

Expansion d'aire et processus d'introductions biologiques en milieu marin : le cas de *Cyclope neritea* (Nassariidae) sur les côtes françaises.

THÈSE

présentée et soutenue publiquement le 4 Juillet 2006

pour l'obtention du grade de

Docteur de l'université de La Rochelle

(spécialité biologie évolutive et écologie)

par

Benoît Simon-Bouhet

Composition du jury

<i>Président :</i>	Guy Bachelet	Station Marine d'Arcachon
<i>Rapporteurs :</i>	Jérôme Goudet	Université de Lausanne
	Serge Gofas	Université de Malaga
<i>Directrice :</i>	Frédérique Viard	Station Biologique de Roscoff
<i>Co-directrice :</i>	Pascale Garcia	Université de La Rochelle

Remerciements

J'ai eu la chance de réaliser cette thèse en co-encadrement : deux directrices, deux laboratoires, deux environnements. . . J'ai donc évidemment beaucoup de monde à remercier, mais hélas, j'ai surtout deux fois plus de chances d'oublier quelqu'un !

Je tiens tout d'abord à exprimer ma gratitude à Myriame Valéro et Gérard Blanchard, qui m'ont ouvert les portes de leurs laboratoires respectifs, à Roscoff et La Rochelle. Durant ces quatre années, ils m'ont accordé leur confiance pour mener à bien ce travail et je les en remercie.

Merci également à messieurs Guy Bachelet, Serge Gofas et Jérôme Goudet d'avoir accepté de faire partie de mon jury de thèse.

Naturellement, ma plus grande reconnaissance va ensuite à mes deux co-directrices, Frédérique Viard et Pascale Garcia, sans qui ce travail n'aurait jamais vu le jour. Elles me suivent et me supportent depuis le stage de D.E.A., au cours duquel j'ai découvert ce qu'était la génétique des populations. Depuis cette époque, elles m'ont énormément appris, et ont su me communiquer leur passion de la recherche. Au-delà de leur expertise dans la discipline, c'est, j'en suis convaincu, leurs qualités humaines qui ont fait de ces quatre années une expérience enrichissante et particulièrement enthousiasmante. Merci infiniment à toutes les deux pour le temps et l'énergie que vous m'avez consacrés au cours de ces quelques années de travail acharné.

Si cette thèse s'est bien déroulée et a pu arriver à son terme, c'est également grâce à toutes les "*petites mains*" qui m'ont accompagné, à Roscoff et La Rochelle. A la technique, applaudissez Vanessa Becquet et Claire Daguin ! Elles ont été là chaque fois que j'ai eu besoin d'elles, c'est-à-dire. . . fréquemment ! Leur aide à la paillasse a été inestimable, et connaissant leur efficacité et leur rigueur, je me suis reposé sur leurs compétences bien (trop ?) souvent. Merci à toutes les deux de m'avoir aidé quand j'en ai eu besoin.

Qu'il me soit ici permis de remercier un homme qui a beaucoup de métier métier (oui oui, deux fois métier). Il aurait pu être boulanger, rémouleur ou encore bourrelier. . . Mais non. Heureusement pour moi (et pour beaucoup d'autres), Jacques Pigeot a choisi de faire de la Science. Ou plutôt DES Sciences, car ses connaissances et compétences sont diverses, variées, multiples, que dis-je, encyclopédiques ! Jacques m'a été d'une aide indispensable, donnant de son temps sans compter au cours de quelques "*manips*" terrain mémorables, notamment en Espagne et au Portugal. Toujours sur le coup pour dénicher, à marée basse, les sites où ma bestiole était susceptible de se cacher, et à marée haute, les bouges (pas toujours) fréquentables où la Sangria coule à flot et où la morue est servie. . . bien salée ! Malgré ton grand âge, tu as su garder ton âme d'enfant (un enfant très au fait des "*choses*"

de la vie, j'en conviens!). Merci à toi donc, mon maître en tout (!), pour la multitude de conseils, tes nombreux coups de mains et ta bonne humeur contagieuse!

Ont également contribué à cette thèse, deux personnes que je ne connaissais pas avant de découvrir la Station Biologique de Roscoff et qui sont très rapidement devenues des amis sincères. Couple improbable (la Belle et la Bête!) mais envié, ceux qui les connaissent les auront reconnus : il s'agit bien sûr de Gaëlle et Xavier! Dès le début, ils m'ont accueilli chez eux comme leur "*fls adoptif*" (même si je ne suis pas sûr de vouloir d'un père comme Xavier!) et ont su rendre moins cruel l'éloignement de ma douce et tendre lorsque venant de la chaude et ensoleillée La Rochelle, je débarquais dans la froideur pluvieuse du Finistère Nord (j'entends d'ici les Bretons hurler à la calomnie!). Merci à tous les deux, pour tout, et surtout parce que vous êtes comme vous êtes : ne changez rien.

Je tiens également à remercier Nathalie et Jean-François pour leur amitié et leur soutien. S'ils ne sont pas directement à l'origine de cette thèse, c'est tout de même un peu grâce à eux si j'en suis là aujourd'hui. Dès le D.E.U.G. (eux diraient au berceau!), ils m'ont dit, avec le bon sens qui les caractérise : "Mais Benoît! Qu'est-ce que tu fais à Poitiers si tu veux faire de la biologie marine?... Tu perds ton temps, y a pas la mer! Viens plutôt à La Rochelle!" Influençable comme j'étais, je les ai écoutés. Pendant un moment, j'ai cru que leur insistance cachait quelque chose : peut-être ont-ils des parts dans l'immobilier Rochelais, ou un intéressement quelconque aux inscriptions de l'université? Mais non. Ces gens-là ne sont pas comme ça (en dehors de la mainmise de Nathalie sur la cagnote du café bien sûr!), ils ne voulaient que mon bien. Merci donc à tous les deux pour votre bienveillance et votre amitié.

Merci également à la fine équipe des sportifs de l'institut : Thierry, Paco, Florence, Gilles, Denis... Que ce soit au badminton ou pour "*Mon légionnaire*", ça fait du bien la rigolade entre midi et deux! Merci aussi aux doctorants de Roscoff et de La Rochelle (Marie, Mélanie, Thomas, Pierre-Yves, JC, Oliv', Élise, Benoît, Karolina...) qui, à l'occasion d'un pot, d'un échange de créneau de TD, ou d'une permutation de dernière minute sur le planning du Li-Cor, ont rendu cette thèse plus vivable.

Comme le chômeur en fin de droits, le doctorant en fin de thèse est souvent paumé, au bord de la dépression et parfois même, dans la précarité! Dans ces moments difficiles, ce sont bien sûr les amis et la famille qui, sans le savoir, lui remontent le moral. A ce titre, je tiens en particulier à remercier les copains de lycée ainsi que leurs tendres moitiés (par ordre alphabétique, pour ne fâcher personne!) : Alex, Cécile, Céline, Étienne, Fabienne, Matthieu, Mawy-Cézil, Nico, Patrice et Tifaine. Jérôme et Laurence, ainsi que tous les amis plongeurs avides d'épaves et de réseau karistique (Pierre, Jean-Claude, Marc, Thomas, Flo, Vincent...) ont aussi joué un grand rôle : rien de tel qu'une semaine de

plongée entre amis pour se vider la tête (même si ces derniers temps, la fréquence de ces fameuses semaines de décompression est en chute libre).

Ma famille proche a également été d'un soutien sans faille. Ma belle-famille tout d'abord, toujours aux petits soins avec moi. Peut-être arriveront-ils un jour à retenir le nom de la bestiole qui m'a occupé pendant plus de quatre ans ! Mes parents bien sûr, qui, chaque fois que je rentrais au pays me demandaient l'air inquiet : "Alors, cette thèse... Ça avance ?" sans trop vraiment comprendre de quoi il pouvait bien s'agir ! Ma sœur et son mari, toujours inquiets (certainement à juste titre) de ma santé physique et morale... Leur adorable Léna, dont le moindre éclat de rire suffit à effacer bien des soucis. Mon frère, épicurien devant l'éternel (sait-il seulement ce que ça veut dire, lui qui a fait des études de sport ?!!), qui m'a dit un jour depuis l'Australie : "To bigorneau to be... That's the question!". Si ça c'est pas de la culture ! Merci à tous pour les ondes positives que vous m'avez envoyées durant tout ce temps, et merci tout simplement d'être là, même à distance.

Et enfin, merci à toi Delphine. Même si tu n'as jamais tenu de pipette, enfilé de cuissardes ou fait tourner un logiciel de statistiques, cette thèse est autant la tienne que la mienne. Merci pour tes sacrifices et ton amour. Cette fois ça y est, ta patience va enfin être récompensée : le début de la suite commence...

"Mais à qui qu'ô vous sert mes pauv' drôles ?..."
(Trad. : mais à quoi cela vous sert-il mes pauvres enfants ?)
Odette Triboire, communication personnelle

*"L'Homme de Science le sait bien, lui, que seule la Science a pu, au fil des siècles,
lui apporter l'horloge pointeuse et le parcmètre automatique,
sans lesquels il n'est pas de bonheur terrestre possible."*
P. Desproges, Vivons heureux en attendant la mort

Table des matières

Remerciements	i
Table des figures	xiii
Liste des tableaux	xvii
Liste des Encadrés	xix
 Introduction générale	 1
 1 L'aire de distribution des espèces	 3
1.1 Généralités	3
1.2 Expansion et limites d'aires : particularités en milieu marin	5
1.2.1 Dispersion et expansion	5
1.2.2 Barrières géographiques	7
1.2.3 Établissement	9
 2 L'influence de l'homme sur l'étendue des aires de distribution	 11
2.1 Changements climatiques et déplacements directionnels	11
2.2 Réductions : pollutions et destruction des habitats	13
2.3 Les invasions biologiques : un cas particulier d'expansion	14
 3 Des situations peu claires ; le cas atypique de <i>Cyclope neritea</i>	 17
3.1 Espèce native ou introduite ? Un statut parfois ambigu	17
3.2 Le modèle d'étude : <i>Cyclope neritea</i>	18
3.2.1 Biologie de l'espèce	18
3.2.2 Aire de distribution, problématique	21

Partie I	<i>Cyclope neritea</i> en limite d'aire : la façade atlantique de la péninsule ibérique	25
	Pourquoi étudier cette zone ?	27
4	Instabilité des populations de <i>Cyclope neritea</i> en limite d'aire	29
4.1	Rappel des processus en cause	29
4.2	L'étude comparative	30
4.2.1	Technique d'échantillonnage	30
4.2.2	Mise en œuvre dans le cadre de l'étude comparative	33
4.2.3	Résultats et discussion	35
5	Identification de <i>Cyclope neritea</i> et <i>Cyclope pellucida</i>	43
5.1	Variations morphologiques et définition des espèces cibles	43
5.2	Le marqueur d'espèce	47
	Conclusions	51
Partie II	Phylogéographie dans l'aire "<i>naturelle</i>" de <i>Cyclope neritea</i>	53
	Introduction	55
6	Méthodes statistiques mises en œuvre	61
6.1	Diversité génétique	62
6.2	Arbres et réseaux	63
6.3	Analyses hiérarchiques, isolement par la distance	65
6.3.1	Marqueurs microsatellites	65
6.3.2	Marqueur mitochondrial	67
6.3.3	Isolement génétique par la distance géographique	67
6.4	Modèles de regroupements ou " <i>clustering</i> "	69
6.4.1	Principe	69
6.4.2	Paramètres utilisés	70

7	Résultats	73
7.1	Comparaison entre diversité mitochondriale et nucléaire	73
7.2	Une structure génétique marquée	78
7.2.1	Distribution géographique des haplotypes et divergence mitochondriale	78
7.2.2	Analyses hiérarchiques et isolement par la distance	81
7.2.3	Arbres de populations	84
7.3	Inférences bayésiennes	87
7.3.1	Structure génétique inférée avec le logiciel Structure	87
7.3.2	Structure génétique inférée avec le logiciel Geneland	87
8	Discussion	93
8.1	Profil génétique d'une espèce à faible pouvoir dispersif	93
8.2	Processus historiques	94
8.3	Processus contemporains	95
8.4	Apports de l'approche comparative	97
	Conclusions	99
Partie III Statut des populations de <i>Cyclope neritea</i> en limite d'aire et sur la façade atlantique française		101
Historique de l'expansion récente de l'aire de distribution de <i>Cyclope neritea</i>		103
9	Méthodes statistiques mises en œuvre	107
9.1	Diversité et structure génétique	107
9.2	Isolement par la distance	108
9.3	Similitudes entre populations	108
9.4	Profils de " <i>mélanges</i> " et simulations des coalescences	109
9.5	Étude à micro-échelle	114
10	Résultats	117
10.1	Analyses mitochondriales à l'échelle globale	117
10.1.1	Diversité moléculaire	117

10.1.2	Distribution géographique des haplotypes des populations françaises.	118
10.2	Analyses nucléaires à macro-échelle	124
10.2.1	Diversité génétique	124
10.2.2	Structure génétique	124
10.3	Analyses nucléaires à micro-échelle spatiale	129
11	Discussion	137
11.1	Un scénario d'introductions récurrentes depuis de multiples sources . .	137
11.1.1	Schéma d'introduction	137
11.1.2	Vecteurs d'introduction	139
11.2	Le cas des populations en limite d'aire	140
11.3	Combinaison de deux facteurs : introductions humaines et dispersion naturelle	143
	Conclusions	147
	Conclusion générale, perspectives	149
12	Expansion et déplacement des aires de distribution	151
13	Perspectives	157
13.1	Histoire évolutive de <i>Cyclope neritea</i>	157
13.2	Succès reproducteur et capacité d'adaptation : approches expérimentales	158
13.3	Développement d'outils théoriques pour tester des scénarios mixtes . .	159
	Bibliographie	163
Annexes		181
A	Compléments méthodologiques	183
A.1	Méthodes non spécifiques aux marqueurs mis en œuvre	184
A.1.1	Extraction de l'ADN génomique total	184
A.1.2	La Réaction de Polymérisation en Chaîne : principe	185

A.2	Mise au point d'un marqueur mitochondrial	188
A.2.1	Amorces universelles	188
A.2.2	Séquençage	192
A.2.3	Définition d'un couple d'amorces spécifiques	193
A.3	Mise au point de marqueurs microsatellites	195
A.3.1	Construction d'une banque enrichie	195
A.3.2	Séquençage et sélection des locus utilisables	197
B	Données moléculaires	203
B.1	Fréquences alléliques par locus et par Population	204
B.2	Données mitochondriales brutes	213
C	Présentation du programme réalisé sous Delphi.	217
D	Articles publiés	223
D.1	Marine Ecology Progress Series	224
D.2	Molecular Ecology Notes	238
D.3	Molecular Ecology	242

Table des figures

1.1	Exemples de dispersion par diffusion.	6
1.2	Principaux courants de surface des eaux océaniques mondiales.	8
2.1	Oscillations Nord Atlantique.	12
2.2	Évolution de l'indice NAO depuis 1864.	13
3.1	Spécimens actifs de <i>Cyclope neritea</i>	19
3.2	Aire de distribution de <i>Cyclope neritea</i> telle qu'elle était reconnue jusqu'en 1950.	21
1	Les trois zones de l'aire de distribution de <i>Cyclope neritea</i>	28
4.1	Exemples de sites où <i>Cyclope neritea</i> est présente en forte densité.	32
4.2	Attraction de <i>Cyclope neritea</i> par une carcasse.	33
4.3	Test de la collecte par la méthode du cadrat.	33
4.4	Sites examinés entre Gibraltar et le Pays Basque français.	34
4.5	Vue de détail de la baie de Setubal.	35
4.6	Vues de détail de la baie de Santander.	36
5.1	Variabilité de la couleur de la coquille chez <i>Cyclope neritea</i>	45
5.2	Différences morphologiques des coquilles de <i>Cyclope neritea</i> , <i>Cyclope donovani</i> et <i>Cyclope pellucida</i>	46
5.3	Migration sur gel d'agarose de produits de PCR réalisées sur <i>Cyclope neritea</i> et <i>Cyclope pellucida</i> avec des amorces universelles.	47
5.4	Migration sur gel d'agarose de produits de PCR réalisées sur <i>Cyclope neritea</i> et <i>Cyclope pellucida</i> avec des amorces spécifiques de <i>Cyclope neritea</i>	48
6.1	Localisation des treize sites d'échantillonnage de l'aire "naturelle" de <i>Cyclope neritea</i>	61

7.1	Nombre d'allèles et richesse allélique dans les populations de l'aire "naturelles" de <i>Cyclope neritea</i>	77
7.2	Réseau présentant les relations phylogénétiques entre les haplotypes de <i>Cyclope neritea</i>	79
7.3	Distribution des haplotypes dans les populations "naturelles" de <i>Cyclope neritea</i>	80
7.4	Profil d'isolement génétique par la distance géographique mis en évidence dans l'aire "naturelle" sur les données microsatellites.	83
7.5	Profil d'isolement génétique par la distance géographique mis en évidence dans l'aire "naturelle" sur les données mitochondriales.	84
7.6	Arbre des populations "naturelles" construit avec les données haplotypiques.	85
7.7	Arbre des populations "naturelles" construit avec les données microsatellites.	86
7.8	Inférence du nombre de classes grâce aux simulations réalisées par Structure	88
7.9	Probabilités postérieures individuelles d'appartenance aux classes identifiées par Structure	89
7.10	Distribution du nombre de classes inférées par Geneland	90
7.11	Cartes des probabilités postérieures d'appartenance aux 7 classes identifiées par Geneland	91
7.11	<i>Suite</i> . Cartes des probabilités postérieures d'appartenance aux 7 classes identifiées par Geneland	92
9.1	Distribution "match-mismatch" théorique d'une population introduite.	110
9.2	Étapes de la reconstruction d'une généalogie de gènes pour un locus haploïde.	113
9.3	Localisation des sites d'échantillonnage de la Mar Menor en vue de l'étude de la structure génétique à micro-échelle.	115
10.1	Réseau présentant les relations phylogénétiques entre les haplotypes de <i>Cyclope neritea</i>	119
10.2	Distribution des haplotypes dans les huit populations étudiées.	120
10.3	Arbre de populations construit avec les données haplotypiques.	121
10.4	Distribution "match-mismatch" et tests K	122
10.5	Isolement génétique par la distance géographique entre les cinq populations françaises récemment détectées.	123
10.6	Arbre de populations construit avec les données microsatellites.	128
10.7	Résultats des tests d'assignation des individus des populations FAM.	130
10.8	Arbre de populations présentant les similitudes entre les 13 populations utilisées pour étudier la micro-structure.	133

10.9	Analyse Factorielle des Correspondances.	134
10.9	<i>suite</i> . Analyse Factorielle des Correspondances.	135
10.9	<i>suite et fin</i> . Analyse Factorielle des Correspondances.	136
11.1	Résultats partiels produits par Geneland en vue de déterminer le statut des populations en limite d'aire.	142
12.1	Influences humaines sur la biodiversité et le changement global.	152
12.2	Modèle conceptuel de la dynamique de l'aire de distribution d'une espèce. .	153
A.1	Les étapes de la Réaction de Polymérisation en Chaîne	186
A.2	Évolution théorique du pourcentage de copies d'ADN de taille souhaitée dans le milieu réactionnel au cours d'une réaction de polymérisation en chaîne.	187
A.3	Exemple de résultats d'amplifications obtenu pour COI avec un gradient de température.	190
A.4	Exemple de PCR ne donnant pas les résultats escomptés.	191
A.5	Exemple de chromatogramme.	193
A.6	Résultats de PCR réalisée avec les amorces Cy2 et CyR.	194
A.7	Représentation schématique d'un fragment d'ADN de <i>Cyclope neritea</i> ob- tenu par séquençage d'un clone positif.	197
A.8	Exemple d'image obtenue après migration d'un locus microsatellite sur gel de polyacrylamide (15 individus).	198
A.9	Principales caractéristiques des microsatellites testés au niveau de la popu- lation.	200
A.10	Capture d'écran du logiciel Saga.	201
C.1	Exemple d'affichage de la fenêtre de saisie dans le cas d'une simulation impliquant une population panmictique de taille constante.	218
C.2	Exemple d'affichage de la fenêtre de saisie dans le cas d'une simulation impliquant une population subissant un goulot d'étranglement.	219
C.3	Indices statistiques calculés dans le cas d'une simulation simple.	220
C.4	Exemple de résultats obtenus lorsque l'utilisateur souhaite calculer l'inter- valle de confiance à 95% d'un indice statistique.	220
C.5	Diagramme schématique présentant le mode de fonctionnement du pro- gramme de coalescence. D'après Muller (2002).	221

Liste des tableaux

4.1	Comparaison des données de terrain (“ <i>présence 2002-2004</i> ”) avec les données bibliographiques (présence antérieure).	37
5.1	Position des sites variables et nature des substitutions nucléotidiques entre deux haplotypes de <i>Cyclope neritea</i> et l’haplotype de <i>Cyclope pellucida</i> . . .	49
1	Comparaison de quelques caractéristiques moyennes de trois marqueurs nucléaires couramment utilisés	56
7.1	Diversité génétique des populations de l’aire “ <i>naturelle</i> ” de <i>Cyclope neritea</i>	74
7.2	Diversité génétique des 4 locus microsatellites pour chaque population de l’aire “ <i>naturelle</i> ”.	75
7.3	Déséquilibres de liaison pour l’ensemble des populations de l’aire “ <i>naturelle</i> ”, entre chaque paire de locus.	76
7.4	F —statistiques hiérarchiques.	81
7.5	Valeurs de $\hat{F}_{IS}$ et probabilités associées.	82
10.1	Diversité génétique des populations de <i>Cyclope neritea</i>	118
10.2	Diversité des 4 locus microsatellites pour les populations des groupes 1, 2 et 3.	125
10.3	Structure génétique au sein des populations des groupes 1 et 2.	126
10.4	Valeurs de $\hat{F}_{IS}$ pour chaque population étudiée.	127
10.5	Valeurs de $\hat{F}_{IS}$ et probabilités associées pour chaque population des cinq sites étudiés.	129
10.6	$\hat{F}_{ST}$ par paires de populations et valeurs de probabilité associées.	131
A.1	Description des différentes amorces testées	189
A.2	Conditions de PCR testées avec les différents couples d’amorces universelles.	190
A.3	Résultats qualitatifs obtenus pour chaque locus testé.	191

B.1	Fréquences alléliques du locus 2C4 pour chaque populations étudiée.	205
B.2	Fréquences alléliques du locus 3C4 pour chaque populations étudiée.	206
B.3	Fréquences alléliques du locus 2H10b pour chaque populations étudiée. . .	207
B.4	Fréquences alléliques du locus 2H10b pour chaque populations étudiée (<i>suite et fin</i>).	208
B.5	Fréquences alléliques du locus 1A10 pour chaque populations étudiée. . . .	209
B.6	Fréquences alléliques du locus 1A10 pour chaque populations étudiée (<i>suite</i>).210	
B.7	Fréquences alléliques du locus 1A10 pour chaque populations étudiée (<i>suite</i> <i>et fin</i>).	211
B.8	Moyennes et écart-types de N_{all} , H_O et H_E pour chaque population.	212
B.9	Position des sites variables des 39 haplotypes de <i>Cyclope neritea</i>	214
B.10	Nombre d'haplotypes par populations.	215

Liste des Encadrés

6.1	Principe des statistiques F . D'après Dupont (2004).	66
6.2	Principe des statistiques Φ	68
9.1	L'indice D de Tajima.	111

Introduction générale

1

L'aire de distribution des espèces

Le travail réalisé au cours de cette thèse s'intègre dans la problématique générale de l'étude de l'expansion géographique récente de l'aire de distribution d'une espèce. Dans cette introduction, nous présenterons quelques généralités concernant la définition de l'aire de distribution des espèces, nous verrons quels sont les facteurs susceptibles d'influencer son évolution et ses limites, en particulier en milieu marin. Notre modèle biologique ayant potentiellement été soumis aux conséquences des actions anthropiques, nous montrerons ensuite en quoi les activités humaines peuvent influencer les mouvements d'aires de distribution. Enfin, nous présenterons notre modèle biologique : le gastéropode marin *Cyclope neritea*.

1.1 Généralités

Lomolino *et al.* (2005) rappellent un principe fondamental de la biogéographie¹ de la façon suivante : la vie, à la surface du globe, semble répartie de façon non aléatoire et hautement prédictible. La diversité des espèces rencontrées est par exemple plus grande sur une île de superficie importante que sur une île de taille réduite, dans les régions tropicales que dans les milieux tempérés. De même, la végétation en zone montagneuse change de façon ordonnée des plaines aux sommets, tout comme les organismes marins sont répartis de façon non aléatoire du littoral aux fosses abyssales.

Plusieurs auteurs ont ainsi tenté d'identifier des règles générales susceptibles d'expliquer la répartition des espèces à la surface du globe. Willig *et al.* (2003) ont suggéré que les facteurs biotiques et abiotiques limitent l'aire de distribution de la plupart des espèces de la même manière : les interactions biotiques tendraient à limiter la distribution

¹Étude de la distribution géographique des organismes

et l'abondance des espèces aux basses latitudes, alors que les facteurs abiotiques seraient prépondérants aux hautes latitudes (Dobzhansky, 1950; Mac Arthur, 1972, cités dans Brown *et al.*, 1996). En milieu marin intertidal, des hypothèses similaires ont été émises : les facteurs abiotiques, liés à l'exposition aux stress physiques entre les marées fixeraient la limite supérieure des aires de distribution alors que la prédation et la compétition interspécifique fixeraient leur limite inférieure (Connell, 1975). Le même schéma a été proposé pour plusieurs autres gradients écologiques, tels que les gradients d'élévation ou d'aridité en milieu terrestre, ou les gradients de profondeur en milieu marin : dans une direction le long du gradient, la distribution des espèces est limitée par un stress physique croissant, alors que dans l'autre direction, la limitation est le fruit du nombre et de l'impact croissant des "*ennemis biologiques*" (Brown *et al.*, 1996). La "*règle de Rapoport*", indiquant que la largeur (en degrés de latitude) de l'aire de distribution d'une espèce augmente avec la latitude, est intimement liée à ces gradients et plusieurs auteurs ont étendu cette règle à la relation entre taille de l'aire de distribution et altitude en milieu terrestre (Stevens, 1992), ou profondeur en milieu marin (Smith et Brown, 2002).

Toutefois, de nombreux contre-exemples contredisent ces règles d'apparente simplicité à grande échelle. L'universalité de la règle de Rapoport est ainsi remise en cause de plus en plus fréquemment. Rohde *et al.* (1993), montrent par exemple que cette règle ne s'applique pas aux poissons téléostéens. De même, Gaston *et al.* (1998) et Gaston et Chown (1999) démontrent en quoi cette règle, loin d'être générale pourrait n'être qu'un phénomène local (le gradient de taille d'aire de distribution, assez bien visible entre 40 et 50 degrés de latitude nord, reste à mettre clairement en évidence partout ailleurs). Plusieurs règles énoncées afin d'expliquer des schémas écologiques *a priori* généraux, se sont finalement avérées n'être que des cas particuliers. C'est notamment le cas de la règle de Bergmann sur les gradients de tailles individuelles en fonction de la latitude (Hawkins et Lawton, 1995), de la règle de Cope sur la tendance des lignées à évoluer vers des tailles corporelles importantes (Jablonski, 1997), ou encore de la règle de Hutchinson, stipulant que des espèces de même niveau trophique, occupant la même niche écologique, ne peuvent coexister que s'il existe un ratio de taille corporelle d'au moins 1,3 entre les espèces considérées (Eadie *et al.*, 1987; Simberloff et Boecklen, 1981).

