2013 ; Hoban et al., 2016 ; Pratlong et al., 2018) et d'autre part, l'étude d'un nombre limité de comparaisons entre environnements peut biaiser les résultats obtenus. Ainsi dans le cas de *E. zosteroides*, la forte différenciation observée entre la population la plus profonde et les autres pourrait être liée à un contact entre des lignées génétiques bien différenciées pour des raisons historiques. Une compréhension plus complète de l'évolution de l'adaptation locale nécessiterait d'étudier un plus grand nombre de populations en environnements contrastés (Lotterhos & Whitlock, 2015). La détection d'un même locus impliqué dans ces différentes comparaisons entre environnements serait un argument important sur son implication dans l'adaptation locale à l'image de ce qui a été observé chez l'épinoche *Gasterosteus aculeatus* avec le locus *Eda* (DeFaveri et al., 2011). Cependant des événements de recombinaison entre SNPs identifiés par séquençage RAD et les locus sous sélection font que le signal peut ne pas être conservé pour toutes les paires de comparaisons entre environnements. Plus généralement, et comme mentionné en introduction, la faible couverture du génome issue des approches RAD peut limiter la puissance de recherche de locus sous sélection. Une

meilleure couverture du génome, ainsi qu'un génome bien assemblé et annoté, seraient utiles pour tester la présence d'îlots génomiques de différenciation impliqués dans l'adaptation locale. Enfin, il faut là encore tenir compte de la possibilité d'un couplage entre des barrières exogènes (environnementales) et endogènes (isolement reproductif) dans l'interprétation des locus outliers (Bierne et al., 2011).

### 4.1.2 Réponses physiologiques de *Cystoseira sensu lato* à un stress radiatif

La profondeur des populations est un facteur intéressant à considérer pour l'étude de l'adaptation locale chez les Fucales du genre *Cystoseira sensu lato*, notamment car la réponse des individus à un stress radiatif est contrastée selon qu'ils soient présents en limite supérieure ou inférieure de distribution. Par exemple, Betancor et al. (2015) ont montré que la réponse physiologique de *Cystoseira humilis* Schousboe ex Kützing était différente quand les individus étaient étudiés en limite supérieure (-0,05 m) ou inférieure (-3,5 m) de distribution des cuvettes du littorale. En effet, les individus en limite supérieure sont caractérisés par une concentration élevée de composés photoprotecteurs, tels que les phénols et une forte activité antioxydante ($EC_{50}$). Le paramètre $EC_{50}$ est ici une mesure de la quantité d'antioxydant dans les composés phénoliques des algues brunes nécessaire pour réduire 50 % de DPPH (2, 2-diphenyl-1-picrylhydrazyil). En complément de l'étude de Betancor et al. (2015), des transplantations croisées entre ces deux profondeurs seraient ici utiles pour tester si les individus situés en limite inférieure de distribution possèdent des mécanismes de défense, similaires à ceux évoluant constamment en limite supérieure. Dans cette optique, l'expérience en jardin commun mis en place par Celis-Plá et al. (2014) sur *Ericaria selaginoides* (synonyme de *Cystoseira tamarascifolia*) a montré l'existence d'une réponse différentielle pour les individus, de la zone intertidale, prélevés, dans la limite haute (-0.5 m) et basse (-2 m) de distribution lors d'une exposition commune à -0.2 m en condition *in situ* d'irradiance contrôlée. Des lignes de flottaisons équipées de cylindres plastiques ont été employées afin de contrôler 100% ou 70% de l'irradiance de surface (PAB). Il ressort de cette étude que les individus de -2 m montrent une augmentation de composés phénoliques, de pigments accessoires et un accroissement de l'activité antioxydante ($EC_{50}$) en condition d'irradiance $100\%_{PAB}$ qui n'était pas détectée lors de leur exposition à $70\%_{PAB}$. Ces résultats indiqueraient une acclimatation rapide de *E. selaginoides* à de fortes conditions d'irradiance.

Bien que la distribution bathymétrique de *E. zosteroides* (de 15 à 50 mètres de profondeur) soit très différente de celle des autres taxons de l'étage de l'infralittoral supérieur, la réponse physiologique des individus à un stress radiatif / thermique en limite supérieure de distribution est à explorer compte tenu du contexte climatique actuel du bassin méditerranéen. Ainsi, l'augmentation expérimentale de la température de 20 à 24°C affecte négativement l'installation et la germination des zygotes de *E. zosteroides* par comparaison à des conditions de contrôle à 16°C (Capdevila et al., 2019). Cette expérimentation a été conduite à partir d'individus prélevés entre 20 et 25 mètres de profondeur dans le nord-ouest de la Méditerranée, ce qui ne permet donc pas d'évaluer si la réponse des zygotes à la température est conservée voire même amplifiée dans les populations les plus profondes, il aurait fallu *ad minima* tester 13°C qui est la température constante de la Méditerranée sous la thermocline.

L'intégration de ces résultats dans un modèle de projection intégrale indique néanmoins que la résilience des populations à une perturbation pourrait grandement diminuer dans un scénario de réchauffement climatique, compatible avec cette gamme de températures (Capdevila et al., 2019). Par ailleurs, ces projections doivent tenir compte de la capacité d'une population à s'adapter de façon dynamique à de nouvelles conditions environnementales, soit par adaptation génétique, soit par acclimatation individuelle.

### 4.1.3 Evaluer la réponse physiologique des populations d'*E. zosteroides*

Compte tenu des possibilités d'adaptation locale de cette espèce, je propose en perspective des résultats du chapitre 2, un protocole expérimental de transplantations croisées afin d'étudier la diversité adaptative des populations. Spécifiquement, cette expérimentation devrait permettre d'évaluer la réponse physiologique des populations à un stress radiatif et thermique et ainsi confronter les résultats attendus à ceux du signal de différenciation porté par les 12 SNPs outliers.

Pour cela, le site du tombant de l'Armoire, présentement étudié (chapitre 2), a été choisi. Ce site est un laboratoire *in situ* pour l'étude de l'adaptation locale en raison du fait que *E. zosteroides* y est distribuée de manière continue entre 22 et 40 mètres de profondeur, avec une très forte pente. En effet, la profondeur génère des conditions environnementales hétérogènes, notamment en termes de température (Bensoussan et al., 2010), d'irradiance et de disponibilité en nutriments auxquelles sont exposées les populations. Les profils

thermiques obtenus pour l'année 2019 (année des transplantations) dans un site proche du site d'étude nous indiquent que les individus à 20 et 40 mètres sont exposés à des conditions thermiques très contrastées (Figure 4.1A). À 20 mètres de profondeur la température peut atteindre 21°C de manière régulière en période estivale alors qu'à 40 mètres cette dernière dépassait occasionnellement 16°C à la même période. De plus, on s'attend à ce que l'irradiance soit plus forte en limite supérieure qu'en limite inférieure du tombant. Il en résulte que la population en limite supérieure du tombant (20-22 m) est exposée en période estivale aux eaux plus chaudes de la couche de mélange alors que celle en limite inférieure (38-40 m) bénéficient au cours d'une année de conditions thermiques qui sont davantage stables. Le deuxième intérêt à la sélection de ce site réside dans le fait que l'analyse de la structure génétique à des locus neutres permet de mettre en évidence que la différenciation génétique des individus de la partie haute et basse tombant est suffisamment élevée pour les considérer comme appartenant à deux populations séparées. Enfin, les individus génotypés par RAD-seq (ARM20 et ARM40, chapitre 2) ont été marqués individuellement ; ce qui nous permet donc de transplanter des individus pour lesquels nous avons accès à l'information génomique.

L'objectif de cette expérimentation est d'évaluer si les individus de la limite inférieure du tombant mobilisent des mécanismes de photoprotection, et d'acclimatation une fois leur exposition à des conditions stressantes en limite supérieure, ou inversement si ces derniers présentent des signes physiologiques de mal-adaptation, voire de mortalité. L'hypothèse d'adaptation locale serait ici testée par une comparaison de la réponse des résidents par rapport aux non-résidents (Kawecki & Ebert, 2004). Il est à noter que la vérification de ce critère au sein d'une même génération peut correspondre soit à l'expression différentielle d'un même génotype (plasticité phénotypique) soit à des différences génétiques évoluant sous l'action de sélection divergente (adaptation locale), voire plus probablement à l'interaction entre ces deux mécanismes.

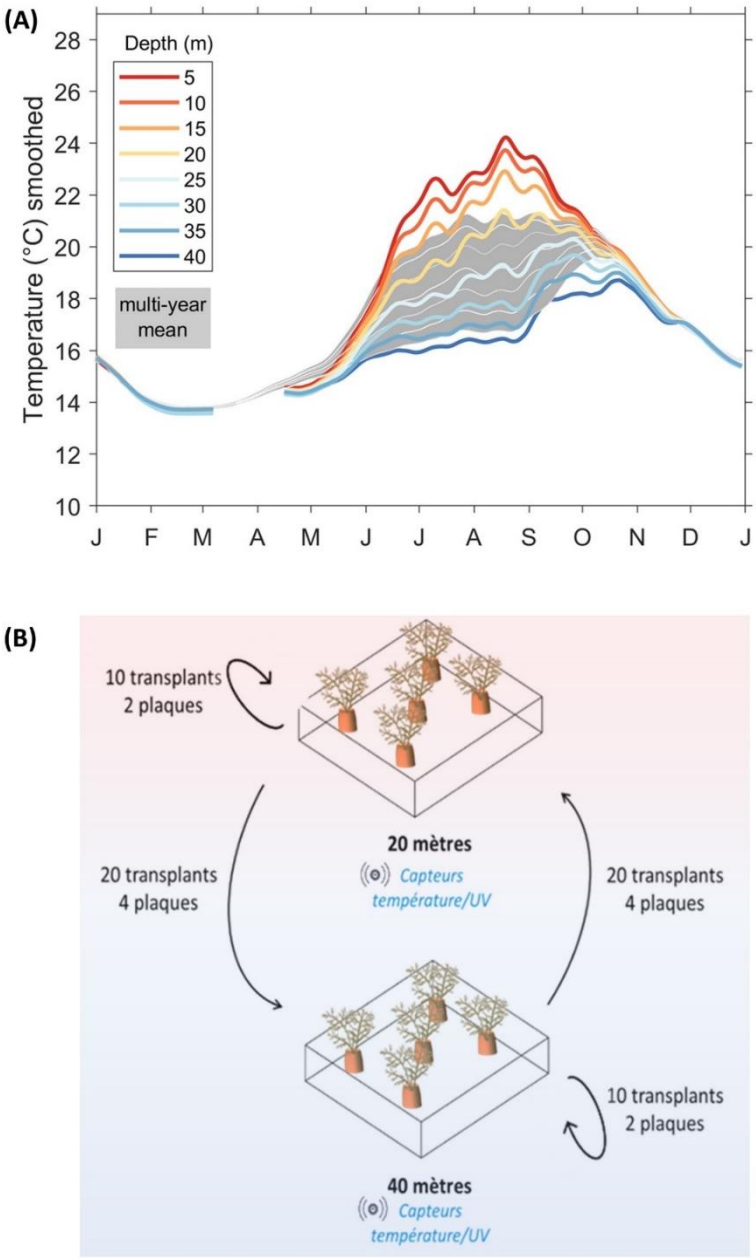



**Figure 4.1**. (A) profil thermique de l'année 2019 au Cap Sicié (baie de Toulon), la zone la plus proche du site de l'Armoire pour laquelle ce type de données était disponible (T-MEDNet ; www.t-mednet.org). Chaque courbe lissée matérialise les températures de 5 à 40 mètres de profondeur. La température est exprimée en degré Celsius (°C) sur l'axe des ordonnées et les mois de l'année sont indiqués sur l'axe des abscisses. La moyenne multi-annuelle a été calculée sur les années 2014-2020. (B) protocole de transplantations croisées des populations d'*Ericaria zosteroides* du tombant de l'Armoire (baie de Toulon).

***Protocole de transplantations croisées***

Vingt individus de la partie supérieure du tombant (20 mètres) ont été transplantés sur la partie basse de ce dernier (40 mètres), et inversement pour ceux de 40 mètres (Figure 4.1B). Les individus ont transplanté sur des supports en béton et fixés avec de la colle époxy

non toxique, déjà utilisée pour des transplantations de *Ericaria amentacea*. Afin de tenir compte d'effets propres à l'expérimentation, 10 individus ont été auto-transplantés à leurs profondeurs respectives, ceux-ci servant de témoins négatifs pour apprécier la réponse physiologique des transplants. Chaque support de transplantation, d'une longueur et d'une largeur de 35 cm pour 10 cm de hauteur permet d'accueillir cinq individus. Au total, ce sont 10 supports qui ont été conçus en béton armé. Ce matériau a été choisi pour sa résistance aux conditions hydrodynamiques, notamment des périodes de houle, ceci pour permettre une expérimentation sur une durée de plusieurs mois. Des capteurs de température à haute fréquence (Hobo Tidbit v2) ont été intégrés aux plaques des transplants ; la température est enregistrée toutes les heures à chacune des deux profondeurs.

Une fois le prélèvement des individus effectué en plongée sous-marine, ceux-ci ont été remontés à bord des embarcations, fixés immédiatement sur les supports avant d'être de nouveaux immergés à leurs profondeurs respectives. Les supports ont été déposés à même le fond dans les parties hautes et basses du tombant (Figure 4.1B). J'ai considéré plusieurs paramètres physiologiques pour évaluer la réponse des transplants à la transplantation dans un nouvel environnement du point de vue de la température et de la luminosité :

- Le taux de survie des transplants par analogie à celui des témoins négatifs.
- Le dosage des composés phénoliques (PC), à partir du protocole employé par Celis-Plá et al. (2016) qui repose sur un dosage des phénols par la méthode de Folin, puis d'une mesure de l'absorbance à 760 nm.
- La mesure de l'activité antioxydante ($EC_{50}$) à partir des composées phénoliques. Le protocole employé est identique à celui de Celis-Plá et al. (2016).

L'expérimentation devait durer de mai 2019 à mai 2020, notamment pour inclure les conditions stressantes de la saison estivale. Par ailleurs, Il était prévu de mesurer ces paramètres physiologiques lors d'un suivi trimestriel. Or, dès le premier suivi en juillet 2019 nous avons constaté que les trois quarts des supports de la zone des 20 mètres avaient été retournés, malgré le poids des supports en béton. La cause la plus probable est la malveillance des plongeurs. En effet, si le tombant de l'Armoire est un laboratoire *in situ* nous paraissant prometteur pour l'étude de l'adaptation locale, il est aussi un site de plongée très fréquenté en raison de la diversité d'espèces associées à cette forêt sous-marine, dont des espèces

emblématiques telles que le mérou brun (*Epinephelus marginatus*). Par ailleurs, les suivis trimestriels planifiés ont été stoppés par la situation sanitaire de l'année 2020. Malgré tout, ce protocole expérimental *in situ* ouvre la perspective à d'autres travaux d'étude du potentiel adaptatif des populations des forêts de Fucales de l'étage circalittoral.

## 4.2 Inférence des processus démographiques

Dans la perspective de déterminer les tailles efficaces et les flux de gènes passés et actuels de *Ericaria zosteroides* et *Laminaria rodriguezii*, les données génomiques employées dans cette thèse pourront encore être valorisées. Ces inférences démographiques sont essentielles pour aborder les interrogations soulevées par les chapitres 2 et 3 de ce manuscrit : à quand remonte la divergence évolutive des populations ? Ont-elles subi une réduction ou une augmentation de taille efficace ? Le niveau de flux de gènes a-t-il toujours été aussi limité qu'il l'est actuellement ? Cette ambition de reconstruire l'histoire évolutive de ces deux espèces est par ailleurs une des manières de préciser la biogéographie historique des kelps et Fucales du bassin méditerranéen.

### 4.2.1 Estimation de tailles efficaces à la génération actuelle

L'analyse de la structure génétique des populations de *E. zosteroides* (chapitre 2) et *L. rodriguezii* (chapitre 3) apporte une image des effets indissociables des flux de gènes et de la dérive génétique. En perspective de ces analyses, les marqueurs du séquençage RAD pourraient nous permettre d'estimer la taille efficace actuelle des populations (notée *Ne*). On définit la taille efficace d'une population comme étant la taille d'une population « idéale » de Wright-Fisher où la dérive génétique aurait la même intensité que dans la population étudiée. La taille efficace d'une population est un indicateur pertinent pour la conservation des espèces car elle peut renseigner sur l'efficacité de la sélection, le degré de flux gène ou encore le rythme de perte de diversité génétique (Charlesworth, 2009 ; Frankham et al., 2014). De plus, *Ne* est un paramètre facilement compréhensif par des gestionnaires et l'opinion publique, ce qui faciliterait, en théorie, une prise de décision instantanée lorsqu'un déclin inhabituel des effectifs efficaces est suspecté. A ce titre, Laikre et al. (2020) encouragèrent la démocratisation de *Ne* dans les stratégies de conservations afin d'accorder davantage d'attention au potentiel évolutif des populations. Cela permettrait d'adopter des mesures de

sauvetages génétiques, telle que les flux de gènes assistés lorsque la valeur minimale $Ne \leqslant 500$ est atteinte. Bien que ce seuil soit discutable, il permettrait de limiter les effets délétères de la consanguinité et de la dérive génique et ainsi maintenir la population considéré dans un état génétiquement viable (Jamieson & Allendorf, 2012).

***Suivi du nombre effectif de géniteurs par cohorte de Ericaria zosteroides***

Chez les Fucales longévives *Ne* est difficilement estimable en raison du fait que plusieurs cohortes persistent des dizaines d'années dans une population. Ainsi pour ces populations disposant d'une structure en âges, il peut être plus intéressant de considérer le nombre effectif de géniteurs par cycle de reproduction (au sens nombre d'événement reproductifs) (noté *Nb*) plutôt que considérer la taille efficace d'une population (*Ne*). Par exemple, l'analyse de la distribution de la taille des individus de *E. zosteroides* (Annexe B, Figure B2) montre l'existence de plusieurs cohortes dans un même échantillon, celles-ci sont notamment apparentes dans le site de la Roche percée (RP, deux cohortes) et du Cap Caveau (CAV, trois cohortes) avec des individus âgés de plus de 10 ans. Pour ces populations, le suivi de *Nb* sur trois ou quatre cycles de reproduction permettrait la détection précoce d'une réduction des effectifs au cours des générations (Luikart et al., 2021). Par exemple, les simulations de Antao et al. (2011) ont montré qu'une réduction de la taille efficace de 600 à 50 individus entre la première et la deuxième génération serait détectable après une génération par la méthode du déséquilibre de liaison (Hill, 1981), implémentée dans le modèle *LDNE* (Waples & Do, 2008 ; voir ci-dessous). Cela a été constaté dans 80% des cas et avec un échantillon de 25 individus et 20 microsatellites (10 allèles par locus). De même, ce modèle conserverait suffisamment de puissance pour détecter un déclin moins abrupt de *Ne* (réduction de $Ne \simeq 33\%$ entre deux générations).

La mise en place d'un suivi par cohorte nécessite cependant d'identifier, avec certitude, l'appartenance des individus à chaque génération. Malheureusement et en raison des difficultés d'échantillonner à ces profondeurs, cela n'a pu être fait dans le cadre de cette thèse pour *E. zosteroides*. En perspective, un échantillonnage par classe de taille (mesure de l'axe primaire) peut être envisagé pour la mise en place de ce suivi. Par exemple, en considérant séparément, le stade de recrues (< 1 cm), des individus de la première génération (2-5 cm), de ceux plus âgés (8-11 cm). Cela dit, il serait nécessaire de s'assurer que ces classes de tailles appartiennent à des générations différentes car la croissance variable des individus

d'une même génération est envisageable en présence de compétition intra et inter-spécifique pour les ressources. Pour cela il suffirait de marquer *in situ* les individus et de les suivre plusieurs années comme cela a été fait à Scandola (Ballesteros et al., 2009) mais malheureusement sans faire de prélèvements génétiques associés.

En ce qui concerne les marqueurs génétiques à employer, les séquences RAD disponibles et le séquençage *de novo* du génome d'un des individus, nous offrent l'opportunité d'isoler rapidement et à un moindre coût un panel de microsatellites chez cette espèce. Pour cela des outils comme MiddRAD-seq data (Qin et al., 2017) sont à privilégier. Une fois le génotypage d'un nombre d'individus conséquent par cohorte établi, l'algorithme implémenté dans *agestrucnb* (Luikart et al., 2021) permettrait de tester spécifiquement si le nombre de géniteurs a diminué aux cours des générations. Par exemple, certaines populations de la réserve de Scandola sont constituées d'individus atteignant 25 cm, ceux-ci seraient âgés de plus 50 ans alors que d'autres de moins d'un centimètre sont les témoins d'un recrutement récent (Ballesteros et al., 2009). Dans ces conditions le suivi de *Nb* des recrues et des individus les plus âgés pourrait permettre d'analyser si le nombre de géniteurs a diminué par comparaison à une situation antérieure d'approximativement 50 ans.

***Estimations de tailles efficaces par le déséquilibre de liaison (LDNE)***

Actuellement, plusieurs méthodes existent pour estimer *Ne* et *Nb* à partir des marqueurs génétiques sans nécessairement avoir besoin d'effectuer un échantillonnage à deux temps différents (Luikart et al., 2010). Certaines méthodes ont été spécifiquement développées sous un modèle d'évolution compatible avec des microsatellites (*onesamp* ; Tallmon et al., 2008) alors que d'autres, moins spécifiques, utilisent le déséquilibre de liaison (*LDNE* ; Waples & Do, 2008) ou l'excès d'hétérozygotes (*Nb_HetEx* ; Zhdanova & Pudovkin, 2008). Par exemple, le modèle d'inférence démographique implémentée dans *onesamp* a permis d'estimer la taille efficace de populations fragmentées de la gorgone *Paramuricea clavata* aux îles Baléares (Espagne) (Arizmendi-Mejía et al., 2015). Chez cette espèce longévive, *Ne* serait compris entre 50-135 individus selon les sites, avec des valeurs minimales pour les populations isolées, fortement différenciées et génétiquement appauvries. L'existence de populations isolées et génétiquement appauvries est une situation analogue à celle de *E. zosteroides*, notamment de la population hautement différenciée de la Roche percée (RP) en baie de Marseille (chapitre 2). Parmi les différentes méthodes existantes

d'estimations de tailles efficaces, la méthode *LDNE* de Waples & Do (2008) est l'une des plus couramment employée quand une seule génération est échantillonnée. Le principe associé étant que la dérive génétique génère des associations non aléatoires entre les allèles à différent locus (déséquilibre de liaison) à mesure que la taille efficace diminue (Hill, 1981). Ainsi sous certaines hypothèses simplificatrices de neutralité sélective, de population fermée et de générations discrètes, le déséquilibre de liaison renseigne sur l'effectif efficace d'une population (Waples & Do, 2008).

***Les hypothèses du modèle LDNE sont-elles compatibles à la biologie des kelps et Fucales ?***

Il se pose la question de la fiabilité du modèle *LDNE* pour l'estimation de *Ne* ou *Nb* lorsque certaines de ces hypothèses ne sont pas respectées. En effet, compte tenu des traits d'histoire de vie de *E. zosteroides* et de *L. rodriguezii*, différents facteurs sont susceptibles de générer du déséquilibre de liaison et ainsi, biaiser les estimations du modèle : par exemple, la reproduction entre individus apparentés, voir même l'autofécondation. Les $F_{IS}$ positifs observés chez les espèces monoïques du genre *Cystoseira sensu lato* (Robvieux, 2013 ; Thibaut et al., 2016b ; Buonomo et al., 2017 ; Medrano et al., 2020 ; et chapitre 2 pour *E. zosteroides*) pourraient être la conséquence de ces modes de reproduction fermées. Néanmoins, la structure en âges est aussi attendue pour générer des $F_{IS}$ positifs, dont du déséquilibre de liaisons, tous deux par un effet Wahlund (Johnson & Black, 1984 ; Teixeira et al., 2016 ; Almeida et al., 2017). L'effet Wahlund est la conséquence de considérer comme une même population des échantillons avec des fréquences alléliques différentes. Par ailleurs, en ce qui concerne spécifiquement *L. rodriguezii* (chapitre 3) ou d'autres espèces partiellement clonales, l'absence de ségrégation des allèles par reproduction clonale accentue le déséquilibre de liaison (Krueger-Hadfield et al., 2021). Ainsi, des simulations de populations partiellement clonales, et d'autres hautement autofécondes avec et sans structure en âges seraient nécessaires pour évaluer l'incidence du cycle de vie de ces espèces sur l'estimation de *Ne* et *Nb* par le modèle *LDNE*.

À cette fin, les récents travaux de Luikart et al. (2021) apportent des éléments de réponse. Les simulations qui ont été faites de *Nb* ne se limitent plus à une population « idéale » mais considèrent des populations avec des taux de longévité, de survie ou encore de fécondité variables. Certaines de ces simulations considèrent même des populations dont la valeur donnée à ces paramètres est compatible avec les traits d'histoire de vie d'algues

brunes, comme la Fucales *Ascophyllum nodosum* (Waples et al., 2013). De plus, ces simulations ont permis d'évaluer le biais induit par la présence de plusieurs cohortes dans un même échantillon. Les résultats de Luikart et al. (2021) montrent que l'estimation du nombre de géniteur par génération (*Nb*) est en moyenne biaisée de 3 à 15% par la structure en âges des individus et cela pour différents traits d'histoire de vie et une taille réelle de *Nb* de 50-100 géniteurs par génération. Dans une configuration où *Nb/Ne*⩾ 1, pour *A. nodosum* (*Nb/Ne* = 1,26 ; Waples et al., 2013), le regroupement de deux à trois cohortes dans un même échantillon surestimerait *Nb* de l'ordre de 30% à 40%, tel que cela était constaté par Waples et al. (2014).

***Estimation de tailles efficaces par le modèle LDNE avec les marqueurs RAD-seq***

Augmenter le nombre de locus par une approche de type RAD-seq peut être bénéfique pour accroitre la précision de l'estimation du nombre de géniteurs par génération (*Nb*) (Luikart et al., 2021) et plus généralement, la taille efficace des populations (*Ne*). Néanmoins, en dessous d'un certain seuil, 180 SNPs apporteraient le même niveau de précisions que 20 microsatellites avec 10 allèles chacun (Waples & Do, 2010). Par ailleurs, augmenter le nombre de locus, augmente à la fois le risque de cibler des régions sous sélection et d'autres en déséquilibre de liaison par proximité physique sur le génome. Cela nécessite donc d'effectuer un traitement en amont du génotypage pour considérer ces effets, et d'évaluer les conséquences du traitement sur le signal biologique. Les premières analyses de simulations de taille efficaces avec des données RAD-seq (Nunziata & Weisrock, 2018) indiqueraient que les biais du RAD-seq sont semblables à ceux constatés avec l'utilisation de microsatellites (Waples & Do, 2010) ou d'un nombre limité de SNPs (Luikart et al., 2021). Ces biais sont notamment associés à la taille efficace réelle de la population avec une difficulté d'estimer *Ne* ou *Nb* quand la taille efficace est élevée, à l'inclusion des allèles rares dans les analyses et à la taille des échantillons (Tallmon et al., 2010 ; Waples & Do, 2010).