Ainsi, un autre principe fondateur de la biogéographie, connu sous le nom de "*Loi de Buffon*", montrant à quel point identifier des règles générales peut être hasardeux, souligne que des régions aux caractéristiques environnementales similaires mais isolées géographiquement présentent des assemblages d'espèces différents. Essayer d'identifier des facteurs clés influençant la taille, la forme, la structure ou les limites précises de l'aire de distribution des espèces, s'avère donc extrêmement complexe et relève souvent plus du cas

particulier que de la généralité (Gaston, 1990). Gaston (1996) montre par exemple que la distribution de la taille de l'aire de distribution des espèces d'un taxon donné n'est jamais symétrique mais qu'elle est au contraire toujours biaisée du côté des grandes aires de distribution. Autrement dit, au sein d'un groupe taxonomique donné, la majorité des espèces aura une aire de distribution naturelle réduite, tandis que quelques unes seulement auront une aire de distribution vaste. Les efforts à fournir pour réaliser des inventaires exhaustifs seront d'autant plus importants que la proportion d'espèces ayant des aires de distribution réduites sera grande. Chaque espèce étudiée est donc particulière, et sa distribution dépendra, entre autres facteurs, de son milieu de vie (terrestre, marin, dulçaquicole), de ses capacités de dispersion et de ses capacités adaptatives et compétitives.

1.2 Expansion et limites d'aires : particularités en milieu marin

Les individus d'une espèce dont l'aire de distribution est en expansion sont capables : (i) de se déplacer vers de nouvelles zones, (ii) de supporter des conditions potentiellement défavorables au cours de leurs déplacements et de franchir d'éventuelles barrières (physiques, géographiques...) et (iii) d'établir durablement de nouvelles populations à leur arrivée (Holt, 2003). Chacune de ces trois étapes fait appel à des propriétés biologiques et des stratégies particulières.

1.2.1 Dispersion et expansion

L'une des particularités de nombreuses espèces marines concerne leur fort potentiel dispersif, notamment grâce au relargage dans les masses d'eaux de gamètes, d'oeufs ou de larves en très grande quantité (Strathmann et Strathmann, 1982; Palumbi, 1994). Par exemple, la production de cellules gamétiques dépasse la production de cellules somatiques chez *Mytilus edulis* à l'âge adulte (Rodhouse *et al.*, 1986) et *Ostrea virginica* peut perdre jusqu'à 65% de sa masse corporelle après le relargage de 15 à plus de 110 millions de gamètes dans le milieu marin (Galtsoff, 1930; Buroker, 1983).

Toutes les espèces ne possèdent pas pour autant dans leur cycle de vie des phases de dispersion aussi efficaces et trois principaux modes de dispersion sont généralement distingués (Pielou, 1979; Birks, 1989) :

1. la dispersion par sauts : envoi à longue distance de migrants généralement en faible nombre.

2. la diffusion : expansion naturelle et graduelle dans le prolongement de l'aire de distribution de l'espèce (voir par exemple figure 1.1). Ce processus se distingue de la dispersion par saut par sa lenteur ainsi que par son mode d'expansion continue (et non pas ponctuelle) et souvent à grande échelle géographique.
3. la dispersion séculaire : processus proche de la dispersion par diffusion mais encore plus lent, se mettant donc en place sur des périodes beaucoup plus longues (de l'ordre de la centaine ou du millier de générations), au cours desquelles les espèces ont généralement le temps d'évoluer.

La mise en place de ces différentes stratégies dépend en partie des capacités et propriétés intrinsèques de l'espèce. Roy *et al.* (2001) ont par exemple montré que les bivalves marins ayant eu la capacité de déplacer leur aire de distribution en réponse aux fluctuations climatiques du pléistocène sont majoritairement ceux dont les tailles corporelles sont les plus importantes.

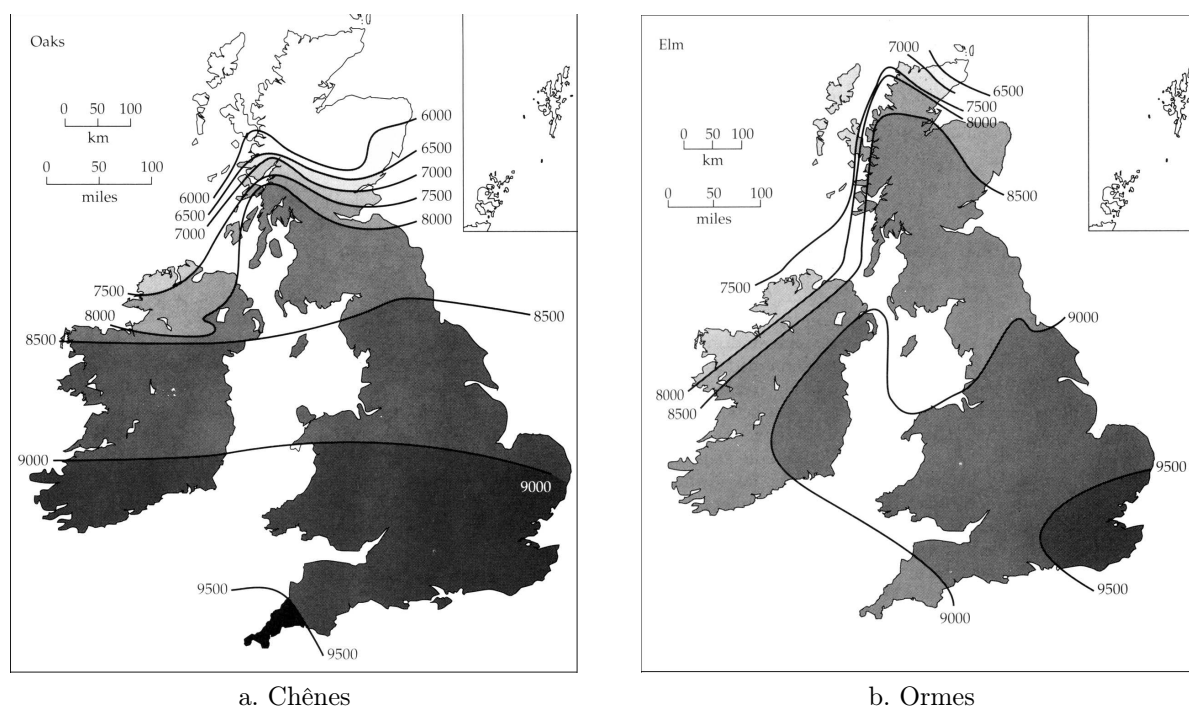


FIG. 1.1: Exemples de dispersion par diffusion. La figure 1.1a représente la colonisation de la Grande Bretagne par les chênes *Quercus* spp. L'expansion géographique s'est étendue de -9 500 à -6 000, ce qui, compte tenu de la durée de vie longue de cet arbre, place cette expansion dans la catégorie des dispersions par diffusion. La figure 1.1b présente, pour la même période, l'expansion des ormes *Ulmus* spp. D'après Birks (1989).

Le mode de dispersion varie également : la dispersion est majoritairement passive pour la plupart des organismes marins possédant un stade gamétique ou larvaire planctonique,

toutefois, elle peut être active pour les plus grandes espèces (cétacés, requins, poissons prédateurs, tortues...) qui ont par conséquent souvent des aires de distribution vastes mais discontinues (voir par exemple le cas des tortues vertes étudiées par Bowen *et al.*, 1992, ou celui des migrations trans-océaniques des requins blancs mâles, Pardini *et al.*, 2001).

Enfin, à côté des propriétés intrinsèques des espèces, Davis *et al.* (1998) soulignent le fait que la distribution des espèces dépend non seulement de facteurs abiotiques mais aussi des interactions avec les autres espèces. Ainsi, prédire les mouvements d'espèces en rapport, par exemple, avec les fluctuations climatiques, peut s'avérer délicat dans la mesure où ces changements sont susceptibles de modifier les rapports existants entre les espèces.

1.2.2 Barrières géographiques

Si la notion de barrière géographique a une réalité concrète en milieu terrestre (montagnes, cours d'eau, bras de mer...), elle reste souvent impalpable en milieu marin (Palumbi, 1994). En effet, les facteurs pouvant limiter la dispersion des espèces en milieu marin paraissent à première vue, sinon absents, du moins peu contraignants du fait de la communication entre les masses d'eaux. Pourtant, les contraintes abiotiques sont nombreuses : température, salinité, quantité de lumière, présence de nutriments, pH, pression... sont autant de facteurs susceptibles d'influencer les mouvements d'espèces marines et le positionnement des limites d'aire de distribution des espèces (Palumbi, 1994).

Les caractéristiques physiques des masses d'eau sont, dans une certaine mesure, prédictibles, car elles dépendent de gradients géographiques et de processus physiques bien identifiés (voir par exemple les conséquences des marées sur les écosystèmes littoraux, Nixon, 1988). Ainsi, la pénétration des rayonnements solaires dans les masses d'eau est un processus physique bien connu, qui délimite plusieurs zones de profondeurs en fonction des longueurs d'ondes pouvant y pénétrer et influence les espèces algales qui peuvent y vivre (*e.g.* algues vertes, puis brunes et enfin rouges). De même, l'absorption rapide des rayons lumineux en milieu marin implique que seule la couche superficielle des océans est chauffée, le transfert de chaleur vers les couches plus profondes nécessitant des courants de convection. La densité des masses d'eau dépendant de la température, une stratification thermique stable existe et limite les déplacements d'espèces (Valiela, 1995).

De même, la circulation verticale et horizontale des courants océaniques est complexe mais relativement bien documentée (voir par exemple la carte des courants de surface présentée à la figure 1.2, ou les facteurs influençant la circulation de surface à une échelle

plus régionale, Molinari *et al.*, 1990; Reverdin *et al.*, 1994). Elle est, entre autres facteurs, sous l'influence de la force de Coriolis, de la forme des traits de côtes continentaux, des gradients de salinité et de densité des masses d'eaux (directement dépendants des rayonnements différentiels reçus par les masses d'eaux suivant la latitude d'une part et de la balance entre apports et évaporation d'eau douce d'autre part) et, plus localement, de la marée (Nixon, 1988). Ces courants sont à la fois un moyen de dispersion passive à longue distance pour de nombreuses espèces (possédant, au cours de leur cycle de vie, un stade planctonique) et une barrière à la dispersion pour les espèces incapables de franchir les gradients physiques associés à ces courants (Gaylord et Gaines, 2000, montrent même que ces courants, indépendamment des gradients physiques freinant la dispersion, peuvent contraindre les espèces par le simple jeu des mouvements de masses d'eau).

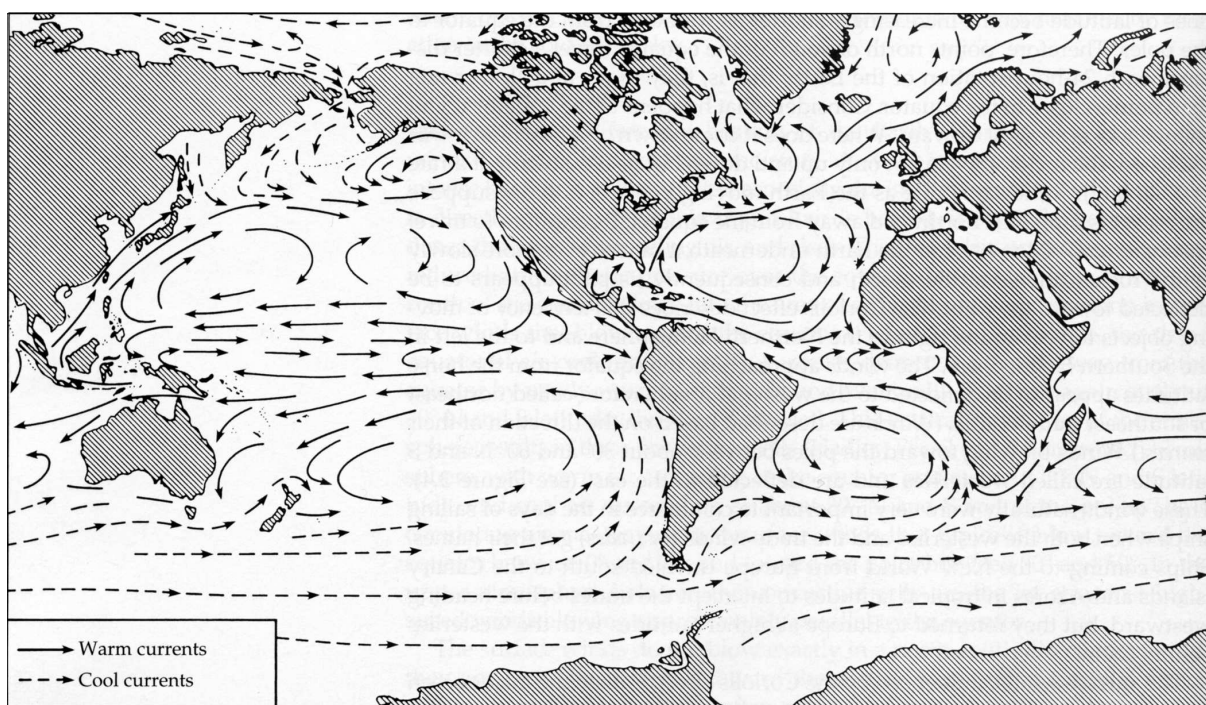


FIG. 1.2: Principaux courants de surface des eaux océaniques mondiales. Les courants chauds sont figurés par les flèches pleines, et les courants froids par les flèches pointillées. Notez le contre-courant équatorial pacifique : le faible courant qui suit l'équateur d'ouest en est, à l'opposé des gyres, et qui se renforce certaines années pour causer le phénomène El Niño (voir section 2.1). Source : Lomolino *et al.* (2005)

En zone intertidale, la plupart de ces gradients sont exacerbés par le balancement des marées : la température, la salinité, le pH ou l'éclairement peuvent en effet fortement varier en l'espace de quelques heures. Ces fluctuations sont autant de contraintes que les espèces intertidales doivent supporter et représentent des barrières à la dispersion pour nombre d'entre elles (Bertness et Leonard, 1997).

Enfin, la nature des barrières limitant la dispersion des espèces dépend elle aussi fortement des limitations intrinsèques des espèces. Willis (1974) a par exemple montré que plusieurs espèces d'oiseaux Formicariidae, après avoir disparu de l'île de Barro Colorado (Panama) ne l'ont pas recolonisée alors qu'ils n'auraient eu que quelques centaines de mètres à parcourir pour franchir les eaux du lac Gatun qui les en séparaient. L'incapacité comportementale de ces oiseaux forestiers à traverser des espaces ouverts en serait la cause. Les limitations physiologiques des espèces sont également à prendre en compte : la très grande majorité des espèces aquatiques vivent soit en milieu marin, soit en milieu dulçaquicole et leurs capacités d'osmorégulation ne leur permettent pas de survivre longtemps dans un milieu qui n'est pas le leur.

1.2.3 Établissement

Dans le cas de la dispersion par sauts, les espèces ayant réussi à franchir les barrières à la dispersion doivent ensuite s'établir dans leur nouveau milieu. Plusieurs facteurs, tels que la sélection par l'habitat, les relations inter-spécifiques ou les stratégies de reproduction, sont ici d'une importance capitale (Lomolino *et al.*, 2005). Même les espèces ayant un mode de dispersion passif sont capables, dans une certaine mesure, de sélectionner un habitat favorable à leur établissement (c'est le cas de nombreuses larves planctoniques d'invertébrés marins fixés, Connell, 1975). Toutefois, les migrants réussissant à franchir les barrières à la dispersion et qui recrutent sur un substrat favorable à leur arrivée dans un nouvel environnement sont souvent peu nombreux et le hasard joue un rôle important : la présence d'un prédateur qui prélève quelques migrants avant ou aussitôt après leur établissement peut suffire à mettre en péril la pérennité de la nouvelle population (voir l'étude expérimentale de Carr et Hixon, 1995, sur quelques espèces de poissons coralliens). Ainsi, la plupart des tentatives d'installation de nouvelles espèces échouent, ce qui est également le cas des espèces introduites accidentellement par l'homme : la règle des dixièmes (ou "*tens rule*") de Williamson (1996) indique que seule une espèce sur dix parvient à réussir chaque étape du processus d'invasion biologique (introduction, établissement, intégration, prolifération). Il est maintenant admis que plus les tentatives seront nombreuses et impliqueront un nombre élevé d'individus, plus les migrants auront de grandes chances de s'établir durablement. Ainsi, Clarke (1971) a montré que l'introduction du cerf européen en Nouvelle Zélande n'a réussi qu'après trente tentatives infructueuses ! D'une façon plus générale, Forsyth et Duncan (2001) ont montré que les espèces d'ongulés établies avec succès en Nouvelle Zélande sont le résultat de l'introduction d'un nombre d'individus égal ou supérieur par rapport aux espèces proches dont l'introduction a échoué.

Dans le cas de la dispersion par diffusion ou de la dispersion séculaire, la situation est tout autre. La création de nouvelles populations en bordure de l'aire de distribution d'une espèce nécessite (i) l'envoi de migrants depuis le cœur de l'aire de distribution vers la périphérie et (ii) le développement d'adaptations aux conditions locales qui, par définition, ne sont pas optimales pour le développement normal des individus de cette espèce. Deux processus s'affrontent alors :

- d'une part, l'évolution d'adaptations locales dans le nouveau milieu
- d'autre part, l'arrivée régulière de migrants "*non adaptés*", provenant du cœur de l'aire de distribution de l'espèce qui freine ou empêche totalement la fixation de caractères avantageux dans le nouvel environnement.

C'est le mécanisme de "*migration contre sélection*". Les flux de gènes étant toujours orientés du centre vers la périphérie de l'aire de distribution, les populations en limite d'aire souffrent la plupart du temps de maladaptation et sont ainsi freinées dans leur expansion. Kirkpatrick et Barton (1997) soulignent que dans certaines conditions, la modification minime d'un paramètre environnemental (biotique ou abiotique) peut entraîner un changement radical de l'équilibre migration-sélection, permettant ainsi à une espèce ayant une aire de distribution limitée de se disperser rapidement.

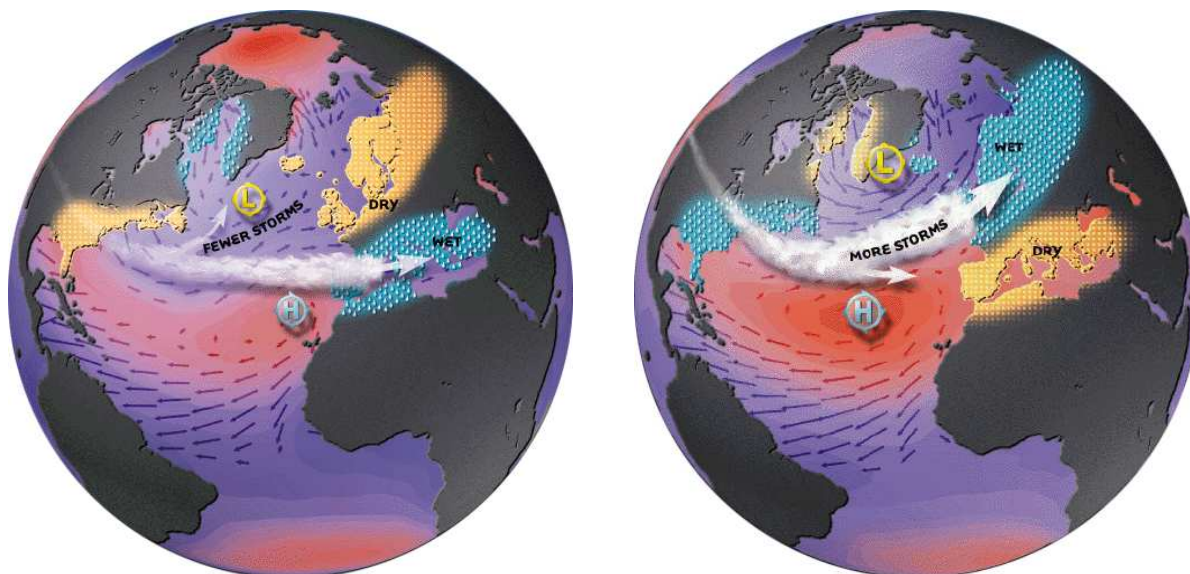
Les populations situées en limite d'aire de répartition sont donc particulièrement instables et sensibles aux fluctuations environnementales. Par conséquent, leur pérennité dépendra en partie d'évènements stochastiques et/ou catastrophiques. Or, les évènements catastrophiques naturels ayant un impact fort sur la distribution des espèces à la surface du globe sont rares (extinctions majeures, glaciations, fermeture du passage Atlantique/Pacifique en Amérique Centrale...). A l'opposé, les évènements catastrophiques impliquant l'homme et ayant des conséquences néfastes directes sur les communautés (Vitousek *et al.*, 1997) sont de plus en plus fréquents, l'homme étant responsable de la plupart des altérations évolutives et des extinctions d'espèces récentes (Palumbi, 2001).

2

L'influence de l'homme sur l'étendue des aires de distribution

2.1 Changements climatiques et déplacements directionnels

Le terme “*changement global*” désigne l'ensemble des modifications physiques, environnementales et écologiques profondes causées par les actions humaines. L'empreinte la plus profonde et probablement la plus durable concerne le “*réchauffement global*” et ses conséquences, déjà perceptibles (voir à ce sujet l'abondante littérature publiée par l'Intergovernmental Panel on Climate Change, dont IPCC, 2001). Les émissions humaines massives de gaz à effet de serre dans l'atmosphère terrestre font par exemple augmenter la fréquence et l'intensité des phénomènes ENSO (El Niño Southern Oscillation) influençant le système climatique global (Timmermann *et al.*, 1999). Dans l'Atlantique nord, le même phénomène d'oscillations climatiques pluriannuelles existe et porte le nom d'Oscillation Nord Atlantique (NAO, voir figure 2.1). La mesure de l'indice NAO, dont la valeur dépend des différences de pressions atmosphériques (mesurées au niveau de la mer) entre les Açores et l'Islande, permet de déterminer quel régime climatique affectera une partie de l'Amérique du Nord, l'Atlantique Nord, l'Europe et l'Afrique du Nord. Comme l'illustre la figure 2.1a, les années dont l'indice NAO est négatif seront marquées par des hivers rigoureux, plus froids et secs que la normale dans le nord de l'Europe, alors que le sud de l'Europe, la Méditerranée et le nord de l'Afrique recevront la majorité des précipitations. A l'inverse, un indice NAO positif sera le signe d'un hiver plus doux dans le nord de l'Europe, marqué cependant par des précipitations accrues et des tempêtes plus fréquentes (figure 2.1b).



a. Indice NAO négatif

b. Indice NAO positif

FIG. 2.1: Illustration des conséquences des “Oscillations Nord Atlantique” (North Atlantic Oscillations, NAO) sur la distribution des tempêtes et des précipitations en Amérique du Nord, en Europe et dans les pays du Maghreb. Source : <http://www.ldeo.columbia.edu/NAO> par Martin Visbeck.

Tout comme El Niño, le phénomène d'Oscillations Nord Atlantique est également affecté par le réchauffement global (donc par les émissions humaines de gaz à effet de serre). Des séries temporelles ont en effet permis de montrer que depuis 1864, l'indice NOA alterne régulièrement entre des états positifs et négatifs (voir figure 2.2). Mais depuis le début des années 1980, cet indice est subitement entré dans une phase positive d'une intensité et d'une longévité jamais observée par le passé (Hurrell, 1995), phase qui se poursuit encore aujourd'hui.

Ce phénomène à grande échelle se traduit par des modifications à plus petite échelle. Ainsi, au cours de cette période, la température des eaux de surface du Golfe de Gascogne (situé dans l'Atlantique nord-est) a augmenté en moyenne de 0,6 à 0,8 °C par décennie, soit selon les estimations, une augmentation de 1,8 à 2,4 °C depuis les années 70 (IPCC, 2001). Or, un changement même modéré des conditions climatiques peut s'avérer avoir des conséquences marquées sur les écosystèmes. Taylor *et al.* (2002) montrent en effet que des perturbations environnementales subtiles peuvent être grandement amplifiées par les communautés biologiques, notamment en raison de la non-linéarité de la réponse de ces communautés aux modifications environnementales. L'augmentation moyenne de la température mondiale au cours des 100 dernières années, moins spectaculaire que celle observée dans le Golfe de Gascogne depuis les années 70 (0.6 °C), devrait cependant se

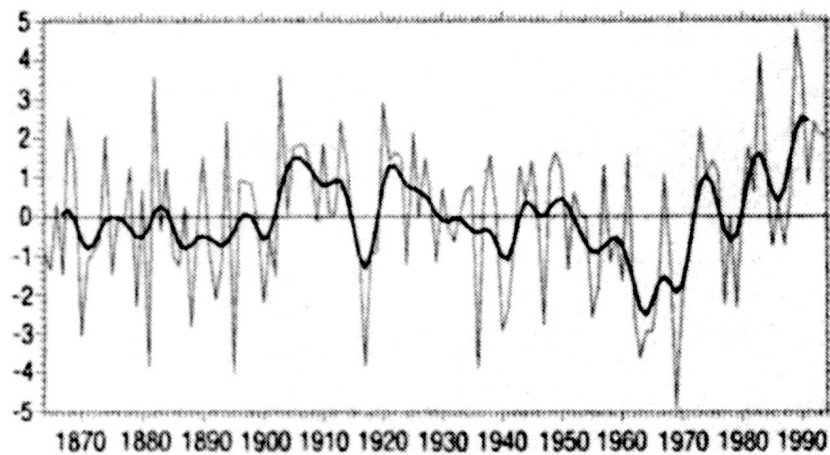


FIG. 2.2: Évolution de l'indice NAO depuis 1864. D'après Hurrell (1995)

poursuivre et s'accroître (Root *et al.*, 2003).

Les anomalies climatiques observées au cours des dernières décennies affectent déjà les communautés, la physiologie, la phénologie et la distribution des espèces (Hughes, 2000). Une étude de méta-données réalisée sur 143 espèces a ainsi révélé que des espèces allant des mollusques aux mammifères et des plantes herbacées aux végétaux supérieurs ont vu leur aire de distribution modifiée sur cette période en rapport avec les variations de température. Dans plus de 80% des cas, les déplacements observés sont cohérents avec les attendus théoriques, compte tenu des contraintes physiologiques des espèces (Root *et al.*, 2003). Parmesan et Yohe (2003) ont réalisé une étude similaire sur un panel de 1700 espèces et ont montré que les réponses biologiques sont en accord avec les prédictions de changements climatiques (déplacement de l'aire de distribution des espèces vers les pôles ou vers l'altitude, apparitions plus précoces d'évènements marquant l'arrivée du printemps comme, par exemple, le retour d'espèces migratrices...). Ces déplacements d'espèces ne sont pas des phénomènes locaux, et sont, au contraire observés, aussi bien sur les côtes rocheuses du Pacifique (Barry *et al.*, 1995) que de l'Atlantique (Mieszkowska *et al.*, 2006). Enfin, des modifications majeures de la structure, l'abondance, la distribution et la composition spécifique des communautés planctoniques marines sont déjà nettement visibles en Atlantique Nord (Hays *et al.*, 2005). Ces communautés étant à la base des réseaux trophiques, la plupart des espèces marines peuvent potentiellement en être affecté.