Malgré les imprécisions du modèle auquel on doit s'attendre compte tenu des traits d'histoires de vie des espèces étudiées dans cette thèse et des différents points soulevés précédemment, j'ai souhaité valoriser une seconde fois les données du RAD-seq pour fournir une première estimation du nombre de géniteurs par génération (*Nb*) des populations de *Ericaria zosteroides*. Bien que les résultats attendus soient préliminaires, ces estimations pourraient être bénéfiques pour la conservation des populations et à la mise en place de

mesures de gestions appliquées pour cette espèce dans l'aire du Parc National des Calanques (Marseille, ouest de la Provence) et plus généralement en Méditerranée. Par ailleurs, cette analyse pourrait permettre d'évaluer si la divergence évolutive du site particulier de la Roche percée (RP) est associée à une hausse de la dérive génétique et si la population établie depuis quelques générations sur la conduite Altéo (ALT) dispose d'une taille efficace réduite par effet fondateur. En raison de la présence de plusieurs cohortes, indissociables dans les échantillons, l'estimation de *Ne* est à interpréter comme le nombre de géniteurs par génération (*Nb*) dans les analyses qui vont suivre.

***Estimation du nombre de géniteurs par génération des populations d'Ericaria zosteroides***

L'estimation du nombre de géniteurs par génération (*Nb*) a été effectuée en employant la méthode *LDNE* de Waples & Do (2008). Les résultats ont été comparés avec ceux attendus par la méthode *Nb_HetEx* de Zhdanova & Pudovkin (2008) qui considère que l'excès en hétérozygotes parmi les descendants est dû à l'allofécondation. Ces deux méthodes sont implémentées dans NeEstimator v2.0 (Do et al., 2014). La méthode *Nb_HetEx* serait notamment adaptée pour les populations du tombant de l'Armoire en baie de Toulon car l'allofécondation a été avancée comme une hypothèse plausible à l'explication des valeurs négatives de $F_{IS}$ significatives dans ce site. Pour rappel, cette population (ARM20) est très dense (moyenne de 11 ± 0.62 ind.m$^{-2}$) en comparaison à celle de la Roche percée (RP) en baie de Marseille (moyenne de 1.23 ± 0.25 ind.m$^{-2}$). Nous pouvons vraisemblablement admettre que le nombre de géniteurs est plus important dans une population très dense que dans une population peu dense. Par conséquent, la probabilité que la syngamie ait lieu entre des gamètes appartement à des individus distincts (allofécondation) serait plus élevée à ARM20 qu'a RP. Les estimations de *Ne* ont été faites en considérant un sous échantillon aléatoire de 16 individus par site sans distinction possible des cohortes. Ceux-ci ont été génotypés à 10 755 SNPs par RAD-seq, après exclusion des outliers (chapitre 2). L'effet des allèles rares a été évalué en répliquant les estimations de *Ne* à différent seuils de fréquences alléliques (*PCrit*, 0,01 ; 0,05). Le paramètre *PCrit* correspond à la fréquence minimale d'un allèle pour qu'il soit inclus dans les analyses.

***Convergence des méthodes et effets des allèles rares***

Les résultats de la méthode *Nb_HetEx* sont à considérer avec précaution car ceux-ci sont en principe très imprécis par comparaison à d'autres modèles (Nomura, 2008 ; Wang

,2009). Néanmoins, les estimations avec *Nb_HetEx* sont d'un ordre de grandeur comparable à celles de l'analyse du déséquilibre de liaison, notamment pour une des populations du tombant de l'Armoire (ARM40) et de la réserve de Scandola (SC) en Corse (Table 4.1). En ce qui concerne, la méthode *LDNE*, les estimations de taille efficace avec les allèles rares (*Pcrit* = 0,01) peuvent être jusqu'à 11 fois supérieures à une situation où ils sont exclus (*Pcrit* = 0,05). Cet effet est particulièrement visible pour la population du Cap Caveau (CAV) où l'estimation de *Ne* diminue de 5240 à *Pcrit* = 0,01 à 458 individus à *Pcrit* = 0,05. La surestimation de la taille efficace par les allèles rares est conforme aux prédictions de Waples & Do (2010) pour des marqueurs très polymorphes (un maximum de 10 allèles par locus) et à celles de Nunziata & Weisrock (2018) pour des marqueurs RAD-seq (3000-4000 SNPs). Par ailleurs, la taille limite des échantillons (16 individus par site) pourrait être à l'origine des valeurs infinies des intervalles de confiance à 95%, notamment dans une configuration où *Pcrit* = 0,01.

***Ericaria zosteroides, un nombre limité de géniteurs ?***

En considérant le modèle *LDNE* et en excluant les allèles rares (Pcrit = 0,05), le nombre de géniteurs serait inférieur à 100 individus par site à l'exception près des populations du Cap Caveau (CAV) et de Tiboulen de Maïre (TIB) pour lesquelles *Ne* serait de 400-500 individus (Table 4.1). Le contraste entre CAV et TIB et le reste des sites est intriguant compte tenu du fait que TIB est une population de très grande taille, l'une des plus grande du Parc National des Calanques (Bonhomme et al., 2014) alors que le nombre de géniteurs très restreint à la conduite Altéo (*Ne* = 39) serait compatible avec l'hypothèse d'un effet fondateur lors de la colonisation du site par un nombre vraisemblablement limité de géniteurs. Par ailleurs, la lignée divergente de la Roche percée (RP) ne semble pas avoir un nombre de géniteurs par génération plus limité que d'autres sites. Ce résultat est néanmoins à interpréter avec prudence, car pour une espèce longévive comme *E. zosteroides*, le mélange de plusieurs cohortes dans un échantillon, comme constaté à RP, est un facteur attendu pour surestimer l'estimation de *Ne* (Waples et al., 2014).

En perspective, il serait nécessaire de reproduire ces analyses en sous-échantillonnant le jeu de donnés à plusieurs reprises, de sorte à évaluer la robustesse de ces estimations pour plusieurs tailles d'échantillons (n = 5 ; 10 ; 15 ; 20) et différents sets de SNPs, rangés par ordre décroissant de l'index $F_{ST}$ (500 ; 1000 ; 2000 ; 3000 etc.), ceci jusqu'à l'obtention d'un plateau où l'estimation de *Nb* se stabilise. Par ailleurs, ces estimations sont faites en supposant que

les populations sont isolées ; or, nous avons constaté que les flux de gènes sont probablement différents selon les populations considérées. Ainsi, il serait intéressant d'estimer simultanément la taille efficace et les flux de gènes, avec des méthodes d'inférences évolutives.

**Table 4.1**. Estimations de tailles efficaces (*Ne*) des populations d'*Ericaria zosteroides* (10 755 SNPs) en employant la méthode *LDNE* et *Nb_HetEx* à deux seuils de fréquences alléliques minimales (*Pcrit = 0,05 et 0,01*). Les intervalles de confiance à 95% de *Ne* paramétrique et Jackknife sont indiqués.

| **Population** | | ***LDNE*** *(Pcrit)* | | ***Nb_HetEx*** *(Pcrit)* | |
|---|---|---|---|---|---|
| | | *(0,05)* | *(0,01)* | *(0,05)* | *(0,01)* |
| Cap Caveau (CAV) | *Ne* | 458 | 5240 | ∞-∞ | ∞-∞ |
| | *IC paramétrique* | ∞-∞ | 53-53 | ∞-∞ | ∞-∞ |
| | *IC JackKnife* | ∞-∞ | 47-60 | - | - |
| Tiboulen (TIB) | *Ne* | 538 | 2478 | 189 | 706 |
| | *IC paramétrique* | 85-88 | ∞-∞ | 57-∞ | 105-∞ |
| | *IC JackKnife* | 51-233 | ∞-∞ | - | - |
| Le Veyron (VEY) | *Ne* | 71 | 115 | ∞-∞ | ∞-∞ |
| | *IC paramétrique* | 45- 45 | ∞-∞ | ∞-∞ | ∞-∞ |
| | *IC JackKnife* | 30-80 | ∞-∞ | - | - |
| La Roche percée (RP) | *Ne* | 60 | 95 | ∞-∞ | ∞-∞ |
| | *IC paramétrique* | 73-77 | 165-178 | ∞-∞ | ∞-∞ |
| | *IC JackKnife* | 44-210 | 75-∞ | - | - |
| Conduite Alteo (ALT) | *Ne* | 39 | 71 | ∞-∞ | ∞-∞ |
| | *IC paramétrique* | 34-34 | ∞-∞ | ∞-∞ | ∞-∞ |
| | *IC JackKnife* | 21-71 | ∞-∞ | - | - |
| La Ciotat (LC) | *Ne* | 43 | 90 | ∞-∞ | ∞-∞ |
| | *IC paramétrique* | 33-33 | 29-29 | 111-∞ | 467- ∞ |
| | *IC JackKnife* | 17-113 | 21-42 | - | - |
| Les Oursinières (OUR) | *Ne* | 122 | 231 | ∞-∞ | ∞-∞ |
| | *IC paramétrique* | 481-571 | ∞-∞ | ∞-∞ | ∞-∞ |
| | *IC JackKnife* | 90-∞ | 190-∞ | - | - |
| L'Armoire_20 (ARM20) | *Ne* | 83 | 108 | 28 | 43 |
| | *IC paramétrique* | 111-115 | ∞-∞ | 21-40 | 32-64 |
| | *IC JackKnife* | 58-729 | ∞-∞ | - | - |
| L'Armoire_40 (ARM40) | *Ne* | 44 | 66 | 38 | 52 |
| | *IC paramétrique* | 104-112 | ∞-∞ | 24-85 | 34-118 |
| | *IC JackKnife* | ∞-∞ | ∞-∞ | - | - |
| Scandola (SC) | *Ne* | 42 | 83 | 60 | 80 |
| | *IC paramétrique* | 26-27 | ∞-∞ | 35-219 | 47-296 |
| | *IC JackKnife* | 15-66 | ∞-∞ | - | - |

### 4.2.2 Vers la modélisation de l'histoire évolutive des populations

Des analyses plus poussées, basées sur des modèles d'inférences démographiques pourraient permettre d'apporter des précisions sur la cause d'un déclin d'une population, en estimant notamment de manière conjointe les variations de *Ne* et des flux de gènes. C'est le potentiel offert par les outils relevant de la modélisation démographique, telles que les approches ABC pour « *approximate Bayesian computation* » (Boitard et al., 2016), MSMC pour « *sequential Markovian coalescent methods* » (Terhorst et al., 2017). Leur principe réside dans la modélisation de plusieurs scénarios démographiques en considérant une large gamme de paramètres (taux de migration, tailles efficaces, processus de sélection etc.) pour retenir ceux qui maximisent la vraisemblance des données observées en ce qui concerne spécifiquement les méthodes de maximum de vraisemblance. Parmi les nombreuses études émergentes dans ce domaine, l'une d'elle d'intérêt pour la conservation, modélise l'existence d'un goulot d'étranglement qui se serait produit il y a quelques générations chez l'otarie de Kerguelen *Arctocephalus gazella*. Chez cette espèce, intensément chassée au cours des 18 et 19$^{ème}$ siècles en Antarctique, l'approche ABC a permis de reconstruire l'évolution de la taille efficace sur plusieurs générations et ainsi rechercher la signature génétique d'une réduction passée des effectifs (Hoffman et al., 2011). Les paramètres les plus parcimonieux indiquèrent l'existence un goulot d'étranglement survenu il y a 11 générations, marqué par une forte réduction de la taille efficace ($Ne \simeq 39$) puis une récupération contemporaine à $Ne \simeq 178\,000$. L'originalité de cette étude a été, entre autres, de tester plusieurs modèles démographiques, incluant un modèle nul de taille constante et de comparer l'estimation des paramètres avec des locus nucléaires (microsatellites) et mitochondriaux (SNPs), cela pour différents modèles de mutations.

***Laminaria rodriguezii, des populations reliques ?***

L'une des questions centrales pour l'étude et la conservation de *Laminaria rodriguezii* a été de se demander si les populations éparses que l'on observe actuellement sont les reliques d'une répartition ancestrale plus vaste. Plusieurs approches permettraient de considérer cette question, la première consisterait à modéliser *Ne* sur plusieurs générations pour détecter les signatures d'une réduction contemporaine des effectifs. L'autre approche, complémentaire de celle-ci, consisterait à évaluer si des flux de gènes historiques ont eu lieu parmi les populations étudiées et à quelle période ils remontent. En l'état actuel, la clonalité

partielle et les conséquences qu'elle engendre dans le génome de cette espèce, est un défi à de telles investigations. Néanmoins, l'information génomique dont nous disposions a été l'opportunité de reconstituer le spectre des fréquences alléliques (SFS) des populations pour, dans un travail futur, évaluer la robustesse des modèles d'inférence démographique en présence de clonalité partielle.

***Focus sur le spectre des fréquences alléliques (SFS)***

Le spectre des fréquences alléliques (SFS) est un résumé à l'échelle du génome relatant le nombre d'occurrences de l'allèle ancestral à une fréquence donnée pour chaque échantillon. Lorsque l'allèle ancestral n'est pas identifié, une des solutions consiste à considérer l'allèle minoritaire (MAF) comme l'allèle dérivé (on parle de SFS replié). Dans un scénario démographique incluant plusieurs populations, le SFS est élargi à plusieurs dimensions (SFS-2D) au lieu d'être considéré en une seule dimension (SFS 1D).

Devant les données décrites dans le chapitre 3, une question émerge : à l'échelle de l'est de la Provence, pourrait-on observer les traces de flux de gènes passés entre les populations du Banc du Magaud et celle du Cap Camarat ? Pour rappel, ces deux sites sont proches de quelques dizaines de kilomètres. Afin de considérer cette question, la première étape consisterait à employer les données RAD-seq (4077 SNPs) pour générer le SFS-2D de ces populations (Figure 4.2). Pour cet exemple, le SFS-2D a été construit à partir de l'allèle minoritaire en sous échantillonnant neuf individus diploïdes pour chaque site, soit un maximum de 18 occurrences pour un allèle dans chacun des sites. La lecture de ce SFS-2D est la suivante, chaque ligne indique la fréquence à laquelle un certain allèle se produit dans la population du Cap Camarat (*p2*) tandis que les colonnes indiquent la fréquence d'un certain allèle dans celle du Banc du Magaud (*p1*). Le SFS est ici représenté avec une représentation en log. Pour la première cellule de cette matrice ($p1_{f(0)}$ ;$p2_{f(0)}$) l'interprétation à donner est la suivante : 932 allèles (log(x)+1 $\simeq$ 2,97) ne sont ni présents dans la population du Banc du Magaud, ni dans celle du Cap Camarat. En effet, ce sont donc des allèles privés à la population de Bonifacio (sud de la Corse) qui n'est pas considérée dans ce cas. Pour la cellule ($p1_{f(2)}$ ;$p2_{f(7)}$), celle-ci nous indique que seul six allèles (log(x) +1 $\simeq$ 0,84) surviennent à deux reprises dans la population *p1* du Banc du Magaud et à sept reprises dans la population *p2* du Cap Camarat. Chez cette espèce disposant d'une structure génétique à courte distance, la très grande majorité des allèles partagés entre *p1* et *p2* survient à très basse fréquence (Figure 4.2). Une

fois l'information génétique résumée par une statistique, telle que le SFS, la deuxième étape consiste à définir plusieurs scénarios démographiques. Dans cet exemple, on considère par exemple des paramètres associés à l'isolement actuel de deux populations ($N_1$ et $N_2$) issus à un temps de génération ($t_s$) d'une population ancestrale ($N_A$), avec un scénario sans flux de gènes et un autre avec flux de gènes historique et actuel (Figure 4.2). En optant pour une méthode d'inférence démographique, telle que celle implémentée dans *fastsimcoal2* (Excoffier et al., 2013), le programme modélise des SFS attendus pour ces différents scénarios démographiques. La finalité sera de choisir le scénario le plus probable compte tenu de la vraisemblance des SFS attendus à celui observé. C'est une des solutions à envisager, en perspective de cette thèse, pour reconstruire l'histoire évolutive des populations de *L. rodriguezii*.

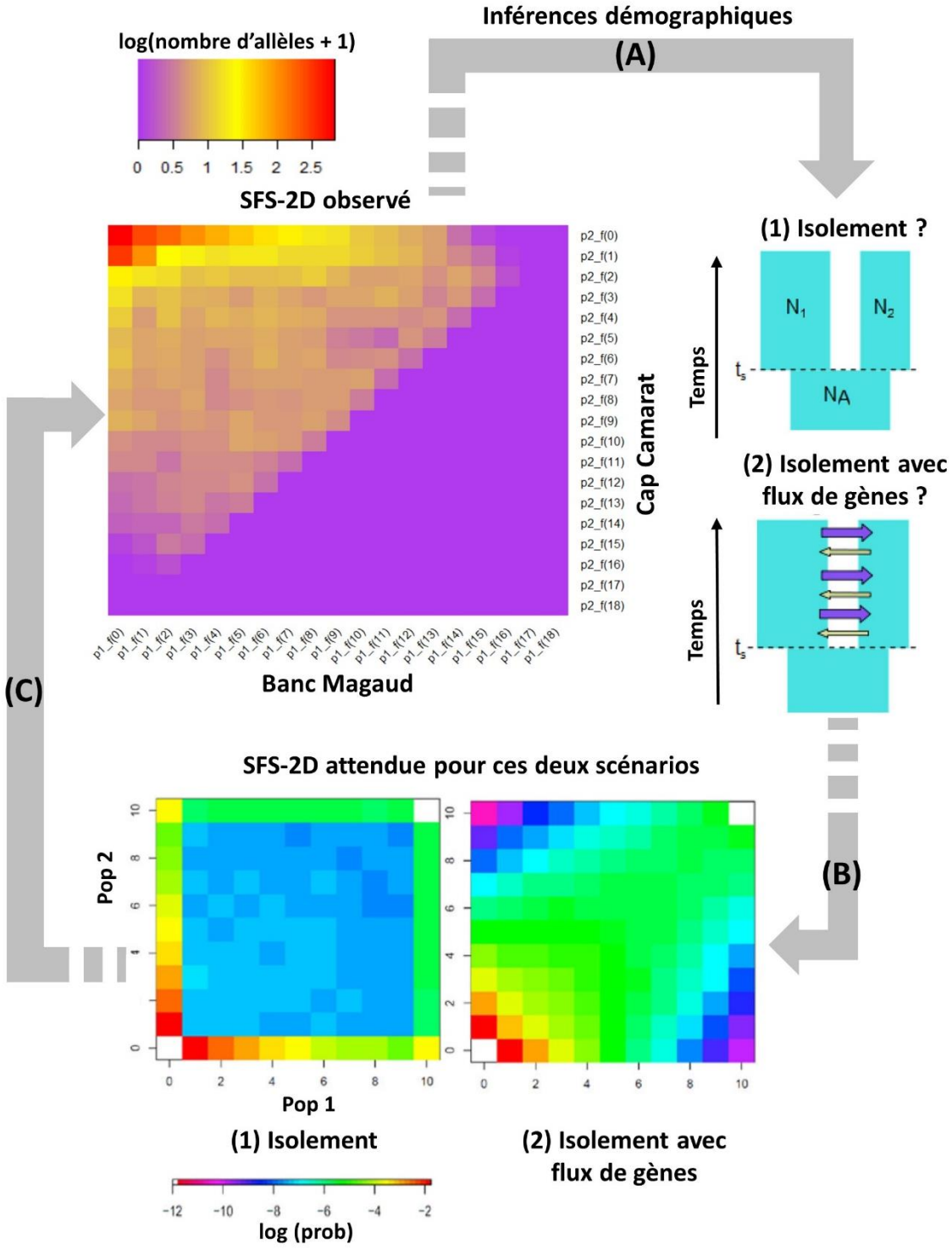



**Figure 4.2**. Représentation schématique de l'inférence démographique (flux de gènes et tailles efficaces passés) à partir du spectre des fréquences alléliques (SFS) et d'une méthode de maximum de vraisemblance. Pour cet exemple, le SFS-2D (observé) a été généré à partir de deux populations de *Laminaria rodriguezii* et de 4077 SNPs à l'aide du script easySFS.py (www.github.com/isaacovercast/easySFS), (A) deux scénarios démographiques sont ensuite envisagés, (B) chacun d'eux est modélisé pour générer le SFS-2D attendu correspondant, (C) le modèle retenu sera celui qui maximise la vraisemblance des données observées.

## 4.3 Conclusion générale

À travers cette thèse nous avons constaté :

**Chapitre 2 (*Ericaria zosteroides*)**

- L'existence d'une structure génétique ($F_{ST}$ = 0,09 ; 10 755 SNPs) à courte distance, incluant la différenciation significative des individus entre la partie haute et basse du tombant de l'Armoire (baie de Toulon, ouest de la Provence), ceux-ci étant séparés par un dénivelé de 20 mètres.

- Le rôle prépondérant des courants de fond et de la connectivité océanographique dans la structure génétique des populations, en dépit du faible potentiel de dispersion de cette espèce par les zygotes.

- La présence d'une population hautement différenciée et génétiquement appauvrie en Provence (RP, baie de Marseille) pour laquelle la différenciation génétique est corrélée à un patron d'isolement par les courants marins.

- Le rôle de la dérive de rameaux fertiles, vecteurs de flux de gènes, pour expliquer la colonisation d'un nouveau site (ALT, Cassis) par un ou des individus originaires d'une zone géographique située à 9 km de distance (LC, La Ciotat).

- Douze locus outliers, susceptibles d'être impliqués dans la réponse adaptative à la profondeur. Les patrons de différenciation génétique pour ces locus sont corrélés à la température.

**Chapitre 3 (*Laminaria rodriguezii*)**

- L'existence d'une structure génétique locale et régionale ($F_{ST}$ = 0,28 ; 4077 SNPs), avec un isolement par la distance dès les 20 premiers mètres. A cette échelle du mètre, les différences génétiques restent suffisamment limitées pour que des individus soient associés à une même population.

- L'impact de la clonalité sur la diversité génétique à l'échelle du génome des individus, avec notamment plusieurs formes de distribution de $F_{IS}$ variable.

- La détection de quatre lignées clonales d'une taille de deux à sept individus avec une occurrence variable dans trois des quatre sites d'études.

- L'existence de certains individus hautement différenciés de leur population d'origine (forte proportion d'allèles privés et de génotypes hétérozygotes), questionnant sur l'âge de ces lignées ou une possible hybridation.

***Implication des résultats obtenus pour la gestion et la conservation des populations***

L'identification d'une structure génétique à courte distance, et la présence de dèmes très peu connectés ont des implications importantes pour la conservation d'*Ericaria zosteroides* et *Laminaria rodriguezii*. Ainsi, chaque population est à considérer comme une entité unique à préserver. Ces deux espèces sont inscrites dans la convention de Berne et Barcelone mais malgré ce statut, leur protection effective n'a jamais été transcrite en droit français (ni dans aucun pays européen d'ailleurs). À l'image des « *paper parks »* ; terme désignant la création d'une Aire Marine protégée sur plan législatif qui ne dispose d'aucun moyen et d'aucune mesures concrètes pour parvenir à ses objectifs de conservation, nous pourrions qualifier le statut protection de ces espèces comme une « *paper protection* ». Ainsi, je propose un ensemble de mesures à mettre en place pour leur conservation :

1 Les populations étudiées ici sont en grande majorité situées dans une aire marine protégée, et l'une des causes de leur déclin est l'arrachage des individus par les filets de pêches. Nous proposons aux gestionnaires des AMPs que *E. zosteroides* et *L. rodriguezii* soient inscrites dans les déclarations de captures des pêcheurs, cela pour suivre et évaluer sur du long terme l'impact de cette activité. Les réserves naturelles de Cerbère Banyuls, de Scandola, des Bouches de Bonifacio et les parcs nationaux de Port-Cros sont concernés par *E. zosteroides* et seulement Port-Cros et Bonifacio pour *L. rodriguezii.*

2 Pour les populations les plus vulnérables, telles que celle de la Roche percée (RP) en baie de Marseille pour *E. zosteroides*, nous pensons que rétablir ou établir des flux de gènes grâce à des individus provenant d'autres populations serait intéressant. Pour cela, l'une des solutions consiste à transplanter des fragments de thalles fertiles et les disposer à proximité des sites cibles des transplantations. Cette technique a été utilisée pour la colonisation des récifs artificiels du Prado en baie de Marseille par *E. zosteroides* et pour réintroduire *Ericaria mediterranea* après que l'espèce avait disparu dans l'Aire Marine Protégée de la Côte Agathoise (travaux en cours). À l'inverse des précédentes actions, celles-ci devront être faites en connaissance de la structure génétique des populations afin de prendre en considération et minimiser les risques de dépression hybride.

3 Plus généralement ce manuscrit souligne la nécessité d'intégrer le potentiel évolutif des espèces dans les mesures de conservation. Des suivis temporels sur plusieurs générations de la diversité génétique seront bénéfiques pour évaluer la stabilité des populations et considérer leur conservation à long terme. Par ailleurs, mon équipe de recherche prône une approche écosystémique de la conservation, incluant tous les compartiments trophiques associés à ces espèces ingénieurs. Il est indéniable que le maintien du potentiel évolutif de ces populations passe par une gestion intégrée des écosystèmes.

# Références

## ~A~


Adjeroud, M., Guérécheau, A., Vidal-Dupiol, J., Flot, J.-F., Arnaud-Haond, S., & Bonhomme, F. (2014). Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: a multi-scale analysis in an insular, fragmented reef system. *Marine Biology*, *161*(3), 531–541. doi: 10.1007/s00227-013-2355-9

Agapow, P.-M., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, *1*(1–2), 101–102. doi: 10.1046/j.1471-8278.2000.00014.x

Alberto, F., Gouveia, L., Arnaud-Haond, S., Perez-Llorens, J. L., Duarte, C. M., & Serrão, E. A. (2005). Within-population spatial genetic structure, neighbourhood size and clonal subrange in the seagrass *Cymodocea nodosa*. *Molecular Ecology*, *14*(9), 2669–2681. doi: 10.1111/j.1365-294X.2005.02640.x

Alberto, F., Raimondi, P. T., Reed, D. C., Watson, J. R., Siegel, D. A., Mitarai, S., … Serrão, E. A. (2011). Isolation by oceanographic distance explains genetic structure for *Macrocystis pyrifera* in the Santa Barbara Channel. *Molecular Ecology*, *20*(12), 2543–2554. doi: https://doi.org/10.1111/j.1365-294X.2011.05117.x

Ali, S., Soubeyrand, S., Gladieux, P., Giraud, T., Leconte, M., Gautier, A., … Enjalbert, J. (2016). Cloncase: Estimation of sex frequency and effective population size by clonemate resampling in partially clonal organisms. *Molecular Ecology Resources*, *16*(4), 845–861. doi: 10.1111/1755-0998.12511

Almeida, S. C., Nicastro, K. R., Zardi, G. I., Pearson, G. A., Valero, M., & Serrão, E. A. (2017). Reproductive strategies and population genetic structure of *Fucus* spp. Across a northeast Atlantic biogeographic transition. *Aquatic Living Resources*, *30*, 16. doi: 10.1051/alr/2017012

Alter, S. E., Tariq, L., Creed, J. K., & Megafu, E. (2021). Evolutionary responses of marine organisms to urbanized seascapes. *Evolutionary Applications*, *14*(1), 210–232. doi: https://doi.org/10.1111/eva.13048

Anderson, E. C., Waples, R. S., & Kalinowski, S. T. (2008). An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(7), 1475–1486. doi: 10.1139/F08-049

Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: Standard methods are upwardly biased. *Molecular Ecology Resources*, *10*(4), 701–710. doi: 10.1111/j.1755-0998.2010.02846.x

Andrew, S. (2010). FastQC: a quality control tool for high throughput sequence data. (Version 0.11.7). Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Andrews, K. R., Adams, J. R., Cassirer, E. F., Plowright, R. K., Gardner, C., Dwire, M., … Waits, L. P. (2018). A bioinformatic pipeline for identifying informative SNP panels for parentage assignment from RADseq data. *Molecular Ecology Resources*, *18*(6), 1263–1281. doi: 10.1111/1755-0998.12910

Antao, T., Pérez-Figueroa, A., & Luikart, G. (2011). Early detection of population declines: High power of genetic monitoring using effective population size estimators. *Evolutionary Applications*, *4*(1), 144–154. doi: https://doi.org/10.1111/j.1752-4571.2010.00150.x

Antiste, A. (2020). Étude de la diversité génétique et clonale au sein des radeaux de sargasses pélagiques à partir de marqueurs microsatellites. Rapport de Master, Aix Marseille Université.

Araújo, R. M., Assis, J., Aguillar, R., Airoldi, L., Bárbara, I., Bartsch, I., … Sousa-Pinto, I. (2016). Status, trends and drivers of kelp forests in Europe: an expert assessment. *Biodiversity and Conservation*, *25*(7), 1319–1348. doi: 10.1007/s10531-016-1141-7

Ardehed, A., Johansson, D., Schagerström, E., Kautsky, L., Johannesson, K., & Pereyra, R. T. (2015). Complex spatial clonal structure in the macroalgae *Fucus radicans* with both sexual and asexual recruitment. *Ecology and Evolution*, *5*(19), 4233–4245. doi: 10.1002/ece3.1629

Arenas, R., González, C. F., Ordás, J. M. R., Fernández, E., & Haya, D. (1995). Growth and reproductive strategies of *Sargassum muticum* (Yendo) Fensholt and *Cystoseira nodicaulis* (Whit.) Roberts. *Scientia Marina*, *59*(Extra 1), 1–8.

Arizmendi-Mejía, R., Linares, C., Garrabou, J., Antunes, A., Ballesteros, E., Cebrian, E., … Ledoux, J.-B. (2015). Combining genetic and demographic data for the conservation of a Mediterranean marine habitat-forming species. *PloS One*, *10*(3), e0119585. doi: 10.1371/journal.pone.0119585

Arnaud-Haond, S., Alberto, F., Teixeira, S., Procaccini, G., Serrão, E. A., & Duarte, C. M. (2005). Assessing genetic diversity in clonal organisms: Low diversity or low resolution? Combining power and cost efficiency in selecting markers. *Journal of Heredity*, *96*(4), 434–440. doi: 10.1093/jhered/esi043

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van De Vliet, M. S., Alberto, F., … Serrão, E. A. (2007a). Vicariance patterns in the Mediterranean Sea: east–west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. *Journal of Biogeography*, *34*(6), 963–976. doi: 10.1111/j.1365-2699.2006.01671.x

Arnaud-Haond, S., Duarte, C. M., Alberto, F., & Serrão, E. A. (2007b). Standardizing methods to address clonality in population studies. *Molecular ecology*, 16(24), 5115-5139. doi: 10.1111/j.1365-294X.2007.03535.x

Arnaud-Haond, S., Stoeckel, S., & Bailleul, D. (2020). New insights into the population genetics of partially clonal organisms: when seagrass data meet theoretical expectations. *Molecular Ecology*, mec.15532. doi: 10.1111/mec.15532

Assis, J., Serrão, E. A., Claro, B., Perrin, C., & Pearson, G. A. (2014). Climate-driven range shifts explain the distribution of extant gene pools and predict future loss of unique lineages in a marine brown alga. *Molecular Ecology*, *23*(11), 2797–2810. doi: https://doi.org/10.1111/mec.12772

Assis, J., Coelho, N. C., Lamy, T., Valero, M., Alberto, F., & Serrão, E. Á. (2016). Deep reefs are climatic refugia for genetic diversity of marine forests. *Journal of Biogeography*, *43*(4), 833–844. doi: https://doi.org/10.1111/jbi.12677

Assis, J., Berecibar, E., Claro, B., Alberto, F., Reed, D., Raimondi, P., & Serrão, E. A. (2017). Major shifts at the range edge of marine forests: The combined effects of climate changes and limited dispersal. *Scientific Reports*, *7*(1), 1–10. doi: 10.1038/srep44348

Assis, J., Araújo, M. B., & Serrão, E. A. (2018a). Projected climate changes threaten ancient refugia of kelp forests in the North Atlantic. *Global Change Biology*, *24*(1), e55–e66. doi: https://doi.org/10.1111/gcb.13818

Assis, J., Serrão, E. Á., Coelho, N. C., Tempera, F., Valero, M., & Alberto, F. (2018b). Past climate changes and strong oceanographic barriers structured low-latitude genetic relics for the golden kelp Laminaria ochroleuca. *Journal of Biogeography*, *45*(10), 2326–2336. doi:

https://doi.org/10.1111/jbi.13425

Avia, K., Coelho, S. M., Montecinos, G. J., Cormier, A., Lerck, F., Mauger, S., … Boudry, P. (2017). High-density genetic map and identification of QTLs for responses to temperature and salinity stresses in the model brown alga *Ectocarpus*. *Scientific Reports*, *7*(1), 1–15. doi: 10.1038/srep43241

**~B~**

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*, *3*(10), e3376. doi: 10.1371/journal.pone.0003376

Ballenghien, M., Faivre, N., & Galtier, N. (2017). Patterns of cross-contamination in a multispecies population genomic project: detection, quantification, impact, and solutions. *BMC Biology*, *15*(1), 25. doi: 10.1186/s12915-017-0366-6

Ballesteros, E. (1990). Structure and dynamics of the community of *Cystoseira zosteroides* (Turner) C. Agardh (Fucales, Phaeophyceae) in the Northwestern Mediterranean. *Scientia Marina*, *54*(3), 217–229.

Ballesteros, E. (2006). Mediterranean coralligenous assemblages: a synthesis of present knowledge. *Oceanography and Marine Biology: An Annual Review*, *44*, 123–195.

Ballesteros, E., Garrabou, J., Hereu, B., Zabala, M., Cebrian, E., & Sala, E. (2009). Deep-water stands of *Cystoseira zosteroides* C. Agardh (Fucales, Ochrophyta) in the Northwestern Mediterranean: Insights into assemblage structure and population dynamics. *Estuarine, Coastal and Shelf Science*, *82*, 477–484.

Balloux, F., Lehmann, L., & De Meeûs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics*, *164*(4), 1635–1644.

Barrier, N., Petrenko, A. A., & Ourmières, Y. (2016). Strong intrusions of the Northern Mediterranean Current on the eastern Gulf of Lion: Insights from in-situ observations and high resolution numerical modelling. *Ocean Dynamics*, *66*(3), 313–327. doi: 10.1007/s10236-016-0921-7

Bartsch, I., Wiencke, C., Bischof, K., Buchholz, C. M., Buck, B. H., Eggert, A., … Wiese, J. (2008). The genus *Laminaria sensu lato*: Recent insights and developments. *European Journal of Phycology*, *43*(1), 1–86. doi: 10.1080/09670260701711376

Bartsch, I., Paar, M., Fredriksen, S., Schwanitz, M., Daniel, C., Hop, H., & Wiencke, C. (2016). Changes in kelp forest biomass and depth distribution in Kongsfjorden, Svalbard, between 1996–1998 and 2012–2014 reflect Arctic warming. *Polar Biology*, *39*(11), 2021–2036. doi: 10.1007/s00300-015-1870-1

Becheler, R., Masson, JP., Arnaud-Haond, S., Halkett, F., Mariette, S., Guillemin, ML., … Stoeckel, S. (2017). ClonEstiMate, a Bayesian method for quantifying rates of clonality of populations genotyped at two-time steps. *Molecular Ecology Resources*, *17*(6), e251–e267. doi: 10.1111/1755-0998.12698

Becheler, R., Guillemin, M.-L., Stoeckel, S., Mauger, S., Saunier, A., Brante, A., Destombe, C., & Valero, M. (2020). After a catastrophe, a little bit of sex is better than nothing: Genetic consequences of a major earthquake on asexual and sexual populations. *Evolutionary Applications*. 13, 2086–2100. https://doi.org/10.1111/eva.12967.

Belkin, I. M. (2009). Rapid warming of large marine ecosystems. *Progress in Oceanography*, *81*(1–4), 207–213. doi: 10.1016/j.pocean.2009.04.011

Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., … Bernatchez, L.

(2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, *25*(20), 5073–5092. doi: 10.1111/mec.13811

Bensoussan, N., Romano, J.-C., Harmelin, J.-G., & Garrabou, J. (2010). High resolution characterization of northwest Mediterranean coastal waters thermal regimes: To better understand responses of benthic communities to climate change. *Estuarine, Coastal and Shelf Science*, *87*(3), 431–441. doi: 10.1016/j.ecss.2010.01.008

Bermejo, R., Chefaoui, R. M., Engelen, A. H., Buonomo, R., Neiva, J., Ferreira-Costa, J., … Serrão, E. A. (2018). Marine forests of the Mediterranean-Atlantic *Cystoseira tamariscifolia* complex show a southern Iberian genetic hotspot and no reproductive isolation in parapatry. *Scientific Reports*, *8*(1), 1–13. doi: 10.1038/s41598-018-28811-1

Betancor, S., Domínguez, B., Tuya, F., Figueroa, F. L., & Haroun, R. (2015). Photosynthetic performance and photoprotection of *Cystoseira humilis* (Phaeophyceae) and *Digenea simplex* (Rhodophyceae) in an intertidal rock pool. *Aquatic Botany*, *121*, 16–25. doi: 10.1016/j.aquabot.2014.10.008

Bierne, N., Welch, J., Loire, E., Bonhomme, F., & David, P. (2011). The coupling hypothesis: Why genome scans may fail to map local adaptation genes. *Molecular Ecology*, *20*(10), 2044–2072. doi: https://doi.org/10.1111/j.1365-294X.2011.05080.x

Bierne, N., Roze, D., & Welch, J. J. (2013). Pervasive selection or is it…? Why are FST outliers sometimes so frequent? *Molecular Ecology*, *22*(8), 2061–2064. doi: https://doi.org/10.1111/mec.12241

Bijlsma, R., & Loeschcke, V. (2012). Genetic erosion impedes adaptive responses to stressful environments. *Evolutionary Applications*, *5*(2), 117–129. doi: https://doi.org/10.1111/j.1752-4571.2011.00214.x

Billard, E., Serrão, E., Pearson, G., Destombe, C., & Valero, M. (2010). *Fucus vesiculosus* and *spiralis* species complex: A nested model of local adaptation at the shore level. *Marine Ecology Progress Series*, *405*, 163–174. doi: 10.3354/meps08517

Billot, C., Engel-Gautier, C., Rousvoal, S., Kloareg, B., & Valero, M. (2003). Current patterns, habitat discontinuities and population genetic structure: The case of the kelp *Laminaria digitata* in the English Channel. *Marine Ecology Progress Series*, *253*, 111–121. doi: 10.3354/meps253111

Birky, C. W. (1996). Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics*,144(1), 427-437.

Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Modelling directional spatial processes in ecological data. *Ecological Modelling*, *215*(4), 325–336. doi: 10.1016/j.ecolmodel.2008.04.001

Blanfuné, A. (2016). Le changement global en Méditerranée Nord Occidentale: Forêt de Cystoseires, de Sargasses, encorbellement à *Lithophyllum* et bloom d'Ostreopsis (Thèse de doctorat, Aix-Marseille). Aix-Marseille. Retrieved from http://www.theses.fr/2016AIXM4024

Blanfuné, A., Boudouresque, C. F., Verlaque, M., & Thibaut, T. (2016). The fate of *Cystoseira crinita*, a forest-forming Fucale (Phaeophyceae, Stramenopiles), in France (North Western Mediterranean Sea). *Estuarine, Coastal and Shelf Science*, *181*, 196–208. doi: 10.1016/j.ecss.2016.08.049

Blanfuné, A., Boudouresque, C. F., Verlaque, M., & Thibaut, T. (2019). The ups and downs of a canopy-forming seaweed over a span of more than one century. *Scientific Reports*, *9*(1), 1–

10. doi: 10.1038/s41598-019-41676-2

Boero, F. (2010). The study of species in the era of biodiversity: A tale of stupidity. *Diversity*, *2*(1), 115–126. doi: 10.3390/d2010115

Boitard, S., Rodríguez, W., Jay, F., Mona, S., & Austerlitz, F. (2016). Inferring population size history from large samples of genome-wide molecular data-an approximate Bayesian computation approach. *PloS Genetics*, *12*(3), e1005877. doi: 10.1371/journal.pgen.1005877

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Bonhomme, P., Goujard, A., Javel, A., Grondin, J., & Boudouresque, C. F. (2014). Unexpected artificial-reef-like effect due to a Mediterranean pipeline and the conservation of two circalittoral emblematic species: *Centrostephanus longispinus* and *Cystoseira zosteroides*. *In Proceedings of the 2nd Mediterranean Symposium on the Conservation of Coralligenous and Other Calcareous Bio-Concretions*, 43–48.

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2020). Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Molecular Ecology*, *29*(22), 4274–4279. doi: https://doi.org/10.1111/mec.15501

Boscari, E., Abbiati, M., Badalamenti, F., Bavestrello, G., Benedetti-Cecchi, L., Cannas, R., … Congiu, L. (2019). A population genomics insight by 2b-RAD reveals populations' uniqueness along the Italian coastline in *Leptopsammia pruvoti* (Scleractinia, Dendrophylliidae). *Diversity and Distributions*, *25*(7), 1101–1117. doi: 10.1111/ddi.12918

Boudouresque, C.-F. (1971). Contribution à l'étude phytosociologique des peuplements algaux des côtes varoises. *Vegetatio*, *22*(1), 83–184. doi: 10.1007/BF01955721

Boudouresque, C.-F., & Perret, M. (1977). Inventaire de la flore marine de Corse (Lubrecht&Cramer Limited, Vol. 25). Retrieved from https://mawujemymuvake.travelingartsfiesta.com/inventaire-de-la-flore-marine-de-corse-book-17420vc.php

Boudouresque, C.-F., Blanfuné, A., Harmelin-Vivien, M., Personnic, S., Ruitton, S., Thibaut, T., & Verlaque, M. (2017). Where seaweed forests meet animal forests: The examples of macroalgae in coral reefs and the Mediterranean coralligenous ecosystem. In *Marine Animal Forests* (pp. 369–396). Cham: Springer International Publishing. doi: 10.1007/978-3-319-21012-4_48

Boudouresque, C.-F., Blanfuné, A., Ruitton, S., & Thibaut, T. (2020). Macroalgae as a tool for coastal management in the Mediterranean Sea. In O. Konur (Ed.), *Handbook of Algal Science, Technology and Medicine* (pp. 277–290). Elsevier and Academic Press. doi: 10.1016/B978-0-12-818305-2.00017-6

Bracco, A., Liu, G., Galaska, M. P., Quattrini, A. M., & Herrera, S. (2019). Integrating physical circulation models and genetic approaches to investigate population connectivity in deep-sea corals. *Journal of Marine Systems*, *198*, 103189. doi: 10.1016/j.jmarsys.2019.103189

Bringloe, T. T., Starko, S., Wade, R. M., Vieira, C., Kawai, H., Clerck, O. D., … Verbruggen, H. (2020). Phylogeny and Evolution of the Brown Algae. *Critical Reviews in Plant Sciences*, *39*(4), 281–321. doi: 10.1080/07352689.2020.1787679

Brown, J. W., & Sorhannus, U. (2010). A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): Substantive underestimation of putative fossil ages. *PloS One*, *5*(9), e12759. doi: 10.1371/journal.pone.0012759

Bruno de Sousa, C., Cox, C. J., Brito, L., Pavão, M. M., Pereira, H., Ferreira, A., … Varela, J. (2019).

Improved phylogeny of brown algae *Cystoseira* (Fucales) from the Atlantic-Mediterranean region based on mitochondrial sequences. *PloS One*, *14*(1), e0210143. doi: 10.1371/journal.pone.0210143

Buonomo, R., Assis, J., Fernandes, F., Engelen, A. H., Airoldi, L., & Serrão, E. A. (2017). Habitat continuity and stepping-stone oceanographic distances explain population genetic connectivity of the brown alga *Cystoseira amentacea*. *Molecular Ecology*, *26*(3), 766–780. doi: https://doi.org/10.1111/mec.13960

Buonomo, R., Chefaoui, R. M., Lacida, R. B., Engelen, A. H., Serrão, E. A., & Airoldi, L. (2018). Predicted extinction of unique genetic diversity in marine forests of *Cystoseira* spp. *Marine Environmental Research*, *138*, 119–128. doi: 10.1016/j.marenvres.2018.04.013

**~C~**

Cahill, A. E., Jode, A. D., Dubois, S., Bouzaza, Z., Aurelle, D., Boissin, E., … Chenuil, A. (2017). A multispecies approach reveals hot spots and cold spots of diversity and connectivity in invertebrate species with contrasting dispersal modes. *Molecular Ecology*, *26*(23), 6563–6577. doi: https://doi.org/10.1111/mec.14389

Cánovas-Molina, A., Montefalcone, M., Bavestrello, G., Masmoudi, M. B., Haguenauer, A., Hammami, P., … Aurelle, D. (2018). From depth to regional spatial genetic differentiation of *Eunicella cavolini* in the NW Mediterranean. *Comptes Rendus Biologies*, *341*(9), 421–432. doi: 10.1016/j.crvi.2018.09.002

Capdevila, P., Linares, C., Aspillaga, E., Navarro, L., Kersting, D. K., & Hereu, B. (2015). Recruitment patterns in the Mediterranean deep-water alga *Cystoseira zosteroides*. *Marine Biology*, *164*(6), 1665–1674.

Capdevila, P. (2017). Life history, population dynamics and conservation of underwater Mediterranean forests: Insights from the long-lived alga *Cystoseira zosteroides*. doi: 10.13140/RG.2.2.13192.47360

Capdevila, P., Linares, C., Aspillaga, E., Riera, J. L., & Hereu, B. (2018). Effective dispersal and density-dependence in mesophotic macroalgal forests: Insights from the Mediterranean species *Cystoseira zosteroides*. *PloS One*, *13*(1), e0191346. doi: 10.1371/journal.pone.0191346

Capdevila, P., Hereu, B., Salguero-Gómez, R., Rovira, G., Medrano, A., Cebrian, E., … Linares, C. (2019). Warming impacts on early life stages increase the vulnerability and delay the population recovery of a long-lived habitat-forming macroalga. *Journal of Ecology*, *107*(3), 1129–1140. doi: https://doi.org/10.1111/1365-2745.13090

Cariou, M., Duret, L., & Charlat, S. (2013). Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecology and Evolution*, *3*(4), 846–852. doi: https://doi.org/10.1002/ece3.512

Carpenter, S. R., & Lodge, D. M. (1986). Effects of submersed macrophytes on ecosystem processes. *Aquatic Botany*, *26*, 341–370. doi: 10.1016/0304-3770(86)90031-8

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3: Genes, Genomes, Genetics*, *1*(3), 171–182. doi: 10.1534/g3.111.000240

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. doi: https://doi.org/10.1111/mec.12354

Catchen, J. M., Hohenlohe, P. A., Bernatchez, L., Funk, W. C., Andrews, K. R., & Allendorf, F. W.

(2017). Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Molecular Ecology Resources*, *17*(3), 362–365. doi: 10.1111/1755-0998.12669

Celis-Plá, P. S. M., Martínez, B., Quintano, E., García-Sánchez, M., Pedersen, A., Navarro, N., … Korbee, N. (2014). Short-term ecophysiological and biochemical responses of *Cystoseira tamariscifolia* and *Ellisolandia elongata* to environmental changes. *Aquatic Biology*, *22*, 227–243. doi: 10.3354/ab00573

Celis-Plá, P. S. M., Bouzon, Z. L., Hall-Spencer, J. M., Schmidt, E. C., Korbee, N., & Figueroa, F. L. (2016). Seasonal biochemical and photophysiological responses in the intertidal macroalga *Cystoseira tamariscifolia* (Ochrophyta). *Marine Environmental Research*, *115*, 89–97. doi: 10.1016/j.marenvres.2015.11.014

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1). doi: 10.1186/s13742-015-0047-8

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195–205. doi: 10.1038/nrg2526

Coleman, M. A., & Brawley, S. H. (2005). Are life history characteristics good predictors of genetic diversity and structure? A case study of the Intertidal alga *Fucus spiralis* (heterokontophyta; Phaeophyceae). *Journal of Phycology*, *41*(4), 753–762. doi: https://doi.org/10.1111/j.0022-3646.2005.04136.x

Coleman, M. A., Gillanders, B. M., & Connell, S. D. (2009). Dispersal and gene flow in the habitat-forming kelp, *Ecklonia radiata*: relative degrees of isolation across an east–west coastline. *Marine and Freshwater Research*, *60*(8), 802–809. doi: 10.1071/MF08268

Coleman, M. A., Roughan, M., Macdonald, H. S., Connell, S. D., Gillanders, B. M., Kelaher, B. P., & Steinberg, P. D. (2011). Variation in the strength of continental boundary currents determines continent-wide connectivity in kelp. *Journal of Ecology*, *99*(4), 1026–1032. doi: 10.1111/j.1365-2745.2011.01822.x

Coleman, M. A., & Wernberg, T. (2018). Genetic and morphological diversity in sympatric kelps with contrasting reproductive strategies. *Aquatic Biology*, *27*, 65–73. doi: 10.3354/ab00698

Coleman, M. A., Clark, J. S., Doblin, M. A., Bishop, M. J., Kelaher, B. P., Coleman, M. A., … Kelaher, B. P. (2019). Genetic differentiation between estuarine and open coast ecotypes of a dominant ecosystem engineer. *Marine and Freshwater Research*, *70*(7), 977–985. doi: 10.1071/MF17392

Coll, M., Piroddi, C., Albouy, C., Lasram, F. B. R., Cheung, W. W. L., Christensen, V., … Pauly, D. (2012). The Mediterranean Sea under siege: Spatial overlap between marine biodiversity, cumulative threats and marine reserves. *Global Ecology and Biogeography*, *21*(4), 465–480. doi: https://doi.org/10.1111/j.1466-8238.2011.00697.x

Coscia, I., Wilmes, S. B., Ironside, J. E., Goward-Brown, A., O'Dea, E., Malham, S. K., … Robins, P. E. (2020). Fine-scale seascape genomics of an exploited marine species, the common cockle *Cerastoderma edule*, using a multimodelling approach. *Evolutionary Applications*, *13*(8), 1854–1867. doi: https://doi.org/10.1111/eva.12932

Couceiro, L., Robuchon, M., Destombe, C., & Valero, M. (2013). Management and conservation of the kelp species *Laminaria digitata*: Using genetic tools to explore the potential exporting role of the MPA "Parc naturel marin d'Iroise." *Aquatic Living Resources*, *26*(2), 197–205. doi: 10.1051/alr/2012027

Coyer, J. A., Peters, A. F., Stam, W. T., & Olsen, J. L. (2003). Post-ice age recolonization and differentiation of *Fucus serratus* L. (Phaeophyceae; Fucaceae) populations in Northern Europe. *Molecular Ecology*, *12*(7), 1817–1829. doi: https://doi.org/10.1046/j.1365-294X.2003.01850.x

Cruaud, A., Gautier, M., Galan, M., Foucaud, J., Sauné, L., Genson, G., … Rasplus, J.-Y. (2014). Empirical assessment of RAD sequencing for interspecific phylogeny. *Molecular Biology and Evolution*, *31*(5), 1272–1274. doi: 10.1093/molbev/msu063

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, *25*(7), 410–418. doi: 10.1016/j.tree.2010.04.001

**~D~**

Dalongeville, A., Andrello, M., Mouillot, D., Lobreaux, S., Fortin, M.-J., Lasram, F., … Manel, S. (2017). Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. *Evolutionary Applications*, *11*(8), 1437–1447.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. doi: 10.1093/bioinformatics/btr330

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499–510. doi: 10.1038/nrg3012

Dayton, P. K. (1985). Ecology of kelp communities. *Annual Review of Ecology and Systematics*, *16*(1), 215–245. doi: 10.1146/annurev.es.16.110185.001243

De Jode, A., David, R., Haguenauer, A., Cahill, A. E., Erga, Z., Guillemain, D., … Chenuil, A. (2019). From seascape ecology to population genomics and back. Spatial and ecological differentiation among cryptic species of the red algae *Lithophyllum stictiforme*/*L. cabiochiae*, main bioconstructors of coralligenous habitats. *Molecular Phylogenetics and Evolution*, *137*, 104–113. doi: 10.1016/j.ympev.2019.04.005

De Meester, L., Gómez, A., Okamura, B., & Schwenk, K. (2002). The Monopolization Hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta Oecologica*, *23*(3), 121–135. doi: 10.1016/S1146-609X(02)01145-1

De Meeûs, T., Lehmann, L., & Balloux, F. (2006). Molecular epidemiology of clonal diploids: A quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and Evolution*, *6*(2), 163–170. doi: 10.1016/j.meegid.2005.02.004

De Meeûs, T., Prugnolle, F., & Agnew, P. (2007). Asexual reproduction: Genetics and evolutionary aspects. *Cellular and Molecular Life Sciences*, *64*(11), 1355–1372. doi: 10.1007/s00018-007-6515-2

De Meeûs, T. (2012). Initiation à la génétique des populations naturelles: Applications aux parasites et à leurs vecteurs. IRD Editions.