2.2 Réductions : pollutions et destruction des habitats

A des échelles de temps plus rapides que le changement climatique, de nombreuses actions humaines ont pour effet de restreindre l'étendue de l'aire de distribution des espèces.

Cela est notamment visible en zone côtière : d'après Cohen *et al.* (1997), entre 37 et 60% de la population mondiale vit à moins de 100 kilomètres du littoral. Les rejets urbains contribuent donc fortement à la pollution des milieux aquatiques (Vitousek *et al.*, 1997), ayant pour conséquence majeure l'eutrophisation récurrente en zone littorale (Nixon, 1995). Les activités maritimes (plaisance, transports commerciaux) sont également source de pollutions. Outre les événements catastrophiques du type marée noire, ayant un impact ponctuel fort sur de nombreuses espèces (voir par exemple Jewett *et al.*, 1999), les rejets chroniques de polluants, même en faible quantité, peuvent avoir des conséquences importantes. Ainsi, le tributylétain (TBT), présent dans les peintures anti-salissures des coques de bateaux ou d'ouvrages immergés, a été largement utilisé en France jusqu'en 1982, date où son utilisation a été interdite sur les navires de moins de 25 mètres (voir la page 3 du Journal Officiel C 176 du 14.7.1982). Le TBT, qui perturbe le système endocrinien de nombreuses espèces marines, est responsable de phénomènes d'imposex (masculinisation des individus femelles), notamment chez les gastéropodes (Bryan *et al.*, 1986, 1989; Gooding *et al.*, 1999). Cette substance, même à des concentrations très faibles dans le milieu (0,5 ng/L), peut également être responsable d'une sensibilité accrue aux maladies (Anderson *et al.*, 1996), de défauts de calcification chez certains bivalves (Alzieu, 1998) et a entraîné le déclin de populations entières du gastéropode *Nucella lapillus* (Bryan *et al.*, 1986).

Enfin, la destruction (ou la fragmentation) des habitats par l'homme est considérée comme la première cause de diminution de la biodiversité mondiale et même les espèces dominantes des écosystèmes à l'heure actuelle semblent, à terme, menacées (Tilman *et al.*, 1994). Ainsi, les marais côtiers qui jouent un rôle de tampon entre les zones marines et terrestres ont été largement altérés : à l'échelle mondiale, environ 50% des zones de mangroves ayant été transformées ou détruites par les activités humaines (Vitousek *et al.*, 1997). De même, la construction de marinas, zones portuaires, digues, enrochements, modifie souvent profondément la courantologie et la sédimentologie des zones impliquées. Les espèces exigeantes en terme de substrat disparaissent ainsi purement et simplement de ces zones (Walker et Kendrick, 1998).

2.3 Les invasions biologiques : un cas particulier d'expansion

Certaines actions humaines ont tendance à supprimer des barrières naturelles permettant à des espèces incapables de les franchir spontanément d'étendre leur aire de

distribution. La mise en contact d'espèces ne l'ayant jamais été auparavant ou ayant été séparées depuis des milliers, voire des millions d'années est ainsi provoquée par l'homme. Par exemple, sans l'ouverture du canal de Suez et les migrations dites "*Lessepsiennes*" (migrations actives ou passives d'espèces depuis la Mer Rouge vers la Méditerranée via le Canal de Suez), de nombreuses espèces aujourd'hui en contact ne se seraient probablement jamais rencontrées (Por, 1971).

De même, les introductions biologiques peuvent avoir pour conséquence l'augmentation de l'aire de distribution d'une espèce. Une introduction biologique se définit par l'apparition soudaine et imprévisible dans un écosystème d'une espèce provenant d'une aire géographique généralement lointaine (Carlton, 1987, 1996b; Williamson, 1996). L'augmentation de l'aire de distribution de l'espèce introduite se fait la plupart du temps au détriment de l'aire de distribution des espèces locales. En effet, à ce terme d'introduction biologique, la littérature préfère généralement celui d'invasion biologique, en particulier lorsque l'espèce introduite s'avère nuisible, plus précisément quand elle représente un danger pour la biodiversité naturelle, le fonctionnement des écosystèmes, l'économie ou la santé humaine (Myers *et al.*, 2000).

Le rythme de ces invasions s'est considérablement accéléré au cours des dernières décennies (Williamson, 1996; Cohen et Carlton, 1998; Ruiz *et al.*, 1997) en raison notamment de l'intensification des échanges transcontinentaux, maritimes ou terrestres (mise en place de corridors d'invasions Ricciardi et MacIsaac, 2000; Carlton, 1987), ainsi que de l'exploitation intensive des écosystèmes (Sakai *et al.*, 2001) avec par exemple des introductions volontaires pour les cultures ou les élevages, et l'introduction accidentelle d'espèces associées à ces espèces cultivées. Tous les milieux et la plupart des taxons sont concernés (Reise *et al.*, 1999) :

- en milieu terrestre sont par exemple introduits des végétaux (*e.g.* *Helianthus annuus*, *Senecio vulgaris*, *Iris nelsonii*, *Tragopogon mirus*, *Lonicera japonica*, *Betula pendula* Abbott, 1992; Keane et Crawley, 2002), des insectes (*e.g.* *Ceratitis capitata*, *Linepithema humile* Davies *et al.*, 1999; Tsutsui *et al.*, 2000), ainsi que des oiseaux, amphibiens et mammifères (Shigesada et Kawasaki, 1997; D'Antonio et Kark, 2002)...
- en milieu dulçaquicole, le cas de la colonisation des grands lacs américains par des espèces européennes est particulièrement frappant. Selon Ricciardi et MacIsaac (2000), 145 espèces exogènes y ont été introduites au cours des dernières décennies parmi lesquelles des poissons (*e.g.* *Neogobius melanostomus*, *Proterorhinus marmoratus*, *Alosa aestivalis*), des crustacés (*e.g.* *Echinogammarus ischnus*) et des mollusques (*e.g.* *Dreissena polymorpha*, *Dreissena bugensis*, *Potamopyrgus antipodarum*).

- en milieu marin, des échantillons de plancton prélevés en Oregon dans les eaux de ballast de navires en provenance du Japon, contenaient un total de 367 taxons (Carlton et Geller, 1993). Wasson *et al.* (2001) ont recensé 56 espèces d'invertébrés exotiques (crustacés, mollusques, cnidaires, spongiaires, plathelminthes, annélides) dans la baie d'Elkhorn Slough (États-Unis). D'après Gouletquer *et al.* (2002), au moins 104 espèces ont été introduites le long des côtes françaises de l'Atlantique et de la Manche. De même, selon Carlton et Geller (1993), le nombre d'espèces exogènes présentes dans les estuaires nord américains s'élèverait à 525 (tous taxons confondus : crustacés, chélicérates, échinodermes, chordés, bryozoaires, annélides, plathelminthes némerthes, mollusques, nématodes, cnidaires, radiolaires, dinoflagellés, diatomées, chlorophytes, rhodophytes...).

Everett (2000) estime que 50 000 espèces exogènes, tous milieux confondus, sont établies aux États-Unis, induisant un coût annuel à près de 137 milliards de dollars. L'ampleur des invasions biologiques et leurs répercussions sur les espèces du point de vue évolutif (Mooney et Cleland, 2001; Palumbi, 2001; Grosholz, 2002) sont telles que ce phénomène est aujourd'hui considéré comme la deuxième cause de diminution de la biodiversité mondiale après la fragmentation et la destruction des habitats (Coblentz, 1990; Cohen et Carlton, 1998; Shea et Chesson, 2002). De plus, la liste des espèces exogènes recensées est largement sous estimée. La plupart des études concerne en effet des espèces introduites de taille macroscopique ou pour lesquelles la systématique est clairement définie. Pour cette raison, un grand nombre de taxons ou de phylums ne sont pas, ou peu, étudiés (Ruiz *et al.*, 1997; Reise *et al.*, 1999). Ainsi, les algues unicellulaires, les protozoaires ou les petits métazoaires sont, par exemple, sous-représentés dans les inventaires d'espèces invasives.

Si pour les espèces déplacées par l'homme, l'introduction représente une expansion géographique, les conséquences potentielles sur les écosystèmes n'ont rien de comparable avec celles d'une expansion naturelle. Les invasions biologiques sont en effet un facteur rapide d'érosion de la biodiversité et d'homogénéisation des biotopes à l'échelle mondiale, tandis que les expansions naturelles sont généralement progressives, rythmées par des processus géologiques (Futuyama, 1986), se font à l'échelle des communautés et sont un moteur de l'évolution (Williamson, 1996; Mooney et Cleland, 2001; Palumbi, 2001). Il est donc important de pouvoir distinguer introductions humaines et expansions naturelles.

3

Des situations peu claires ; le cas atypique de *Cyclope neritea*

3.1 Espèce native ou introduite ? Un statut parfois ambigu

Lorsqu'une espèce est repérée dans une zone éloignée de son aire d'origine, et que son apparition est récente, le statut d'espèce introduite par l'homme peut être établi de façon quasi-certaine. Cette situation correspond à la très grande majorité des invasions biologiques recensées et étudiées.

Il existe toutefois des situations où la détermination d'un statut clair pour une espèce nouvellement repérée dans une zone n'est pas sans poser quelques problèmes. Ainsi, sans informations historiques concernant la présence ou l'absence d'une espèce donnée dans une zone particulière, comment savoir si l'espèce en question est native de cette zone ou non ? Considérons par exemple une espèce occupant deux zones distinctes et isolées géographiquement. L'espèce a-t-elle été récemment introduite dans l'une des deux aires qu'elle occupe, ou est-elle au contraire native de l'ensemble de la zone, la discontinuité de l'aire de distribution actuelle étant apparue à la suite de la disparition de l'espèce en certains endroits ? Ces espèces, dont l'origine et le statut sont incertains (*i.e.* espèces natives ou introduites), sont appelées “*espèces cryptogéniques*” par Carlton (1996a).

Les exemples d'espèces cryptogéniques sont nombreux. C'est le cas par exemple du diptère nuisible *Bactrocera depressa*, présent de longue date au Japon et détecté en 1974 en Corée. La comparaison des inventaires faunistiques actuels et anciens suggèrent que cette espèce a été introduite récemment en Corée, probablement depuis le Japon. Cependant, les résultats d'analyses génétiques semblent incompatibles avec ces inventaires et renforcent

plutôt l'hypothèse de la présence historique de l'espèce à l'état cryptique (Mun *et al.*, 2003).

La même question est toujours en débat concernant la présence du gastéropode *Littorina littorea* des deux côtés de l'Atlantique Nord. La présence de cette espèce native des côtes européennes n'a été relevée le long des côtes nord américaines, que ponctuellement avant 1840, et est très peu documentée. Si l'on suppose que cette espèce était présente en Amérique du Nord avant l'arrivée des premiers colons européens, son aire de distribution devait être relativement restreinte (*e.g.* limitée à certaines zones du Canada) et son abondance particulièrement faible compte tenu du petit nombre de coquilles retrouvées dans les gisements archéologiques (voir Vermeij, 1982, et les références qui y sont citées). Or, cette espèce a vu son aire de distribution nord américaine s'étendre rapidement vers le sud depuis la seconde moitié du XIX^e siècle. Cette expansion est-elle le fruit d'une recolonisation naturelle tardive depuis les côtes européennes à la suite du dernier maximum glaciaire du pléistocène, d'une introduction humaine récente (comme semble le penser Carlton, 1999), ou l'espèce a-t-elle toujours été présente sur les côtes américaines (les analyses génétiques et les simulations réalisées par Wares *et al.*, 2002, montrant par exemple que l'espèce était présente il y a au moins 8000 ans, et que son expansion récente doit être le fruit de modifications environnementales) ?

Un autre cas de figure concerne l'expansion rapide et récente d'une espèce dans le prolongement de son aire de distribution. Le caractère instable des populations situées en bordure de l'aire de distribution d'une espèce (en raison du processus de “*migration contre sélection*” décrit dans la partie 1.2.3), souvent associé à une absence de données historiques concernant la présence ou l'absence de l'espèce en question compliquent souvent l'interprétation. Le gastéropode marin *Cyclope neritea*, notre modèle biologique, relève précisément de cette situation.

3.2 Le modèle d'étude : *Cyclope neritea*

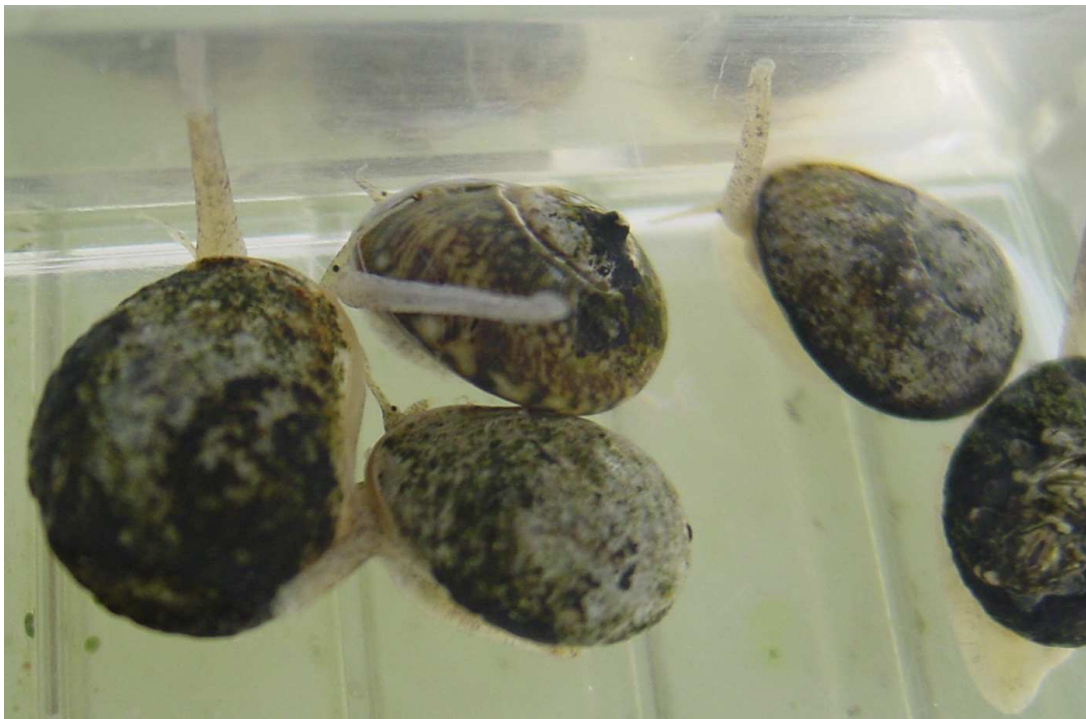
3.2.1 Biologie de l'espèce

Cyclope neritea (L., 1758) est un petit gastéropode nassariidae² marin (10–15mm) à coquille aplatie (voir figure 3.1) inféodé aux substrats sablo-vaseux caractéristiques des zones littorales abritées (Tardy *et al.*, 1985). Cette espèce euryhaline vit en eaux peu profondes, dans les lagunes, les estuaires abrités et, dans la partie atlantique de son aire

²Position systématique : Mollusca, Gastropoda, Prosobranchia, Apogastropoda, Caenogastropoda, Buccinoidea, Nassariidae



a. Dans leur milieu naturel



b. En aquarium

FIG. 3.1: Spécimens actifs de *Cyclope neritea*.

de distribution, en zone intertidale (voir l'aire de distribution de *Cyclope neritea* telle qu'elle était reconnue jusque dans les années 1950 à la figure 3.2).

Son cycle de vie est entièrement benthique (Gomoiu, 1964; Boulhic et Tardy, 1986b) : les femelles adultes (qui atteignent, comme les mâles, la maturité sexuelle à environ un an et vivent jusqu'à plus de deux ans, Le Roux, 1994) déposent des capsules contenant un unique embryon sur des substrats durs, de préférence près de l'interface eau/sédiment afin que les capsules soient bien oxygénées (sur ses congénères, ou sur les valves de *Cerastoderma edule*, cette espèce ayant un siphon relativement court, elle ne s'enfonce que peu profondément dans le sédiment, Boulhic et Tardy, 1986b). En aquarium, les œufs sont ainsi quasi-systématiquement déposés sur les parois des bacs, à l'interface air/eau (observation personnelle). La période de ponte s'étend sur presque toute l'année, avec un ou plusieurs pics printaniers et une légère diminution à la fin de l'automne et au cours des mois d'hiver (Le Roux, 1994). A l'issue d'une période d'incubation de 36 jours (Boulhic et Tardy, 1986b), l'éclosion laisse apparaître des juvéniles benthiques autonomes qui semblent s'alimenter pendant quelques jours sur le reste de leurs réserves vitellines et qui acquièrent rapidement le même mode d'alimentation que leurs parents : la nécrophagie.

Cyclope neritea, qui vit la plupart du temps enfouie dans le premier centimètre du sédiment, a des récepteurs sensoriels particulièrement développés qui lui permettent de repérer rapidement et à distance ses proies potentielles. Les individus de cette espèce sont en outre capable de se désenfouir en quelques secondes seulement pour se diriger vers leurs proies. Ce gastéropode particulièrement véloce pourrait ainsi, dans certaines conditions, être plus rapide que la nasse réticulée *Nassarius reticulatus* et constituer un compétiteur efficace dans les zones où leur distribution se chevauche (Bachelet *et al.*, 2004, article présenté en annexe D.1). Par ailleurs, en l'absence de proie, les individus de cette espèce semblent également capables de brouter le film microalgal se développant à la surface du sédiment, du moins les plus jeunes individus dont le système sensoriel n'est pas encore tout à fait fonctionnel dans les premiers jours suivant l'éclosion (Boulhic et Tardy, 1986a).

Certaines caractéristiques biologiques de *Cyclope neritea* impliquent, théoriquement, de faibles capacités de dispersion. Cette espèce n'ayant pas de stade larvaire planctonique, la dispersion rapide à grande distance ne doit donc pas être très fréquente. De même, *Cyclope neritea* étant inféodée aux zones sablo-vaseuses, les discontinuités, en terme de substrat le long des côtes où cette espèce est présente, doivent freiner les échanges entre populations ainsi que la dispersion naturelle. Ce dernier point explique probablement en partie pourquoi cette espèce ne présente pas une distribution continue le long des côtes de son aire naturelle mais se trouve plutôt à l'état de "*patches*" (ou îlots de populations), au sein desquels les densités individuelles peuvent atteindre des niveaux impressionnants

(de 300 à plus de 1600 individus par mètre carré ! Tardy *et al.*, 1985) alors qu'elle reste invisible quelques centaines de mètres plus loin (observation personnelle).

3.2.2 Aire de distribution, problématique

L'aire de distribution "*historique*" de *Cyclope neritea* comprend les côtes de la Mer Noire, de la Méditerranée, et les côtes atlantiques du Maroc, du sud du Portugal et de l'Espagne (Sauriau, 1991, voir figure 3.2). Des populations discontinues et éphémères ont été observées en Atlantique des côtes portugaises jusqu'au Pays Basque français, Socoa étant en 1950 la limite septentrionale reconnue de cette espèce (Sauriau, 1991; Kisch, 1950).



FIG. 3.2: Aire de distribution de *Cyclope neritea* telle qu'elle était reconnue jusqu'en 1950.

Depuis 1976, plusieurs populations sont apparues graduellement du sud vers le nord, le long des côtes de l'Atlantique et de la Manche (Bachelet *et al.*, 1980; Tardy *et al.*, 1985; Pigeot, 1988; Sauriau, 1989; Le Roux *et al.*, 1988; Simon-Bouhet *et al.*, 2006). En moins

de trente ans, cinq importantes baies ont ainsi été colonisées, entre le Bassin d’Arcachon et la Baie de Morlaix (nous reviendrons plus en détail sur l’aire de distribution de *Cyclope neritea* dans les parties suivantes). Ainsi, la zone géographique récemment colonisée se trouve dans le prolongement direct de l’aire de distribution historique. De telles zones sont connues pour abriter des populations potentiellement instables (voir section 1.2.3), pouvant apparaître spontanément, subsister à faible densité et disparaître naturellement. Il est par conséquent difficile de savoir si *Cyclope neritea* est présente de longue date dans cette zone mais est restée indétectée jusqu’en 1976, ou si au contraire les populations découvertes sont réellement implantées depuis peu, ou encore si l’augmentation des densités, et donc la découverte de *Cyclope neritea*, est le fruit de l’action conjointe de ces deux processus. Sur les côtes atlantiques françaises, cette espèce correspond donc tout à fait à la définition d’espèce cryptogénique de Carlton (1996a).

A supposer que l’espèce soit arrivée récemment sur les côtes françaises, l’expansion rapide vers le nord peut avoir plusieurs causes.

1. Comme nous l’avons vu à la section 2.1, le réchauffement global est responsable du déplacement des limites de l’aire de distribution de nombreuses espèces. Ces déplacements ont lieu, soit vers l’altitude pour certaines espèces terrestres, soit en direction des pôles (Parmesan et Yohe, 2003). Or, le réchauffement des eaux de surface du Golfe de Gascogne a été extrêmement rapide depuis les années 70 (de 0,6 à 0,8 °C par décennie, IPCC, 2001). Il n’est donc pas exclu que *Cyclope neritea*, suite à un relâchement des contraintes environnementales ayant freiné les migrations naturelles vers le nord jusque dans les années 70, ait été en mesure d’étendre naturellement son aire de distribution.
2. Toutefois, compte tenu des capacités de dispersion intrinsèques limitées de *Cyclope neritea*, et des nombreuses discontinuités de substrat peu favorables à l’espèce (en particulier le long des côtes landaises), une hypothèse alternative doit être envisagée : tous les sites français où l’espèce a été repérée depuis 1976 sont des bassins de production conchylicole importants (Pigeot, 1988; Sauriau, 1989), et *Cyclope neritea* est pratiquement toujours observée à proximité de parcs ostréicoles, parfois en très forte densité. Il est donc possible que les transferts humains d’espèces cultivées (*e.g.* huîtres, palourdes...), entre l’aire d’origine de ce gastéropode et les sites récemment colonisés, soient à l’origine de l’installation durable de *Cyclope neritea* sur les côtes françaises.

La situation atypique de *Cyclope neritea* nous donne donc l’opportunité d’étudier un cas particulier d’expansion en limite d’aire. Clarifier le statut des populations récemment découvertes de cette espèce (*i.e.* présence ancestrale *vs.* récente, introduction acciden-

telle *vs.* expansion naturelle) devrait nous permettre de mieux comprendre les processus en action en bordure d'aire de distribution. Pour cela, nous avons principalement utilisé une approche d'écologie moléculaire. En complément d'approches plus classiques (en particulier des suivis de terrain), les outils moléculaires et analytiques de la génétique des populations, en étudiant la distribution de la variabilité génétique au niveau intra et inter-populationnel, sont susceptibles d'apporter des indications sur l'histoire démographique et évolutive de ces populations (Sakai *et al.*, 2001).

Ce travail de thèse s'articule autour de trois parties distinctes répondant chacune à des objectifs précis :

1. une étude sera menée dans la zone historiquement reconnue comme étant la limite de distribution naturelle de l'espèce, c'est à dire, la zone au sein de laquelle des populations "*instables*" ont déjà été enregistrées par le passé. Cette partie de l'étude (partie I) concernera toute la façade atlantique de la péninsule ibérique, du Détroit de Gibraltar au Pays Basque français. Dans la mesure où parler d'expansion naturelle ou d'introduction nécessite d'avoir une bonne connaissance des populations potentiellement sources de migrants, le but principal de cette partie sera de clarifier la situation des populations dans cette zone et d'identifier les mécanismes récents ayant pu influencer la distribution et la stabilité des populations s'y trouvant. Nous montrerons que la détermination précise de la limite de l'aire géographique de cette espèce a, par le passé, été hasardeuse, en raison, d'une part, de la grande instabilité des populations de cette zone, et d'autre part, de la difficulté à effectuer des inventaires faunistiques pertinents (notamment à cause de confusions possibles entre plusieurs espèces du genre *Cyclope*). Ce dernier point nous amènera à présenter la mise au point d'un marqueur moléculaire diagnostique de notre espèce.
2. une analyse des populations situées dans l'aire naturelle de l'espèce sera ensuite conduite (partie II). Par une approche phylogéographique et de génétique des populations, basée sur des marqueurs mitochondriaux et microsatellites développés au cours de cette thèse, nous testerons l'hypothèse d'une forte structure génétique au sein de l'aire de distribution naturelle de l'espèce, structure génétique attendue compte tenu des capacités de dispersion limitées de *Cyclope neritea*.
3. enfin, la partie III concernera l'étude des populations françaises récemment découvertes le long des côtes de l'Atlantique et de la Manche ainsi que les populations dont le statut de population "*naturelle*" n'aura pu être établi avec certitude (il s'agit en particulier de populations situées près de la frontière franco-espagnole du Pays Basque). Cette étude nous permettra de préciser le statut de ces populations à la lumière des résultats obtenus dans les deux parties précédentes

Cette thèse reposant en grande partie sur les techniques de la biologie moléculaire, l'annexe A regroupe les descriptions de l'ensemble des méthodes et protocoles employés pour la mise au point et l'utilisation en routine des marqueurs moléculaires.

Sont également présentés en annexe, les données génétiques brutes obtenues avec les différents marqueurs moléculaires dans l'ensemble des populations étudiées (annexe B), le principe de fonctionnement et quelques captures d'écran d'un programme de simulation en partie développé et utilisé au cours de cette thèse (annexe C), ainsi que les articles scientifiques publiés au cours de ce travail de thèse (annexe D).

Première partie

Cyclope neritea en limite d'aire : la
façade atlantique de la péninsule
ibérique

Pourquoi étudier cette zone ?

L'un des objectifs principaux de cette thèse est d'évaluer si des transports humains de *Cyclope neritea* sur les côtes atlantiques françaises peuvent expliquer son apparition depuis les années 70. Or, pour étudier les processus d'introductions biologiques, en particulier les schémas d'introductions primaires³, il est nécessaire de pouvoir analyser les populations "*supposées introduites*" en regard des populations "*supposées naturelles*" et situées dans les régions potentiellement sources. Ce constat nous a amenés à nous pencher sur l'aire de distribution de *Cyclope neritea* qui jouxte la région d'apparition récente des populations françaises, c'est-à-dire les façades atlantiques de l'Espagne et du Portugal (voir figure 1).

La première population détectée sur le littoral atlantique français depuis les années 50 (*i.e.* le bassin d'Arcachon en 1976, Bachelet *et al.*, 1980) se situe en effet à moins de 200 kilomètres de la frontière franco-espagnole, zone historiquement reconnue comme la limite septentrionale de *Cyclope neritea* en Atlantique (Sauriau, 1991). Les limites d'aires étant souvent des zones où les populations se caractérisent par leur instabilité (Brown *et al.*, 1996), l'une des questions est donc celle du statut des populations situées en limite apparente d'aire de distribution. En effet, comment parler d'expansion si les limites naturelles antérieures à l'expansion ne sont pas clairement identifiées ? Ou au contraire, pourquoi faire l'hypothèse d'introductions humaines s'il n'est pas clairement établi que *Cyclope neritea* ne pouvait se trouver à l'état naturel dans les sites récemment colonisés ? Un descriptif précis de cette zone est donc indispensable dans la mesure où l'expansion naturelle, si elle a eu lieu, pourrait trouver sa source dans ces populations.

Dans cette partie, nous tenterons donc dans un premier temps de déterminer si les populations de la zone historiquement reconnue comme étant la limite septentrionale de l'aire de distribution de *Cyclope neritea* sont instables. De telles études sont rares pour les espèces marines à développement direct, et nous verrons, grâce à des comparaisons entre la bibliographie traitant de cette distribution et à notre propre expérience de terrain que

³Introduction initiale d'individus en provenance de l'aire d'origine de l'espèce, par opposition aux introductions secondaires, impliquant le transport d'individus issus de populations déjà installées dans l'aire considérée.



FIG. 1: Les trois zones de l'aire de distribution de *Cyclope neritea* : (i) l'aire de distribution naturelle "historique" de l'espèce (trait de côte surligné en jaune) qui comprend notamment les côtes de la Mer Noire et de la Méditerranée, (ii) la zone d'expansion récente (naturelle et/ou anthropique) le long des côtes françaises de l'Atlantique et de la Manche Occidentale (trait de côte surligné en rouge) et (iii), la zone intermédiaire comprise entre Gibraltar et le Pays Basque français, correspondant à la zone instable en limite de l'aire de distribution historique (trait de côte surligné en bleu). Ne disposant pas d'informations précises sur la présence historique de *Cyclope neritea* sur les côtes marocaines, cette zone figure en blanc sur le carte.

la situation dans cette zone, loin d'être figée, évolue au contraire en permanence.

Dans un second temps, nous nous attacherons à distinguer morphologiquement *Cyclope neritea* d'une autre espèce du même genre : *Cyclope pellucida*. Nous verrons que dans certains sites, les deux espèces cohabitent et qu'il est parfois difficile de les distinguer, au moins pour les plus jeunes individus. La discrimination de ces deux espèces est un point critique pour déterminer avec précision la limite de l'aire de distribution de *Cyclope neritea*. Nous présenterons donc la mise au point d'une méthode moléculaire permettant de faire la distinction entre les deux espèces.

4

Instabilité des populations de *Cyclope neritea* en limite d'aire

4.1 Rappel des processus en cause

Les forces évolutives opposées (migration d'une part et sélection d'autre part) qui s'affrontent en limite d'aire de répartition orientent généralement les flux de gènes nets au sein de l'aire de distribution d'une espèce du centre vers la périphérie (Kirkpatrick et Barton, 1997). Deux cas de figure peuvent alors être rencontrés (Holt, 2003) :

1. les conditions environnementales en limite d'aire sont relativement stables dans le temps. Dans ce cas, l'arrivée régulière de migrants dans les populations périphériques a tendance à contrer les processus d'adaptations locales aux conditions limites régnant en bordure d'aire de distribution.
2. les conditions environnementales sont fluctuantes en limite d'aire. Dans ce cas, les populations périphériques ne doivent leur maintien qu'à l'envoi régulier de migrants depuis le centre de l'aire de distribution, dans la mesure où des adaptations locales n'ont pas le temps de se mettre en place.

Dans un cas comme dans l'autre, les populations situées en limite d'aire souffrent de maladaptation, ce qui explique en partie leur instabilité.

Cyclope neritea ne devrait pas échapper à cette règle d'instabilité des populations en limite d'aire. Au contraire : de par sa biologie, cette instabilité devrait même être exacerbée. Comme nous l'avons vu, *Cyclope neritea* est (i) un gastéropode à développement direct (ii) inféodé à un type de substrat particulier. Sa capacité de dispersion à grande distance est donc moins importante que celle d'organismes ayant un stade pélagique au cours de leur cycle de vie (Boulhic et Tardy, 1986b). De même, les espèces plus tolérantes que *Cy-*

clope neritea en terme de substrat auront une distribution moins discontinue. Autrement dit, chez *Cyclope neritea*, l'envoi de migrants vers les populations situées à la périphérie de l'aire de distribution doit être rare comparativement à des espèces plus “mobiles” et “tolérantes”. Ainsi, si l'installation de rares migrants échoue, la zone cible pourra alors être difficile et lente à recoloniser. Par contre, si l'établissement de migrants dans la nouvelle zone est un succès, la majorité des juvéniles recrutés localement seront issus des premiers migrants (Boulhic et Tardy, 1986b) et, au cours du temps, des processus d'adaptations locales pourront se mettre en place. A terme, la densité de cette population pourra augmenter rapidement (Tardy *et al.*, 1985, ont par exemple observé des densités de plus de 1600 individus par mètre carré dans la population de Rivedoux (Île de Ré) détectée pour la première fois 3 ans plus tôt). Ceci reste vrai tant que les conditions environnementales restent stables. Or, dans le Golfe de Gascogne, ces conditions ont drastiquement changé en quelques décennies, la meilleure illustration étant probablement l'augmentation moyenne de la température des eaux de surface de près de 2 ° C depuis les années 1970 (IPCC, 2001).

4.2 L'étude comparative

L'étude de la zone comprise entre le sud du Portugal et le Pays Basque français a été conduite par la confrontation :

1. des données bibliographiques concernant *Cyclope neritea*. L'analyse bibliographique repose en grande partie sur un article de revue de Sauriau (1991) rapportant la distribution de *Cyclope neritea*. L'auteur, en s'appuyant notamment sur plusieurs inventaires faunistiques réalisés entre 1900 (Fischer, 1899; Hidalgo, 1917) et 1987 (Borja, 1987), examine au sein de ces inventaires, la présence ou l'absence de l'espèce le long des côtes atlantiques espagnoles, portugaises et françaises.
2. des données que nous avons obtenues sur le terrain dans cette zone entre 2002 et 2004 au cours de deux campagnes d'échantillonnage. La première a été menée entre Hossegor et Bilbao (décembre 2002) et la seconde entre Bilbao et le Delta de l'Ebre (mars 2004).

L'échantillonnage a été réalisé en suivant le méthode décrite ci-dessous.

4.2.1 Technique d'échantillonnage

Cyclope neritea vit en eaux peu profondes et se cantonne aux zones sablo-vaseuses abritées du médio-littoral des estrans atlantiques, aux lagunes peu profondes ou aux estuaires

abrités (voir figure 4.1). L'espèce étant nécrophage, il suffit généralement, pour s'assurer de sa présence de déposer une proie quelconque (*e.g.* cadavre de crépidule, d'oursin, de bivalve, ou de poisson...) dans une flaque, un léger courant d'eau de l'estran à marée basse, ou dans quelques dizaines de centimètres d'eau en Méditerranée. Les capacités sensorielles ("*odorat*") particulièrement développées de *Cyclope neritea* ainsi qu'une vitesse peu commune pour un gastéropode, lui permettent en quelques minutes de détecter la proie, de se désenfouir du sédiment et de se diriger vers elle (voir figures 4.2).

L'efficacité de cette méthode de collecte extrêmement rapide a été testée afin de déterminer si elle ne favorisait pas le ramassage d'un type d'individus particulier : deux autres méthodes d'échantillonnage ont ainsi été comparées à cette technique.

1. tout d'abord, afin de s'assurer que la collecte, habituellement effectuée en quelques minutes ne sélectionne pas uniquement les individus les plus rapides, un cadrat de 4 mètres carrés interdisant l'entrée ou la sortie des individus a été réalisé, et un poisson mort y a été placé (voir figure 4.3). Tous les individus attirés par la proie ont été collectés, et ce pendant 90 minutes.
2. ensuite, afin de vérifier que la collecte habituelle ne sélectionne pas uniquement les individus les plus gros ou ceux ayant le meilleur "*odorat*", une zone équivalente à la surface du cadrat a été entièrement tamisée (sur maille de 500 μm) afin de collecter l'ensemble des individus présents.

Les mesures de longueur des individus obtenus par les trois techniques sont comparables : un test de Student (95%) ne peut rejeter l'hypothèse d'égalité des moyennes, ni entre la méthode du dépôt d'une proie et la méthode du tamisage ($p = 0,813$), ni entre la méthode du dépôt et la méthode du cadrat ($p = 0,062$). Les effectifs récoltés sont en outre plus importants avec la méthode du dépôt (170 individus récoltés contre 116 et 101 avec le cadrat et le tamisage respectivement). La méthode de collecte après dépôt d'une proie permettant d'obtenir un échantillon représentatif des populations et étant la plus rapide et la plus simple à mettre en œuvre, c'est celle qui a été utilisée pour l'échantillonnage de toutes les populations que nous avons récoltées au cours de cette étude.

Dans les populations denses, de 50 à 100 individus ont été collectés, généralement en quelques minutes sur une surface n'excédant pas quelques mètres carrés. Pour les populations les moins denses, bien que les premiers individus apparaissent généralement dès les toutes premières minutes suivant le dépôt de la proie, près d'une heure d'échantillonnage sur une surface de plusieurs dizaines de mètres carrées a parfois été nécessaire pour récolter une trentaine d'individus.

Enfin, les sites pour lesquels nous avons conclu à l'absence de *Cyclope neritea* entrent dans deux catégories principales :



a. Mar Menor, Espagne (Méditerranée)



b. Caorle, Italie (Adriatique)

FIG. 4.1: Exemples de sites où *Cyclope neritea* est présente en forte densité.

1. Dans les sites présentant un substrat susceptible de convenir à *Cyclope neritea* (milieu sablo-vaseux), l'absence de l'espèce a été confirmée par le dépôt d'une proie pendant 15 minutes sans qu'aucun individu n'ait été observé ainsi que par l'absence de coquilles vides dans les laisses de mer.
2. Dans les sites présentant un substrat très différent de l'habitat préférentiel de *Cyclope neritea* (sables battus trop grossiers, ou au contraire milieux très envasés), l'absence de la faune associée à *Cyclope neritea* était un bon indicateur de l'absence de l'espèce.

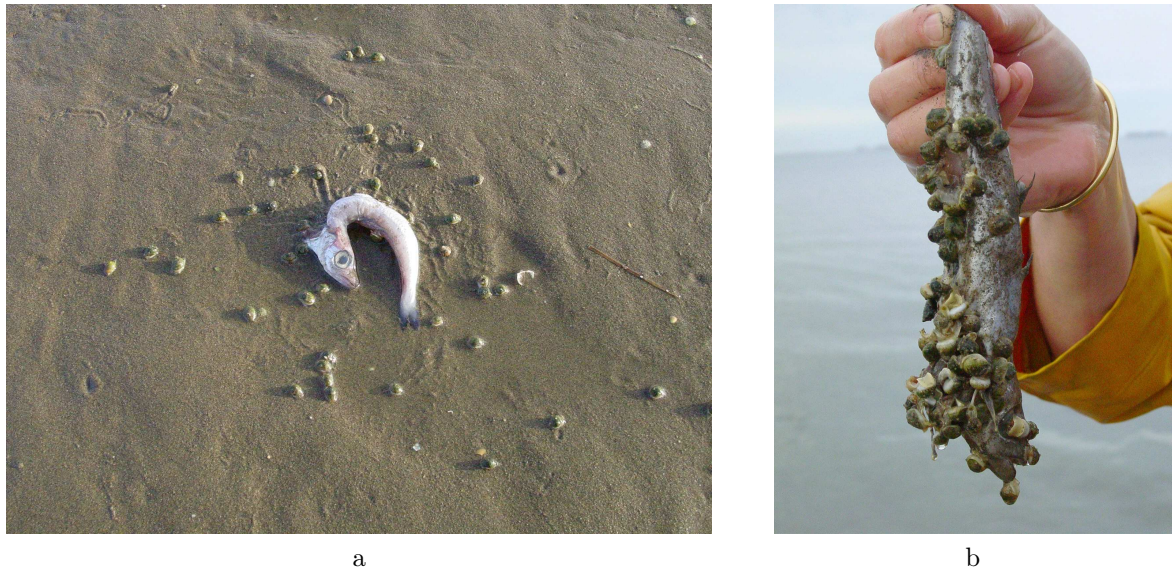


FIG. 4.2: Attraction de *Cyclope neritea* par une carcasse déposée sur le substrat (figure 4.2a) et dans quelques dizaines de centimètres d'eau (figure 4.2b), pendant moins de 3 minutes.



FIG. 4.3: Test de la collecte par la méthode du cadrat.

4.2.2 Mise en œuvre dans le cadre de l'étude comparative

La carte présentée en figure 4.4 indique les différents sites examinés. Dans la plupart de ces sites, nous avons inspecté de nombreuses stations d'échantillonnage, et ce, parti-

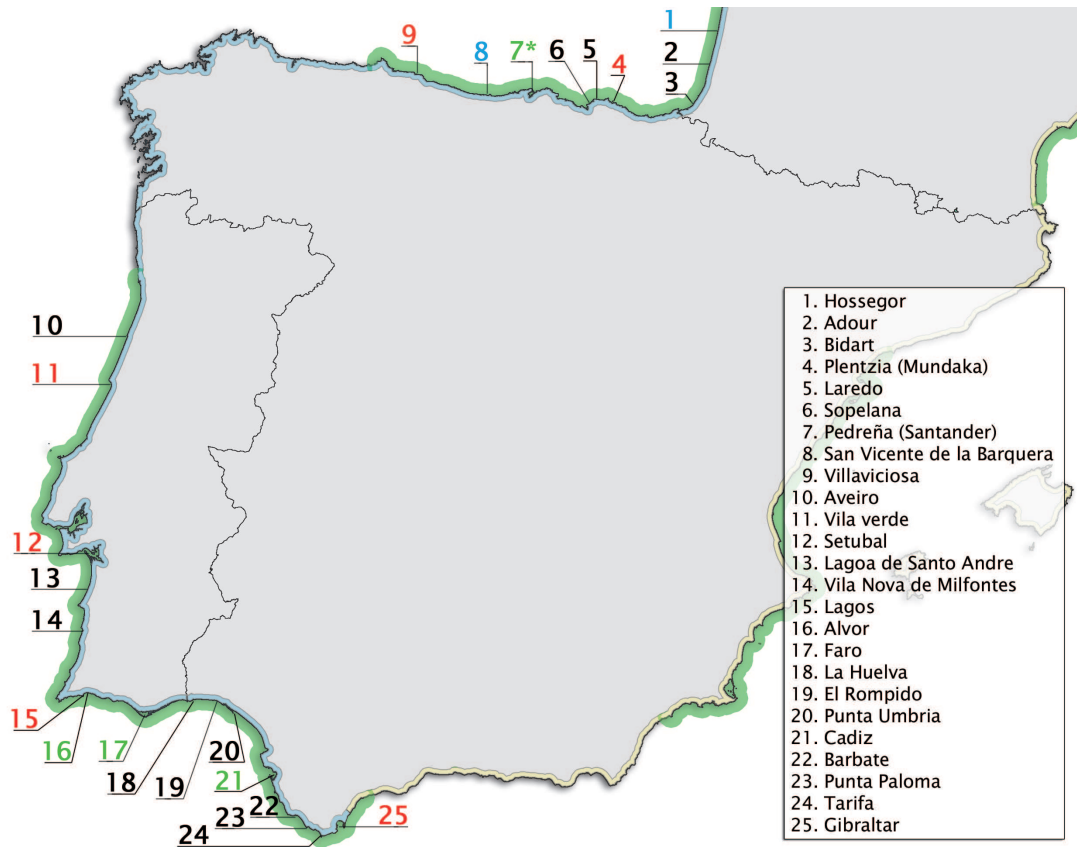


FIG. 4.4: Sites examinés entre Gibraltar et le Pays Basque français. Le code de couleur du trait de côte reprend celui de la carte de situation générale présentée à la figure 1. Le surlignage vert indique les côtes examinées en décembre 2002 et mars 2004. Les chiffres en bleu correspondent aux sites où *Cyclope neritea* a été trouvée pour la première fois au cours de notre échantillonnage. Les chiffres vert marquent ceux où *Cyclope neritea*, présente par le passé, a également été observée au cours de notre prospection (Le 7* indique un site où très peu d'individus vivants ont pu être observés). Les autres chiffres correspondent aux stations où *Cyclope neritea* n'a pas été trouvée : en noir, les sites où l'espèce n'a, à notre connaissance, jamais été observée, donc où elle semble durablement absente ; en rouge, les sites où *Cyclope neritea*, présente par le passé, a disparu au moment de notre échantillonnage.

culièrement dans les zones où *Cyclope neritea* était présente par le passé. Par exemple, trois stations ont été testées sans succès à Mundaka (Espagne) ainsi que huit sites dans la Baie de Sétubal (Portugal, figure 4.5), où *Cyclope neritea* était pourtant abondante jusqu'en 1932 (Fischer, 1899; Nobre, 1932). De même sur quatre sites examinés dans la baie de Santander (Espagne), nous n'avons trouvé *Cyclope neritea* qu'en un seul endroit (à Pedreña, voir figure 4.6). Notons que seuls trois individus vivants ont été trouvés à cet endroit alors que de très nombreuses coquilles vides étaient présentes. Par ailleurs, aucun spécimen n'a été trouvé au niveau du site de Somo, pourtant situé à quelques centaines de mètres seulement de Pedreña et présentant les mêmes caractéristiques sédimentaires.



FIG. 4.5: Vues de détail de la baie de Setubal. Les cercles rouges indiquent les sites où *Cyclope neritea* n'a pas été détectée. Source photographique : logiciel Google™ Earth (<http://earth.google.com/>)

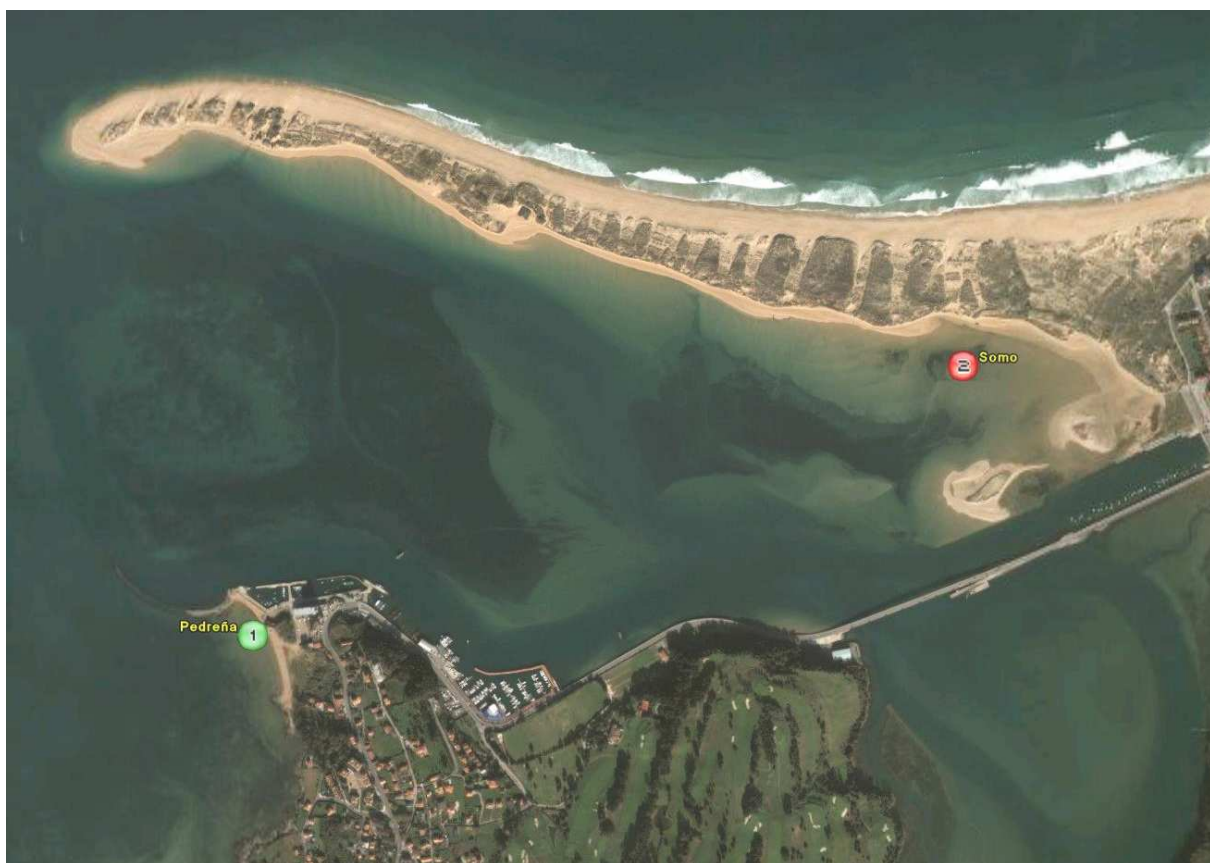
4.2.3 Résultats et discussion

Une synthèse de la comparaison de nos observations avec les données bibliographiques en notre possession est présentée dans le tableau 4.1. Différents cas de figure sont observés.

Sur les 25 sites répertoriés, 13 ont une histoire cohérente avec les observations que nous avons réalisées. Ainsi, seuls deux se caractérisent par la stabilité des populations de *Cyclope neritea*. Il s'agit d'Alvor (16) et de Faro (17). Pour ce dernier site, la population de *Cyclope neritea* connue depuis plus de 100 ans (Fischer, 1899) est toujours présente à des densités importantes (jusqu'à 80 individus par mètre carré à certaines stations d'échantillonnage). La population d'Alvor, remarquée pour la première fois en 1985 (Rodrigues et Quintino, 1991) est, elle aussi, toujours présente. Notons que cette population a été



a. Santander



b. Pedreña et Somo (Baie de Santander)

FIG. 4.6: Vues de détail de la baie de Santander. Le cercle vert indique le seul site où *Cyclope neritea* a été détectée. Source photographique : logiciel Google™ Earth, (<http://earth.google.com/>)

trouvée à proximité directe de parcs ostréicoles à plat, dont le naissain provient en grande partie du Bassin d'Arcachon, et dont une partie des adultes y retourne pour affinage en claires avant commercialisation (ostréiculteur local, communication personnelle). Par ailleurs, *Cyclope neritea* est absente de 11 autres sites dans lesquels sa présence n'a jamais été enregistrée auparavant : soit par manque d'information sur les sites considérés (sites 10, 18-20 et 22-24), soit parce que les inventaires réalisés permettent d'exclure sa présence (sites 5, 6, 13 et 14).

TAB. 4.1: Comparaison des données de terrain ("*présence 2002-2004*") avec les données bibliographiques (présence antérieure). Les numéros de la première colonne correspondent aux sites de la figure 4.4 et sont présentés du nord au sud, en suivant le trait de côte. Pour chaque site, la présence (ou l'absence) de *Cyclope neritea*, actuelle et passée, est indiquée suivant le code établi par Sauriau (1991). "0" : absence, "+" : densité faible (seuls quelques individus récoltés) et "++" : densité forte (jusqu'à plus de 400 individus par mètre carré). Un "-" indique l'absence de données historiques concernant *Cyclope neritea* pour le site considéré. Enfin, pour chaque observation antérieure, l'année ou la période ainsi que la référence bibliographique (principalement tiré de Sauriau, 1991) correspondante sont indiquées.

n°	Sites examinés	Présence 2002-4	Présence antérieure	Année	Référence
1	Hossegor	++	0	1964-1965	Lagardère (1966)
2-3	Adour - Bidart, Socoa (sud Pays Basque français)	0	0 ++ ++ ++ 0 0	avant 1899 1924 1949 1959 1964-1965 avant 1987	Fischer (1899) de Beauchamp (1948) Kisch (1950) Morton (1960) Lagardère (1966) Borja (1987)
4	Plentzia (Mundaka)	0	++ +	1978-1979 Avant 1987	Rallo (1981) Borja (1987)
5-6	Laredo - Sopelana	0	0 0	1978-1979 Avant 1987	Rallo (1981) Borja (1987)
7	Santander	+	++ ++	1900-1910 1978	Hidalgo (1917) Lopez-Cotelo <i>et al.</i> (1982)
8	San Vicente de la Barquera	++	-	-	-
9	Villaviciosa	0	++	1900-1910	Hidalgo (1917)

... suite... →

TAB. 4.1: *suite*. Comparaison des données de terrain (“présence 2002-2004”) avec les données bibliographiques (présence antérieure).

n°	Sites examinés	Présence 2002-4	Présence antérieure	Année	Référence
10	Aveiro	0	-	-	-
11	Vila Verde	0	++	1900-1910	Fischer (1899)
12	Setubal	0	++	1900-1910	Fischer (1899)
			++	Avant 1932	Nobre (1932)
13	Lagoa de Santo Andre	0	0	1985-1989	Sauriau (1991)
14	Vila Nova de Milfontes	0	0	1985-1989	Sauriau (1991)
15	Lagos	0	++	Avant 1932	Nobre (1932)
16	Alvor	++	++	1985	Rodrigues et Quintino (1991)
17	Faro	++	++	1900-1910 Avant 1932	Fischer (1899) Nobre (1932)
18	La Huelva	0	-	-	-
19	El Rompido	0	-	-	-
20	Punta Umbria	0	-	-	-
21	Cadiz	++	++	1900-1910	Hidalgo (1917)
			0	1974-1975	Arias (1976)
22	Barbate	0	-	-	-
23	Punta Paloma	0	-	-	-
24	Tarifa	0	-	-	-
25	Gibraltar	0	++	1900-1910	Hidalgo (1917)

La Galice n’a pas été examinée car *Cyclope neritea* n’y est pas connue actuellement (S. Gofas, communication personnelle). L’espèce est historiquement absente dans cette zone (Figueras, 1956; Vieitez, 1981; Rolán, 1984), mais un faible nombre d’individus y a toutefois été rencontré deux fois par le passé :

- La première fois en 1972 dans la Ria de Arosa (Hernandez Otero et Jimenez Millan, 1972). *Cyclope neritea* semble avoir disparu de ce site par la suite (Penas et Gonzalez, 1983).
- Plus récemment, en 1992, d’autres individus ont été observés sur la plage de O Grove,

très probablement introduits accidentellement associés à des échanges d'huîtres méditerranéennes (Rolán, 1992).

Les sites au sein desquels *Cyclope neritea* est absente actuellement et n'a jamais été enregistrée par le passé, soit en raison d'une absence de données (sites 10, 18-20 et 22-24), soit parce que l'espèce est historiquement absente des sites (sites 5, 6, 13 et 14 ; Rallo, 1981; Borja, 1987; Sauriau, 1991) sont répartis tout au long de la zone étudiée, ce qui confirme une distribution en "patches" de l'espèce, rendant difficile la détermination précise des limites de l'aire de distribution naturelle (Russell et Lindberg, 1988; Gaston, 1994). Notons également que les deux seuls sites au sein desquels la présence durable de *Cyclope neritea* a été relevée sont situés sur la côte sud du Portugal. Or, une étude de Berrou *et al.* (2004), portant sur l'analyse morphologique des espèces du genre *Ocenebra*, a montré que les populations de l'Algarve (sud du Portugal) étaient plus proches des populations méditerranéennes que des populations atlantiques, en raison de l'influence marquée des eaux méditerranéennes dans cette zone. Ainsi, concernant *Cyclope neritea*, la présence durable de populations à cet endroit pourrait être le fruit d'échanges avec le centre de distribution de l'espèce, et ce, malgré les capacités de dispersion limitées de l'espèce.