DeFaveri, J., Shikano, T., Shimada, Y., Goto, A., & Merilä, J. (2011). Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution*, *65*(6), 1800–1807. doi: https://doi.org/10.1111/j.1558-5646.2011.01247.x

Demes, K. W., & Graham, M. H. (2011). Abiotic Regulation of Investment in Sexual Versus Vegetative Reproduction in the Clonal Kelp *Laminaria Sinclairii* (laminariales, Phaeophyceae). *Journal of Phycology*, *47*(3), 463–470. doi: 10.1111/j.1529-8817.2011.00981.x

Deter, J., Juhel, J.-B., Boulanger, E., Guellati, N., Mauron, S., Holon, F., … Hocdé, R. (2020). Gombessa 5 cruise: CTD profiles in western Mediterranean, July 2019 [Data set]. SEANOE. doi: 10.17882/71814

Diehl, N., & Bischof, K. (2021). Coping with a changing Arctic: Mechanisms of acclimation in the brown seaweed *Saccharina latissima* from Spitsbergen. *Marine Ecology Progress Series*, *657*, 43–57. doi: 10.3354/meps13532

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. doi: https://doi.org/10.1111/1755-0998.12157

Dorant, Y., Cayuela, H., Wellband, K., Laporte, M., Rougemont, Q., Mérot, C., … Bernatchez, L. (2020). Copy number variants outperform SNPs to reveal genotype–temperature association in a marine species. *Molecular Ecology*, *29*(24), 4765–4782. doi: https://doi.org/10.1111/mec.15565

Dorken, M. E., & Eckert, C. G. (2001). Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *Journal of Ecology*, *89*(3), 339–350. doi: 10.1046/j.1365-2745.2001.00558.x

Draisma, S. G. A., Ballesteros, E., Rousseau, F., & Thibaut, T. (2010). DNA Sequence data demonstrate the polyphyly of the genus *Cystoseira* and other *Sargassaceae* genera (Phaeophyceae). *Journal of Phycology*, *46*(6), 1329–1345. doi: 10.1111/j.1529-8817.2010.00891.x

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, *196*(3–4), 483–493. doi: 10.1016/j.ecolmodel.2006.02.015

Drinan, D. P., Gruenthal, K. M., Canino, M. F., Lowry, D., Fisher, M. C., & Hauser, L. (2018). Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (*Gadus macrocephalus*). *Evolutionary Applications*, *11*(8), 1448–1464. doi: https://doi.org/10.1111/eva.12639

Druehl, L. D. (1981). The distribution of Laminariales in the North Pacific with reference to environmental influences. Retrieved from https://agris.fao.org/agris-search/search.do?recordID=US19830915543

Duarte, C. M. (2014). Global change and the future ocean: A grand challenge for marine sciences. *Frontiers in Marine Science*, *1*(63). doi: 10.3389/fmars.2014.00063

Duggins, D., Eckman, J. E., Siddon, C. E., & Klinger, T. (2001). Interactive roles of mesograzers and current flow in survival of kelps. *Marine Ecology Progress Series*, *223*, 143–155. doi: 10.3354/meps223143

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can Population Genetic Structure Be Predicted from Life-History Traits? *The American Naturalist*, *169*(5), 662–672. doi: 10.1086/513490

Durrant, H. M. S., Burridge, C. P., Kelaher, B. P., Barrett, N. S., Edgar, G. J., & Coleman, M. A. (2014). Implications of macroalgal isolation by distance for networks of marine protected areas. *Conservation Biology*, *28*(2), 438–445. doi: https://doi.org/10.1111/cobi.12203

Durrant, H. M. S., Barrett, N. S., Edgar, G. J., Coleman, M. A., & Burridge, C. P. (2018). Seascape habitat patchiness and hydrodynamics explain genetic structuring of kelp populations. *Marine Ecology Progress Series*, *587*, 81–92. doi: 10.3354/meps12447

**~E~**

Eaton, D. A. R. (2014). PyRAD: Assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*(13), 1844–1849. doi: 10.1093/bioinformatics/btu121

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, *66*(3), 399–412. doi: 10.1093/sysbio/syw092

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422–433. doi: 10.1038/nrg.2016.58

Ercegović, A. (1960). Quelques traits caractéristiques de la végétation des algues de l'Adriatique. *Acta Botanica Croatica*, *18–19*.

Evankow, A., Christie, H., Hancke, K., Brysting, A., Junge, C., Fredriksen, S., & Thaulow, J. (2019). Genetic heterogeneity of two bioeconomically important kelp species along the Norwegian coast. *Conservation Genetics*, *20*(3), 615–628. doi: 10.1007/s10592-019-01162-8

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PloS Genet*, *9*(10), e1003905. doi: 10.1371/journal.pgen.1003905

**~F~**

Falace, A., Kaleb, S., Fuente, G. D. L., Asnaghi, V., & Chiantore, M. (2018). *Ex situ* cultivation protocol for *Cystoseira amentacea var. Stricta* (Fucales, Phaeophyceae) from a restoration perspective. *PloS One*, *13*(2), e0193011. doi: 10.1371/journal.pone.0193011

Flores-Moya, A. (2012). Warm temperate seaweed communities: A case study of deep water kelp forests from the Alboran Sea (SW Mediterranean Sea) and the Strait of Gibraltar. In C. Wiencke & K. Bischof (Eds.), *Seaweed Biology: Novel Insights into Ecophysiology, Ecology and Utilization* (pp. 315–327). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-28451-9_15

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*(2), 977–993. doi: 10.1534/genetics.108.092221

Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, *10*(6), 1500–1508. doi: 10.1046/j.1523-1739.1996.10061500.x

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. doi: 10.1016/j.biocon.2005.05.002

Frankham, R., Bradshaw, C. J. A., & Brook, B. W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, *170*, 56–63. doi: 10.1016/j.biocon.2013.12.036

Fraser, C. I., Winter, D. J., Spencer, H. G., & Waters, J. M. (2010). Multigene phylogeny of the southern bull-kelp genus *Durvillaea* (Phaeophyceae: Fucales). *Molecular Phylogenetics and Evolution*, *57*(3), 1301–1311. doi: 10.1016/j.ympev.2010.10.011

Fraser, C. I., Spencer, H. G., & Waters, J. M. (2012). *Durvillaea poha* sp. nov. (Fucales, Phaeophyceae): A buoyant southern bull-kelp species endemic to New Zealand. *Phycologia*, *51*(2), 151–156. doi: 10.2216/11-47.1

Fraysse, M., Pinazo, C., Faure, V. M., Fuchs, R., Lazzari, P., Raimbault, P., & Pairaud, I. (2013). Development of a 3D coupled physical-biogeochemical model for the Marseille coastal area (NW Mediterranean Sea): What complexity is required in the coastal zone? *PloS One*, *8*(12), e80012. doi: 10.1371/journal.pone.0080012

Fraysse, M., Pairaud, I., ROSS, O. N., FAURE, V., & Pinazo, C. (2014). Intrusion of Rhone River diluted

water into the Bay of Marseille: Generation processes and impacts on ecosystem functioning. *Journal of Geophysical Research. Oceans*, *119*(10), 6535–6556. doi: 10.1002/2014JC010022

Fredj, G. (1972). Compte rendu de plongée en S. P. 300 sur les fonds à *Laminaria rodriguezii* Bornet de la pointe de Revellata (Corse). *Bulletin Institut Océanographique de Monaco, 71*, 1–42.

Fredriksen, S., Sjøtun, K., Lein, T. E., & Rueness, J. (1995). Spore dispersal in *Laminaria hyperborea* (Laminariales, Phaeophyceae). *Sarsia*, *80*(1), 47–53. doi: 10.1080/00364827.1995.10413579

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*(4), 973–983. doi: 10.1534/genetics.113.160572

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. doi: 10.1111/2041-210X.12382

Fritsch, F. E. (1935). The structure and reproduction of the algae (Cambridge University Press).

**~G~**

Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., … Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, *8*(8), 769–786. doi: 10.1111/eva.12288

Gagnaire, P.-A. (2020). Comparative genomics approach to evolutionary process connectivity. *Evolutionary Applications*, *13*(6), 1320–1334.

Garrabou, J., Coma, R., Bensoussan, N., Bally, M., Chevaldonné, P., Cigliano, M., … Cerrano, C. (2009). Mass mortality in Northwestern Mediterranean rocky benthic communities: Effects of the 2003 heat wave. *Global Change Biology*, *15*(5), 1090–1103. doi: 10.1111/j.1365-2486.2008.01823.x

Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhué, C., Pudlo, P., … Estoup, A. (2013). The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology*, *22*(11), 3165–3178. doi: https://doi.org/10.1111/mec.12089

Giaccone, G. (1972). Struttura, ecologia e corologia dei popolamenti a Laminarie dello stretto di Messina e del mare di Alboran. *Mem. Biol. Mar. Ocean*, *2*, 37–49.

Giaccone, G. (1973). Écologie et chorologie des *Cystoseira* de Méditerranée. *Rapports Commission Internationale Mer Méditerranée*, *22*, 49–50.

Goecke, F., Klemetsdal, G., & Ergon, Å. (2020). Cultivar Development of Kelps for Commercial Cultivation—Past Lessons and Future Prospects. *Frontiers in Marine Science*, *7*, 110. doi: 10.3389/fmars.2020.00110

González de León, S., Herrera, I., & Guevara, R. (2016). Mating system, population growth, and management scenario for *Kalanchoe pinnata* in an invaded seasonally dry tropical forest. *Ecology and Evolution*, *6*(13), 4541–4550. doi: 10.1002/ece3.2219

Gorospe, K. D., Donahue, M. J., & Karl, S. A. (2015). The importance of sampling design: spatial patterns and clonality in estimating the genetic diversity of coral reefs. *Marine Biology*, *162*(5), 917–928. doi: 10.1007/s00227-015-2634-8

Gosselin, T., Anderson, E. C., & Bradbury, I. (2016). Assigner: Assignment Analysis with GBS/RAD Data using R. R package version 0.4. 1. (Version 0.4.1.).

Gotelli, N. J. (1991). Metapopulation models: The rescue effect, the propagule rain, and the core-satellite hypothesis. *The American Naturalist*, *138*(3), 768–776. doi: 10.1086/285249

Goudet, J., Raymond, M., De Meeûs, T., & Rousset, F. (1996). Testing differentiation in diploid populations. *Genetics*, *144*(4), 1933–1940. doi: 10.1093/genetics/144.4.1933

Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184–186. doi: 10.1111/j.1471-8286.2004.00828.x

Graham, M. H., Kinlan, B. P., Druehl, L. D., Garske, L. E., & Banks, S. (2007). Deep-water kelp refugia as potential hotspots of tropical marine diversity and productivity. *Proceedings of the National Academy of Sciences*, *104*(42), 16576–16580. doi: 10.1073/pnas.0704778104

Gregory, T. R., Nicol, J. A., Tamm, H., Kullman, B., Kullman, K., Leitch, I. J., … Bennett, M. D. (2007). Eukaryotic genome size databases. *Nucleic Acids Research*, *35*(Database issue), D332–D338. doi: 10.1093/nar/gkl828

Groupe d'experts intergouvernemental sur l'évolution du climat (2008). Rapport spécial du GIEC sur les conséquences d'un réchauffement planétaire de 1,5 °C (p. 110). Retrieved from https://climat.be/changements-climatiques/changements-observes/rapports-du-giec/2018-rapport-special-sur-la-hausse-de-1-5c

Grulois, D. (2010). Etude de la dispersion et du recrutement à différentes échelles spatiales chez *Undaria pinnatifida*, une macro-algue introduite le long des côtes bretonnes (Thèse de doctorat, Paris 6). Paris 6. Retrieved from https://hal.sorbonne-universite.fr/tel-01111061

Grulois, D., Lévêque, L., & Viard, F. (2011). Mosaic genetic structure and sustainable establishment of the invasive kelp *Undaria pinnatifida* within a bay (Bay of St-Malo, Brittany). *CBM-Cahiers de Biologie Marine*, *52*(4), 485–498.

Guern, M. (1962). Embryologie de quelques espèces du genre *Cystoseira* Agardh 1821 (Fucales). *Vie et Milieu*, *13*(4), 649–679.

Guillemin, M.-L., Faugeron, S., Destombe, C., Viard, F., Correa, J. A., & Valero, M. (2008). Genetic variation in wild and cultivated populations of the haploid–diploid red alga *Gracilaria chilensis*: How farming practices favor asexual reproduction and heterozygosity. *Evolution*, *62*(6), 1500–1519. doi: 10.1111/j.1558-5646.2008.00373.x

Guiry, M., & Guiry, G. (2021). Algaebase. Retrieved February 9, 2021, from https://www.algaebase.org/

Gutenkunst, R., Hernandez, R., Williamson, S., & Bustamante, C. (2010). Diffusion Approximations for Demographic Inference: DaDi. *Nature Precedings*, 1–1. doi: 10.1038/npre.2010.4594.1

Guzinski, J., Ballenghien, M., Daguin-Thiébaut, C., Lévêque, L., & Viard, F. (2018). Population genomics of the introduced and cultivated Pacific kelp *Undaria pinnatifida*: Marinas—not farms—drive regional connectivity and establishment in natural rocky reefs. *Evolutionary Applications*, *11*(9), 1582–1597. doi: https://doi.org/10.1111/eva.12647

Guzinski, J., Ruggeri, P., Ballenghien, M., Mauger, S., Jacquemin, B., Jollivet, C., … Valero, M. (2020). Seascape Genomics of the Sugar Kelp *Saccharina latissima* along the North Eastern Atlantic Latitudinal Gradient. *Genes*, *11*(12), 1503. doi: 10.3390/genes11121503

**~H~**

Halkett, F., Simon, J.-C., & Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, *20*(4), 194–201. doi: 10.1016/j.tree.2005.01.001

Hamrick, J. L., & Godt, M. J. W. (1990). Allozyme diversity in plant species. *Plant Population Genetics, Breeding, and Genetic Resources.*, 43–63.

Hawes, N. A., Taylor, D. I., & Schiel, D. R. (2017). Transport of drifting fucoid algae: Nearshore transport and potential for long distance dispersal. *Journal of Experimental Marine Biology and Ecology*, *490*, 34–41. doi: 10.1016/j.jembe.2017.02.001

Hereu, B., Mangialajo, L., Ballesteros, E., & Thibaut, T. (2008). On the occurrence, structure and distribution of deep-water *Cystoseira* (Phaeophyceae) populations in the Port-Cros National Park (north-western Mediterranean). *European Journal of Phycology*, *43*(3), 263–273. doi: 10.1080/09670260801930330

Herrera, S., Reyes-Herrera, P. H., & Shank, T. M. (2015). Predicting RAD-seq marker numbers across the eukaryotic tree of life. *Genome Biology and Evolution*, *7*(12), 3207–3225. doi: 10.1093/gbe/evv210

Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetics Research*, *38*(3), 209–216. doi: 10.1017/S0016672300020553

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. doi: 10.1093/molbev/msx281

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., … Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*(4), 379–397. doi: 10.1086/688018

Hoban, S., Bruford, M., D'Urban Jackson, J., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P. A., … Laikre, L. (2020). Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. *Biological Conservation*, *248*, 108654. doi: 10.1016/j.biocon.2020.108654

Hoffman, J. I., Grant, S. M., Forcada, J., & Phillips, C. D. (2011). Bayesian inference of a historical bottleneck in a heavily exploited marine mammal. *Molecular Ecology*, *20*(19), 3989–4008. doi: https://doi.org/10.1111/j.1365-294X.2011.05248.x

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet*, *6*(2), e1000862. doi: 10.1371/journal.pgen.1000862

Houghton, J. T., Jenkins, G. J., & Ephraums, J. J. (1990). Climate change. Unitated kingdom. Retrieved from https://www.osti.gov/etdeweb/biblio/6041139

Hughes, T. P., Anderson, K. D., Connolly, S. R., Heron, S. F., Kerry, J. T., Lough, J. M., … Wilson, S. K. (2018). Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science*, *359*(6371), 80–83. doi: 10.1126/science.aan8048

Huson, D. H. (1998). SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics*. *14*(1), 68–73. doi: 10.1093/bioinformatics/14.1.68

Huvé, H. (1955). Présence de *Laminaria rodriguezii* Bornet sur les côtes françaises de la Méditerranée. *Recueil Des Travaux de La Station Marine d'Endoume*, *15*.

**~I~**

Iveša, L., Djakovac, T., & Devescovi, M. (2015). Spreading patterns of the invasive *Caulerpa cylindracea* Sonder along the west Istrian Coast (northern Adriatic Sea, Croatia). *Marine Environmental Research*, *107*, 1–7. doi: 10.1016/j.marenvres.2015.03.008

Iveša, L., Djakovac, T., & Devescovi, M. (2016). Long-term fluctuations in *Cystoseira* populations along the west Istrian Coast (Croatia) related to eutrophication patterns in the northern Adriatic Sea. *Marine Pollution Bulletin*, *106*(1), 162–173. doi: 10.1016/j.marpolbul.2016.03.010

**~J~**

Jamieson, I. G., & Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution*, *27*(10), 578–584. doi: 10.1016/j.tree.2012.07.001

Jaron, K. S., Bast, J., Nowell, R. W., Ranallo-Benavidez, T. R., Robinson-Rechavi, M., & Schwander, T. (2021). Genomic Features of Parthenogenetic Animals. *Journal of Heredity*, *112*(1), 19–33. doi: 10.1093/jhered/esaa031

Jódar-Pérez, A. B., Terradas-Fernández, M., López-Moya, F., Asensio-Berbegal, L., & López-Llorca, L. V. (2020). Multidisciplinary Analysis of *Cystoseira sensu lato* (SE Spain) Suggest a Complex Colonization of the Mediterranean. *Journal of Marine Science and Engineering*, *8*(12), 961. doi: 10.3390/jmse8120961

Johnson, M. S., & Black, R. (1984). The Wahlund effect and the geographical scale of variation in the intertidal limpet *Siphonaria* sp. *Marine Biology*, *79*(3), 295–302. doi: 10.1007/BF00393261

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 1–15. doi: 10.1186/1471-2156-11-94

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. doi: 10.1093/bioinformatics/btr521

Jones, C. G., Lawton, J. H., & Shachak, M. (1994). Organisms as ecosystem engineers. In F. B. Samson & F. L. Knopf (Eds.), *Ecosystem management* (pp. 130–147). New York, NY: Springer. doi: 10.1007/978-1-4612-4018-1_14

Judson, O. P., & Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution*, *11*(2), 41–46. doi: 10.1016/0169-5347(96)81040-8

Jueterbock, A., Coyer, J. A., Olsen, J. L., & Hoarau, G. (2018). Decadal stability in genetic variation and structure in the intertidal seaweed *Fucus serratus* (Heterokontophyta: Fucaceae). *BMC Evolutionary Biology*, *18*(1), 1–12. doi: 10.1186/s12862-018-1213-2

**~K~**

Kain, J. (1979). A view of the genus *Laminaria*. *Oceanography and Marine Biology*, *17*, 101–161.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. doi: 10.1038/nmeth.4285

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. doi: 10.7717/peerj.281

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*. doi: 10.3389/fgene.2015.00208

Kapraun, D. (2005). Nuclear DNA content estimates in multicellular green, red and brown algae: Phylogenetic considerations. *Annals of Botany*, *95*(1), 7–44. doi: 10.1093/aob/mci002

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241. doi: https://doi.org/10.1111/j.1461-0248.2004.00684.x

Keller, L., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230–241. doi: 10.1016/S0169-5347(02)02489-8

Kelly, R. P., & Palumbi, S. R. (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PloS One*, *5*(1), e8594. doi: 10.1371/journal.pone.0008594

Kettenring, K. M., & Mock, K. E. (2012). Genetic diversity, reproductive mode, and dispersal differ

between the cryptic invader, *Phragmites australis*, and its native conspecific. *Biological Invasions*, *14*(12), 2489–2504. doi: 10.1007/s10530-012-0246-5

King, N. G., McKeown, N. J., Smale, D. A., Bradbury, S., Stamp, T., Jüterbock, A., … Moore, P. J. (2020). Hierarchical genetic structuring in the cool boreal kelp, *Laminaria digitata*: Implications for conservation and management. *ICES Journal of Marine Science*, *77*(5), 1906–1913. doi: 10.1093/icesjms/fsaa055

Kinlan, B. P., & Gaines, S. D. (2003). Propagule dispersal in marine and terrestrial environments: A community perspective. *Ecology*, *84*(8), 2007–2020. doi: 10.1890/01-0622

Kinnby, A., Jonsson, P. R., Ortega-Martinez, O., Töpel, M., Pavia, H., Pereyra, R. T., & Johannesson, K. (2020). Combining an ecological experiment and a genome scan show idiosyncratic responses to salinity stress in local populations of a seaweed. *Frontiers in Marine Science*, *7*. doi: 10.3389/fmars.2020.00470

Koffi, M., de Meeûs, T. D., Bucheton, B., Solano, P., Camara, M., Kaba, D., … Jamonneau, V. (2009). Population genetics of *Trypanosoma brucei gambiense*, the agent of sleeping sickness in Western Africa. *Proceedings of the National Academy of Sciences*, *106*(1), 209–214. doi: 10.1073/pnas.0811080106

Krueger-Hadfield, S. A., Guillemin, M.-L., Destombe, C., Valero, M., & Stoeckel, S. (2021). Exploring the genetic consequences of clonality in haplodiplontic taxa. *Journal of Heredity*, *112*(1), 92–107. doi: 10.1093/jhered/esaa063

Krumhansl, K. A., Okamoto, D. K., Rassweiler, A., Novak, M., Bolton, J. J., Cavanaugh, K. C., … Byrnes, J. E. K. (2016). Global patterns of kelp forest change over the past half-century. *Proceedings of the National Academy of Sciences*, *113*(48), 13785–13790. doi: 10.1073/pnas.1606102113

**~L~**

Laikre, L., Hoban, S., Bruford, M. W., Segelbacher, G., Allendorf, F. W., Gajardo, G., … Vernesi, C. (2020). Post-2020 goals overlook genetic diversity. *Science*, *367*(6482), 1083–1085. doi: 10.1126/science.abb2748

Laurin-Lemay, S., Brinkmann, H., & Philippe, H. (2012). Origin of land plants revisited in the light of sequence contamination and missing data. *Current Biology*, *22*(15), R593–R594. doi: 10.1016/j.cub.2012.06.013

Lazure, P., & Dumas, F. (2008). An external–internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). *Advances in Water Resources*, *31*(2), 233–250. doi: 10.1016/j.advwatres.2007.06.010

Le Cam, S., Daguin-Thiébaut, C., Bouchemousse, S., Engelen, A. H., Mieszkowska, N., & Viard, F. (2020). A genome-wide investigation of the worldwide invader *Sargassum muticum* shows high success albeit (almost) no genetic diversity. *Evolutionary Applications*, 13(3), 500-514. doi: 10.1111/eva.12837

Ledoux, J.-B., Garrabou, J., Bianchimani, O., Drap, P., Féral, J.-P., & Aurelle, D. (2010). Fine-scale genetic structure and inferences on population biology in the threatened Mediterranean red coral, *Corallium rubrum*. *Molecular Ecology*, *19*(19), 4204–4216. doi: 10.1111/j.1365-294X.2010.04814.x

Ledoux, J.-B., Aurelle, D., Bensoussan, N., Marschal, C., Féral, J.-P., & Garrabou, J. (2015). 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. doi: https://doi.org/10.1002/ece3.1324

Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, *69*(1), 24.

Lejeusne, C., Chevaldonné, P., Pergent-Martini, C., Boudouresque, C. F., & Pérez, T. (2010). Climate change effects on a miniature ocean: The highly diverse, highly impacted Mediterranean Sea. *Trends in Ecology & Evolution*, *25*(4), 250–260. doi: 10.1016/j.tree.2009.10.009

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189. doi: 10.1016/S0169-5347(02)02497-7

Lenormand, T. (2012). From local adaptation to speciation: Specialization and reinforcement. *International Journal of Ecology*, *2012*, 1–11. doi: 10.1155/2012/508458

Levy, E., Byrne, M., Coates, D. J., Macdonald, B. M., McArthur, S., & van Leeuwen, S. (2016). Contrasting Influences of Geographic Range and Distribution of Populations on Patterns of Genetic Diversity in Two Sympatric *Pilbara Acacias*. *PloS One*, *11*(10). doi: 10.1371/journal.pone.0163995

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Liesner, D., Fouqueau, L., Valero, M., Roleda, M. Y., Pearson, G. A., Bischof, K., … Bartsch, I. (2020). Heat stress responses and population genetics of the kelp *Laminaria digitata* (Phaeophyceae) across latitudes reveal differentiation among North Atlantic populations. *Ecology and Evolution*, *10*(17), 9144–9177. doi: 10.1002/ece3.6569

Locatelli, N., & Drew, J. (2019). Population structure and clonal prevalence of scleractinian corals (*Montipora capitata* and *Porites compressa*) in Kaneohe Bay, Oahu. BioRxiv doi: 10.1101/2019.12.11.860585

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*(5), 1031–1046. doi: https://doi.org/10.1111/mec.13100

Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, *19*(15), 3038–3051. doi: https://doi.org/10.1111/j.1365-294X.2010.04688.x

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2016). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, *17*(2), 142–152. doi: https://doi.org/10.1111/1755-0998.12635

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017). Responsible RAD: Striving for best practices in population genomic studies of adaptation. *Molecular Ecology Resources*, *17*(3), 366–369. doi: 10.1111/1755-0998.12677

Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., & Allendorf, F. W. (2010). Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. *Conservation Genetics*, *11*(2), 355–373. doi: 10.1007/s10592-010-0050-7

Luikart, G., Antao, T., Hand, B. K., Muhlfeld, C. C., Boyer, M. C., Cosart, T., … Waples, R. S. (2021). Detecting population declines via monitoring the effective number of breeders (Nb). *Molecular Ecology Resources*, *21*(2), 379–393. doi: https://doi.org/10.1111/1755-0998.13251

Luttikhuizen, P. C., Heuvel, F. H. M. van den, Rebours, C., Witte, H. J., Bleijswijk, J. D. L. van, & Timmermans, K. (2018). Strong population structure but no equilibrium yet: Genetic connectivity and phylogeography in the kelp *Saccharina latissima* (Laminariales, Phaeophyta). *Ecology and Evolution*, *8*(8), 4265–4277. doi:

https://doi.org/10.1002/ece3.3968

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67–77. doi: 10.1111/1755-0998.12592

Lynch, M. (1984). Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *The Quarterly Review of Biology*, *59*(3), 257–290. doi: 10.1086/413902

## ~M~

Mancuso, F. P., Strain, E. M. A., Piccioni, E., De Clerck, O., Sarà, G., & Airoldi, L. (2018). Status of vulnerable *Cystoseira* populations along the Italian infralittoral fringe, and relationships with environmental and anthropogenic variables. *Marine Pollution Bulletin*, *129*(2), 762–771. doi: 10.1016/j.marpolbul.2017.10.068

Mangialajo, L., Chiantore, M., Susini, M.-L., Meinesz, A., Cattaneo-Vietti, R., & Thibaut, T. (2012). Zonation patterns and interspecific relationships of fucoids in microtidal environments. *Journal of Experimental Marine Biology and Ecology*, *412*, 72–80. doi: 10.1016/j.jembe.2011.10.031

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, *26*(22), 2867–2873. doi: 10.1093/bioinformatics/btq559

Mann, K. H. (1973). Seaweeds: Their productivity and strategy for growth. *Science*, *182*(4116), 975–981.

Mao, X., Augyte, S., Huang, M., Hare, M. P., Bailey, D., Umanzor, S., … Jannink, J.-L. (2020). Population Genetics of Sugar Kelp Throughout the Northeastern United States Using Genome-Wide Markers. *Frontiers in Marine Science*, *7*, 694. doi: 10.3389/fmars.2020.00694

Marshall, D. R., & Weir, B. S. (1979). Maintenance of genetic variation in apomictic plant populations. *Heredity*, 42(2), 159-172. doi: 10.1038/hdy.1979.20

Martins, H., Caye, K., Luu, K., Blum, M. G. B., & François, O. (2016). Identifying outlier loci in admixed and in continuous populations using ancestral population differentiation statistics. *Molecular Ecology*, *25*(20), 5029–5042. doi: 10.1111/mec.13822

Maruki, T., & Lynch, M. (2017). Genotype calling from population-genomic sequencing data. *G3: Genes, Genomes, Genetics*, *7*(5), 1393–1404. doi: 10.1534/g3.117.039008

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*(1), 28–41. doi: 10.1111/1755-0998.12291

McKinney, G. J., Larson, W. A., Seeb, L. W., & Seeb, J. E. (2017a). RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: Comment on Breaking RAD by Lowry et al. (2016). *Molecular Ecology Resources*, *17*(3), 356–361. doi: https://doi.org/10.1111/1755-0998.12649

McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017b). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources*, *17*(4), 656–669. doi: 10.1111/1755-0998.12613

Medrano, A., Hereu, B., Mariani, S., Neiva, J., Pagès-Escolà, M., Paulino, C., … Linares, C. (2020). Ecological traits, genetic diversity and regional distribution of the macroalga *Treptacantha*

*elegans* along the Catalan coast (NW Mediterranean Sea). *Scientific Reports*, *10*(1), 1–10. doi: 10.1038/s41598-020-76066-6

Meloni, M., Reid, A., Caujapé-Castells, J., Marrero, Á., Fernández-Palacios, J. M., Mesa-Coelo, R. A., & Conti, E. (2013). Effects of clonality on the genetic variability of rare, insular species: the case of *Ruta microcarpa* from the Canary Islands. *Ecology and Evolution*, *3*(6), 1569–1579. doi: 10.1002/ece3.571

Millet, B., Pinazo, C., Banaru, D., Pagès, R., Guiart, P., & Pairaud, I. (2018). Unexpected spatial impact of treatment plant discharges induced by episodic hydrodynamic events: Modelling Lagrangian transport of fine particles by Northern Current intrusions in the bays of Marseille (France). *PloS One*, *13*(4), e0195257. doi: 10.1371/journal.pone.0195257

Mineur, F., Arenas, F., Assis, J., Davies, A. J., Engelen, A. H., Fernandes, F., … De Clerck, O. (2015). European seaweeds under pressure: Consequences for communities and ecosystem functioning. *Journal of Sea Research*, *98*, 91–108. doi: 10.1016/j.seares.2014.11.004

Mokhtar-Jamaï, K., Pascual, M., Ledoux, J.-B., Coma, R., Féral, J.-P., Garrabou, J., & Aurelle, D. (2011). From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: The interplay between oceanographic conditions and limited larval dispersal. *Molecular Ecology*, *20*(16), 3291–3305. doi: https://doi.org/10.1111/j.1365-294X.2011.05176.x

Molinari, N., & Guiry, M. D. (2020). Reinstatement of the genera *Gongolaria* Boehmer and *Ericaria* Stackhouse (Sargassaceae, Phaeophyceae). *Notulae Algarum*, *171*, 1–10.

Mooney, K. M., Beatty, G. E., Elsäßer, B., Follis, E. S., Kregting, L., O'Connor, N. E., … Provan, J. (2018). Hierarchical structuring of genetic variation at differing geographic scales in the cultivated sugar kelp *Saccharina latissima*. *Marine Environmental Research*, *142*, 108–115. doi: 10.1016/j.marenvres.2018.09.029

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., … Wilbanks, T. J. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–756. doi: 10.1038/nature08823

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*(1), 2–9. doi: 10.1016/0027-5107(64)90047-8

**~N~**

Nadukkalam Ravindran, P., Bentzen, P., Bradbury, I. R., & Beiko, R. G. (2018). PMERGE: Computational filtering of paralogous sequences from RAD-seq data. *Ecology and Evolution*, *8*(14), 7002–7013. doi: 10.1002/ece3.4219

Næss, T. (2019). Population genetics of *Saccharina latissima* (sugar kelp) in Norway. Master's thesis, The University of Bergen, The University of Bergen. Retrieved from https://bora.uib.no/bora-xmlui/handle/1956/19594

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*(s1), 184–194. doi: https://doi.org/10.1111/j.1755-0998.2011.02987.x

Navarro, L., Ballesteros, E., Linares, C., & Hereu, B. (2011). Spatial and temporal variability of deep-water algal assemblages in the Northwestern Mediterranean: The effects of an exceptional storm. *Estuarine, Coastal and Shelf Science*, *95*(1), 52–58. doi: 10.1016/j.ecss.2011.08.002

Nébavi, F., Ayala, F. J., Renaud, F., Bertout, S., Eholie, S., Moussa, K., … De Meeûs, T. (2006). Clonal population structure and genetic diversity of *Candida albicans* in AIDS patients from Abidjan

(Cote d'Ivoire). *Proceedings of the National Academy of Sciences*, *103*(10), 3663–3668. doi: 10.1073/pnas.0511328103

Neiva, J., Pearson, G. A., Valero, M., & Serrão, E. A. (2012a). Drifting fronds and drifting alleles: Range dynamics, local dispersal and habitat isolation shape the population structure of the estuarine seaweed *Fucus ceranoides*. *Journal of Biogeography*, *39*(6), 1167–1178. doi: https://doi.org/10.1111/j.1365-2699.2011.02670.x

Neiva, J., Pearson, G. A., Valero, M., & Serrão, E. A. (2012b). Fine-scale genetic breaks driven by historical range dynamics and ongoing density-barrier effects in the estuarine seaweed *Fucus ceranoides*. *BMC Evolutionary Biology*, *12*(1), 78. doi: 10.1186/1471-2148-12-78

Neiva, J., Serrão, E. A., Anderson, L., Raimondi, P. T., Martins, N., Gouveia, L., … Pearson, G. A. (2017). Cryptic diversity, geographical endemism and allopolyploidy in NE Pacific seaweeds. *BMC Evolutionary Biology*, *17*(1), 30. doi: 10.1186/s12862-017-0878-2

Neiva, J., Paulino, C., Nielsen, M. M., Krause-Jensen, D., Saunders, G. W., Assis, J., … Serrão, E. A. (2018). Glacial vicariance drives phylogeographic diversification in the amphi-boreal kelp *Saccharina latissima*. *Scientific Reports*, *8*(1), 1–12. doi: 10.1038/s41598-018-19620-7

Neiva, J., Serrão, E. A., Paulino, C., Gouveia, L., Want, A., Tamigneaux, É., … Valero, M. (2020). Genetic structure of amphi-Atlantic *Laminaria digitata* (Laminariales, Phaeophyceae) reveals a unique range-edge gene pool and suggests post-glacial colonization of the NW Atlantic. *European Journal of Phycology*, *55*(4), 517–528. doi: 10.1080/09670262.2020.1750058

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. doi: 10.1093/molbev/msu300

Nicastro, K. R., Assis, J., Serrão, E. A., Pearson, G. A., Neiva, J., Valero, M., … Zardi, G. I. (2020). Congruence between fine-scale genetic breaks and dispersal potential in an estuarine seaweed across multiple transition zones. *ICES Journal of Marine Science*, *77*(1), 371–378. doi: 10.1093/icesjms/fsz179

Nicolle, A., Garreau, P., & Liorzou, B. (2009). Modelling for anchovy recruitment studies in the Gulf of Lions (Western Mediterranean Sea). *Ocean Dynamics*, *59*(6), 953–968. doi: 10.1007/s10236-009-0221-6

Nielsen, M. M., Paulino, C., Neiva, J., Krause-Jensen, D., Bruhn, A., & Serrão, E. A. (2016). Genetic diversity of *Saccharina latissima* (Phaeophyceae) along a salinity gradient in the North Sea–Baltic Sea transition zone. *Journal of Phycology*, *52*(4), 523–531. doi: https://doi.org/10.1111/jpy.12428

Nomura, T. (2008). Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, *1*(3), 462–474. doi: https://doi.org/10.1111/j.1752-4571.2008.00015.x

Norderhaug, K. M., & Christie, H. C. (2009). Sea urchin grazing and kelp re-vegetation in the NE Atlantic. *Marine Biology Research*, *5*(6), 515–528. doi: 10.1080/17451000902932985

Nougué, O., Rode, N. O., Jabbour-zahab, R., Ségard, A., Chevin, L.-M., Haag, C. R., & Lenormand, T. (2015). Automixis in Artemia: Solving a century-old controversy. *Journal of Evolutionary Biology*, *28*(12), 2337–2348. doi: 10.1111/jeb.12757

Nunziata, S. O., & Weisrock, D. W. (2018). Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, *120*(3), 196–207. doi: 10.1038/s41437-017-0037-y

**~O~**

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*(16), 3193-3206. doi: 10.1111/mec.14792

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., & Oksanen, M. J. (2017). Community Ecology Package 'vegan'. Version 2.4-4. *R Package*, *2*, 1–192.

Orellana, S., Hernández, M., & Sansón, M. (2019). Diversity of *Cystoseira sensu lato* (Fucales, Phaeophyceae) in the eastern Atlantic and Mediterranean based on morphological and DNA evidence, including *Carpodesmia* gen. Emend. and *Treptacantha* gen. Emend. *European Journal of Phycology*, *54*(3), 447–465. doi: 10.1080/09670262.2019.1590862

**~P~**

Pageault, L. (2020). Growth and photosynthesis of the endemic mediterranean deep-water kelp *Laminaria rodriguezii* along temperature and irradiance gradient. Master thesis, Aix-Marseille University.

Paine, R. T. (1969). A note on trophic complexity and community stability. *The American Naturalist*, *103*(929), 91–93. doi: 10.1086/282586

Pairaud, I. L., Gatti, J., Bensoussan, N., Verney, R., & Garreau, P. (2011). Hydrology and circulation in a coastal area off Marseille: Validation of a nested 3D model with observations. *Journal of Marine Systems*, *88*(1), 20–33. doi: 10.1016/j.jmarsys.2011.02.010

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, *13*(sp1), 146–158. doi: https://doi.org/10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2

Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D., Castelin, M., … Samadi, S. (2015a). Species are hypotheses: Avoid connectivity assessments based on pillars of sand. *Molecular Ecology*, *24*(3), 525–544. doi: 10.1111/mec.13048

Pante, E., Abdelkrim, J., Viricel, A., Gey, D., France, S. C., Boisselier, M. C., & Samadi, S. (2015b). Use of RAD sequencing for delimiting species. *Heredity*, *114*(5), 450–459. doi: 10.1038/hdy.2014.105

Pardo, C., Guillemin, M.-L., Peña, V., Bárbara, I., Valero, M., & Barreiro, R. (2019). Local Coastal Configuration Rather Than Latitudinal Gradient Shape Clonal Diversity and Genetic Structure of *Phymatolithon calcareum* Maerl Beds in North European Atlantic. *Frontiers in Marine Science*, *6*, 149. doi: 10.3389/fmars.2019.00149

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 637–669. doi: 10.1146/annurev.ecolsys.37.091305.110100

Paterno, M., Schiavina, M., Aglieri, G., Ben Souissi, J., Boscari, E., Casagrandi, R., … Melià, P. (2017). Population genomics meet Lagrangian simulations: Oceanographic patterns and long larval duration ensure connectivity among *Paracentrotus lividus* populations in the Adriatic and Ionian seas. *Ecology and Evolution*, *7*(8), 2463–2479. doi: 10.1002/ece3.2844

Peck, J. R., Yearsley, J. M., & Waxman, D. (1998). Explaining the geographic distributions of sexual and asexual populations. *Nature*, *391*(6670), 889–892. doi: 10.1038/36099

Pérez, A. B. J., Fernández, M. T., López-Moya, F., Berbegal, L. A., & López-Llorca, L. V. (2020). Phylogeny of *Cystoseira sensu lato* from SE Spain suggests a complex colonization of the Mediterranean. *Preprint*. doi: 10.20944/preprints202010.0544.v1

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq:

An Inexpensive Method for *De Novo* SNP Discovery and Genotyping in Model and Non-Model Species. *PloS One*, *7*(5). doi: 10.1371/journal.pone.0037135

Peterson, M. L., Doak, D. F., & Morris, W. F. (2019). Incorporating local adaptation into forecasts of species' distribution and abundance under climate change. *Global Change Biology*, *25*(3), 775–793. doi: https://doi.org/10.1111/gcb.14562

Prada, C., & Hellberg, M. E. (2021). Speciation-by-depth on coral reefs: Sympatric divergence with gene flow or cryptic transient isolation? *Journal of Evolutionary Biology*, *34*(1), 128–137. doi: https://doi.org/10.1111/jeb.13731

Pradal, M.-A., & Millet, B. (2006). Hétérogénéité spatiale du forçage du vent: Fonctionnement des récifs artificiels et circulation des eaux dans la baie sud de Marseille. *Comptes Rendus Biologies*, *329*(7), 541–550. doi: 10.1016/j.crvi.2006.04.002

Pratlong, M., Haguenauer, A., Brener, K., Mitta, G., Toulza, E., Garrabou, J., … Aurelle, D. (2018). Separate the wheat from the chaff: Genomic scan for local adaptation in the red coral *Corallium rubrum*. *BioRxiv-PCI Evolutionary Biology*, 306456. doi: 10.1101/306456

Privé, F., Aschard, H., Ziyatdinov, A., & Blum, M. G. B. (2018). Efficient analysis of large-scale genome-wide data with two R packages: Bigstatsr and bigsnpr. *Bioinformatics*, *34*(16), 2781–2787. doi: 10.1093/bioinformatics/bty185

Pudovkin, A. I., Zaykin, D. V., & Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics*, *144*(1), 383–387. doi: 10.1093/genetics/144.1.383

**~Q~**

Qin, H., Yang, G., Provan, J., Liu, J., & Gao, L. (2017). Using MiddRAD-seq data to develop polymorphic microsatellite markers for an endangered yew species. *Plant Diversity*, *39*(5), 294–299. doi: 10.1016/j.pld.2017.05.008

**~R~**

Rana, S., Valentin, K., Riehl, J., Blanfuné, A., Reynes, L., Thibaut, T., … Glöckner, G. (2021). Analysis of organellar genomes in brown algae reveals an independent introduction of similar foreign sequences into the mitochondrial genome. *Genomics*, *113*(2), 646–654. doi: 10.1016/j.ygeno.2021.01.003

Raymond, M., & Rousset, F. (1995). An exact test for population differentiation. *Evolution*, 49(6), 1280–1283. doi: https://doi.org/10.1111/j.1558-5646.1995.tb04456.x

Reed, D. C., Laur, D. R., & Ebeling, A. W. (1988). Variation in algal dispersal and recruitment: The importance of episodic events. *Ecological Monographs*, *58*(4), 321–335. doi: https://doi.org/10.2307/1942543

Reichel, K., Masson, J.-P., Malrieu, F., Arnaud-Haond, S., & Stoeckel, S. (2016). Rare sex or out of reach equilibrium? The dynamics of FIS in partially clonal organisms. *BMC Genetics*, *17*(1), 76. doi: 10.1186/s12863-016-0388-z

Reynes, L., Thibaut, T., Mauger, S., Blanfuné, A., Holon, F., Cruaud, C., … Aurelle, D. (2021). Genomic signatures of clonality in the deep water kelp *Laminaria rodriguezii*. *Molecular Ecology*. doi: https://doi.org/10.1111/mec.15860

Rius, M., Turon, X., Bernardi, G., Volckaert, F. A. M., & Viard, F. (2015). Marine invasion genetics: From spatio-temporal patterns to evolutionary outcomes. *Biological Invasions*, *17*(3), 869–885. doi: 10.1007/s10530-014-0792-0

Rivetti, I., Fraschetti, S., Lionello, P., Zambianchi, E., & Boero, F. (2014). Global warming and mass mortalities of benthic invertebrates in the Mediterranean Sea. *PloS One*, *9*(12), e115655. doi: 10.1371/journal.pone.0115655

Robinet, T., Roussel, V., Cheze, K., & Gagnaire, P.-A. (2020). Spatial gradients of introgressed ancestry reveal cryptic connectivity patterns in a high gene flow marine fish. *Molecular Ecology*, *29*(20), 3857–3871. doi: https://doi.org/10.1111/mec.15611

Robuchon, M., Le Gall, L., Mauger, S., & Valero, M. (2014). Contrasting genetic diversity patterns in two sister kelp species co-distributed along the coast of Brittany, France. *Molecular Ecology*, *23*(11), 2669–2685. doi: 10.1111/mec.12774

Robvieux, P. (2013). Conservation des populations de *Cystoseira* en régions Provence-Alpes-Côte-d'Azur et Corse (Thèse de doctorat, Université Nice Sophia Antipolis). Université Nice Sophia Antipolis. Retrieved from https://tel.archives-ouvertes.fr/tel-00876899

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. doi: 10.1111/mec.15253

Roger, F., Godhe, A., & Gamfeldt, L. (2012). Genetic diversity and ecosystem functioning in the face of multiple stressors. *PloS One*, *7*(9), e45007. doi: 10.1371/journal.pone.0045007

Ross, O. N., Fraysse, M., Pinazo, C., & Pairaud, I. (2016). Impact of an intrusion by the Northern Current on the biogeochemistry in the eastern Gulf of Lion, NW Mediterranean. *Estuarine, Coastal and Shelf Science*, *170*(5), 1–9. doi: 10.1016/j.ecss.2015.12.022

Rousset, F., & Raymond, M. (1995). Testing heterozygote excess and deficiency. *Genetics*, *140*(4), 1413–1419.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219–1228.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

**~S~**

Salamon, M., Lévêque, L., Ballenghien, M., & Viard, F. (2020). Spill-back events followed by self-sustainment explain the fast colonization of a newly built marina by a notorious invasive seaweed. *Biological Invasions*, *22*(4), 1411–1429. doi: 10.1007/s10530-019-02193-5

Sales, M., Cebrian, E., Tomas, F., & Ballesteros, E. (2011). Pollution impacts and recovery potential in three species of the genus *Cystoseira* (Fucales, Heterokontophyta). *Estuarine, Coastal and Shelf Science*, *92*(3), 347–357. doi: 10.1016/j.ecss.2011.01.008

Santelices, B. (1990). Patterns of reproduction, dispersal and recruitment in seaweeds. *Oceanography and Marine Biology*, *28*, 177–276.

Schoenrock, K. M., Connor, A. M. O., Mauger, S., Valero, M., Neiva, J., Serrão, E. Á., & Krueger-Hadfield, S. A. (2020). Genetic diversity of a marine foundation species, *Laminaria hyperborea* (Gunnerus) Foslie, along the coast of Ireland. *European Journal of Phycology*, *55*(3), 310–326. doi: 10.1080/09670262.2020.1724338

Schwander, T., Henry, L., & Crespi, B. J. (2011). Molecular Evidence for Ancient Asexuality in *Timema* Stick Insects. *Current Biology*, *21*(13), 1129–1134. doi: 10.1016/j.cub.2011.05.026

Schwander, T. (2016). Evolution: The end of an ancient asexual scandal. *Current Biology*, *26*(6), 233–235. doi: 10.1016/j.cub.2016.01.034

Sekhon, J. S. (2008). Multivariate and propensity score matching software with automated balance optimization: The Matching package for R (SSRN Scholarly Paper No. ID 1009044). Rochester, NY: Social Science Research Network.

Serrão, E. A., Pearson, G., Kautsky, L., & Brawley, S. H. (1996). Successful external fertilization in turbulent environments. *Proceedings of the National Academy of Sciences*, *93*(11), 5286–5290. doi: 10.1073/pnas.93.11.5286

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907–917. doi: https://doi.org/10.1111/2041-210X.12700

Shrestha, S., Hu, J., Fryxell, R. T., Mudge, J., & Lamour, K. (2014). SNP markers identify widely distributed clonal lineages of *Phytophthora colocasiae* in Vietnam, Hawaii and Hainan Island, China. *Mycologia*, 106(4), 676–685. doi: 10.3852/13-165

Silberfeld, T., Leigh, J. W., Verbruggen, H., Cruaud, C., de Reviers, B., & Rousseau, F. (2010). A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the "brown algal crown radiation." *Molecular Phylogenetics and Evolution*, *56*(2), 659–674. doi: 10.1016/j.ympev.2010.04.020

Sivertsen, K. (1997). Geographic and environmental factors affecting the distribution of kelp beds and barren grounds and changes in biota associated with kelp reduction at sites along the Norwegian coast. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(12). (Ottawa, Canada). doi: 10.1139/f97-186

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, *16*(1), 393–430.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792. doi: 10.1126/science.3576198

Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, *23*(1), 23–35. doi: 10.1017/S0016672300014634

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., & Neale, D. B. (2013). Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, *9*(4), 901–911. doi: 10.1007/s11295-013-0596-x

Steneck, R. S., Graham, M. H., Bourque, B. J., Corbett, D., Erlandson, J. M., Estes, J. A., & Tegner, M. J. (2002). Kelp forest ecosystems: Biodiversity, stability, resilience and future. *Environmental Conservation*, *29*(4), 436–459. doi: 10.1017/S0376892902000322

Stoeckel, S. (2017). ClonEstiMate, a Bayesian method for quantifying rates of clonality of populations genotyped at two-time steps. *Molecular Ecology Resources*, *17*(6), e251–e267. doi: 10.1111/1755-0998.12698

Stoeckel, S., Grange, J., Fernández-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species—*Prunus avium*. *Molecular Ecology*, *15*(8), 2109–2118. doi: https://doi.org/10.1111/j.1365-294X.2006.02926.x

Stoeckel, S., & Masson, J.-P. (2014). The exact distributions of FIS under partial asexuality in small finite populations with mutation. *PloS One*, *9*(1), e85228. doi: 10.1371/journal.pone.0085228

Stoeckel, S., Porro, B., & Arnaud-Haond, S. (2020). The discernible and hidden effects of clonality on the genotypic and genetic states of populations: Improving our estimation of clonal rates.

*Molecular Ecology Resources*. doi: https://doi.org/10.1111/1755-0998.13316

Susini, M.-L. (2006). Statut et biologie de *Cystoseira amentacea var. Stricta*. (Thèse de doctorat, Université Nice Sophia Antipolis). Université Nice Sophia Antipolis.

Susini, M.-L., Thibaut, T., Meinesz, A., & Forcioli, D. (2007). A preliminary study of genetic diversity in *Cystoseira amentacea* (C. Agardh) Bory *var. Stricta* Montagne (Fucales, Phaeophyceae) using random amplified polymorphic DNA. *Phycologia*, *46*(6), 605–611. doi: 10.2216/06-100.1

**~T~**

Tallmon, D. A., Koyuk, A., Luikart, G., & Beaumont, M. A. (2008). COMPUTER PROGRAMS: Onesamp: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, *8*(2), 299–301. doi: 10.1111/j.1471-8286.2007.01997.x

Tallmon, D. A., Gregovich, D., Waples, R. S., Baker, C. S., Jackson, J., Taylor, B. L., … Schwartz, M. K. (2010). When are genetic methods useful for estimating contemporary abundance and detecting population trends? *Molecular Ecology Resources*, *10*(4), 684–692. doi: https://doi.org/10.1111/j.1755-0998.2010.02831.x

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., & Prins, P. (2015). Sambamba: Fast processing of NGS alignment formats. *Bioinformatics*, *31*(12), 2032–2034. doi: 10.1093/bioinformatics/btv098

Tatarenkov, A., Bergström, L., Jönsson, R. B., Serrão, E. A., Kautsky, L., & Johannesson, K. (2005). Intriguing asexual life in marginal populations of the brown seaweed *Fucus vesiculosus*. *Molecular Ecology*, *14*(2), 647–651. doi: https://doi.org/10.1111/j.1365-294X.2005.02425.x

Tatarenkov, A., Jönsson, R. B., Kautsky, L., & Johannesson, K. (2007). Genetic structure in populations of *Fucus vesiculosus* (Phaeophyceae) over spatial scales from 10 m to 800 km 1. *Journal of Phycology*, *43*(4), 675–685. doi: https://doi.org/10.1111/j.1529-8817.2007.00369.x

Teixeira, S., Pearson, G. A., Candeias, R., Madeira, C., Valero, M., & Serrão, E. A. (2016). Lack of fine-scale genetic structure and distant mating in natural populations of *Fucus vesiculosus*. *Marine Ecology Progress Series*, *544*, 131–142. doi: 10.3354/meps11607

Teixeira, J. C., & Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences*, *118*(10). doi: 10.1073/pnas.2015096118

Tellier, F., Meynard, A. P., Correa, J. A., Faugeron, S., & Valero, M. (2009). Phylogeographic analyses of the 30°S south-east Pacific biogeographic transition zone establish the occurrence of a sharp genetic discontinuity in the kelp *Lessonia nigrescens*: Vicariance or parapatry? *Molecular Phylogenetics and Evolution*, *53*(3), 679–693. doi: 10.1016/j.ympev.2009.07.030

Tellier, F., Tapia, J., Faugeron, S., Destombe, C., & Valero, M. (2011). The *Lessonia nigrescens* species complex (laminariales, phaeophyceae) shows strict parapatry and complete reproductive isolation in a secondary contact zone 1. *Journal of Phycology*, *47*(4), 894–903. doi: https://doi.org/10.1111/j.1529-8817.2011.01019.x

Terhorst, J., Kamm, J. A., & Song, Y. S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, *49*(2), 303–309. doi: 10.1038/ng.3748

Thibaut, T., Pinedo, S., Torras, X., & Ballesteros, E. (2005). Long-term decline of the populations of Fucales (*Cystoseira* spp. and *Sargassum* spp.) in the Albères coast (France, North-western Mediterranean). *Marine Pollution Bulletin*, *50*(12), 1472–1489. doi: 10.1016/j.marpolbul.2005.06.014

Thibaut, T., Blanfuné, A., Markovic, L., Verlaque, M., Boudouresque, C. F., Perret-Boudouresque, M., … Bottin, L. (2014). Unexpected abundance and long-term relative stability of the brown alga *Cystoseira amentacea*, hitherto regarded as a threatened species, in the north-western Mediterranean Sea. *Marine Pollution Bulletin*, *89*(1), 305–323. doi: 10.1016/j.marpolbul.2014.09.043

Thibaut, T., Blanfuné, A., Boudouresque, C.-F., & Verlaque, M. (2015). Decline and local extinction of Fucales in French Riviera: The harbinger of future extinctions? *Mediterranean Marine Science*, *16*(1), 206–224. doi: 10.12681/mms.1032

Thibaut, T., Blanfuné, A., Boudouresque, C.-F., Cottalorda, J.-M., Hereu, B., Susini, M.-L., & Verlaque, M. (2016a). Unexpected temporal stability of *Cystoseira* and *Sargassum* forests in Port-Cros, one of the oldest Mediterranean marine National Parks. *Cryptogamie, Algologie*, *37*(1), 61–90. doi: 10.7872/crya/v37.iss1.2016.61

Thibaut, T., Bottin, L., Aurelle, D., Boudouresque, C.-F., Blanfuné, A., Verlaque, M., … Millet, B. (2016b). Connectivity of populations of the seaweed *Cystoseira amentacea* within the Bay of Marseille (Mediterranean Sea): Genetic structure and hydrodynamic connections. *Cryptogamie, Algologie*, *37*(4), 233–255. doi: 10.7872/crya/v37.iss4.2016.233

Thiel, M., & Gutow, L. (2005). The ecology of rafting in the marine environment. I. The floating substrata. In *Oceanography and Marine Biology: An annual review* (Vol. 42, pp. 181–264).