Les 13 autres sites examinés ont une histoire plus complexe et peuvent être classés dans trois catégories :

1. Les sites où *Cyclope neritea*, jamais repérée auparavant, est maintenant présente. Il s'agit d'Hossegor (site 1), détectée pour la première fois en décembre 2002 malgré l'inventaire réalisé en 1964-65 sur les côtes situées entre Hendaye et Biscarosse par Lagardère (1966), et de San Vicente de la Barquera (site 8) où aucun inventaire antérieur n'a pu être examiné. Bien que l'absence d'inventaire dans une zone ne signifie pas pour autant l'absence de l'espèce dans cette zone, il est intéressant de noter que ces deux populations ont été trouvées dans les sites parmi les plus septentrionaux de l'aire géographique étudiée.
2. Les sites où *Cyclope neritea*, présente par le passé, régresse ou a disparu. Sept sites localisés tout au long de la zone étudiée sont dans ce cas de figure. Si la présence de *Cyclope neritea* au sein de certains de ces sites n'a été relevée qu'une seule fois il y a près de 100 ans (Villaviciosa et Gibraltar en 1900-10, Hidalgo, 1917), d'autres populations, semblent bien établies par le passé mais ont aujourd'hui disparu. C'est le cas à Mundaka, où la diminution de densité, déjà significative entre 1978-79 et 1987 (Rallo, 1981; Borja, 1987) se poursuit par la disparition de *Cyclope neritea*. De même, dans et autour de la baie de Setubal (Cezimbra), où l'espèce était présente en forte densité jusqu'en 1932 aucun individu vivant ou mort n'a pu être collecté,

malgré l'effort d'échantillonnage important (voir figure 4.5) (Fischer, 1899; Nobre, 1932). D'autres populations sont en très nette régression. C'est notamment le cas à Pedreña, dans la Baie de Santander, connue pour abriter *Cyclope neritea* en forte densité (Hidalgo, 1917; Lopez-Cotelo *et al.*, 1982). Nous n'avons pourtant pu y trouver que trois individus vivants en plus de trois heures de recherche. Le grand nombre de coquilles vides, très érodées, trouvées dans les laisses de mer laisse à penser que cette population est en phase de déclin. Les trois autres sites examinés à proximité n'ont pas révélé de *Cyclope neritea* vivantes.

3. Les sites présentant une histoire plus complexe, avec apparitions et disparitions successives de *Cyclope neritea*. C'est notamment le cas de Cadix, où une population importante de *Cyclope neritea* enregistrée en 1900-10 (Hidalgo, 1917) n'a pas été retrouvée en 1974-75 (Arias, 1976), mais est aujourd'hui de nouveau présente à forte densité (plus de 80 individus par mètre carré). Notre propre échantillonnage nous a montré que cette population, bien que dense, était particulièrement peu étendue et concentrée à l'extrémité sud d'une plage de plusieurs kilomètres de long. Il est ainsi probable que l'espèce n'a jamais réellement disparu : la localisation particulière de cette population explique certainement pourquoi l'espèce n'a pas été détectée au cours de l'inventaire réalisé en 1974-75. A l'inverse, les populations françaises du sud du Pays Basque (Adour et Bidart), abondantes entre 1949 et 1959 (Kisch, 1950; Morton, 1960), sont inexistantes avant (Fischer, 1899) et après cette période (Lagardère, 1966; Borja, 1987, observations personnelles). L'hypothèse la plus probable ici est celle d'une population éphémère, apparue en bordure d'aire de distribution.

Pour résumer les informations obtenues, nous pouvons dire que *Cyclope neritea* semble aujourd'hui avoir disparu de toute la façade ouest de l'Espagne et du Portugal, y compris dans des sites où des populations denses de l'espèce étaient historiquement présentes. Les pressions anthropiques exercées sur les écosystèmes et en particulier le développement des aménagements littoraux portugais depuis les années 1970, en sont probablement une cause majeure. Par exemple, la construction de barrages sur le Rio Guadiana, à la frontière sud entre le Portugal et l'Espagne a fortement modifié le régime hydrologique du bassin versant, affectant grandement les espèces piscicoles y vivant (certaines étant même en voie d'extinction ; Collares-Pereira *et al.*, 2000), et la sédimentologie. Autre exemple, l'estuaire du Sado (Baie de Setubal, Portugal), fortement modifié par les actions humaines. D'après Caeiro *et al.* (2002), le développement récent de complexes touristiques sur la péninsule du sud de l'estuaire a généré des modifications de pressions environnementales. Ces auteurs rappellent que l'importante zone industrialo-portuaire du nord de l'estuaire

(présence d'usines thermiques, d'abattoirs, de fabriques de pâte à papier, d'industrie chimique, agroalimentaire...), en s'étendant sans cesse, modifie profondément la qualité de l'eau, les régimes sédimentaires, et affecte donc également les communautés vivant au sein de cette baie. *Cyclope neritea* étant une espèce particulièrement exigeante en terme de substrat, il n'est pas étonnant qu'elle ait disparu de cette zone.

Ainsi, d'une manière générale, nous avons vu que les populations les plus stables de la zone étudiée sont situées au sud de la zone étudiée (populations 16, 17 et 21), probablement en raison de l'influence des eaux méditerranéennes baignant le sud de l'Espagne et du Portugal. Ces populations seront donc étudiées dans la partie II au même titre que les populations du centre de l'aire de distribution naturelle de l'espèce. La façade ouest de la péninsule ibérique est aujourd'hui caractérisée par l'absence de *Cyclope neritea*, probablement en raison de modifications sédimentaires importantes, conséquences d'actions anthropiques au niveau littoral. Enfin, les seules populations récemment détectées dans cette zone ayant un statut indéterminé, sont situées au nord, sur les côtes du Pays Basque français et espagnol (populations 1 et 8, voir figure 4.4). Ces populations (ainsi que la population 7), très éloignées des populations du sud et proches des populations françaises récemment détectées, seront étudiées dans la partie III de ce manuscrit, consacrée à l'étude des populations récemment installées sur les côtes françaises de l'Atlantique et de la Manche. L'étude de ces populations du nord de la zone en limite d'aire (et en particulier de leurs caractéristiques génétiques) devrait permettre de déterminer si elles sont plus proches des populations du sud de l'Espagne ou au contraire plus proches des populations françaises récemment établies.

Identification de *Cyclope neritea* et *Cyclope pellucida*

5.1 Variations morphologiques et définition des espèces cibles

Nous venons de voir que les populations de *Cyclope neritea* du littoral atlantique de la péninsule ibérique sont particulièrement instables. En particulier, des populations jamais détectées par le passé ont été observées sur la côte nord de l'Espagne. D'autres, historiquement bien établies ont totalement disparu (essentiellement sur la côte ouest). Certaines populations, notamment au sud, semblent stables, et enfin, en plusieurs endroits, les populations semblent successivement apparaître puis disparaître (Galice, Cadix, frontière franco-espagnole). Non seulement la dynamique complexe de l'espèce dans cette zone rend les inventaires difficiles à réaliser, mais il apparaît en outre que la détermination de *Cyclope neritea*, du point de vue morphologique, ait pu être source de confusion par le passé. Nous allons montrer en quoi la reconnaissance de *Cyclope neritea* peut être problématique sur le terrain. Il ne s'agit pas ici de réaliser une étude de systématique morphologique, mais de montrer et justifier l'intérêt de l'utilisation de marqueurs moléculaires pour distinguer rapidement des espèces morphologiquement et phylogénétiquement proches.

D'après la “*Check List of European MARine Mollusca*” (CLEMAM, 2006, site internet : <http://www.somali.asso.fr/clemam/>), les synonymes du genre *Cyclope*, utilisés à différentes périodes, sont au moins au nombre de dix : *Cencus*, *Cyclocyrtia*, *Cyclonassa*, *Cyclops*, *Nana*, *Nanina*, *Nannia*, *Neritula*, *Panormela*, *Proneritula*... L'espèce *Cyclope neritea*, initialement décrite par Linné (1758) comme *Buccinum neriteum*, possède quant

à elle plus de 30 synonymes (*e.g.* *Cyclops asterizans*, *Cyclonassa fasciata*, *Cyclonassa vernicata*, *Cyclope neritoidea*, *Cyclope westerlundi*, *Cyclope Kamischiensis*...). Certaines dénominations sont, aujourd'hui encore, utilisées à mauvais escient (voir par exemple Demir, 2003, qui distingue, dans son inventaire des côtes turques, *Cyclope neritea* et *Cyclope westerlundi*), et le statut de *Cyclope donovani* n'est toujours pas clair aujourd'hui (CLEMAM, 2006) : s'agit-il d'une espèce vraie ou d'un autre synonyme de *Cyclope neritea* ?

L'importante variété synonymique existant chez *Cyclope neritea* est le reflet de la grande plasticité morphologique de cette espèce (voir par exemple la figure 5.1 pour un exemple de la variabilité touchant la coloration des coquilles). Selon les sites échantillonnés, outre des colorations variées, des tailles adultes parfois très différentes ont ainsi été observées (longueur moyenne des coquilles des individus prélevés à Rivedoux, Île de Ré : $10,78 \pm 1,70$ mm, et à Eforie, Mer Noire : $13,82 \pm 1,06$ mm). L'importance des variations phénotypiques de *Cyclope neritea* ainsi que la multiplicité des synonymies rend hasardeuse l'identification précise de l'espèce sur le terrain lors d'inventaires faunistiques rapidement menés et peut donc remettre en question leur validité.

En dehors de *Cyclope neritea*, la seule autre espèce reconnue du genre *Cyclope* sur les côtes européennes est *Cyclope pellucida* (Risso, 1826). Bien que morphologiquement différentes, les deux espèces, dont les aires de distribution se chevauchent, peuvent parfois être difficiles à distinguer en particulier pour les individus de petite taille. Comme le montre la figure 5.2, les différences morphologiques (au moins en terme de coloration) sont en effet parfois plus marquées au niveau intra-spécifique qu'au niveau inter-spécifique (des coquilles de *Cyclope neritea*, *Cyclope pellucida* et *Cyclope donovani* sont présentées). Cela renforce l'idée qu'au cours d'inventaires faunistiques, des confusions sont possibles lors de l'identification de spécimens sur le terrain. Lors de notre échantillonnage autour de la péninsule ibérique, nous avons nous-même fait les frais de ces similitudes : à Faro (Portugal), puis, plus tard, dans le Delta de l'Ebre (Espagne), pensant récolter des individus de *Cyclope neritea*, nous avons échantillonné des individus de *Cyclope pellucida*. Ce n'est que plus tard, au laboratoire, lors des analyses génétiques, que nous avons été amenés à réexaminer de près nos échantillons et que nous avons constaté notre erreur. Pourtant au fait des similarités morphologiques entre les deux espèces, nous n'avons pas été capables d'identifier à coup sûr tous les spécimens en notre possession. Nous avons donc décidé de mettre au point un “marqueur moléculaire diagnostique”, afin de pouvoir identifier sans ambiguïté les spécimens collectés sur le terrain.



FIG. 5.1: Variabilité de la couleur de la coquille chez *Cyclope neritea* : quelques spécimens de la population de l'île de Ré présentés en face ventrale.



a. *Cyclope neritea* (14 mm, Portugal)



b. *Cyclope neritea* (13 mm, Turquie)



c. *Cyclope donovani* (9 mm, Portugal)



d. *Cyclope pellucida* (12 mm, Portugal)



e. *Cyclope donovani*



f. *Cyclope pellucida*

FIG. 5.2: Différences morphologiques des coquilles de *Cyclope neritea*, *Cyclope donovani* et *Cyclope pellucida*. Sources photographiques : 5.2a, 5.2b, 5.2c et 5.2e : <http://www.femorale.com.br/>. 5.2d et 5.2f : <http://www.gastropods.com/>

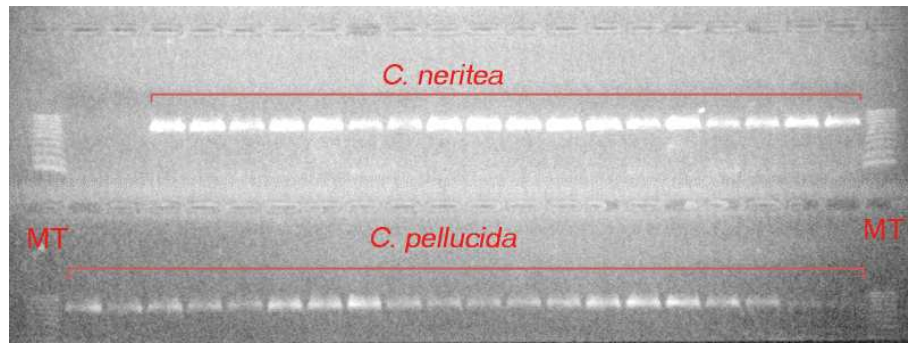


FIG. 5.3: Migration sur gel d'agarose de produits de PCR réalisées sur *Cyclope neritea* et *Cyclope pellucida* avec des amorces universelles. Partie supérieure de l'image : *Cyclope neritea*, partie inférieure : *Cyclope pellucida*. MT : Marqueur de Taille.

5.2 Le marqueur d'espèce

Le protocole suivi est rigoureusement identique à celui présenté dans l'annexe A (compléments méthodologiques), page 188.

Dans un premier temps, des PCR (voir principe section A.1.2 et protocole section A.2) ont été réalisées, pour les deux espèces, avec les amorces universelles mises au point par Folmer *et al.* (1994) :

- HCO2198 : 5'-TAAACTTCAGGGTGACCAAAAAATCA-3'
- LCO1490 : 5'-GGTCAACAAATCATAAAGATATTGG-3'

Ces amorces universelles ont la particularité de permettre l'amplification d'un fragment de plus de 700 paires de bases (bp) du gène codant pour la sous-unité I de la cytochrome oxydase dans une grande variété de taxons. Les résultats de ces PCR sont présentés à la figure 5.3.

Cette photo de gel montre qu'un fragment de taille similaire est amplifié chez *Cyclope neritea* et *Cyclope pellucida*. Le séquençage de quelques individus des deux espèces (voir méthode décrite en annexe A) nous a permis : (i) de redéfinir des amorces spécifiques pour l'une des deux espèces, et (ii) d'estimer, après alignement des séquences, le niveau de différenciation existant entre les deux espèces.

La redéfinition des amorces a été effectuée pour *Cyclope neritea*, car ces nouvelles amorces permettent une amplification spécifique de meilleure qualité qu'avec les amorces universelles, et facilitent donc les analyses en routine au niveau de la "population" qui seront effectuées par la suite (voir parties II et III). Ces amorces ont été choisies grâce au logiciel Primer 3 (Rozen et Skaletsky, 2000) :

- Amorce sens : 5'-GTTAAAATTTTCGATCTGTTA-3' (Cy2)
- Amorce antisens : 5'-GGATTAGTTGGTACAGC-3' (CyR)

Dans un second temps, de nouvelles PCR ont été réalisées sur les deux espèces, en utilisant cette fois-ci les amorces Cy2 et CyR. Les résultats sont présentés à la figure 5.4.

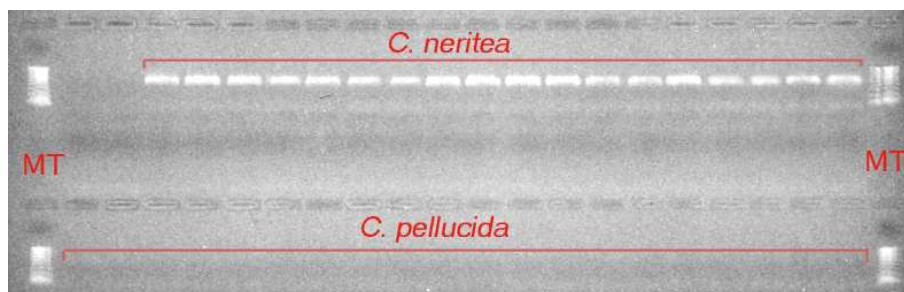


FIG. 5.4: Migration sur gel d'agarose de produits de PCR réalisées sur *Cyclope neritea* et *Cyclope pellucida* avec des amorces spécifiques de *Cyclope neritea*. Partie supérieure de l'image : *Cyclope neritea*, partie inférieure : *Cyclope pellucida*. MT : Marqueur de Taille.

Tous les individus de *Cyclope neritea* qui répondaient positivement avec les amorces universelles définies par Folmer *et al.* (1994) répondent positivement avec les amorces spécifiques, alors qu'aucun individu de *Cyclope pellucida* n'est amplifié. Les amorces spécifiques définies permettent donc bien de différencier facilement les deux espèces.

Le séquençage de cinq individus de *Cyclope pellucida* collectés à Faro a été réalisé en utilisant les amorces universelles (séquences complètes présentées en annexe B). Les cinq individus ont montré le même haplotype (*i.e.* la même séquence nucléotidique). La comparaison avec les séquences de *Cyclope neritea* (533 paires de bases) déjà en notre possession nous a permis de rendre compte de la divergence entre les deux espèces. Le tableau 5.1 indique les différences observées entre l'haplotype de *Cyclope pellucida* et les deux haplotypes de *Cyclope neritea* qui en sont les plus proches. Sur le fragment étudié (COI, 533 bp), 34 mutations séparent *Cyclope pellucida* des individus les moins divergents de *Cyclope neritea*. Comme nous le verrons dans la partie II, ce niveau de divergence est très supérieur à celui existant au sein de populations même très divergentes de *Cyclope neritea*. En outre, sur ces 34 caractères divergents, 25 sont fixés, c'est-à-dire que le caractère présent chez les individus de *Cyclope pellucida* séquencés ne se retrouve chez aucun individu de *Cyclope neritea* (comparaisons réalisées sur l'ensemble du jeu de données mitochondrial obtenu au cours des analyses réalisées dans les parties II et III, soit 652 individus). Ces 25 caractères correspondent à la définition de caractères “apomorphiques”⁴, qui confirme que *Cyclope pellucida* et *Cyclope neritea* sont bien deux espèces vraies.

Notons par ailleurs que la distinction entre les deux espèces peut également se faire sur la base des génotypages utilisant les microsatellites mis au point chez *Cyclope neritea* (voir

⁴Des caractères sont dits “apomorphiques” lorsqu'ils sont présents chez tous les individus appartenant à une espèce donnée et absents chez tous les individus n'en faisant pas partie.

TAB. 5.1: Position des sites variables et nature des substitutions nucléotidiques entre deux haplotypes de *Cyclope neritea* et l'haplotype de *Cyclope pellucida*.

Position	11111122222222333334444445
des sites	123357901466711233799127791577990
variables	1705804219647658105539700789569143
<i>Cyclope neritea</i>	ATAATGTCTAACAGGATGGATGCTAGCTTGATCC
<i>Cyclope neritea</i>	G-----
<i>Cyclope pellucida</i>	ACCGCACTCGTTCAAGCAAGCATCGATAAAGATT

section A.3 en annexe). Toutefois, cet outil est moins facile à mettre en œuvre. Le marqueur mitochondrial est l'outil diagnostique de choix en raison de sa facilité d'utilisation (une PCR et une migration sur gel d'agarose) et de sa robustesse.

Conclusions

Nous avons dans cette première partie confirmé la présence de populations instables de *Cyclope neritea* dans la zone en limite d'aire de distribution de l'espèce et en particulier précisé la localisation de plusieurs types de populations ayant des dynamiques distinctes :

- les populations les plus stables dans le temps sont situées dans la partie sud de la zone étudiée. Les côtes du sud de l'Espagne et du Portugal sont en partie baignées par des eaux méditerranéennes, ce qui pourrait expliquer le maintien dans le temps de ces populations grâce à des échanges durables et stables avec le centre de l'aire de distribution de *Cyclope neritea*. Ces populations seront donc considérées comme faisant partie intégrante de l'aire de distribution naturelle de l'espèce, et seront à ce titre étudiées conjointement avec les autres populations de l'aire naturelle dans la partie II de cette thèse.
- les populations éphémères, sur le déclin, ou détectées pour la première fois au cours de notre échantillonnage sont principalement localisées au nord de la zone étudiée, à proximité immédiate des populations françaises récemment apparues. Il conviendra de déterminer si elles sont naturelles ou installées à la suite d'introductions d'origine humaines. Cette question sera abordée dans la partie III consacrée à l'étude des populations récemment apparues le long des côtes françaises de l'Atlantique et de la Manche.
- enfin, entre la zone sud et la zone nord, les populations de *Cyclope neritea* semblent avoir disparu de la façade atlantique de la péninsule ibérique, y compris dans des baies où l'espèce était présente de longue date et en forte densité. Les actions anthropiques sont certainement en partie responsables de la disparition de ces populations : en modifiant les habitats côtiers et notamment la sédimentologie, l'homme est à l'origine de la perturbation des communautés, et probablement de la disparition des espèces les plus exigeantes en terme de qualité de substrat.

L'instabilité naturelle des populations en limite d'aire, du fait de leur maladaptation, est ici renforcée par les activités humaines. Par ailleurs, les confusions possibles entre deux espèces du même genre, *Cyclope neritea* et *Cyclope pellucida*, ont pu être sources d'erreurs

dans les inventaires faunistiques passés. La mise au point d'un marqueur diagnostique de ces deux espèces a permis de dissiper ces doutes.

La situation des populations situées en limite d'aire de distribution étant plus claire, il s'agit à présent d'étudier en détail l'aire "*naturelle*" de *Cyclope neritea*. Cette étude tentera notamment de mettre en lumière la structure phylogéographique des populations de cette zone à plusieurs échelles, en s'appuyant à la fois sur un marqueur mitochondrial et des marqueurs nucléaires.

Deuxième partie

Phylogéographie dans l'aire “*naturelle*”
de
Cyclope neritea

Introduction

L'étude de l'aire "*naturelle*" : un double objectif

L'étude de la structure génétique dans l'aire dite "*naturelle*" de *Cyclope neritea* est réalisée ici avec deux objectifs principaux.

- Le premier objectif est de reconstituer l'histoire de *Cyclope neritea* dans son aire "*naturelle*" : il s'agit en particulier de vérifier si la structure génétique dans l'aire "*naturelle*" de *Cyclope neritea* est bien conforme à celle attendue pour une espèce à cycle de vie strictement benthique. En effet, les traits d'histoire de vie de cette espèce, et en particulier son cycle à développement direct, sont censés conférer à cette espèce des capacités de dispersion assez limitées. En théorie, cela doit avoir pour effet (i) de favoriser la mise en place d'une structure génétique inter-populations plus marquée (du fait du faible nombre de migrants échangés entre populations, des populations isolées durablement doivent diverger), et (ii) de maintenir une diversité génétique intra-populations plus faible (du fait de l'absence d'apport de nouveaux gènes par immigration) que dans les espèces à fort pouvoir dispersif.

Toutefois, *Cyclope neritea* est une espèce ayant une histoire commune avec l'homme particulièrement ancienne. De nombreux individus du genre *Cyclope* ont par exemple été retrouvés dans des sépultures ou des grottes grecques, italiennes ou françaises datant du paléolithique supérieur (jusqu'à -35 000 ans, Kuhn et Stiner, 1998; Shackleton, 1988; Vanhaeren et D'Errico, 2001). Dans certains de ces sites, *Cyclope neritea* représente l'espèce de mollusque la plus abondante (*e.g.* 8 487 individus de cette espèce sur 15 838 spécimens collectés dans la grotte de Franchtchi en Grèce). Il est peu probable que *Cyclope neritea* ait été récoltée pour la consommation. Deith et Shackleton (1988) soulignent en effet que les individus devaient être collectés morts. Seules les coquilles, certainement utilisées pour la parure ou en tant que monnaie (Kuhn et Stiner, 1998; Vanhaeren et D'Errico, 2001) devaient être recherchées et déplacées au gré des migrations humaines. Plus récemment, les analyses sédimentaires de l'ancien port phénicien de Tyr (Liban), extrêmement actif entre -5500 et -1500,

ont révélé la présence de *Cyclope neritea* en grande quantité tout au long de cette période (Marriner *et al.*, 2005). Les échanges transméditerranéens ont donc, comme aujourd’hui, pu être vecteurs de dispersion. Enfin, de nos jours, les pratiques conchylicoles sont susceptibles de déplacer *Cyclope neritea*, comme en atteste la présence éphémère de plusieurs individus trouvés en Galice en 1992 (Rolán, 1992; Advisory Committee on the Marine Environment, 2001).

- Le second objectif est de vérifier l’existence de signatures génétiques suffisamment claires dans les populations de l’aire “naturelle” de *Cyclope neritea*, afin de pouvoir déterminer les processus à l’origine de l’établissement des populations récemment apparues sur les côtes françaises de l’Atlantique et de la Manche. Si *Cyclope neritea* a été affectée par des mouvements anthropiques par le passé, la structure génétique actuelle des populations “naturelles” a pu en conserver la trace. Dans cette hypothèse, déterminer l’origine des populations françaises récemment découvertes pourrait s’avérer impossible ou difficile, dans la mesure où les populations ne présenteraient alors plus de signatures génétiques distinctes.

Pour répondre à ces deux objectifs, le choix des marqueurs moléculaires est primordial puisque des marqueurs génétiques ayant des caractéristiques différentes (en terme de taux d’évolution, de mode d’hérédité, de ploïdie... voir par exemple le tableau 1) apporteront des informations différentes sur les processus à l’œuvre dans les populations.

TAB. 1: Comparaison de quelques caractéristiques moyennes de trois marqueurs nucléaires couramment utilisés (¹RAPDs : Randomly Amplified Polymorphic DNA). D’après Jarnes et Lagoda (1996)

Caractéristiques	Allozymes	RAPDs ¹	Microsatellites
Codominance	oui	non	oui
Neutralité	douteuse	oui	oui
Stades “conservables”			
embryons/jeunes	rarement	oui	oui
adultes	oui	oui	oui
Nombre d’allèles par locus	1-5	2	1-50
Information moléculaire (structure, mutation)	rarement	rarement	disponible
Coût relatif par individu	1	1	3-4

Le choix des marqueurs moléculaires

Locus mitochondrial

L'ADN mitochondrial (ADN_{mt}) est un outil de choix pour les études phylogéniques ou phylogéographiques (Crozier, 1990), particulièrement au niveau intra-spécifique (Kolbe *et al.*, 2004; Voisin *et al.*, 2005). L' ADN_{mt} présente en effet de nombreuses caractéristiques avantageuses pour ce genre d'étude :

- c'est une molécule circulaire de petite taille (de 14 à 39 Kb) relativement bien connue. L' ADN_{mt} de la plupart des animaux contient les régions suivantes, chacune évoluant à un rythme propre (Wolstenholme, 1992) : une région de contrôle, 2 gènes codant pour de l' ARN_r , 22 pour l' ARN_t et 13 codant pour des protéines (sous unité I, II et III de la cytochrome oxydase, sous unités 6 et 8 de l'ATPase, cytochrome b et 7 sous unités du système NADH réductase).
- l' ADN_{mt} est présent en de nombreuses copies dans les cellules, ce qui facilite donc son extraction et son amplification.
- le génome mitochondrial est haploïde et la transmission de cette molécule est généralement monoparentale (bien que des cas de transmission bi-parentale aient été mis en évidence chez certains bivalves, notamment dans le genre *Mytilus*, Zouros *et al.*, 1992, 1994). Cela garantit l'absence de recombinaison et simplifie les analyses statistiques.
- La taille efficace⁵ des populations de gènes mitochondriaux est quatre fois plus faible que celle des locus nucléaires, en raison de l'hérédité monoparentale associée au caractère haploïde de l' ADN_{mt} . A taux de mutation égal, ce marqueur est ainsi plus sensible à la dérive génétique⁶ que des marqueurs diploïdes à hérédité biparentale, ce qui permet de garder la marque d'évènements démographiques (expansions, goulots d'étranglements⁷...) plus longtemps. La trace laissée par de tels évènements stochastiques sera en effet d'autant plus marquée que la taille efficace des populations considérée sera faible.

⁵Ce concept a été formulé pour la première fois par Wright (1931, 1938) : la taille efficace d'une population est égale au nombre d'individus reproducteurs d'une population idéale (*i.e.* à l'équilibre de Hardy-Weinberg) qui montrerait les mêmes niveaux de variation des fréquences alléliques sous un régime de dérive génétique, ou le même niveau de consanguinité que la population considérée. Plusieurs autres définitions existent. Retenons que ce concept est étroitement lié à la notion de pool génique : la taille efficace d'une population ne sera pas la même si l'on travaille avec un locus diploïde ou haploïde, ou si une partie des individus ne participe pas à la transmission du marqueur génétique que l'on étudie.

⁶La dérive génétique correspond aux fluctuations des fréquences alléliques apparaissant dans les populations par le simple jeu du hasard, indépendamment des mutations, migrations ou de la sélection.

⁷Réductions drastiques de la taille d'une population, s'accompagnant la plupart du temps d'une perte de diversité génétique.

- enfin, bien que peu de données existent sur les taux d'évolution de l'ADN_{mt} des mollusques, cette molécule présente, chez les mammifères, de 0.5 à 1% de substitutions nucléotidiques par million d'années soit, en moyenne, 2 à 10 fois plus que dans le génome nucléaire (Wolstenholme, 1992). Ce taux de mutation élevé permet l'accumulation de polymorphisme au sein des populations, rendant possible l'étude de la diversité et de la structure génétique au niveau intra-spécifique.

Le type de variations génétiques que nous étudierons avec ce marqueur concerne le polymorphisme de séquences (*i.e.* nombre et distribution des substitutions nucléotidiques, insertions et délétions), ce qui nécessitera le séquençage du marqueur choisi chez tous les individus étudiés. Nous parlerons alors d'haplotypes : ce sont des séquences nucléotidiques uniques, observées chez un ou plusieurs individus, et différant des autres haplotypes par un ou plusieurs nucléotides en raison de substitutions, d'insertions et/ou de délétions.

Locus microsatellites

Les microsatellites sont des marqueurs nucléaires à hérédité biparentale, neutres (non soumis à la sélection) et codominants. Ce sont des motifs nucléotidiques souvent courts (entre 2 et 4 paires de bases), répétés en tandem, et hypervariables (les taux de mutations varient en moyenne de $5 \cdot 10^{-5}$ à $5 \cdot 10^{-3}$ pour les microsatellites, contre 10^{-9} pour le reste du génome, Estoup et Angers, 1998). La cause principale de l'hypervariabilité des microsatellites est le “*glissement*” de la polymérase au cours de la réplication de l'ADN, qui a pour effet d'augmenter ou de diminuer le nombre de répétitions du motif de base d'une ou plusieurs répétitions (Levinson et Gutman, 1987; Schlötterer et Tautz, 1992). Le type de variations génétiques que nous étudierons avec ce marqueur concerne le polymorphisme de taille de fragments grâce à l'électrophorèse qui permet la mise en évidence “*d'électromorphes*”, et pas la détection, par séquençage, de mutations au sein des séquences nucléotidiques. Différents types de motifs de base existent et déterminent en partie le mode et le taux d'évolution des locus microsatellites Jarnes et Lagoda (1996) : les motifs purs (*e.g.* CACACACACACACA) évoluent généralement plus rapidement que les motifs composés (*e.g.* CACACACAGAGAGAGA) ou que les motifs discontinus (*e.g.* CACACATTCACATTTCA).

Intérêt de l'approche comparative

Afin de déterminer si dans l'aire “*naturelle*” de *Cyclope neritea*, la structure des populations est conforme à celle attendue chez une espèce incubante ayant des capacités

de dispersion limitées, ou si, au contraire, elle résulte d’actions humaines passées et/ou présentes, l’usage de deux types de marqueurs génétiques présente plusieurs avantages :

1. *Cyclope neritea* étant une espèce à développement direct inféodée à un substrat particulier, les flux géniques entre les populations naturelles de cette espèce sont supposés peu nombreux. Or, dans cette situation précise, l’utilisation exclusive de marqueurs hypervariables peut conduire à sous-estimer la structure génétique réelle existant entre les populations (Balloux *et al.*, 2000; Diaz-Almela *et al.*, 2004; Estoup *et al.*, 2002). En effet, si des taux de mutations élevés sont un avantage pour les études à micro-échelle spatiale et temporelle, ils peuvent en revanche poser problème à plus large échelle du fait de l’homoplasie⁸ de longueur : dans ce cas, deux allèles identiques par nature (taille identique), présents dans deux populations distinctes, ne sont pas forcément hérités d’un même ancêtre commun. L’utilisation conjointe des microsatellites et d’un marqueur mitochondrial permet de s’affranchir de ce biais.
2. en raison de modes d’hérédité distincts entre le génome nucléaire (à hérédité biparentale) et le génome mitochondrial (à hérédité maternelle), la comparaison de marqueurs cytoplasmiques et de marqueurs nucléaires dans les populations étudiées peut permettre de révéler des processus qui seraient restés invisibles par l’étude d’un seul type de marqueur (Asmussen et Basten, 1994) ; cela multiplie les chances d’observer, dans l’aire “naturelle” de *Cyclope neritea*, des signatures génétiques distinctes en fonction des régions ou des populations étudiées. La confrontation des profils obtenus avec les deux types de marqueurs peut par exemple mettre en évidence des taux de migrations différentiels selon les sexes, ou des processus d’introgression sur les lignées maternelles. L’étude cyto-nucléaire peut être ainsi révélatrice (i) du mélange de lignées évolutives divergentes au sein d’une même population, (nous parlerons du phénomène de “mélange”, pour traduire le terme “*d’admixture*” classiquement utilisé), (ii) de structures en méta-populations ou encore (iii) de phénomènes de dérive génétique marquée (Fu et Arnold, 1991, 1992; Arnold *et al.*, 1988; Avise *et al.*, 1990).
3. dans le cas de *Cyclope neritea*, il nous faut étudier à la fois des processus contemporains (*e.g.* éventuels déplacements humains de populations en rapport avec les pratiques aquacoles) et des processus historiques (par exemple les réactions aux

⁸A l’origine, le concept d’homoplasie a été utilisé par les évolutionnistes pour décrire le fait qu’un caractère, présent dans deux espèces, n’est pas dérivé du même caractère présent chez l’ancêtre commun, mais que la similarité observée est due à des phénomènes tels que la convergence, le parallélisme ou la réversion (Estoup *et al.*, 2002). Ainsi, l’homoplasie est considérée par les évolutionnistes comme “*le bruit de fond*” alors que “*l’homologie*” est le signal évolutif réel (Scotland, 1992).

glaciations les plus récentes, tels que d'éventuels déplacements ou fragmentations d'aires de distribution) ayant potentiellement contribué à façonner la structure génétique actuelle des populations naturelles de l'espèce. Toutefois, même les processus historiques que nous cherchons à mettre en évidence restent relativement récents du point de vue évolutif : il s'agit de détecter des variations "*intra-spécifiques*" et non "*inter-spécifiques*". Les taux d'évolution variables des marqueurs employés seront là encore d'une aide précieuse.

Nous allons étudier ici 13 populations de l'aire d'origine reconnue de *Cyclope neritea* (incluant donc les trois populations situées sur la façade atlantique du Sud de la péninsule ibérique et présentées dans la partie précédente). La mise au point des deux types de marqueurs utilisés pour cette étude est détaillée aux sections A.2 et A.3 de l'annexe A. Il s'agit d'un fragment de 503 paires de bases du gène codant pour la sous-unité I de la cytochrome oxydase (gène mitochondrial) et de quatre locus microsatellites (*i.e.* 2C4, 3C4, 2H10b et 1A10).

Méthodes statistiques mises en œuvre

Les treize populations étudiées ici ont pour la plupart été récoltées suivant la méthode décrite dans la section 4.2.1 de la première partie, entre septembre 2002 et septembre 2005. Les populations de Sardaigne, Sicile, Grèce et Mer Noire ainsi que des individus de Faro nous ont été fournis par plusieurs collègues (Davide Tagliapietra, Francesco Patti, Jeroen Jansen et Dragos Micu). La position des sites d'échantillonnage est indiquée à la figure 6.1.

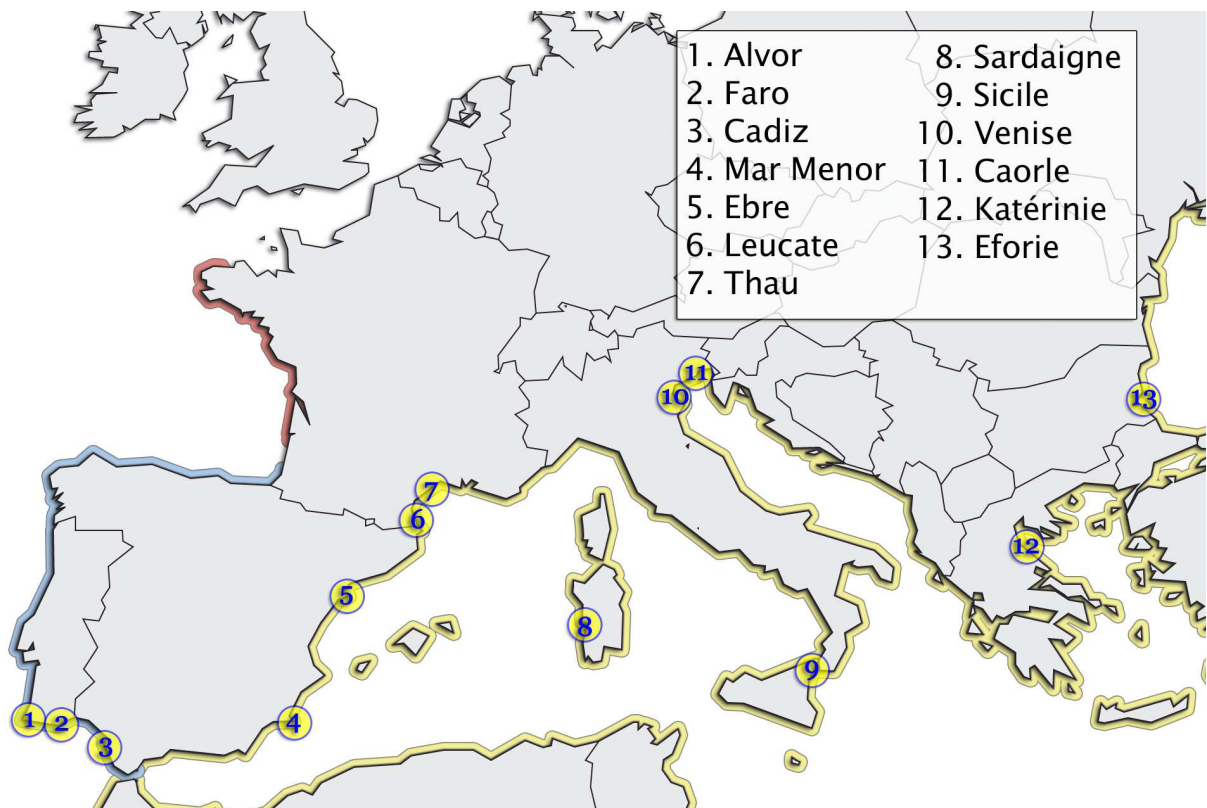


FIG. 6.1: Localisation des treize sites d'échantillonnage de l'aire "naturelle" de *Cyclope neritea*.

L'acquisition des données moléculaires (de l'extraction d'ADN au séquençage et au génotypage) a été réalisée selon les protocoles détaillés dans l'annexe A (voir aussi l'article présenté en annexe D.2 pour la mise au point des marqueurs microsatellites). Nous ne détaillerons ici que la méthodologie mise en œuvre pour les analyses statistiques.

6.1 Diversité génétique

En ce qui concerne les analyses mitochondriales, deux types d'indices de diversité génétique ont été calculés grâce au logiciel DNAsp v.3.51 (Rozas et Rozas, 1999) : des indices mesurant le polymorphisme haplotypique (nombre d'haplotypes⁹ N_H et diversité haplotypique $H_{E.mt}$), et des indices utilisant l'information nucléotidique des séquences (nombre de sites polymorphes S et diversité nucléotidique π). La diversité haplotypique $H_{E.mt}$ est estimée par la formule suivante :

$$\hat{H}_{E.mt} = \frac{n}{n-1} \left(1 - \sum_{i=1}^{N_H} p_i^2 \right)$$

avec n , le nombre de gènes dans l'échantillon (ici égal au nombre d'individus puisque le locus analysé est haploïde) et p_i la fréquence de l'haplotype i . Cet indice définit la probabilité que deux gènes tirés au hasard dans un échantillon soient identiques (Nei, 1987) et prend en compte le nombre d'individus ainsi que la fréquence des haplotypes.

La diversité nucléotidique π est estimée par (Tajima, 1983) :

$$\hat{\pi} = \sum_{i=2}^{N_H} \sum_{j<i} p_i p_j \delta_{ij}$$

N_H est le nombre d'haplotypes observés, p_i et p_j la fréquence des haplotypes i et j respectivement, et δ_{ij} la fraction de différences nucléotidiques entre les séquences i et j (soit le nombre de différences nucléotidiques entre les séquences i et j divisé par le nombre total de paires de bases de la séquence nucléotidique). La diversité nucléotidique mesure donc la divergence nucléotidique moyenne entre toutes les paires de séquences d'un même échantillon.

Ces indices ont été calculés pour chacune des 13 populations étudiées. Les moyennes arithmétiques (et leur écart-type associé) ainsi que les valeurs totales (*i.e.* sur l'ensemble des individus du jeu de données) ont également été calculées.

⁹C'est le nombre de séquences nucléotidiques uniques trouvées dans l'échantillon. Ces haplotypes peuvent différer les uns des autres par des substitutions d'une ou plusieurs bases, ou par des insertions-délétions.

Concernant les microsatellites, le nombre d'allèles N_{all} , la richesse allélique R_{all} ainsi que la diversité génique $H_{E.nuc}$ (*i.e.* hétérozygotie attendue à l'équilibre de Hardy Weinberg¹⁰) ont été calculés à chaque locus, pour chaque population et pour l'ensemble du jeu de données grâce au logiciel **Fstat v2.9.3** (Goudet, 1995). La richesse allélique permet de comparer le nombre d'allèles observé dans des populations ne présentant pas le même effectif. Pour cela, la méthode de raréfaction de Hurlbert (1971) est utilisée afin d'estimer la richesse allélique à un locus donné pour un échantillon de taille fixée. Si pour un locus donné, N gènes sont examinés, le nombre attendu d'allèles différents dans un échantillon de g gènes ($N > g$) est estimé par la formule suivante (Petit *et al.*, 1998) :

$$\hat{R}_{all} = \sum_i \left(1 - \frac{C_g^{N-N_i}}{C_g^N} \right)$$

Avec N_i , le nombre d'occurrences du i^{eme} allèle parmi les N gènes examinés. La diversité génique se définit comme $\hat{H}_{E.mt}$:

$$\hat{H}_{E.nuc} = \frac{n}{n-1} \left(1 - \sum_{i=1}^{N_{all}} p_i^2 \right)$$

avec n , le nombre de gènes dans l'échantillon (ici égal au double du nombre d'individus, soit $2N$, puisque le locus analysé est diploïde) et p_i la fréquence de l'allèle i .

Enfin, l'hypothèse nulle d'indépendance entre les locus a été testée dans **Genepop v3.4** (Raymond et Rousset, 1995) grâce à l'analyse du déséquilibre de liaison par la méthode de Fisher. La procédure implique un processus de Monte Carlo à Chaînes de Markov (processus MCMC) avec pour paramètres : “*Dememorization number*” : 10 000, “*Batch number*” : 5 000, “*Iterations per batch*” : 10 000.

6.2 Arbres et réseaux

En phylogénie inter-spécifique, la construction d'arbres, tels que des arbres de distances minimales (Minimum Spanning Tree), est classiquement effectuée. Les espèces sont alors reliées deux à deux à un ancêtre commun, selon le degré de similitude de leurs séquences nucléotidiques. Certaines des règles de construction de ces arbres ne sont pas valides en

¹⁰La loi de Hardy-Weinberg permet de décrire les relations entre fréquences alléliques et fréquences génotypiques dans une population théorique infinie, au sein de laquelle (i) la panmixie (rencontre des individus au hasard) et la pangamie (rencontre des gamètes au hasard) sont respectées, (ii) il n'y a ni mutations, ni migration, ni sélection sur les allèles considérés, et (iii) les générations ne se chevauchent pas. Dans ces conditions, la diversité génétique de la population se maintient et doit tendre vers un équilibre stable de la distribution génotypique.

phylogénie intra-spécifique. En effet, les haplotypes ancestraux, dont sont issus les haplotypes observés, sont potentiellement toujours présents dans les populations, celles-ci ayant divergé très récemment à l'échelle des temps évolutifs. Ainsi, des haplotypes ancestraux peuvent coexister avec des haplotypes nouveaux et des multifurcations peuvent apparaître dans les arbres (un haplotype ancestral donne naissance à plus de deux haplotypes nouveaux). En outre, différents haplotypes issus d'un haplotype ancestral peuvent eux même dériver l'un de l'autre. Un haplotype donné peut ainsi avoir plusieurs origines potentielles, ce qui a pour conséquence de créer des liens alternatifs entre les différents haplotypes, des structures cycliques pouvant alors apparaître. Dans ce contexte, l'utilisation de réseaux à la place d'arbres est préférable (Posada et Crandall, 2001). Les relations généalogiques entre les haplotypes ont donc été mises en évidence grâce à la construction d'un réseau avec le logiciel **Network v4.0.0.0** (Bandelt *et al.*, 1999), selon la méthode dite de “*median-joining*”, particulièrement adaptée au type d'étude que nous menons (Cassens *et al.*, 2003). La sélection des phylogénies intra-spécifiques les plus probables a été effectuée grâce à un algorithme de maximum de parcimonie, et un unique réseau synthétisant cette information a été représenté. Ce réseau a été interprété en parallèle d'une carte de distribution des haplotypes dans les différentes populations étudiées.

Outre les relations phylogénétiques existant entre les haplotypes observés, les similitudes entre les treize populations étudiées dans l'aire d'origine de l'espèce ont été révélées par la construction d'arbres de populations dans le logiciel **Mega v3.1** (Kumar *et al.*, 2004) sur les données mitochondriales d'une part, et microsatellites d'autre part. Pour cela, une matrice de distances génétiques a tout d'abord été calculée par la méthode de “*Kimura 2-paramètres*” (qui attribue des poids différents aux transitions et au transversion¹¹, Kimura, 1980) pour les données mitochondriales, et par la méthode des distances minimales de Nei (1972) pour les données microsatellites. La reconstruction phylogénique a été réalisée suivant un algorithme de “*neighbour-joining*” (Saitou et Nei, 1987). La réalisation de ces arbres suit donc une approche déterministe (calcul de distances génétiques entre populations). Nous verrons dans la partie consacrée au “*clustering*” qu'une approche probabiliste a également été envisagée et permet d'obtenir d'autres types d'informations.

¹¹Transitions et transversions sont deux types de substitutions nucléotidiques : une transition est le remplacement d'une base purine (A ou G) par une autre purine, ou d'une pyrimidine (T ou C) par une autre pyrimidine, et une transversion (généralement plus rare) est le remplacement d'une purine par une pyrimidine et réciproquement.

6.3 Analyses hiérarchiques, isolement par la distance

6.3.1 Marqueurs microsatellites

Au niveau intra-population, l'écart à l'équilibre de Hardy-Weinberg a été testé dans **Genepop v3.4** (Raymond et Rousset, 1995) par un test exact réalisé pour chaque locus et chaque population. Cette procédure implique un processus de Monte Carlo à Chaînes de Markov (processus MCMC) dont les paramètres utilisés sont les mêmes que ceux utilisés pour le test du déséquilibre de liaison : “*Dememorization number*” : 10 000, “*Batch number*” : 5 000, “*Iteration per batch*” : 10 000.

Afin de quantifier les déviations éventuelles à l'équilibre de Hardy-Weinberg, un estimateur de l'indice F_{IS} (voir encadré 6.1) a été calculé à chaque locus et pour chaque population avec le logiciel **Genepop v3.4**, d'après la méthode de Weir et Cockerham (1984). Ces calculs ($\hat{F}_{IS}$ et valeurs de probabilités associées aux tests exacts) ont également été réalisé pour chaque locus sur l'ensemble des populations et pour chaque population sur l'ensemble des locus. L'analyse des variations alléliques existant à différents niveaux de hiérarchisation a ensuite été réalisée grâce au logiciel **R** (R Development Core Team, 2005) et en particulier au package **hierFstat** (Goudet, 2005). Les niveaux de hiérarchisation impliqués sont les suivants : individus, populations, régions (regroupant plusieurs populations) et aire “*naturelle*” dans son ensemble. Les régions ont été définies en fonction des types mitochondriaux observés, afin de vérifier la cohérence entre données mitochondriales et données nucléaires.

La terminologie classique des F -statistiques (*i.e.* F_{ST} , F_{IS} , F_{IT} , voir encadré 6.1) a été modifiée suivant des notations similaires à celles utilisées par Trouve *et al.* (2005) : les indices *Ind*, *Pop*, *Reg* et *Tot* se rapportent respectivement aux *Individus*, *Populations*, *Régions* et au *Total*. Par exemple, F_{RegTot} est la corrélation entre gènes tirés au hasard au sein des régions par rapport à la corrélation entre gènes tirés au hasard dans l'aire totale. Ainsi, la structure totale observée F_{IndTot} sera le résultat de la structure à plusieurs niveaux : F_{IndPop} , F_{PopReg} et F_{RegTot} . La significativité de la déviation par rapport à zéro a été estimée par un test G , plus adapté qu'un test F classique pour des données multilocus avec des effectifs non identiques entre les populations ou les groupes (Goudet *et al.*, 1996). Ce test a été effectué dans **hierFstat** pour F_{PopReg} (5 000 permutations des individus entre populations d'une même région) et pour F_{RegTot} (5 000 permutations des populations entre les régions), et dans **Fstat** pour F_{IndPop} (5 000 permutations des allèles parmi les individus d'une même population) et pour F_{PopTot} (5 000 permutations des individus parmi les populations et les régions respectivement).

Les statistiques F

Wright (1951) a construit un ensemble d'outils mathématiques, les statistiques F , permettant de décrire la répartition de la variabilité génétique nucléaire entre et au sein des populations. Il a défini les trois indices suivants :

F_{ST} mesure la différenciation génétique entre populations.

F_{IS} mesure la déviation par rapport aux proportions attendues à l'équilibre de Hardy-Weinberg (*i.e.* structure attendue lorsque l'union des gamètes se fait au hasard) au sein de chacune des populations.

F_{IT} mesure la déviation par rapport aux proportions attendues à l'équilibre de Hardy-Weinberg sur l'ensemble des populations.

Ces indices sont liés par la relation $(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$.

Estimation de ces paramètres par la méthode de Weir et Cockerham (1984)

Weir et Cockerham (1984) ont montré la correspondance entre les indices F de Wright et les composantes de la variance des fréquences alléliques dans les populations structurées. Trois nouveaux indices, analogues de F_{ST} , F_{IS} et F_{IT} , peuvent être estimés par le biais d'une analyse de variance :

θ : la corrélation des allèles de différents individus d'une même sous-populations

f : la corrélation des allèles d'un même individu dans une sous-population donnée

F : la corrélation des allèles au sein des individus sur l'ensemble des sous-populations

On considère un modèle à r sous-populations de tailles égales, dérivées d'une population ancestrale à l'équilibre de Hardy-Weinberg, et en équilibre de liaison. Les sous-populations sont supposées n'avoir divergé les unes des autres que sous l'effet de la dérive génétique et de la dérive d'échantillonnage. Pour le cas d'un locus à deux allèles, les estimateurs de θ , f et F sont :

$$\hat{\theta} = \frac{a}{a + b + c} \quad 1 - \hat{f} = \frac{b}{b + c} \quad 1 - \hat{F} = \frac{c}{a + b + c}$$

avec

a : la composante de la variance entre les sous-populations

b : la composante de la variance au sein des populations

c : la composante de la variance au sein des individus

ENCADRÉ 6.1: Principe des statistiques F . D'après Dupont (2004).

6.3.2 Marqueur mitochondrial

Le même type d'analyse a été réalisé sur les données mitochondriales avec le logiciel *Arlequin v2.0* (Schneider *et al.*, 2001). L'analyse de variance moléculaire (AMOVA Excoffier *et al.*, 1992) permet de mesurer la différenciation génétique existant à plusieurs niveaux de hiérarchisation. Dans la mesure où le génome mitochondrial est haploïde, le niveau de variation individuel (qui mesure la corrélation des allèles au sein des individus) n'existe pas. Restent donc les niveaux populations, régions, et aire totale. Les estimateurs $\hat{\Phi}$ (indices de fixation analogues des paramètres F_{ST} , F_{CT} et F_{SC} , Wright, 1951, voir encadré 6.2) ont été calculés afin d'analyser les niveaux de différenciation génétique entre populations au sein de l'aire "naturelle" dans son ensemble ($\hat{\Phi}_{ST}$), entre régions distinctes (groupes de populations) au sein de l'aire "naturelle" ($\hat{\Phi}_{CT}$) et entre populations au sein de ces mêmes régions ($\hat{\Phi}_{SC}$). Les indices Φ prennent en compte non seulement les fréquences haplotypiques observées, mais également les distances nucléotidiques entre les haplotypes (encadré 6.2). Les populations ont été regroupées au sein de cinq régions distinctes entre lesquelles les échanges naturels de migrants sont potentiellement peu nombreux : la première région concerne les populations du bassin méditerranéen occidental (populations de la péninsule ibérique). La seconde région regroupe les populations des lagunes méditerranéennes françaises de Leucate et Thau et la troisième, celles de la Mer Adriatique. Enfin, deux autres régions ont été définies, l'une en Mer Égée et l'autre en Mer Noire. Toutefois, ces régions n'ont pas pu être incluses dans l'analyse, dans la mesure où elles n'abritent qu'une seule population chacune (les procédures de permutation nécessaires au calcul des indices nécessitant un minimum de deux populations par groupe). En outre, les populations insulaires de Sicile et de Sardaigne ont été exclues de l'analyse, car compte tenu de leur localisation géographique, il n'était *a priori* pas possible de les assigner spécifiquement et avec certitude à l'un des groupes précédemment définis. La significativité des indices a été testée dans *Arlequin v2.0* par une procédure de permutation.

6.3.3 Isolement génétique par la distance géographique

L'hypothèse d'un isolement génétique dû à un isolement géographique (IBD) a été testée pour les données mitochondriales et microsatellites dans l'aire "naturelle" de *Cyclope neritea*. Pour cela, un test de Mantel a été réalisé à l'aide du logiciel R (package *ade4*, R Development Core Team, 2005), afin de rechercher, pour chaque type de marqueur, une corrélation positive entre les matrices de distances géographiques et génétiques (5 000 permutations).

Les statistiques Φ

Nous avons vu que la diversité nucléotidique π d'une population s'écrit :

$$\pi = \sum_{ij} p_i p_j \delta_{ij}$$

avec p_i et p_j la fréquence des haplotypes i et j respectivement, et δ_{ij} la fraction de différences nucléotidiques entre les séquences i et j .