Tigano, A. (2020). A population genomics approach to uncover the CNVs, and their evolutionary significance, hidden in reduced-representation sequencing data sets. *Molecular Ecology*, *29*(24), 4749–4753. doi: https://doi.org/10.1111/mec.15665

Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R., & Bird, C. E. (2013). ezRAD: A simplified method for genomic genotyping in non-model organisms. *PeerJ*, *1*, e203. doi: 10.7717/peerj.203

**~V~**

Valero, M., Destombe, C., Mauger, S., Ribout, C., Engel, C. R., Daguin-Thiebaut, C., & Tellier, F. (2011). Using genetic tools for sustainable management of kelps: A literature review and the example of *Laminaria digitata*. *Cahiers de Biologie Marine*, *54*(4), 467.

Vendrami, D. L. J., De Noia, M., Telesca, L., Handal, W., Charrier, G., Boudry, P., … Hoffman, J. I. (2019). RAD sequencing sheds new light on the genetic structure and local adaptation of European scallops and resolves their demographic histories. *Scientific Reports*, *9*(1), 1–13. doi: 10.1038/s41598-019-43939-4

Verdu, C. F., Guichoux, E., Quevauvillers, S., De Thier, O., Laizet, Y., Delcamp, A., … Mariette, S. (2016). Dealing with paralogy in RADseq data: in silico detection and single nucleotide polymorphism validation in *Robinia pseudoacacia*. *Ecology and Evolution*, *6*(20), 7323–7333. doi: 10.1002/ece3.2466

Verlaque, M. (1987). Contributions à l'étude du phytobenthos d'un écosystème photophile thermophile marin en Méditerranée occidentale: Étude structurale et dynamique du phytobenthos et analyse des relations faune - flore (These de doctorat, Aix-Marseille 2). Aix-Marseille 2. Retrieved from https://www.theses.fr/1987AIX22052

Verlaque, M., Boudouresque, C.-F., & Perret-Boudouresque, M. (2019). Mediterranean seaweeds listed as threatened under the Barcelona Convention: A critical analysis. *Scientific Reports. Port-Cros Natl. Park, 33*, 179–214.

Viricel, A., Pante, E., Dabin, W., & Simon-Bouhet, B. (2014). Applicability of RAD-tag genotyping for

interfamilial comparisons: Empirical data from two cetaceans. *Molecular Ecology Resources*, *14*(3), 597–605. doi: https://doi.org/10.1111/1755-0998.12206

Von der Heyden, S. (2017). Making evolutionary history count: Biodiversity planning for coral reef fishes and the conservation of evolutionary processes. *Coral Reefs*, *36*(1), 183–194. doi: 10.1007/s00338-016-1512-2

**~W~**

Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., … Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, *22*(3), 787–798. doi: https://doi.org/10.1111/mec.12023

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., … Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, *416*(6879), 389–395. doi: 10.1038/416389a

Wang, J. (2009). A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, *18*(10), 2148–2164. doi: https://doi.org/10.1111/j.1365-294X.2009.04175.x

Wang, S., Meyer, E., McKay, J. K., & Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods*, *9*(8), 808–810. doi: 10.1038/nmeth.2023

Waples, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, *89*(5), 438–450. doi: 10.1093/jhered/89.5.438

Waples, R. S., & Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*(4), 753–756. doi: https://doi.org/10.1111/j.1755-0998.2007.02061.x

Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, *3*(3), 244–262. doi: https://doi.org/10.1111/j.1752-4571.2009.00104.x

Waples, R. S., Luikart, G., Faulkner, J. R., & Tallmon, D. A. (2013). Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1768), 20131339. doi: 10.1098/rspb.2013.1339

Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, *197*(2), 769–780. doi: 10.1534/genetics.114.164822

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x

Weir, W., Capewell, P., Foth, B., Clucas, C., Pountain, A., Steketee, P., … MacLeod, A. (2016). Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *ELife*, *5*, e11473. doi: 10.7554/eLife.11473

Wernberg, T., & Filbee-Dexter, K. (2019). Missing the marine forest for the trees. *Marine Ecology Progress Series*, *612*, 209–215. doi: 10.3354/meps12867

Wernberg, T., Krumhansl, K., Filbee-Dexter, K., & Pedersen, M. F. (2019). Status and trends for the world's kelp forests. In *World seas: An environmental evaluation* (Academic Press, pp. 57–78). Elsevier. doi: 10.1016/B978-0-12-805052-1.00003-6

White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents

help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1688), 1685–1694. doi: 10.1098/rspb.2009.2214

Wilson, D. J. (2019). The harmonic mean *p* -value for combining dependent tests. *Proceedings of the National Academy of Sciences*, *116*(4), 1195–1200. doi: 10.1073/pnas.1814092116

Wright, D., Bishop, J. M., Matthee, C. A., & Heyden, S. von der. (2015). Genetic isolation by distance reveals restricted dispersal across a range of life histories: Implications for biodiversity conservation planning across highly variable marine environments. *Diversity and Distributions*, *21*(6), 698–710. doi: https://doi.org/10.1111/ddi.12302

**~X~**

Xuereb, A., Benestan, L., Normandeau, É., Daigle, R. M., Curtis, J. M. R., Bernatchez, L., & Fortin, M.-J. (2018a). Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Molecular Ecology*, *27*(10), 2347–2364. doi: 10.1111/mec.14589

Xuereb, A., Kimber, C. M., Curtis, J. M. R., Bernatchez, L., & Fortin, M.-J. (2018b). Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data. *Molecular Ecology*, *27*(24), 5035–5048. doi: https://doi.org/10.1111/mec.14942

Xuereb, A., D'Aloia, C. C., Andrello, M., Bernatchez, L., & Fortin, M.-J. (2020). Incorporating putatively neutral and adaptive genomic data into marine conservation planning. *Conservation Biology*. doi: https://doi.org/10.1111/cobi.13609

**~Z~**

Zardi, G. I., Nicastro, K. R., Canovas, F., Costa, J. F., Serrão, E. A., & Pearson, G. A. (2011). Adaptive traits are maintained on steep selective gradients despite gene flow and hybridization in the intertidal zone. *PloS One*, *6*(6), e19402. doi: 10.1371/journal.pone.0019402

Zhdanova, O. L., & Pudovkin, A. I. (2008). Nb_HetEx: A program to estimate the effective number of breeders. *Journal of Heredity*, *99*(6), 694–695. doi: 10.1093/jhered/esn061

Zhou, X., Sun, F., Xu, S., Fan, G., Zhu, K., Liu, X., … Yang, G. (2013). Baiji genomes reveal low genetic variability and new insights into secondary aquatic adaptations. *Nature Communications*, *4*(1), 1–6. doi: 10.1038/ncomms3708

Žuljević, A., Antolić, B., Nikolić, V., & Isajlović, I. (2011). Review of *Laminaria rodriguezii* records in the Adriatic Sea. *Fifth European Phycological Congress*, 194.

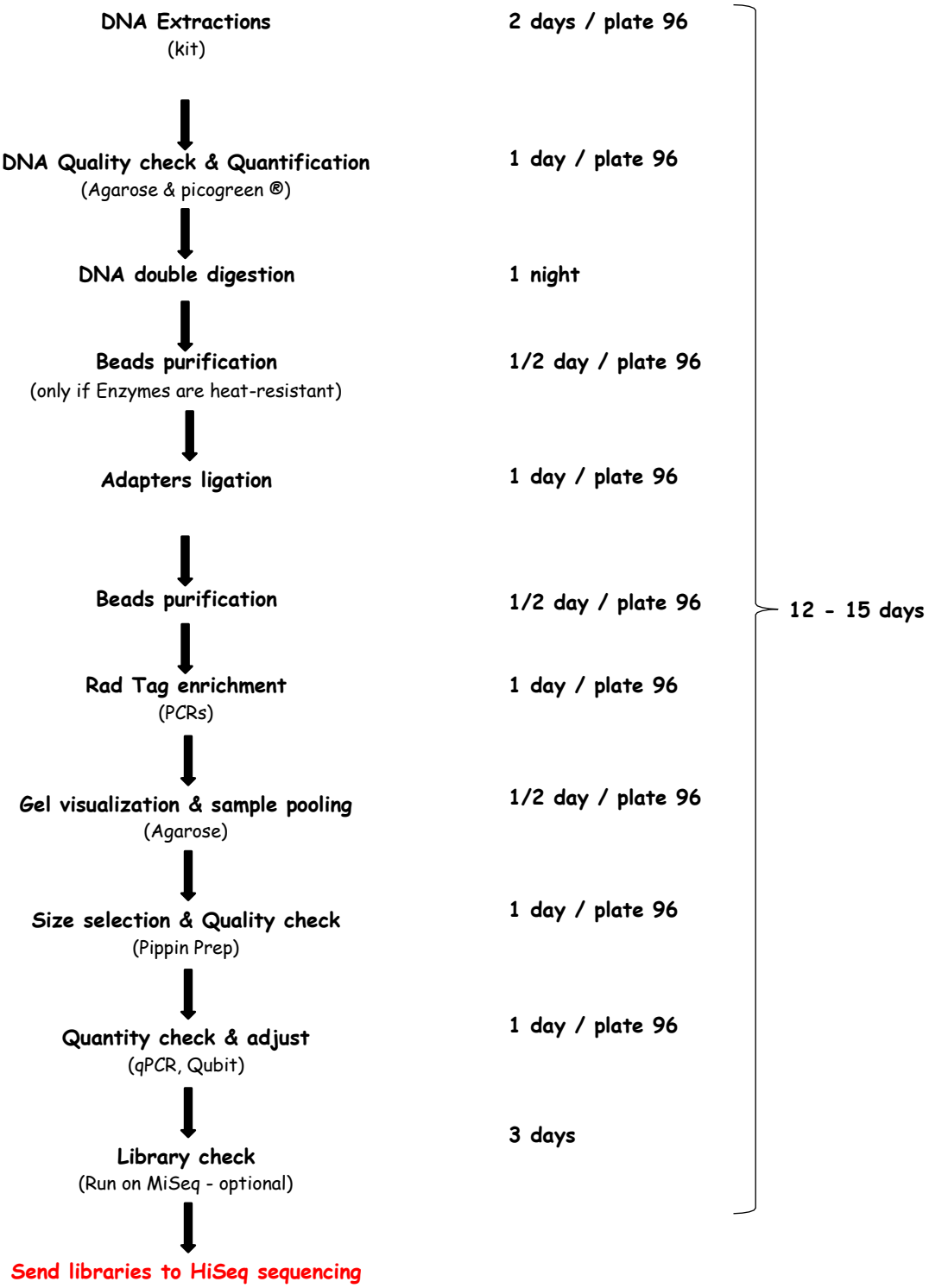
Žuljević, A., Peters, A. F., Nikolić, V., Antolić, B., Despalatović, M., Cvitković, I., … Küpper, F. C. (2016). The Mediterranean deep-water kelp *Laminaria rodriguezii* is an endangered species in the Adriatic Sea. *Marine Biology*, *163*(4), 69. doi: 10.1007/s00227-016-2821-2

# Annexes

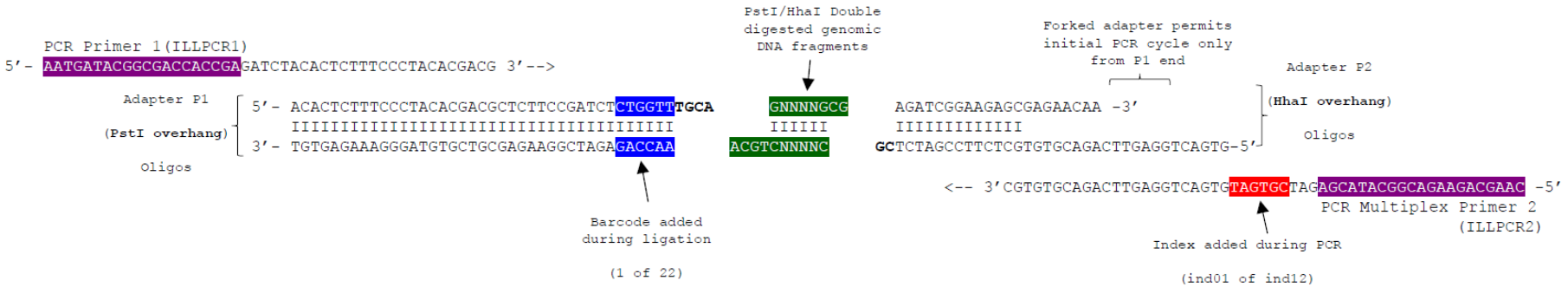
# Annexe A | chapitre 1

**Matériel A.1**. Protocole de séquençage ddRAD-seq (double-digest) employé pour génotyper les populations du Fucales *Ericaria zosteroides* (chapitre 2) et des kelps *Laminaria rodriguezii* et *Laminaria digitata* (chapitre 3). D'après Peterson et al. (2012), adapté et rédigé par Stéphane Mauger (Station Biologique de Roscoff, IRL 3617 biologie évolutive et écologie des algues).

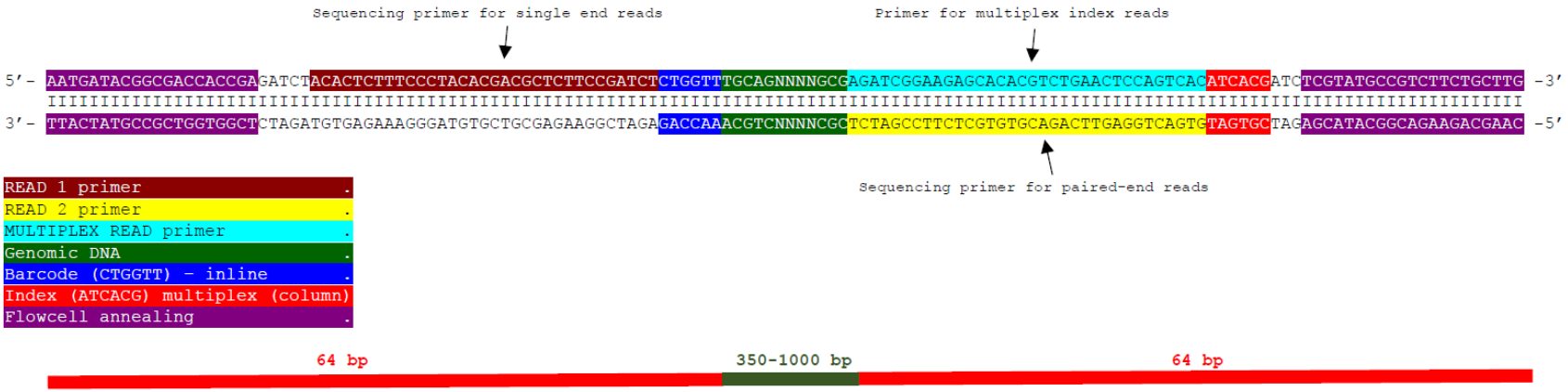
## Overview


DNA Extractions
(kit)
2 days / plate 96
DNA Quality check & Quantification
(Agarose & picogreen ®)
1 day / plate 96
DNA double digestion
1 night
Beads purification
(only if Enzymes are heat-resistant)
1/2 day / plate 96
Adapters ligation
1 day / plate 96
Beads purification
1/2 day / plate 96
Rad Tag enrichment
(PCRs)
1 day / plate 96
Gel visualization & sample pooling
(Agarose)
1/2 day / plate 96
Size selection & Quality check
(Pippin Prep)
1 day / plate 96
Quantity check & adjust
(qPCR, Qubit)
1 day / plate 96
Library check
(Run on MiSeq - optional)
3 days
12 - 15 days
Send libraries to HiSeq sequencing

**Oligos, Adapters, Digested gnomic DNA**




**Final sequencing library**




**BRIEF GLOSSARY**

Adapter: fully or partially double-stranded product of annealing two oligos. Adapters are ligated to genomic DNA at restriction enzyme cut sites in order to add barcodes and common PCR priming sequences.

Barcode: short DNA sequence downstream of the sequencing primer annealing region of an adapter. Used to resolve products of different ligation reactions (usually separate individuals) after sequencing pooled libraries.

Fragment: section of genomic DNA resulting from restriction enzyme cleavage.

Index: short DNA sequence introduced during PCR amplification of the final library that uniquely identifies products of that PCR reaction. Used combinatorially with Adapter P1 barcodes to resolve multiplexed sample pools.

Library: a collection of sequencing-competent fragments

## DNA extraction and quantification

### I.1- DNA extraction

DNA was extracted using Nucleospin 96 Plant II® extraction kit from Macherey Nagel.

Between 10 and 20 mg of dried tissue conserved in silica gel was used.

The samples were milled using beads with a ball mill Retsch M300

Cell lysis was carried out by adding 500µl of PL1 buffer and 10 µl of RNase A, mixing and remaining at room temperature for 4 hours (instead of heating according to manufacturer protocol)

From there, the manufacturer protocol was followed and genomic DNAs were eluted in 120 µl of elution buffer (2 elution steps with 60µl)

### I.2- Ethanol precipitation (optional)

**Reagents Needed:**

1 3 M sodium acetate pH 5.2
2 DNA
3 100% ethanol

**Protocol**

1 Measure the volume of the DNA sample.
2 Add 1/10 volume of sodium acetate, pH 5.2, (final concentration of 0.3 M)
- These amounts assume that the DNA is in TE only; if DNA is in a solution containing salt, adjust salt accordingly to achieve the correct final concentration.
3 Mix well.
4 Add 2 to 2.5 volumes of cold 100% ethanol (calculated after salt addition).
5 Mix well.
6 Place on ice or at -20 °C for Overnight.
7 Spin 15 min 11 000 RPM 4°C in a microfuge.
8 Carefully decant supernatant.
9 Add 1 ml 70% ethanol. Mix. Spin 3 min 11 000 RPM 4°C. Carefully decant supernatant.
10 Air dry or briefly vacuum dry pellet.
11 Resuspend pellet in the appropriate volume (40µl) of TE or water.

### I.3- DNA Quality check on gel and quantification

5µl of genomic DNA was deposited on a 1.5% agarose gel to check for degradation DNA was quantified using **PicoGreen®** dsDNA quantitation assay (Invitrogen) or Quanti-Fluor®dsDNA system (Promega)

#### I.3.1-Principe

The Picogreen® reagent is a fluorescent dye ultrasensitive nucleic acid to quantify dsDNA in solution Measuring the concentration of nucleic acid is by measuring absorbance at □ = 260 nm (absorbance at 260 nm = 0.1 corresponds to 5µg / ml DNA). There is no difference in the extent between the DNA and RNA. Using a plate reader in fluorescence can be measured from 250pg / ml DNA.

#### I.3.2- Equipment

H2O distilled DNase/RNase free
Kit Quant-it Picogreen® Micro-plate UV

#### I.3.3-Preparation for ONE plate (in chemical hood)

Put aluminum foil on the bench (he used to pack the plate)

1 Preparation of buffer TE 1X

Dilute 1:20 the bufferTE 20x included in the kit.
10,450ml d'$H_20$ DNase/RNase free + 550µl TE 20X □ 11ml TE 1X (store at 4°C). Prepare an aliquot D-day

**2** Preparation of DNA solution at 5000pg/µl **for ONE range**

Dilute 1:20 the DNA solution at 100ng/µl included in the kit. 28,5µl TE 1X + 1,5µl DNA solution (100ng/µl) → 30µl at 5000pg/µl

3 Preparation of the standard range 0pg/µl to 1000pg/µl

Dilution in TE 1X
Vortex and centrifuge each tube

| [Final] | [Stock] | Vol Stock | Vol TE 1X | Volsolfinale |
|---|---|---|---|---|
| **pg/µl** | pg/µl | µl | µl | µl |
| **1000** | 5000 | 14 | 56 | 70 |
| **500** | 5000 | 7 | 63 | 70 |
| **250** | 5000 | 3 | 57 | 60 |
| **100** | 1000 | 7 | 63 | 70 |
| **50** | 500 | 7 | 63 | 70 |
| **10** | 100 | 6 | 54 | 60 |
| **5** | 50 | 6 | 54 | 60 |
| **0** | 0 | 0 | 60 | 60 |

The reference range is usable for about 7 days, after the DNA is degraded.

4 Preparation of picogreen® 1X (away from light) Dilute 1:200 the picogreen® 200X solution included in the kit.

4,975ml TE 1X + 25µl picogreen® 200X → 5 ml picogreen® 1X (or 2,4875ml TE 1X + 12,5µl picogreen® 200X → 2,5 ml picogreen® 1X) Mix up and down

5 Distribution in the plate picotitrée ½ volume

The plate's picotitrée UV reading are stored in the black cabinet sequencing platform.
Be careful never to touch the bottom of the plate (if proof, read error with the plate reader).
For the standard range, distribute 50µl of each finale concentration (or 25µl of each finale concentration)

For samples to be tester, distribute:
-49µl of TE 1X (or 24µl of TE 1X)
- 1µl of sample

Vortex, centrifuge and distribute 50µl of Picogreen® 1X in all wells (or 25µl of Picogreen® 1X in all wells). Mixer and turn into the well without touching the bottom of the plate and refit the tip along the well to remove the last drop without bubble.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 1000pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| B | 500 pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| C | 250 pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| D | 100 pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| E | 50 pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| F | 15 pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| G | 5 pg/µl | test | test | test | test | test | test | test | test | test | test | test |
| H | 0 pg/µl | test | test | test | test | test | test | test | test | test | test | test |

Close the foil on the UV plate (away from light) and incubate at room temperature for 20 minutes. Read the plate (on Safire)

## Library construction

### II.1- Double digestion

#### II.1.1-DNA normalization (Optional)

Genomic DNAs were normalized at a concentration of 2,5ng/µl with nuclease-free water

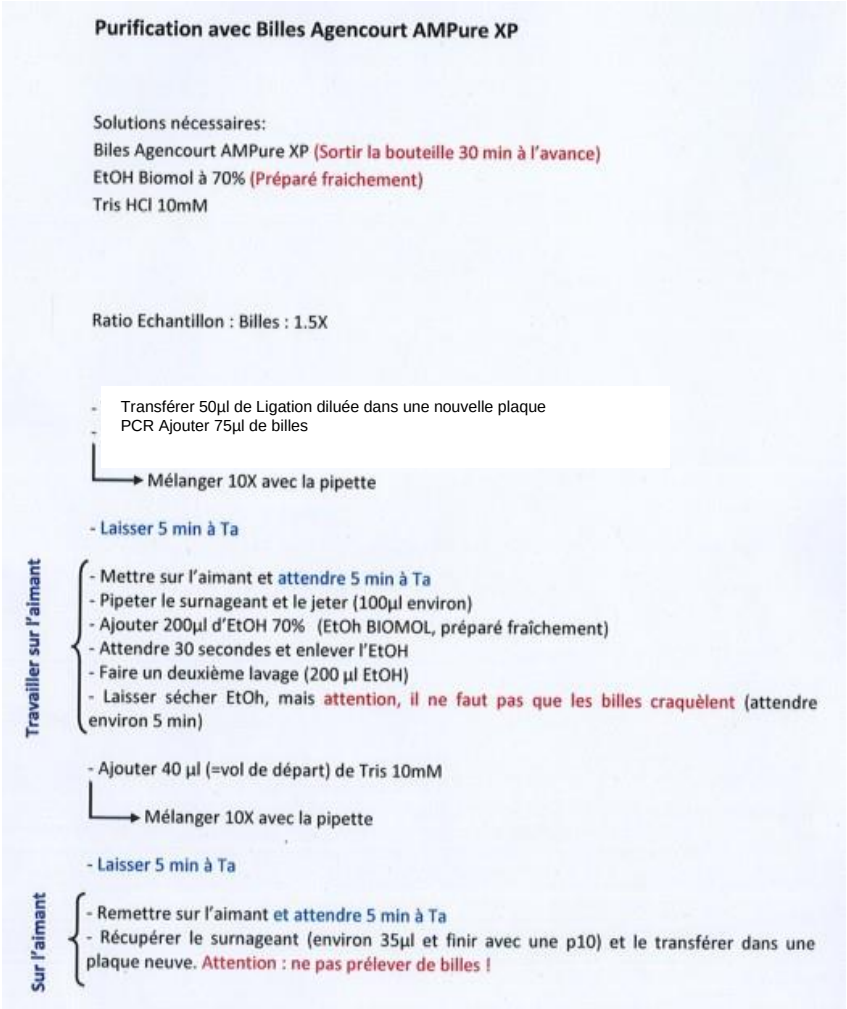
#### II.1.2-Digestion protocol

Consult product information for your chosen restriction enzymes regarding digest conditions.

Double digest 100-1000 ng of high quality genomic DNA with selected restriction enzymes, in a 20-50ul reaction volume. Use a digestion buffer appropriate for both enzymes. (Note: This protocol is expected to work with as little as 100ng of genomic DNA, though we typically digest between 200-500 ng of DNA.)