Tout comme il est intéressant de partitionner la diversité génétique au sein et entre les populations lorsque nous utilisons l'information de fréquences alléliques, il est intéressant de partitionner la diversité génétique entre et au sein de populations lorsque nous utilisons l'information des séquences nucléotidiques.

Notons p_{ik} , la fréquence du i^{eme} haplotype dans la k^{ieme} population. Alors,

$$p_i = \frac{1}{K} \sum_{k=1}^K p_{ik}$$

est la fréquence moyenne de l'haplotype i parmi l'ensemble des populations, avec K , le nombre total de populations. Il est alors possible de définir deux nouveaux indices :

$$\begin{aligned} \pi_t &= \sum_{ij} p_i p_j \delta_{ij} \\ \pi_s &= \frac{1}{K} \sum_{k=1}^K \sum_{ij} p_{ik} p_{jk} \delta_{ij} \end{aligned}$$

π_t est la diversité nucléotidique de l'ensemble des populations, et π_s est la diversité nucléotidique moyenne de chacune des populations. Alors,

$$\Phi_{ST} = \frac{\pi_t - \pi_s}{\pi_t}$$

est un analogue direct du F_{ST} de Wright concernant la diversité nucléotidique. De la même façon que pour les statistiques F , l'analyse moléculaire de variance peut inclure plusieurs niveaux de hiérarchie :

Φ_{ST} : différenciation entre l'ensemble des populations

Φ_{SC} : différenciation entre les populations au sein des régions

Φ_{CT} : différenciation entre les régions

Ces trois indices sont liés, comme dans le cas des statistiques F , par la relation

$$(1 - \Phi_{ST}) = (1 - \Phi_{CT})(1 - \Phi_{SC})$$

ENCADRÉ 6.2: Principe des statistiques Φ .

Le même estimateur de distance génétique a été utilisé pour les deux types de mar-

queur, $\hat{F}_{ST}$, calculé pour chaque paire de population et servant à déterminer la distance génétique définie par $\frac{\hat{F}_{ST}}{1-\hat{F}_{ST}}$ (Rousset, 1997). La longitude de chaque population a été utilisée pour calculer les distances géographiques entre paires de populations. Le choix d'utiliser la longitude, plutôt que la distance linéaire ou la distance en suivant le trait de côte, repose notamment sur le fait que la dispersion des populations de l'aire “naturelle” couvre moins 8 ° de latitude, mais s'étend sur près de 40 ° de longitude. L'axe longitudinal est donc potentiellement le plus discriminant.

6.4 Modèles de regroupements ou “clustering”

L'utilisation d'analyses hiérarchiques de la distribution des fréquences alléliques (telles que les statistiques F ou Φ), présente un inconvénient majeur : la nécessité de définir des groupes (*e.g.* populations, régions ...) *a priori*. Un ensemble de théories et techniques statistiques a été récemment développé afin de rechercher les regroupements d'individus les plus pertinents, en utilisant objectivement les données moléculaires à disposition, et parfois l'information de localisation géographique des individus collectés. Avec les données microsatellites obtenues dans les différentes populations “naturelles” de *Cyclope neritea*, nous avons mis en œuvre deux de ces méthodes de regroupement (ou “clustering”) utilisant des analyses bayésiennes. Il s'agit, sur la base des données génétiques, d'identifier un certain nombre de classes¹², c'est-à-dire des groupes d'individus ou de populations constituant des unités panmictiques à l'équilibre de Hardy-Weinberg. Nous avons tout d'abord utilisé le logiciel **Structure** v2.1 mis au point par Pritchard *et al.* (2000) et amélioré par Falush *et al.* (2003), puis le package **Geneland** (Guillot *et al.*, 2005b) fonctionnant sous R.

6.4.1 Principe

Dans leur récent article sur le chevreuil *Capreolus capreolus*, Coulon *et al.* (2006) comparent ces deux méthodes. Comme dans toutes les analyses bayésiennes, les principes de l'analyse sont les suivants : (i) il existe, pour chaque paramètre inconnu (principalement le nombre de classes et les fréquences alléliques au sein des classes inférées) une distribution statistique *a priori* et (ii) une fonction de vraisemblance relie ces paramètres aux génotypes observés. Les distributions *a priori* sont censées refléter ce qui est connu des paramètres non observés, mais également les incertitudes au sujet de ces paramètres. Pour

¹²Afin d'éviter toute confusion, le terme “classe” sera ici préféré à “population” et à “région”, le premier correspondant plutôt aux sites de prélèvements, et le second désignant les groupes de populations identifiés sur la base des ressemblances mitochondriales.

les deux logiciels, les paramètres inconnus sont inférés par un processus de Monte Carlo à Chaînes de Markov (MCMC).

Dans **Structure**, la distribution nulle des classes est supposée uniforme. Toutes les solutions de regroupement de populations sont donc considérées comme équiprobables. Avec cette approche, les coordonnées spatiales des individus ne sont utilisées qu'*a posteriori*, au moment de porter l'information de la classe des individus inférée par le programme, sur une carte, et ne sont donc pas utilisées au cours de l'analyse. Cela peut poser problème, dans la mesure où la plupart du temps, une différenciation entre populations n'est attendue que si ces populations sont séparées du fait de leur distance géographique ou par des barrières (biotiques, abiotiques...). C'est pourquoi Guillot *et al.* (2005a) ont proposé un modèle prenant en compte le fait que la différenciation génétique est censée apparaître dans des zones géographiques distinctes. Leur modèle inclut donc les coordonnées spatiales des individus dans l'analyse. En outre, le nombre de classes K , qui doit être fixé *a priori* pour chaque simulation dans **Structure**, est ici considéré comme un paramètre inconnu, et sera donc traité par le processus MCMC sans aucune approximation.

Les deux approches semblent pertinentes ici dans la mesure où d'une part, *Cyclope neritea* est une espèce incubante pour laquelle une forte structure phylogéographique est attendue, et d'autre part, les actions anthropiques sont susceptibles d'abolir les barrières naturelles aux flux de gènes, la localisation géographique des sites échantillonnés pouvant alors être trompeuse quant à l'origine des individus récoltés. Il sera donc intéressant de comparer les résultats obtenus par les deux méthodes de regroupement.

6.4.2 Paramètres utilisés

Structure

Le nombre de classes K est un paramètre fixé du modèle. Il faut donc réaliser plusieurs simulations et pour chacune, fixer K . Le critère de choix de ce modèle, pour déterminer la meilleure valeur de K , est l'estimation de la probabilité postérieure des données sachant K , notée $Pr(X|K)$ (Pritchard *et al.*, 2000). Cette valeur, appelée " $\ln P(D)$ " dans **Structure** est obtenue en calculant, à chaque étape du MCMC, le logarithme de la vraisemblance du jeu de données. A la fin de la simulation, la moyenne de ces valeurs est réalisée et la moitié de la variance est soustraite afin d'obtenir le critère de choix : $\ln P(D)$. Le nombre réel de classes est généralement identifié comme celui produisant la valeur de $\ln P(D)$ maximale (Zeisset et Beebee, 2001; Vernesi *et al.*, 2003).

Toutefois, l'utilisation de ce critère est sujette à caution. Evanno *et al.* (2005) ont en effet montré que le nombre de classes inféré par cette méthode était souvent surestimé.

Leurs simulations ont en particulier révélé que, lorsque la valeur de K réelle est atteinte, $\text{Ln } P(D)$ (appelé par ces auteurs $L(K)$) ne représente pas toujours un maxima évident. Les valeurs de $L(K)$ obtenues pour les K supérieurs, continuent au contraire de croître légèrement ou forment un plateau, et les variances entre les simulations augmentent. C’est pourquoi ces auteurs ont proposé d’utiliser un indice différent qui montre un pic clair à la réelle valeur de K : le taux de variation de second ordre de $L(K)$ nommé ΔK .

$$\Delta K = \frac{m(|L''(K)|)}{s[L(K)]}$$

avec m la moyenne, s l’écart-type et

$$|L''(K)| = |L'(K + 1) - L'(K)|$$

$$L'(K) = L(K) - L(K - 1)$$

soit

$$\Delta K = \frac{m(|L(K + 1) - 2L(K) + L(K - 1)|)}{s[L(K)]}$$

La valeur de K retenue sera donc celle pour laquelle ΔK est maximal. Nous avons ainsi testé des valeurs de K comprises entre 1 et 10, en réalisant chaque simulation six fois afin de s’assurer de la cohérence des résultats. Les paramètres de simulation sont les suivants : “burn-in”¹³ : 50 000 étapes, nombre de répétition de la Chaîne de Markov : 150 000. Dans la mesure où les populations de *Cyclope neritea* peuvent avoir subi l’influence des échanges humains, un modèle avec mélange de classes (phénomène de “mélange” : mélange de lignées évolutives divergentes au sein de la même population, ou ici, au sein de la même classe) a été considéré, le taux de mélange pouvant être différent pour chaque classe. Avec ce modèle, chaque individu est susceptible d’avoir hérité de gènes provenant de plusieurs classes distinctes. Enfin, dans la mesure où les populations isolées sont censées évoluer indépendamment les unes des autres, les fréquences alléliques inférées dans les différentes classes peuvent être différentes.

Geneland

Ce package repose sur les attendus suivants :

1. le nombre de classes est inconnu et toutes les valeurs, comprises entre 1 et la borne maximale fixée par l’utilisateur, sont équiprobables.

¹³La terminologie employée correspond à celle du logiciel.

2. les classes sont distribuées dans l'espace suivant une tessellation¹⁴ de Poisson-Voronoi. Lorsque le nombre de polygones est inférieur à 5 000, Kumar et Kurtz (1993) ont montré que les paramètres topologiques de ce type de tessellation (nombre et longueur des arêtes, périmètre et surface des polygones...) pouvaient être approximés par une distribution log-normale.
3. chaque classe est supposée à l'équilibre de Hardy-Weinberg.
4. les fréquences alléliques au sein de chaque classe sont inconnues et sont traitées comme des variables aléatoires suivant le modèle de Dirichlet¹⁵ (Falush *et al.*, 2003).

Avec ce logiciel, tous les paramètres, y compris K , sont recherchés simultanément par le processus MCMC. Toutefois, pour des raisons pratiques (voir Guillot *et al.*, 2005a), il est conseillé de réaliser une première simulation pour inférer K , puis de relancer l'algorithme avec K fixé à la valeur précédemment définie, afin d'estimer les autres paramètres (en particulier, l'assignation des individus aux classes identifiées). Nous avons lancé 5 simulations (afin de s'assurer de la cohérence des résultats) en autorisant K à varier entre 1 et 26 (26 correspondant au double du nombre de populations, il est fort probable que le nombre réel de classes dans le jeu de données est inférieur à cette valeur) avec les paramètres suivants : 150 000 itérations, taux maximal du processus de Poisson : 430 (soit, conformément aux recommandations de Guillot *et al.* (2005b), une valeur égale au nombre d'individus testés), nombre maximum de noyaux de la tessellation de Poisson-Voronoi : 1 290 (soit trois fois le taux maximal du processus de Poisson). Nous avons ensuite réalisé 5 simulations en fixant K à la valeur inférée au cours des simulations précédentes, les autres paramètres étant identiques. Les probabilités postérieures d'appartenance des individus aux classes identifiées ont enfin été calculées, et la précision des cartes de probabilité a été fixée, selon les axes latitudinaux et longitudinaux des graphiques, à 50 et 250 pixels respectivement.

¹⁴Une "tessellation" est créée par le recouvrement complet d'une surface par un tamis de polygones.

¹⁵Généralisation multivariée de la distribution beta.

7

Résultats

7.1 Comparaison entre diversité mitochondriale et nucléaire

Les indices de diversité génétique calculés sur les données mitochondriales sont présentés dans le tableau 7.1 (la localisation géographique des sites étudiés est présentée à la figure 7.3).

Nous avons séquencé 533 paires de bases d'un fragment du gène codant pour la sous-unité I de la cytochrome oxydase chez 410 individus répartis en 13 populations. Le nombre d'individus par population est compris entre 25 et 40 (sauf en Sicile : 7 individus). Le nombre moyen d'haplotypes observés par population est de 2,8 alors que 23 haplotypes ont été identifiés sur l'ensemble des populations. La majorité des haplotypes a en effet été observée dans une minorité de populations (Katérinie, Venise, Caorle : 9, 6 et 5 haplotypes respectivement), alors que 7 populations sur 13 sont monomorphes (*i.e.* le même haplotype a été identifié chez tous les individus de chacune de ces populations).

De même, le nombre de sites ségrégeants total s'élève à 43 alors que le nombre moyen par population n'est que de 6,4. En dehors des populations monomorphes ne présentant évidemment pas de sites variable, seules les populations de Katérinie et de Sicile présentent une valeur de S identique à leur nombre d'haplotypes, ce qui indique des haplotypes peu divergents du point de vue évolutif au sein de ces populations. Les autres populations présentent un profil contraire : par exemple, 13 sites ségrégeants séparent les 3 haplotypes de la population de Leucate, ou les 5 haplotypes de Caorle en présentent 21.

Les fortes différences observées entre les valeurs moyennes et totales des paramètres $\hat{H}_{E.mt}$ et $\hat{\pi}$ reflètent les situations contrastées observées avec N_H et S . Ainsi, si en moyenne, chaque population semble peu diversifiée ($\hat{H}_{E.mt.moy} = 0,231$, $\hat{\pi}_{moy} = 0,224 \cdot 10^{-2}$), au

TAB. 7.1: Diversité génétique des populations de l'aire "naturelle" de *Cyclope neritea*. (A) : Océan Atlantique, (M) : Mer Méditerranée, (Ad) : Mer Adriatique, (E) : Mer Égée et (N) : Mer Noire. Les indices présentés sont : N : nombre d'individus analysés, N_H : nombre d'haplotypes, S : nombre de sites variables, $\hat{H}_{E.mt}$: diversité haplotypique, $\hat{\pi}$: diversité nucléotidique. La moyenne et l'écart-type (S.D.) de ces indices sont indiqués, ainsi que le total, c'est-à-dire les indices calculés sur l'ensemble des individus.

Population	Localisation (Mer-Océan)	N	N_H	S	$\hat{H}_{E.mt}$	$\hat{\pi} \cdot 10^2$
1. Alvor	Algarve (A)	39	1	0	0,000	0,000
2. Faro	Ria Formosa, Algarve (A)	31	1	0	0,000	0,000
3. Cadiz	Andalousie (A)	36	1	0	0,000	0,000
4. Mar Menor	Murcie (M)	37	4	13	0,527	0,216
5. Ebre	Catalogne (M)	33	1	0	0,000	0,000
6. Leucate	Languedoc-Roussillon (M)	37	3	13	0,156	0,255
7. Thau	Languedoc-Roussillon (M)	29	1	0	0,000	0,000
8. Sardaigne	Golfe d'Oristano (M)	25	1	0	0,000	0,000
9. Sicile	Détroit de Messine (M)	7	2	2	0,571	0,214
10. Venise	Veneto (Ad)	40	6	25	0,603	1,251
11. Caorle	Veneto (Ad)	38	5	21	0,543	0,769
12. Katérinie	Grèce (E)	32	9	9	0,607	0,211
13. Eforie	Roumanie (N)	26	1	0	0,000	0,000
<i>Moyenne</i>		31,5	2,8	6,4	0,231	0,224
<i>(S.D.)</i>		(8,8)	(2,8)	(6,4)	(0,283)	(0,377)
<i>Total</i>		410	23	43	0,744	1,521

niveau global, en revanche, une très forte diversité est observée ($\hat{H}_{E.mt.tot} = 0,744$, $\hat{\pi}_{tot} = 1,521 \cdot 10^{-2}$), indiquant des différences génétiques marquées entre les populations. La plupart des populations monomorphes sont situées à l'ouest de la zone étudiée alors que les populations les plus polymorphes sont situées à l'est (Caorle et Venise présentent les plus fortes valeurs de S , et Katérinie présente le plus grand nombre d'haplotypes).

Cette situation, bien que moins tranchée, reste toutefois bien visible avec les marqueurs microsatellites. Les indices de diversité génétique, calculés après génotypage des quatre locus étudiés, sont présentés dans le tableau 7.2.

Le nombre total d'individus génotypés est légèrement supérieur au nombre d'individus dont le fragment du gène COI a été séquencé (430 contre 410 respectivement). Tous les individus séquencés ont toutefois été génotypés. Les locus ne présentent pas de déséquilibre

TAB. 7.2: Diversité génétique des 4 locus microsatellites pour chaque population de l'aire "naturelle". N : nombre d'individus, N_{all} : nombre d'allèles, $\hat{R}_{all}$: richesse allélique, $\hat{H}_{E,nuc}$: diversité génétique. L'écart-type associé à chaque calcul de moyenne multilocus est indiqué entre parenthèses.

Population	N	2C4			3C4			2H10b			1A10			Multilocus		
		N_{all}	$\hat{R}_{all}$	$\hat{H}_{E,nuc}$	N_{all}	$\hat{R}_{all}$	$\hat{H}_{E,nuc}$	N_{all}	$\hat{R}_{all}$	$\hat{H}_{E,nuc}$	N_{all}	$\hat{R}_{all}$	$\hat{H}_{E,nuc}$	$\bar{N}_{all}$	$\bar{\hat{R}}_{all}$	$\bar{\hat{H}}_{E,nuc}$
1. Alvor	39	4	2,78	0,546	4	3,88	0,674	9	6,34	0,829	11	6,62	0,808	7,0	4,92	0,714
														(3,6)	(1,88)	(0,132)
2. Faro	31	4	2,89	0,561	4	3,60	0,648	16	9,23	0,880	6	5,49	0,810	7,5	5,30	0,725
														(5,7)	(2,84)	(0,146)
3. Cadiz	36	5	2,75	0,543	5	3,86	0,609	8	6,10	0,832	10	7,04	0,852	7,0	4,94	0,709
														(2,5)	(1,98)	(0,156)
4. Mar Menor	37	4	3,95	0,737	5	4,12	0,648	12	8,08	0,849	11	7,88	0,860	8,0	6,01	0,774
														(4,1)	(2,28)	(0,100)
5. Ebre	33	5	3,38	0,514	5	4,39	0,706	14	7,83	0,857	18	10,45	0,924	10,5	6,51	0,750
														(6,6)	(3,25)	(0,182)
6. Leucate	37	6	4,41	0,661	6	4,20	0,611	11	5,89	0,698	17	8,44	0,800	10,0	5,73	0,693
														(5,2)	(1,95)	(0,080)
7. Thau	29	6	4,41	0,692	6	5,17	0,792	13	8,09	0,882	20	11,37	0,942	11,3	7,26	0,827
														(6,7)	(3,16)	(0,109)
8. Sardaigne	25	6	4,31	0,622	4	3,59	0,697	9	6,35	0,788	10	7,07	0,828	7,3	5,53	0,734
														(2,8)	(1,65)	(0,092)
9. Sicile	10	4	4,00	0,761	7	7,00	0,847	9	8,58	0,839	12	11,20	0,933	8,0	7,69	0,845
														(3,4)	(3,01)	(0,070)
10. Venise	38	7	4,67	0,713	6	5,28	0,772	20	9,63	0,895	25	10,95	0,905	14,5	7,63	0,821
														(9,5)	(3,12)	(0,094)
11. Caorle	40	5	4,13	0,638	7	5,37	0,809	20	9,92	0,891	32	13,18	0,956	16,5	8,15	0,824
														(12,6)	(4,18)	(0,138)
12. Katérinie	40	11	6,79	0,831	13	7,12	0,847	18	9,79	0,904	31	12,99	0,954	18,3	9,17	0,884
														(9,0)	(2,88)	(0,056)
13. Eforie	35	6	4,12	0,666	6	4,01	0,455	13	6,04	0,592	28	13,57	0,964	13,3	6,94	0,669
														(10,4)	(4,52)	(0,215)
Moyenne	33,1	5,6	4,05	0,653	6,0	4,74	0,701	13,2	7,84	0,826	17,8	9,71	0,887	10,65	6,58	0,767
(S.D.)	(8,3)	(1,9)	(1,05)	(0,095)	(2,4)	(1,19)	(0,112)	(4,2)	(1,54)	(0,08)9	(8,8)	(2,76)	(0,063)	(7,14)	(2,89)	(0,130)
Total	430	20		0,786	18		0,826	47		0,933	66		0,948	37,75		0,873

de liaison (voir tableau 7.3), ce qui garantit leur indépendance statistique.

TAB. 7.3: Déséquilibres de liaison pour l'ensemble des populations de l'aire "naturelle", entre chaque paire de locus. Les déséquilibres de liaison entre paires de populations (non présentés) sont tous non significatifs.

Locus	χ^2	<i>d.d.l.</i>	probabilité <i>p</i>
2C4 - 3C4	16,211	28	0,963
2C4 - 2H10b	21,421	30	0,874
3C4 - 2H10b	30,887	28	0,322
2C4 - 1A10	20,856	30	0,892
3C4 - 1A10	11,852	28	0,997
2H10b - 1A10	11,053	30	0,999

Le schéma général observé avec les données mitochondriales (*i.e.* grandes disparités selon les populations étudiées et diversité génétique supérieure dans les populations situées à l'est de la zone étudiée) se retrouve ici. Ainsi, pour un locus donné, la diversité génétique peut varier du simple au double (*e.g.* , pour le locus 3C4, $\hat{H}_{E.nuc}$ vaut 0,455 pour la population d'Eforie, et 0,847 pour les populations de Katérinie et de Sicile) et le nombre d'allèles du simple au quintuple (pour 1A10, entre 6 et 32 allèles ont été observés selon les populations). Comme avec les données mitochondriales, les populations les plus diversifiées sont trouvées à l'est de la zone étudiée : pour chaque locus, le plus grand nombre d'allèles est presque systématiquement trouvé dans les populations 10 à 13 (Venise, Caorle, Katérinie et Eforie). Comme le montre la figure 7.1, ce phénomène est indépendant de la taille des populations puisque le nombre d'allèles par population et la richesse allélique de chaque population varient dans le même sens (valeurs calculées sur les données multilocus). La régression linéaire réalisée montre même un meilleur coefficient de corrélation quand toutes les populations sont ramenées à une taille similaire ($r^2 = 0,603$, $p = 0,002$ pour $\hat{R}_{all}$ et $r^2 = 0,517$, $p = 0,006$ pour N_{all}).

En outre, pour chaque locus, le nombre total d'allèles est nettement supérieur au nombre d'allèles observé dans chacune des 13 populations étudiées. Par exemple, pour le locus le moins variable (3C4), entre 4 et 13 allèles ont été observés dans les populations (moyenne : 6 allèles par population) alors que 18 allèles ont été identifiés au total. La situation observée localement est donc très différente de la situation observée à l'échelle globale.

Les deux types de marqueurs apportent ici une information concordante : (i) les populations présentent des profils de diversité génétique différents selon les régions (faible

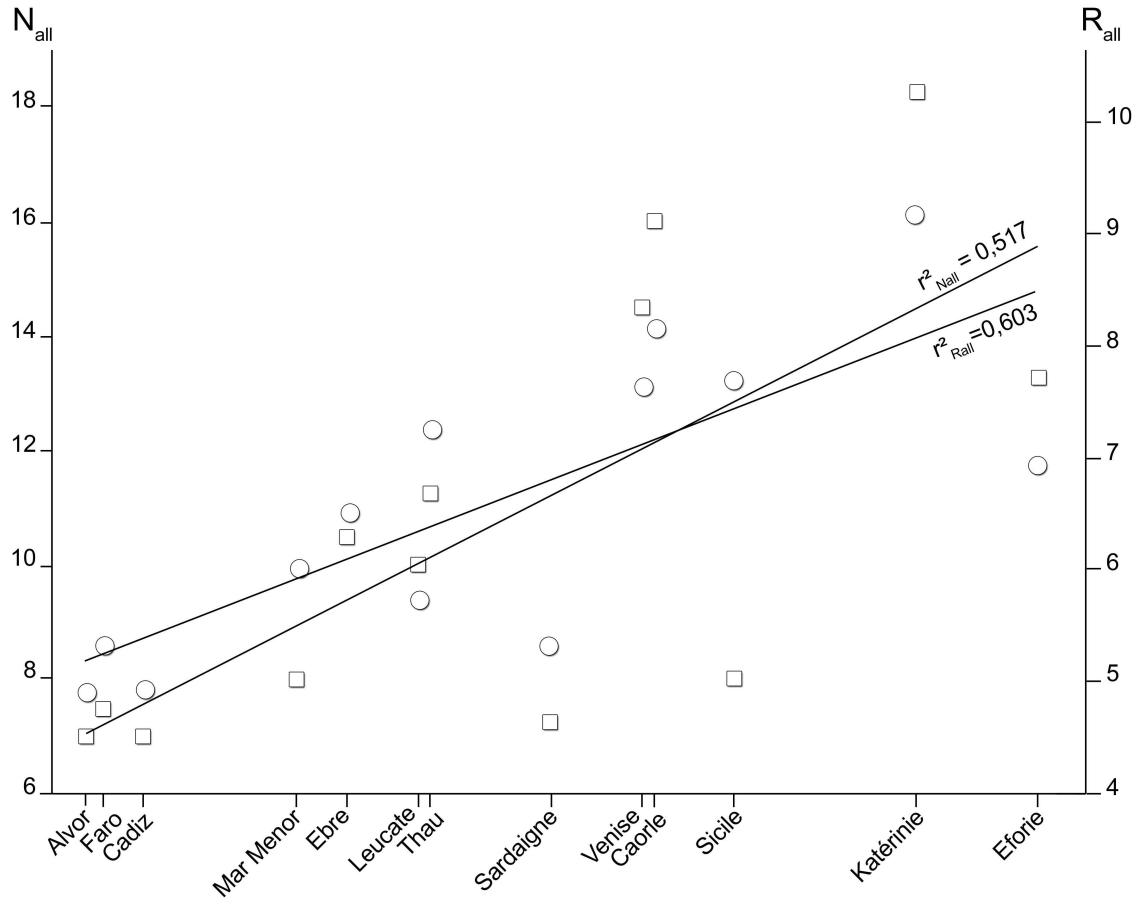


FIG. 7.1: Nombre d'allèles et richesse allélique dans les populations de l'aire "naturelle" de *Cyclope neritea*. Les carrés correspondent au nombre d'allèle N_{all} dans chaque population (axe de gauche) et les cercles à la richesse allélique $\hat{R}_{all}$ (axe de droite). Afin de mettre en évidence le gradient de diversité génétique est-ouest, les populations ont été placées sur l'axe des abscisses en fonction de leur longitude. Pour chaque série, une droite de régression a été tracée et le coefficient de corrélation est indiqué (les deux régressions sont significatives : $p = 0,006$ et $0,002$ pour N_{all} et $\hat{R}_{all}$ respectivement).

à l'ouest et fortes à l'est) (ii) la diversité locale intra-population est nettement inférieure à celle observée à l'échelle de l'ensemble du jeu de données, ce qui laisse présager une structure génétique importante au sein de l'aire "*naturelle*" de *Cyclope neritea*.