*PstI x HhaI* digestion in a 96 microplate at **37°C for 3 to 10h.**
0.5 µl *PstI* (20U/µl)
0.5 µl *HhaI* (20U/µl)
5 µl CutSmart buffer 10X
40µl DNA (100ng of genomic DNA per sample minimum) 4 µl $H_2O$
50 µl total volume

Run the digestion as appropriate for the chosen REs. To ensure complete digestion, we run our double digests for minimum 3 hours at 37°C, holding at 4°C.
Check on the Biolabs® website if the enzymes can be inactivate by heat. YES ? Heat
Yes -> 65°C for 20 min
No -> Do a purification using AMPure XP beads :



## II.2- P1 and P2 adapters ligation

Ligation efficiency depends on the total number of adapters and the number of fragment ends in a digested sample. We find that a 2- to 10- fold excess of adapters to complementary sticky ends produces efficient ligations. To assist in calculation of dilutions from annealed stocks ("Anneal Adapters" above) and ligation conditions, it is useful to save a copy (File->Make a copy) of our ligation molarity calculator in your own google docs account, and edit the fields shaded in green to suit your experiment depending on whether you are estimating adapter ligation concentrations from an in silico experiment (first worksheet) or from single/double enzyme cuts run on a Bioanalyzer (second worksheet). Note that annealed adapters in the molarity calculator are set to dummy values of 4uM (4 pmol/ul) concentration corresponding to the 10-fold dilution of the 40uM annealed adapter stock solution from the annealing reaction described above. If a different annealed adapter stock concentration is used, edit values in row
16 of this spreadsheet accordingly. The ligation conditions for a 40ul reaction are listed below; however, larger ligation reaction volumes can be used if needed and the reagents should be modified accordingly.

| | | | Input appropriate values in green shaded fields |
|---|---|---|---|
| | | | |
| Initial DNA mass (ug) | 0,3 | | mass of double-digested genomic DNA in ligation |
| | P1 | P2 | |
| cut frequency (bp) | 2600 | 250 | average distance between sites for each enzyme (from genome |
| fragment mass (g/mole) | 1716000 | 165000 | |
| sample mass (g) | 3,00E-07 | 3,00E-07 | |
| fragments/sample (moles) | 1,75E-13 | 1,82E-12 | |
| ends/sample (moles) | 3,50E-13 | 3,64E-12 | |
| ends/sample (pmoles) | 0,35 | 3,64 | |
| | | | |
| target adapter fold excess | 6 | 6 | molar excess of adapter to genomic DNA ends |
| target adapter/sample (pmoles) | 2,10 | 21,82 | |
| target adapter volume/rxn (ul) | 5 | 5 | desired amount of working adapter stock to add to ligation rxn |
| adapter working conc (pmol/ul) | 0,42 | 4,36 | |
| annealed adapter conc (pmol/ul) | 4 | 4 | concentration of annealed adapter stock |
| fold dilution to working stock | 9,53 | 0,92 | |
| volume of working stock to make (ul) | 75, | 1 000, | desired volume of working adapter stock |
| **annealed adapter stock (ul)** | **7,87** | **1 090,91** | |
| **1x annealing buffer (ul)** | **67,13** | **-90,91** | |
| | | | |
| **working stock to add to ligation (ul)** | **5** | **5** | |

## II.2.1-Adapters preparation Regents Needed:

**Annealing buffer stock (10X):**

| | |
|---|---|
| 100 mM Tris HCl, pH 8 | 5 mL tris 1M |
| 500 mM NaCl | 5 mL NaCl 5M |
| 10 mM EDTA | 1 mL EDTA 0.5M |
| H2O | 39 mL |
| | 50 mL of Annealing Buffer |

Single-stranded oligos need to be annealed with their appropriate partner before ligation. We provide sequences for 48 uniquely barcoded adapter P1 oligo pairs (oligos P1_PstI_x.F and P1_PstI_x.R) and common adapter P2 (oligos P2_HhaI.F and P2_HhaI.R), see ddRAD oligo table.
To create Adapter P1 and P2, combine each oligo Forward with its complementary oligo Reverse in a 1:1 ratio in working strength annealing buffer (final buffer concentration 1x) for a total annealed adapter concentration of 40µM (for example, if purchased oligos are resuspended to an initial concentration of 100µM, use 40µl oligo Forward, 40µl oligo Reverse, 10µl 10x annealing buffer and 10µl nuclease-free water).

Double stranded P1 (PstI) barcoded adapters preparation (40µM) – (1 to 22 tubes)
40µl P1_PstI_x.F (100µM) + 40µl P1_PstI_x.R (100µM) + 10µl Annealing buffer + 10µl $H_2O$

Double stranded P2 (HhaI) no-barcoded adapter preparation (40µM) – (1 tube)
40µl P2_HhaI.F (100µM) + 40µl P2_HhaI.R (100µM) + 10µl Annealing buffer + 10µl $H_2O$

In a thermocycler : 97.5°C - 2.5 min
96°C – 1 min x 25 (-3°C per cycle)
4°C infinity

**II.2.2- Ligation protocol**

Dilute 1:10 all the adapters P1 (PstI) with annealing buffer (40µM to 4µM). Combine the appropriate amounts of working stock dilution of adapters P1 (5µl), and digested DNA (35µl).

**One P1 adapter per line**.

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P1_PstI_02→ | A | dna-01 | dna-02 | dna-03 | dna-04 | dna-05 | dna-06 | dna-07 | dna-08 | dna-09 | dna-10 | dna-11 | dna-12 |
| P1_PstI_03→ | B | dna-13 | dna-14 | dna-15 | dna-16 | dna-17 | dna-18 | dna-19 | dna-20 | dna-21 | dna-22 | dna-23 | dna-24 |
| P1_PstI_04→ | C | dna-25 | dna-26 | dna-27 | dna-28 | dna-29 | dna-30 | dna-31 | dna-32 | dna-33 | dna-34 | dna-35 | dna-36 |
| P1_PstI_05→ | D | dna-37 | dna-38 | dna-39 | dna-40 | dna-41 | dna-42 | dna-43 | dna-44 | dna-45 | dna-46 | dna-47 | dna-48 |
| P1_PstI_06→ | E | dna-49 | dna-50 | dna-51 | dna-52 | dna-53 | dna-54 | dna-55 | dna-56 | dna-57 | dna-58 | dna-59 | dna-60 |
| P1_PstI_07→ | F | dna-61 | dna-62 | dna-63 | dna-64 | dna-65 | dna-66 | dna-67 | dna-68 | dna-69 | dna-70 | dna-71 | dna-72 |
| P1_PstI_08→ | G | dna-73 | dna-74 | dna-75 | dna-76 | dna-77 | dna-78 | dna-79 | dna-80 | dna-81 | dna-82 | dna-83 | dna-84 |
| P1_PstI_09→ | H | dna-85 | dna-86 | dna-87 | dna-88 | dna-89 | dna-90 | dna-91 | dna-92 | dna-93 | dna-94 | dna-95 | dna-96 |

Separately, create a master mix of T4 DNA ligase for 1 plate (100 reactions):

50 µL - P2 adapter (HhaI 40µM)
600µL – buffer T4 ligase 10X
40µL – T4 ligase (400U/µl)
1310µL – nuclease-free water

2000µL

Distribute 20µL of mix per well and incubate at room temperature (minimum 16°C) for 6h minimum.

## II.3-Purification using AMPure XP beads

Clean the double digest with AMPure XP beads following the provided protocol, using a magnetic plate (SPRIPlate Super Magnet Plate) or rare earth magnets and a tube rack (see photo) to separate beads from the solution. (Note: The Ampure XP protocol recommends using a volume of beads equal to 1.8X the volume of the solution being cleaned; we have had good success using 1.5X volume of beads, and recommend this ratio for all steps prior to PCR amplification).

## II.4- Rad Tag enrichment

Primer mix preparation (One tube per Index):

5µl IllPCR_1 (100µM) + 5µl IllPCR_2_indexed (100µM) + 90µl nuclease-free water.

**One IllPCR_2_index per column**.

| | | IIIPCR2 Ind01 | IIIPCR2 Ind02 | IIIPCR2 Ind03 | IIIPCR2 Ind04 | IIIPCR2 Ind05 | IIIPCR2 Ind06 | IIIPCR2 Ind07 | IIIPCR2 Ind08 | IIIPCR2 Ind09 | IIIPCR2 Ind10 | IIIPCR2 Ind11 | IIIPCR2 Ind12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| P1_PstI_02→ | A | dna-01 | dna-02 | dna-03 | dna-04 | dna-05 | dna-06 | dna-07 | dna-08 | dna-09 | dna-10 | dna-11 | dna-12 |
| P1_PstI_03→ | B | dna-13 | dna-14 | dna-15 | dna-16 | dna-17 | dna-18 | dna-19 | dna-20 | dna-21 | dna-22 | dna-23 | dna-24 |
| P1_PstI_04→ | C | dna-25 | dna-26 | dna-27 | dna-28 | dna-29 | dna-30 | dna-31 | dna-32 | dna-33 | dna-34 | dna-35 | dna-36 |
| P1_PstI_05→ | D | dna-37 | dna-38 | dna-39 | dna-40 | dna-41 | dna-42 | dna-43 | dna-44 | dna-45 | dna-46 | dna-47 | dna-48 |
| P1_PstI_06→ | E | dna-49 | dna-50 | dna-51 | dna-52 | dna-53 | dna-54 | dna-55 | dna-56 | dna-57 | dna-58 | dna-59 | dna-60 |
| P1_PstI_07→ | F | dna-61 | dna-62 | dna-63 | dna-64 | dna-65 | dna-66 | dna-67 | dna-68 | dna-69 | dna-70 | dna-71 | dna-72 |
| P1_PstI_08→ | G | dna-73 | dna-74 | dna-75 | dna-76 | dna-77 | dna-78 | dna-79 | dna-80 | dna-81 | dna-82 | dna-83 | dna-84 |
| P1_PstI_09→ | H | dna-85 | dna-86 | dna-87 | dna-88 | dna-89 | dna-90 | dna-91 | dna-92 | dna-93 | dna-94 | dna-95 | dna-96 |

**PCR mix** (one mix per Index for one plate 96)

| Product | Volume (µl) for 1 PCR | PCR Mix (µl) | final [C] or quantity | | stock [C] | |
|---|---|---|---|---|---|---|
| Nb reactions | 1 | 10 | | | | |
| Primer mix | 1,36 | 13,6 | 170 | nM | 5 | µM |
| Buffer | 8,00 | 80,0 | 1 | X | 5 | X |
| High GC enhanced | 8,00 | 80,0 | 1 | X | 5 | X |
| dNTP | 0,32 | 3,2 | 200 | µM | 25 | mM |
| $H_2O$ | 13,92 | 139,2 | | | | |
| Taq | 0,40 | 4,0 | 0,8 | units | 2 | units/µl |
| Adapted-ligated DNA | 8 | | 10 | ng | | |
| Total vol. Mix | 40 | 320,0 | | | | |

Distribute 10 µl of the PCR preparation (DNA+Mix) in 4 replicates plates
PCR at **98°C for 30s, [98°C for 20s, 60°C for 30s, 72°C for 40s] x 15 cycles, 72°C for 10 min**

The first part of the enrichment phase by PCR is made in limiting conditions in terms of primer and DNTP concentration to normalize the amplification process.
It is completed by a final step by adding extra primers and DNTP Pool
the 4 replicates in one plate

Final cycle mix (One tube per Index):

**One well**

0,80 µl Q5 buffer (5X) 0,32 µl
DNTP (**25 mM each**)

4 µl total volume

**10 wells (8 barcodes – one plate)**

8,00 µl Q5 buffer (5X) 3,20 µl
DNTP (**25 mM each**)

26,80 µl Primer mix (5µM each)

40 µl total volume

Distribute 4 µl of the Final cycle mix in each well.
PCR final cycle at **98°C for 3 min, 60°C for 2 min, 72°C for 12 min**

## II-5. Gel visualization and sample pooling

5µl of each sample was visualized on a 1.5% agarose gel and mixed to reach an equimolar pool (for example 5µl of each sample in one tube or if all sample have same concentration we can pool all PCR in one tube)

## II.6- Size selection and quality check

The 30µl size selection step was carried out using a Pippin prep of pool was deposited on the Pippin cassette 1.5% agarose, 250bp – 1.5kb and selected for a size range of 400-800bp.
The size selection step was checked using a High Sensitivity DNA Analysis Kits on a Bioanalyser. **Before size selection with 1µl of PCR pool diluted 1:10 and after size selection with 1µl of pure PCR pool.**

## II.7- Quantity check

Qubit (ng/µL) and qPCR (nM).
The library will be diluted with TE buffer to the ready-to-run concentration (10µl at 4nM for MiSeqV3 and 10µl (minimum)

# Annexe B | chapitre 2

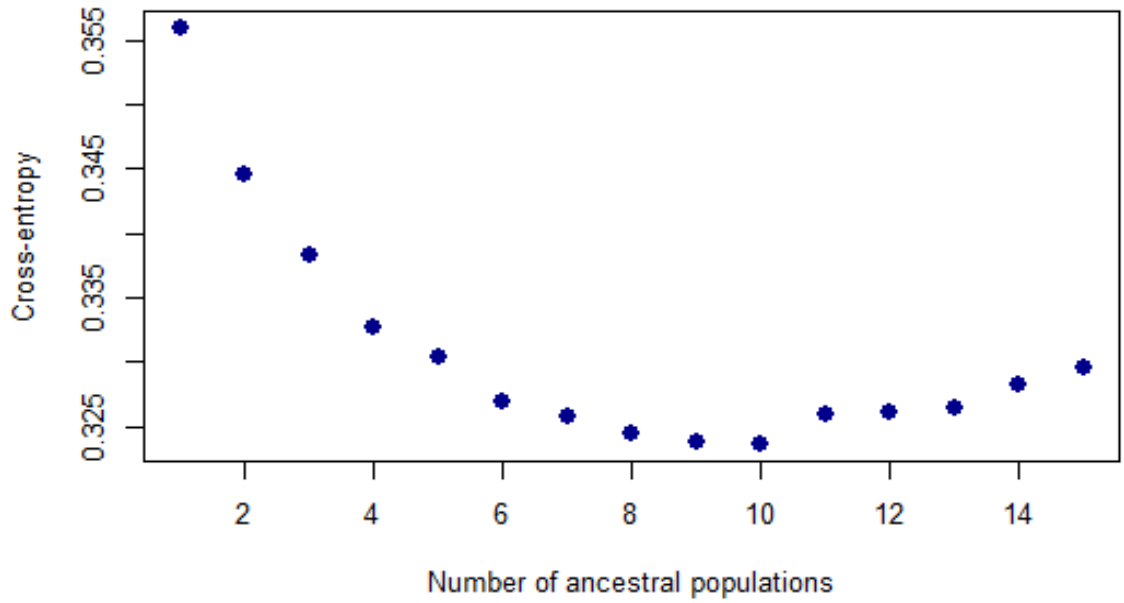



**Figure B.1**. Variation in cross-entropy according to the number K of ancestral populations in sNMF for *Ericaria zosteroides*.

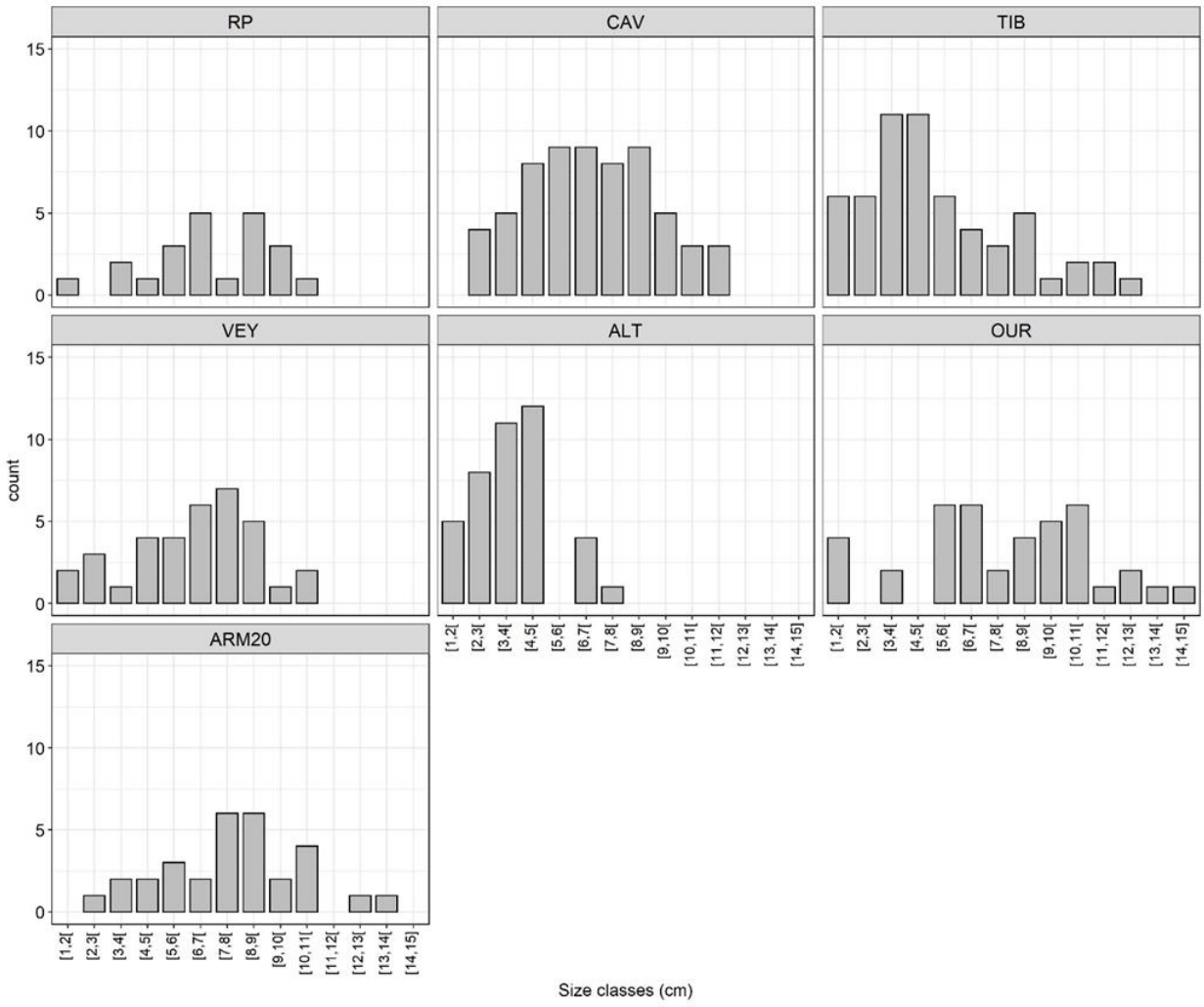



**Figure B.2**. Distribution of the size of the main axis (cm) in the sampled populations; the sizes were distributed into fourteen discrete classes.

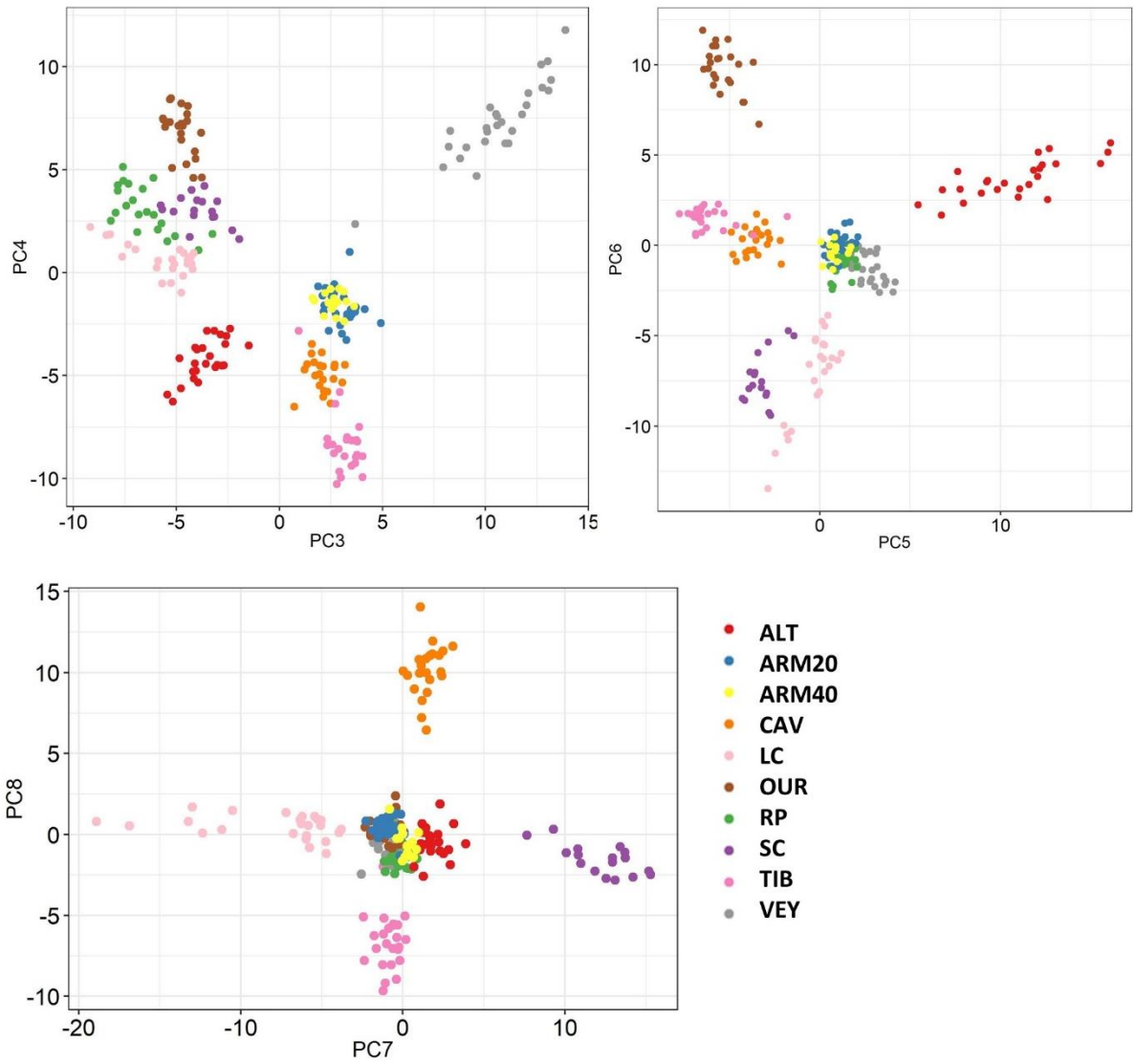



**Figure B.3**. Additional axis from PC3 to PC8 (PC3: 12%; PC4: 11%; PC5: 10%; PC6: 9%; PC7: 9%; PC8: 7%) of the PCA conducted on the neutral set of 10 755 SNPs and 229 individuals.

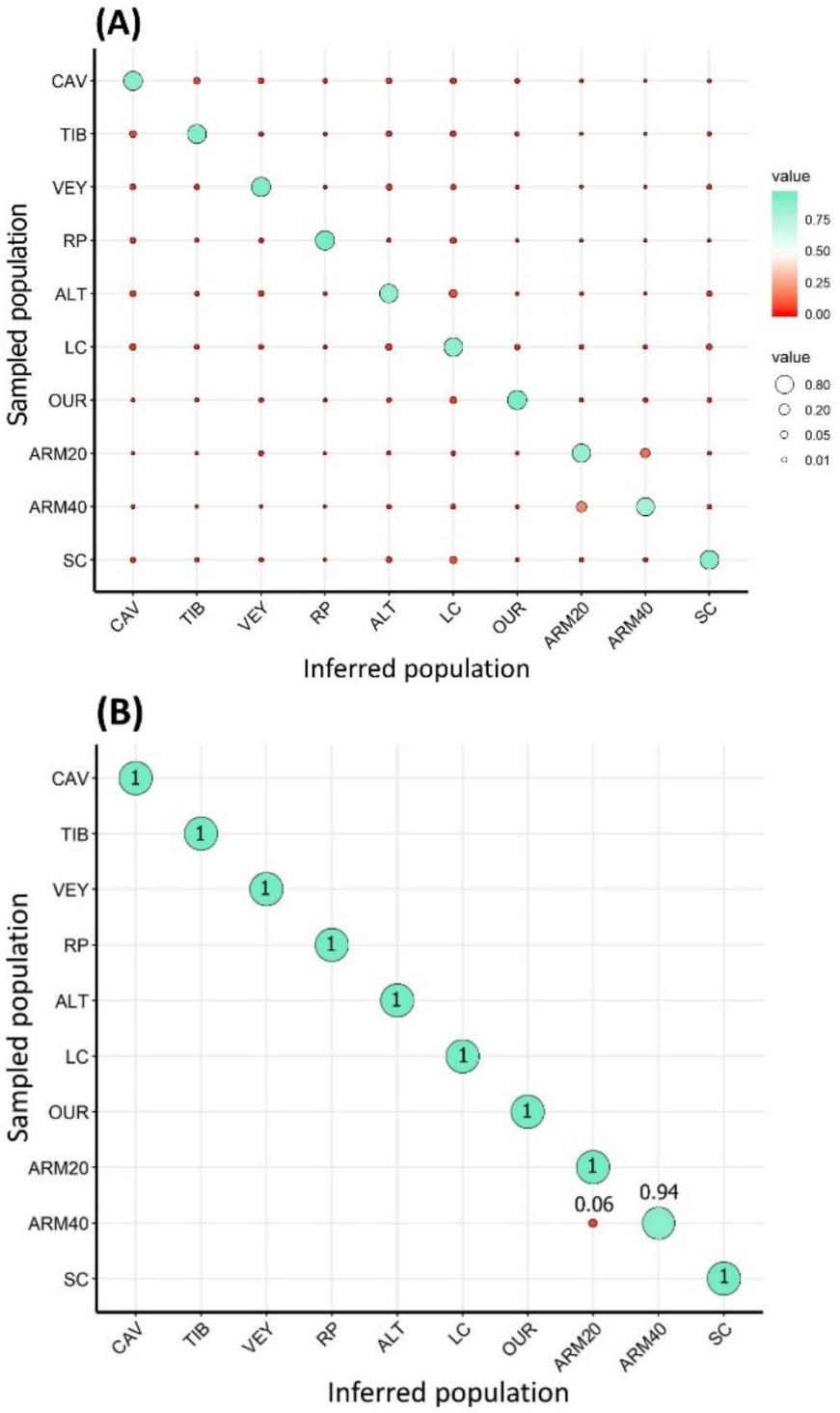



**Figure B.4**. Assignment tests over the different scales of the study. For each population (Sampled population, y-axis), the percentage of individuals assigned (Inferred population, x-axis) to their respective sources (green circle) or other putative sources (red circle). The size of circles is proportional to the percentage of assignment, (A) assignment tests restricted to the top 100 ranked SNPs (decreasing overall $F_{ST}$), (B) assignment tests for all neutral markers (10 755 SNPs).

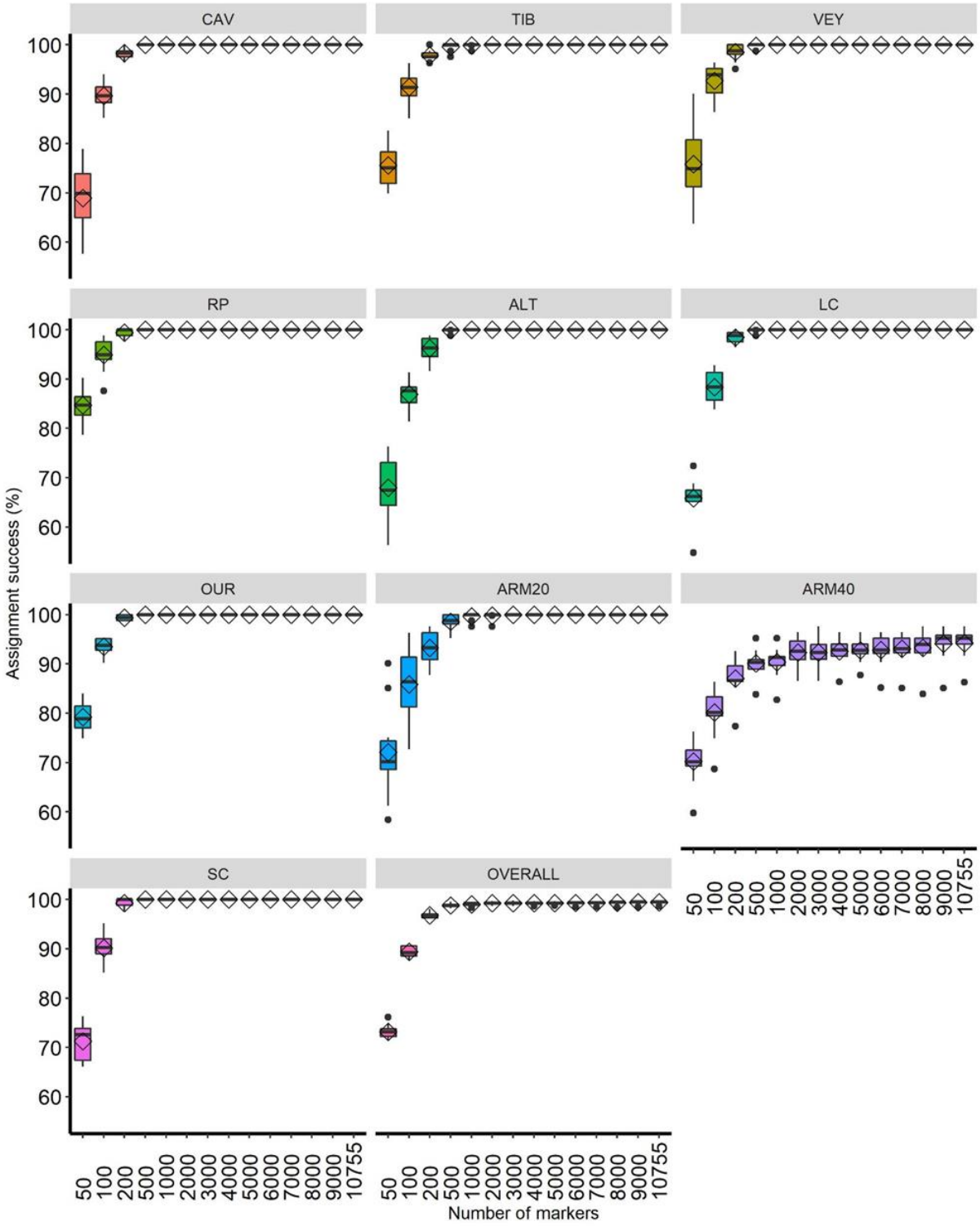



**Figure B.5**. Assignment success (%) at the population level and over populations (Overall) as a function of the number of SNPs ranked based on decreasing $F_{ST}$ values.

**Table B.1**. Quality criteria applied to build the neutral and outlier dataset of *Ericaria zosteroides*.

| Filtering SNP candidates | SNP counts | Individual counts |
|---|---|---|
| Initial catalog of 314 538 loci | 721 374 | 271 |
| **Locus genotyped**<br>> 80% of individuals per population<br>> 80% of study sites<br>**SNP threshold**<br>heterozygosity observed < 50% | 184 723 | 271 |
| **Per-individual quality control filters**<br>missing genotypes < 25%<br>heterozygosity observed < 3 SD from the mean<br>**Per-SNP quality control filters**<br>missing genotypes < 5%<br>Minor Allele Frequency (MAF) > 1% | 21 606 | 248 |
| **Individual relatedness filter**<br>KING-robust estimates = 0.354 | 21 606 | 229 |
| **Linkage Disequilibrium (LD) between SNPs**<br>LD threshold for clumping (on MAF); $r^2 > 0.3$ | 11 067 | 229 |
| **Genome scan filter** | | |
| Putatively neutral | 10 755 | 229 |
| Putatively under divergent selection | 12 | |

For each filtering step, the number of SNPs passing the filter (SNPs count), as well as the number of individuals (individuals count), are reported. The catalog of 314 538 loci was initially built using the Stacks v.2.4 bioinformatics pipeline, with the draft genome of *E. zosteroides* used as a reference.

**Table B.2**. Comparisons in the distribution of the size of the main axis between populations using the Kolmorov-Sirmov pairwise test.

| | CAV | VEY | ALT | RP | TIB | OUR |
|---|---|---|---|---|---|---|
| **VEY** | **9.6e-01** | | | | | |
| **ALT** | 1.4e-08 | 3.5e-06 | | | | |
| **RP** | **1.0e+00** | **7.7e-01** | 1.9e-06 | | | |
| **TIB** | 3.5e-04 | 2.9e-02 | 3.3e-02 | 1.1e-02 | | |
| **OUR** | **1.7e-01** | 5.0e-02 | 9.5e-10 | 3.7e-01 | 1.5e-04 | |
| **ARM20** | **2.7e-01** | **3.0e-01** | 4.9e-08 | 6.2e-01 | 1.6e-03 | **7.8e-01** |

Non-significant P-values are in bold type (p>0.05).

**Table B.3**. Comparisons in genetic diversity (expected heterozygosity among SNPs) between populations using the Pairwise Wilcoxon Test.

| | ALT | ARM20 | ARM40 | CAV | LC | OUR | RP | SC | TIB |
|---|---|---|---|---|---|---|---|---|---|
| **ARM20** | 5.7e-44 | | | | | | | | |
| **ARM40** | 6.1e-54 | **3.3e-02** | | | | | | | |
| **CAV** | **8.6e-01** | 6.5e-48 | 2.6e-62 | | | | | | |
| **LC** | 2.8e-14 | 1.4e-93 | 1.5e-102 | **1.0e+00** | | | | | |
| **OUR** | **3.6e-03** | 2.5e-54 | 3.5e-70 | **1.0e+00** | **1.0e+00** | | | | |
| **RP** | 5.6e-178 | 4.7e-64 | 1.0e-36 | 1.4e-186 | 2.8e-259 | 5.4e-195 | | | |
| **SC** | 1.0e-06 | 3.9e-69 | 1.5e-75 | 2.0e-05 | **1.0e+00** | **2.0e-02** | 1.9e-262 | | |
| **TIB** | **1.0e+00** | 5.3e-28 | 1.9e-38 | 2.4e-08 | 9.3e-13 | 6.8e-09 | 2.1e-147 | 1.9e-16 | |
| **VEY** | **3.5e-03** | 8.8e-30 | 1.3e-41 | 7.9e-04 | 2.4e-26 | 4.0e-07 | 2.0e-152 | 3.1e-14 | **1.0e+00** |

P-values were adjusted for multiple testing with the Bonferroni method. Non-significant P-values are in bold type (p>0.001).

# Annexe C | chapitre 3

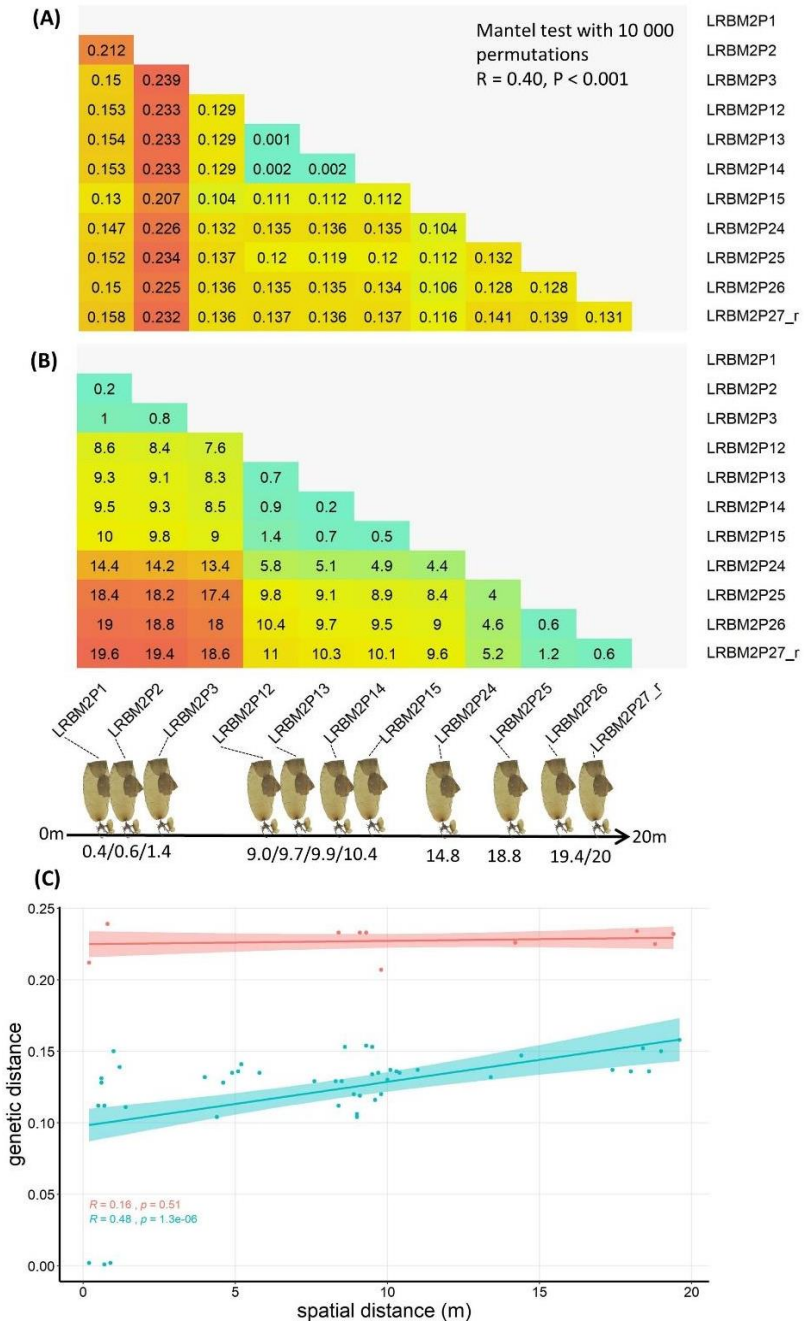



**Figure C.1**. Spatial analyses at Banc Magaud_2. (A) the pairwise genetic distances were those previously used to assess clonal identity and (B) the spatial distance matrix (m) was generated from the positions of individuals within the transect. Mantel tests between spatial and genetic distances were significant (R = 0.40, P < 0.001) after 10 000 permutations. LRBM2P12, LRBM2P13, and LRBM2P14 belong to the same clone, while LRBM2P2 is one of the outliers highlighted in the study. (C) The linear regression between genetic and spatial distances by including (red) or excluding (blue) pairwise comparisons with the outlier.

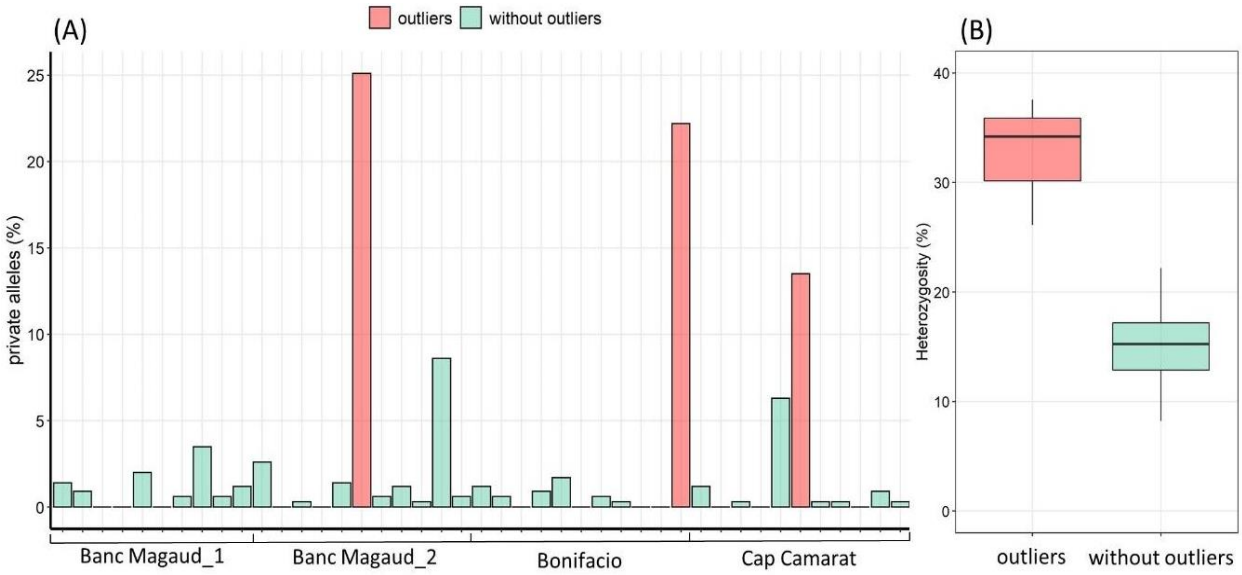



**Figure C.2**. Genomic patterns of the three outliers as comparison to other individuals of *Laminaria rodriguezii*. (A) Percentage of private alleles per individual and (B) percentage of heterozygosity per individual.

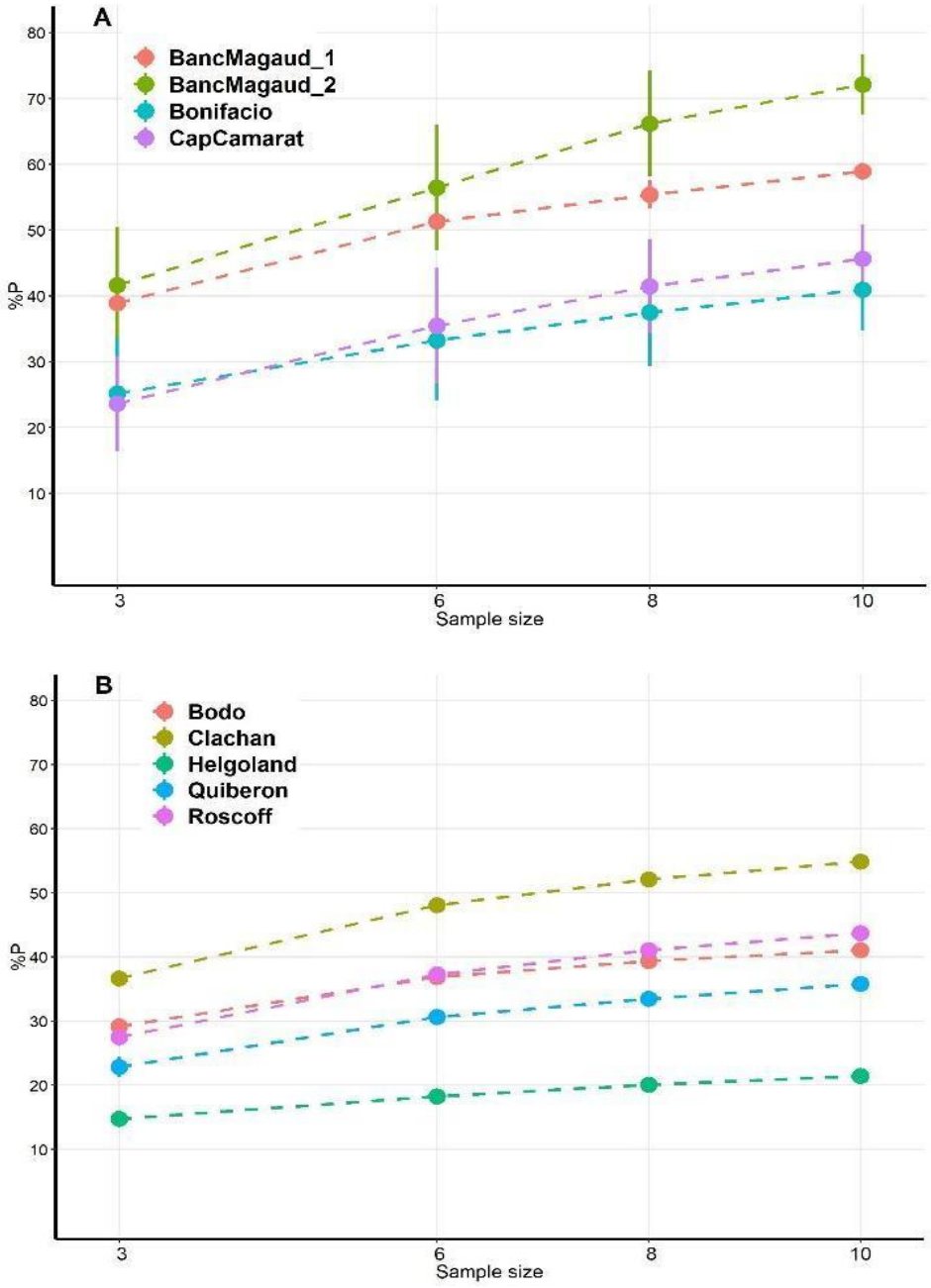



**Figure C.3**. Average percentage of polymorphic loci (%P) within populations of (A) *Laminaria rodriguezii* and (B) *Laminaria digitata* obtained using a random sampling approach implemented in a custom bash script (available at https://github.com/rebecca-cj/revegetation). The percentage of polymorphic loci was calculated for each sample size from 100 random resamples. Standard error from the mean from replicates is represented as error bars.

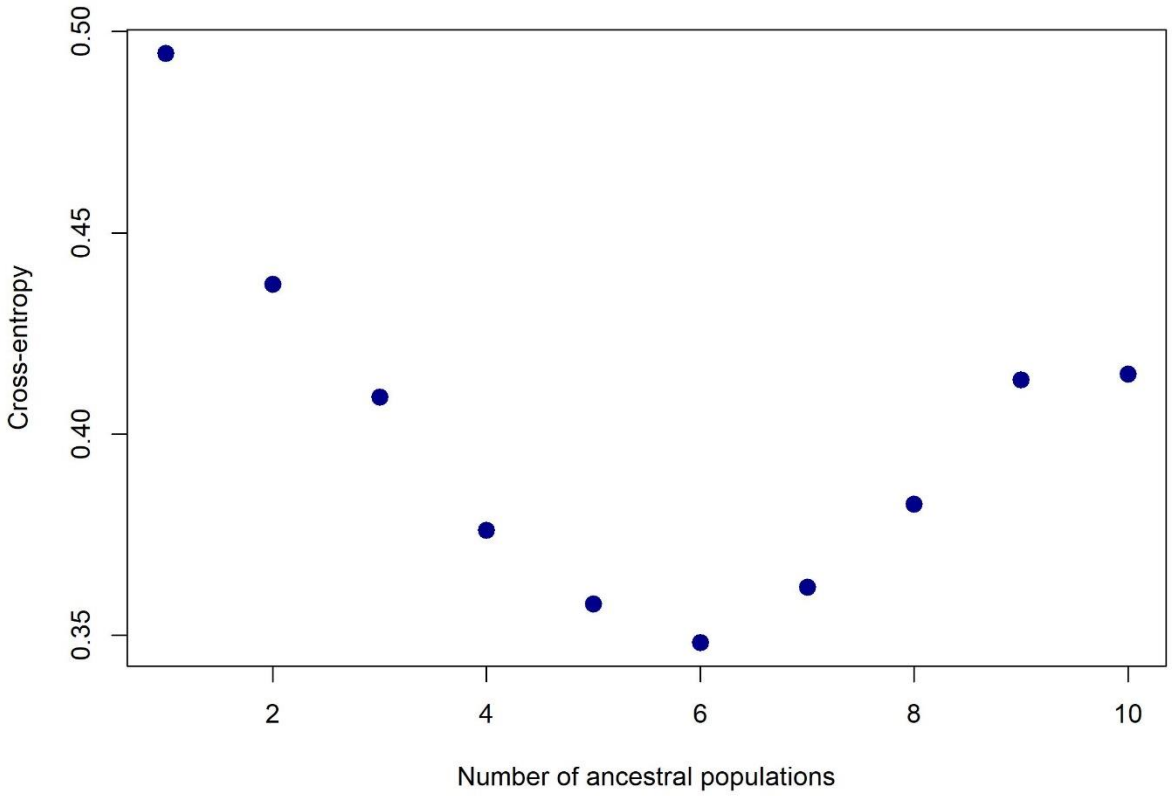



**Figure C.4**. variation in cross-entropy according to the number K of ancestral populations in sNMF for *Laminaria rodriguezii*.

**Table C.1**. Genetic variation within populations of *Laminaria digitata*.

| Population | Label | Lat. | Long. | N | $H_o$ | $H_e$ | $F_{IS}$ | $\overline{r_d}$ |
|---|---|---|---|---|---|---|---|---|
| Bodo | BOD | 67.2765 | 14.5747 | 23 | 0.14 | 0.18 | **0.23** | **0.02** |
| Clachan | CLA | 56.3171 | -5.5832 | 25 | 0.09 | 0.12 | **0.22** | **0.02** |
| Helgoland | HEL | 54.1779 | 7.8926 | 20 | 0.09 | 0.15 | **0.43** | **0.20** |
| Quiberon | QUI | 48.7275 | 4.0079 | 23 | 0.11 | 0.13 | **0.12** | **0.01** |
| Roscoff | ROS | 47.4700 | -3.0914 | 25 | 0.13 | 0.19 | **0.28** | **0.03** |

N, number of individuals successfully genotyped, Ho, observed heterozygosity, He, expected heterozygosity, $F_{IS}$, average fixation index, $\overline{r_d}$, average multilocus linkage disequilibrium. $F_{IS}$ and $\overline{r_d}$ values highlighted in bold indicate significant heterozygote deficit within populations (P-val < 0.001) and significant linkage disequilibrium (P-val < 0.001), respectively.

**Table C.2**. Genotypic richness (R) and Pareto (β) indexes estimated for populations of *Laminaria rodriguezii*.

| Population | β | R |
|---|---|---|
| Banc Magaud_1 | *NA* | 1.00 |
| Banc Magaud_2 | 2.00 | 0.80 |
| Cap Camarat | 3.32 | 0.90 |
| Bonifacio | 0.56 | 0.20 |

No β value is reported at Banc Magaud_1 due to the absence of clonal lineages.

## Connectivité et structure génétiques des populations d'*Ericaria zosteroides* (Fucales) et *Laminaria rodriguezii* (Laminariales) des côtes françaises

Mots clés : diversité génétique, forêts marines, Phaeophyceae, connectivité des populations, clonalité

Les forêts d'algues brunes (Phaeophyceae) ; ordres des Laminariales (kelp) et des Fucales (*Cystoseira, Sargassum, Ericaria, Gongolaria, Fucus* pour la Méditerranée) sont soumises à des combinaisons de pressions d'origine anthropique. La réponse des forêts marines est néanmoins hautement variable selon les espèces et les populations considérées, mais également selon les causes de leur déclin. Le devenir de ces écosystèmes dépendra de la capacité des populations de ces espèces à évoluer dans cet environnement changeant.

A travers ma thèse, j'ai étudié les facteurs susceptibles de contrôler la structure et la diversité génétiques des forêts marines en me focalisant sur deux espèces endémiques du bassin méditerranéen, la Fucales *Ericaria zosteroides* et le kelp des milieux profonds *Laminaria rodriguezii*. Ces espèces peuvent former de vastes forêts marines, d'une grande valeur écologique et socio-économique. Elles sont caractérisées par une distribution éparse et un faible potentiel de dispersion, avec des conséquences potentielles sur leur réponse au changement global.

Dans une première partie, je me suis intéressé à la connectivité des populations par une méthode de séquençage partiel du génome (RAD-seq). L'analyse de la structure génétique indique une forte différenciation génétique à de courtes distances et une connectivité réduite, que ce soit pour les populations de *E. zosteroides* ou *L. rodriguezii*. En couplant structure génétique et modélisation des courants marins, j'ai étudié le rôle de la dispersion des propagules en prenant comme cas d'étude, les populations de *E. zosteroides* en Provence. Les résultats obtenus montrent que la structure génétique est davantage prédite par les courants marins que par l'isolement spatial. Cela supporte l'hypothèse que la dérive de rameaux fertiles peut être un vecteur de flux de gènes. Par ailleurs, le RAD-seq a permis de détecter des locus outliers qui pourraient être impliqués dans l'adaptation locale à la profondeur.

Dans une deuxième partie, je me suis intéressé au mode de reproduction du kelp *L. rodriguezii*, cette espèce étant l'une des rares du genre *Laminaria* à se reproduire par reproduction végétative et sexuée. La clonalité partielle, en interaction avec d'autres forces évolutives, telles que la dérive génétique et les mutations contrôlent l'évolution de la diversité génétique. Ici, j'ai employé le RAD-seq pour étudier à l'échelle du génome les signatures de la clonalité. Les résultats obtenus confirment l'impact de la clonalité sur la diversité génomique de *L. rodriguezii*, avec des niveaux variables selon les populations. La comparaison avec une espèce congénérique à reproduction sexuée (*Laminaria digitata*) montre que d'autres facteurs, notamment la dérive génétique peuvent avoir des effets confondants.

Outre les questionnements fondamentaux soulevés par ces résultats, l'existence d'une forte structure génétique à courte distance pour des populations éparses, a des implications importantes pour la conservation de ces espèces.

## Connectivity and genetic structure of populations of *Ericaria zosteroides* (Fucoids) and *Laminaria rodriguezii* (Laminariales) along the French coast

Keywords: genetic diversity, marine forests, Phaeophycae, population connectivity, clonality

Marine forests of brown algae (Phaeophyceae); orders Laminariales (kelp) and Fucoids (the Mediterranean *Cystoseira, Sargassum, Ericaria, Gongolaria, Fucus*) are impacted by a combination of anthropic pressures. The response of marine forests is nevertheless highly variable according to the species and populations considered, but also according to the causes of the decline. The fate of these ecosystems will depend on the adaptive abilities of populations facing this changing environment.

Through my PhD, I studied the factors likely to control the genetic structure and diversity of marine forests by focusing on two endemic species of the mediterranean basin, the Fucoid *Ericaria zosteroides* and the deep-sea kelp *Laminaria rodriguezii*. These species can form vast marine forests and are of great ecological and socio-economic value. There are characterised by their sparse distribution and their low dispersal abilities with potential consequences on the ability of their populations to face global change.

In the first part, I studied the connectivity of populations by a method of reduced representation sequencing of genomes (RAD-seq). The analysis of the genetic structure indicates a strong genetic differentiation at short distances and reduced connectivity for both *E. zosteroides* and *L. rodriguezii* populations. By coupling genetic structure and modeling of marine currents, I studied the role of propagules dispersal, taking as a study case, the populations of *E. zosteroides* in Provence. The results indicate that the genetic structure is predicted by ocean currents rather than spatial isolation. This finding supports the hypothesis that the drift of fertile zygotes can take part to gene flow in this species. Besides, the RAD-seq was the opportunity to detect outlier loci that could be involved in local adaptation to depth.

In the second part, I studied the mode of reproduction of the kelp *L. rodriguezii*, this species being one of the few of the genus *Laminaria* to reproduce by vegetative and sexual reproduction. The partial clonality, in interaction with other evolutionary forces, such as genetic drift and mutations, controls the evolution of genetic diversity. Here, I used RAD-seq to study the genome-wide signatures of clonality. The results obtained indicate an impact of clonality on the genomic diversity of *L. rodriguezii*, with contrasting levels depending on the populations. The comparison with a congeneric species with only sexual reproduction (*Laminaria digitata*) allowed discussing other potentially confounding effects, such as the effect of genetic drift.

In addition to the fundamental questions raised by these results, the existence of a strong genetic structure at short distances for populations divided into habitat patches has important implications for the conservation of these species.