7.2 Une structure génétique marquée

7.2.1 Distribution géographique des haplotypes et divergence mitochondriale

Le réseau d'haplotypes de la figure 7.2 représente les relations phylogénétiques entre les haplotypes identifiés dans l'ensemble du jeu de données acquis au cours de cette thèse. Il inclut non seulement les populations de l'aire "*naturelle*" étudiées dans ce chapitre, mais également les populations françaises récemment détectées ainsi que les populations de statut indéterminé, qui feront l'objet d'une étude spécifique exposée dans la partie III. Nous avons choisi d'intégrer l'ensemble des données obtenues au cours de cette thèse afin d'obtenir le réseau le plus informatif et le mieux résolu possible. Ce réseau a ainsi été construit sur la base d'un jeu de données représentant 652 individus.

Bien que le réseau tracé représente l'ensemble des haplotypes de notre jeu de données, les résultats présentés ici ne concerneront que les populations de l'aire "*naturelle*". Au sein des populations "*naturelles*", trois haplotypes majoritaires ont été identifiés : il s'agit de H_1 , H_{18} et H_{22} (188, 64 et 47 individus sur 410 respectivement, soit une fréquence totale de 0,73). Dans une moindre mesure, les haplotypes H_{30} , H_{33} et H_{17} ont été trouvés en 26, 20 et 19 copies respectivement (soit une fréquence de 0,16). Sur 23 haplotypes identifiés dans l'aire "*naturelle*" de l'espèce, ces 6 haplotypes représentent donc près de 90% des séquences.

Le réseau se compose de cinq branches distinctes reliées entre elles par un groupe d'haplotypes dont les relations généalogiques sont non résolues (comme en témoigne la présence de boucles au centre du réseau, tous les "*chemins évolutifs*" possibles étant équiprobables). C'est également à cette zone centrale qu'est relié l'unique haplotype de *Cyclope pellucida* identifié dans la partie I. Au sein de chacune des cinq branches, les divergences nucléotidiques sont faibles (entre 1 et 6 mutations) alors qu'elles sont nettement plus importantes entre les branches (de 11 à 20 substitutions). Trois des cinq branches (jaune, rouge et verte) présentent une phylogénie en étoile (ou phylogénie "*star-like*"), généralement caractéristique de populations en expansion.

La distribution des haplotypes dans le réseau est en accord avec leur distribution géographique dans les populations de l'aire "*naturelle*". La figure 7.3 présente la distribution

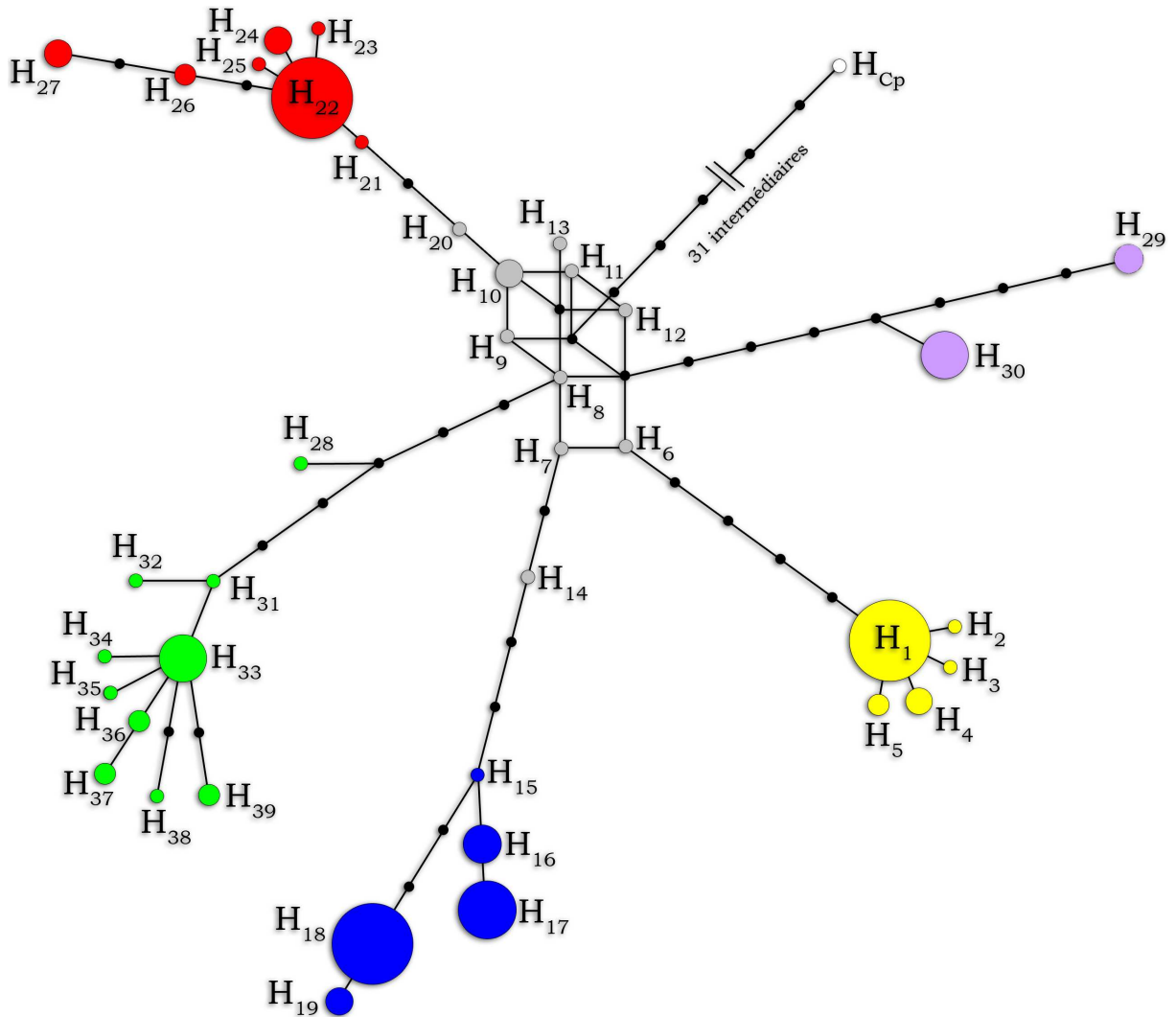


FIG. 7.2: Réseau présentant les relations phylogénétiques entre les haplotypes de *Cyclope neritea*. Chaque trait représente une substitution nucléotidique et les points noirs sur les branches représentent des haplotypes intermédiaires manquants. La taille des cercles est proportionnelle à la fréquence des haplotypes. Chaque couleur correspond à une lignée évolutive distincte. Tous les haplotypes identifiés dans notre jeu de données (*i.e.* aire “naturelle” de *Cyclope neritea*, populations françaises récemment enregistrées et populations au statut indéterminé, $N = 652$) ont été utilisés pour représenter le réseau le plus informatif possible. Ainsi, 16 haplotypes présentés au sein de ce réseau n’ont pas été trouvés dans les populations “naturelles” de *Cyclope neritea*. Il s’agit des haplotypes H_3 , H_4 , H_6 à H_{15} , H_{19} à H_{21} et H_{23} . L’unique haplotype identifié chez *Cyclope pellucida* (extrêmement divergent des haplotypes de *Cyclope neritea*) est indiqué en blanc.

géographique des haplotypes au sein des populations de l’aire “naturelle” de *Cyclope neritea*. Plusieurs régions peuvent être distinguées, chacune correspondant à une “lignée” (ou groupe d’haplotypes proches) identifiée par le réseau d’haplotypes. La couleur des haplotypes est la même sur la carte et le réseau. Ainsi, les populations de la péninsule

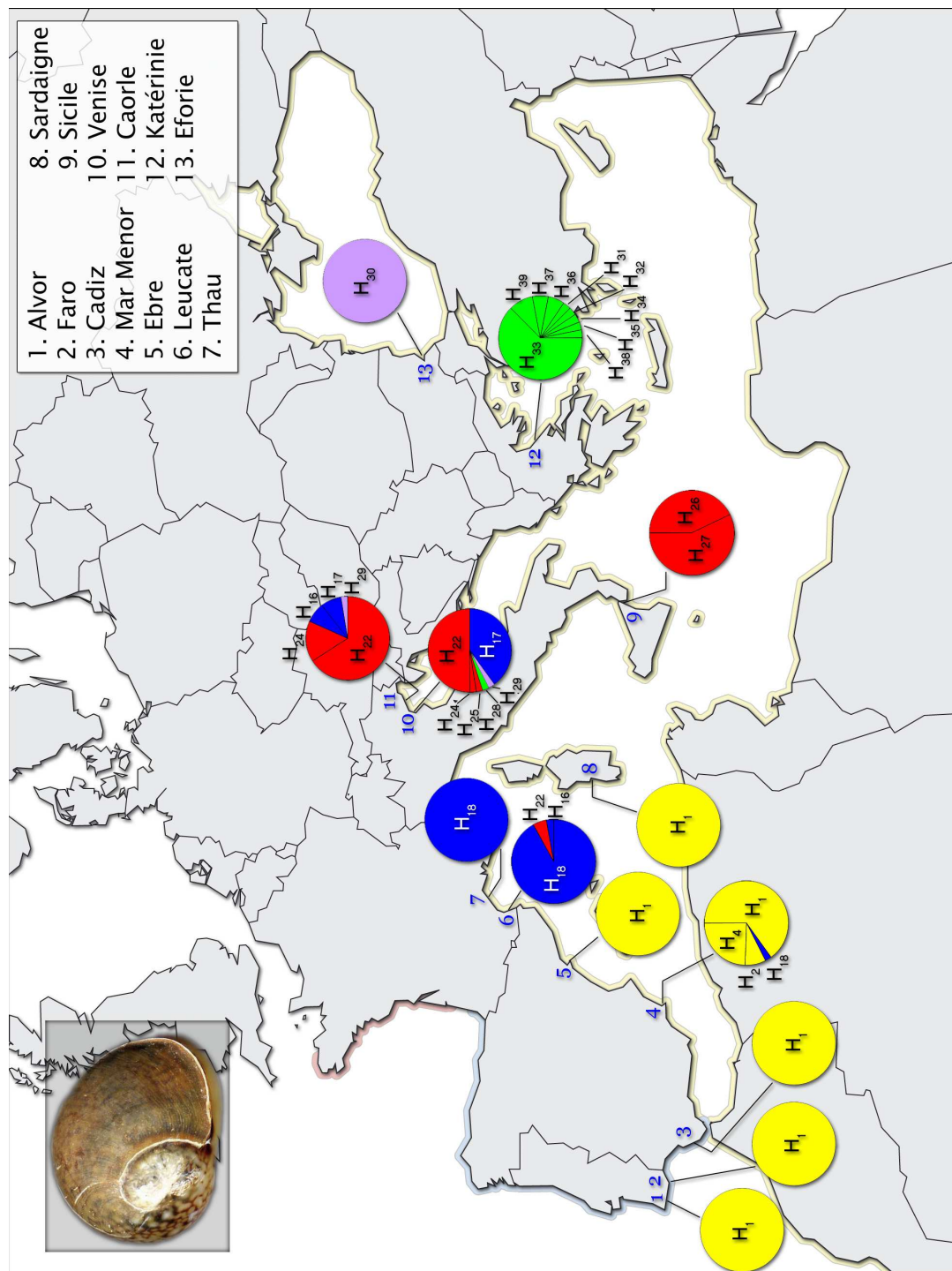


FIG. 7.3: Distribution des haplotypes dans les populations "naturelles" de *Cyclope neritea*. Le nom et la couleur des haplotypes sont les mêmes que ceux utilisés pour le réseau de la figure 7.2.

ibérique et de la Sardaigne présentent toutes l'haplotype majoritaire H_1 et sont quasi-monomorphes. Les populations méditerranéennes des côtes françaises (Leucate et Thau) semblent elles-aussi quasi-monomorphes et se caractérisent par la présence exclusive du second haplotype majoritaire : H_{18} . La population de Sicile et celles de l'Adriatique sont caractérisées par la présence du troisième haplotype majoritaire (H_{22}). Les populations de Venise et Caorle ont toutefois des profils plus complexes et sont plus diversifiées que les populations de l'ouest de la zone. Enfin, les populations de Katérinie et d'Eforie, représentent chacune une lignée évolutive distincte : Katérinie est la seule population où des haplotypes de la lignée figurée en vert sont présents et avec 9 haplotypes, c'est la population la plus diversifiée de la zone étudiée. La population d'Eforie est elle-aussi la seule représentante de la lignée d'haplotypes figurée en violet, mais au contraire de la population de Katérinie, elle est monomorphe.

7.2.2 Analyses hiérarchiques et isolement par la distance

L'analyse moléculaire de variance (AMOVA) réalisée sur les données mitochondriales a permis de mettre en évidence une forte et significative structure générale ($\hat{\Phi}_{ST} = 0,868$, $p = 0,009$) entre les trois régions définies à la section 6.3.2. Cette structure est principalement due à l'importante différenciation génétique existant entre les trois régions ($\hat{\Phi}_{CT} = 0,854$, $p < 10^{-4}$), alors que la différenciation génétique entre populations d'une même région, bien que significative, est beaucoup plus faible ($\hat{\Phi}_{SC} = 0,091$, $p < 10^{-4}$).

Cette forte structure générale, moins visible au niveau de l'ADN nucléaire ($\hat{F}_{ST} = 0,130$, voir aussi le tableau 7.4) reste toutefois présente. La structure globale ($F_{IndTot} = 0,207$) se partage quasiment de façon égale entre les trois niveaux de hiérarchie suivants : $F_{IndPot} = 0,088$, $F_{PopReg} = 0,076$ et $F_{RegTot} = 0,059$ (diagonale du tableau 7.4, les valeurs observées sont toutes hautement significatives $p < 10^{-3}$).

TAB. 7.4: F -statistiques hiérarchiques. La seconde valeur de la première ligne correspond par exemple au F_{PopTot} .

	Région	Population	Individu
Total	0,059	0,130	0,207
Région	-	0,076	0,157
Population	-	-	0,088

Au niveau de la population, l'étude des $\hat{F}_{IS}$ révèle des situations tranchées selon les populations considérées (tableau 7.5). Sur les 13 populations étudiées, 7 s'écartent significativement de l'équilibre de Hardy-Weinberg (Faro, Mar Menor, Leucate, Thau, Venise,

TAB. 7.5: Valeurs de $\hat{F}_{IS}$ et probabilités associées (indiquées entre parenthèses). Les valeurs de $\hat{F}_{IS}$ significatives sont indiquées en gras.

	2C4	3C4	2H10b	1A10	Multilocus
1. Alvor	-0,222 (0,177)	-0,065 (0,158)	-0,076 (0,475)	-0,111 (0,967)	-0,112 (0,367)
2. Faro	0,010 (0,548)	0,126 (0,373)	0,120 (0,004)	0,276 (0,004)	0,145 (0,001)
3. Cadiz	-0,023 (0,684)	-0,186 (0,389)	0,266 (0,025)	0,087 (0,344)	0,060 (0,144)
4. Mar Menor	0,450 (0,000)	-0,085 (0,130)	0,204 (0,046)	0,225 (0,033)	0,208 (0,000)
5. Ebre	0,057 (0,053)	-0,159 (0,133)	0,045 (0,435)	0,049 (0,511)	0,000 (0,117)
6. Leucate	0,182 (0,163)	-0,018 (0,344)	0,109 (0,029)	0,444 (0,000)	0,193 ($<10^{-4}$)
7. Thau	0,198 (0,093)	-0,176 (0,499)	0,140 (0,049)	0,214 (0,001)	0,094 (0,001)
8. Sardaigne	0,099 (0,117)	0,139 (0,479)	-0,067 (0,134)	0,082 (0,082)	0,059 (0,063)
9. Sicile	0,606 (0,024)	0,213 (0,593)	0,166 (0,311)	0,036 (0,262)	0,242 (0,096)
10. Venise	0,193 (0,001)	-0,030 (0,240)	0,078 (0,165)	-0,050 (0,443)	0,043 (0,006)
11. Caorle	0,257 (0,060)	-0,074 (0,900)	0,181 (0,078)	0,208 (0,003)	0,140 (0,004)
12. Katérinie	0,025 (0,154)	-0,033 (0,390)	-0,051 (0,848)	0,065 (0,182)	0,001 (0,321)
13. Eforie	0,586 (0,000)	-0,006 (0,500)	-0,255 (0,997)	-0,037 (0,087)	0,055 ($<10^{-4}$)
Total	0,169 ($<10^{-4}$)	-0,040 (0,356)	0,065 (0,001)	0,148 ($<10^{-4}$)	0,087 (0,000)

Caorle et Eforie) dans le sens d'un déficit en hétérozygotes ($\hat{F}_{IS}$ positifs et significatifs, compris entre 0,043 et 0,208, tableau 7.5). Des profils différents sont également observés suivant le locus considéré. Ainsi, 3C4 ne présente par exemple aucune valeur de $\hat{F}_{IS}$ significative; les populations de Faro, Leucate et Thau présentent des valeurs de $\hat{F}_{IS}$ significatives uniquement pour deux locus (2H10b et 1A10); enfin, seule la population de la Mar Menor présente des valeurs de $\hat{F}_{IS}$ significatives sur trois des quatre locus. Ceci traduit des déficits en hétérozygotes pouvant avoir plusieurs causes : allèles nuls, régimes de reproduction consanguins, effet Wahlund. Ces différentes causes potentielles seront discutées dans la section 8.3 de cette partie.

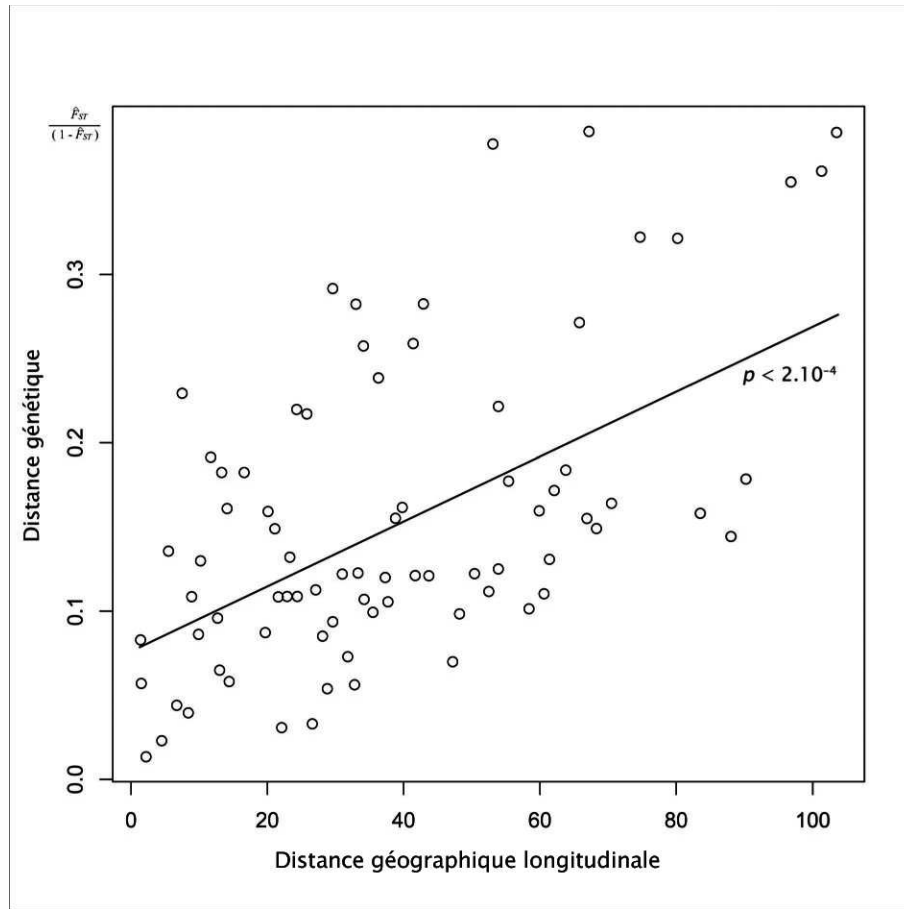


FIG. 7.4: Profil d'isolement génétique par la distance géographique mis en évidence dans l'aire "naturelle" sur les données microsatellites. La droite de régression indique la corrélation positive entre distance génétique ($\frac{\hat{F}_{ST}}{1 - \hat{F}_{ST}}$) et distances "longitudinales" mise en évidence par le test de Mantel hautement significatif ($p < 2.10^{-4}$).

Ainsi, les marqueurs microsatellites ont révélé une structure particulière au sein de certaines populations. En revanche, au niveau de la structure globale, l'effet "région" est moins visible avec les données microsatellites qu'avec les données mitochondriales. La

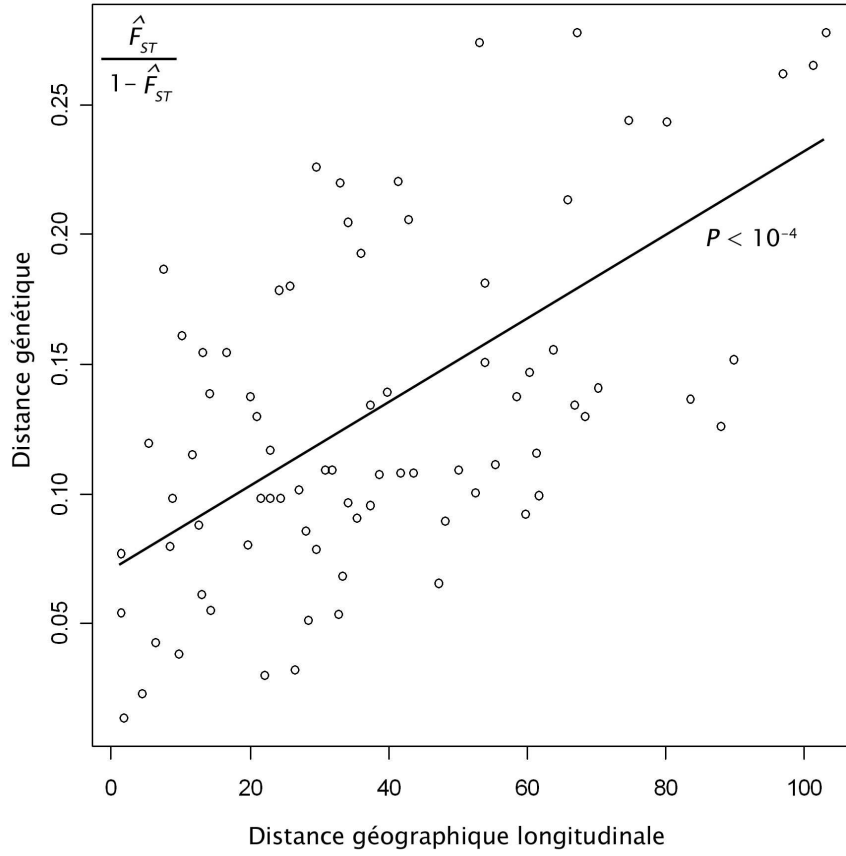


FIG. 7.5: Profil d'isolement génétique par la distance géographique mis en évidence dans l'aire "naturelle" sur les données mitochondriales. La droite de régression indique la corrélation positive entre distance génétique ($\frac{\hat{F}_{ST}}{1 - \hat{F}_{ST}}$) et distances "longitudinales" mise en évidence par le test de Mantel hautement significatif ($p < 10^{-4}$).

structure à grande échelle est toutefois bien réelle : la différenciation génétique est-ouest est continue et hautement significative pour les deux types de marqueurs ($p < 2.10^{-4}$ et $p < 2.10^{-4}$ pour le marqueur mitochondrial et les marqueurs nucléaires respectivement.), comme en attestent les profils d'isolement génétique par la distance géographique et les résultats des tests de Mantel présentés aux figures 7.4 et 7.5.

7.2.3 Arbres de populations

Afin de visualiser les éventuelles similitudes entre les différentes populations de l'aire d'origine de *Cyclope neritea*, deux arbres de populations ont été réalisés. Quelle que soit l'information utilisée pour les construire (données mitochondriales, figure 7.6, ou

nucléaires, figure 7.7), ces arbres de populations s'accordent sur :

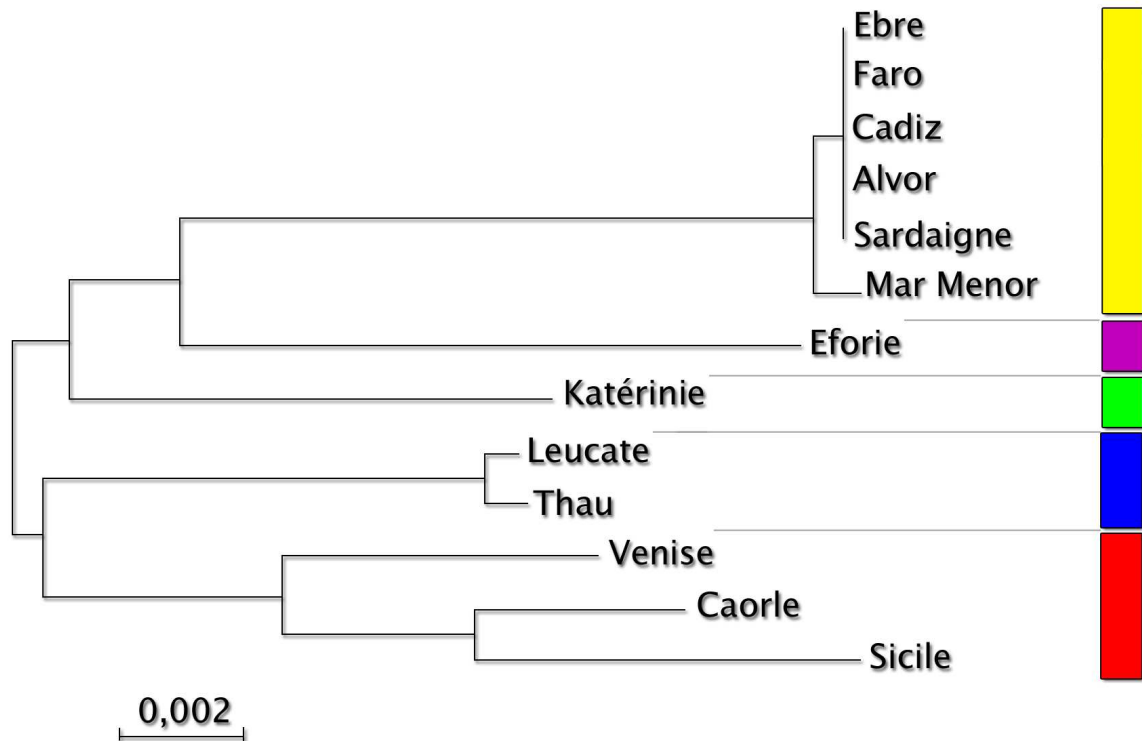


FIG. 7.6: Arbre des populations “naturelles” construit avec les données haplotypiques. Les couleurs à droite correspondent aux lignées évolutives majoritaires dans chacune des populations.

1. une forte convergence de l'information apportée par les deux types de marqueurs, malgré des modes d'hérédité distincts et des taux d'évolution probablement très différents.
2. l'évidence de similitudes régionales au sein de l'aire “naturelle” de l'espèce. Les populations sont pour la plupart regroupées en fonction de leur type mitochondrial majoritaire, la forte convergence entre structure génétique et géographique décrite par les figures 7.2 et 7.3 se retrouve donc ici.

En revanche, la position de la population de la Mar Menor n'est pas la même dans les deux arbres : si elle fait partie intégrante du groupe ibérique à la figure 7.6, elle s'en distingue à la figure 7.7 et vient s'intercaler entre les deux populations des côtes méditerranéennes françaises.

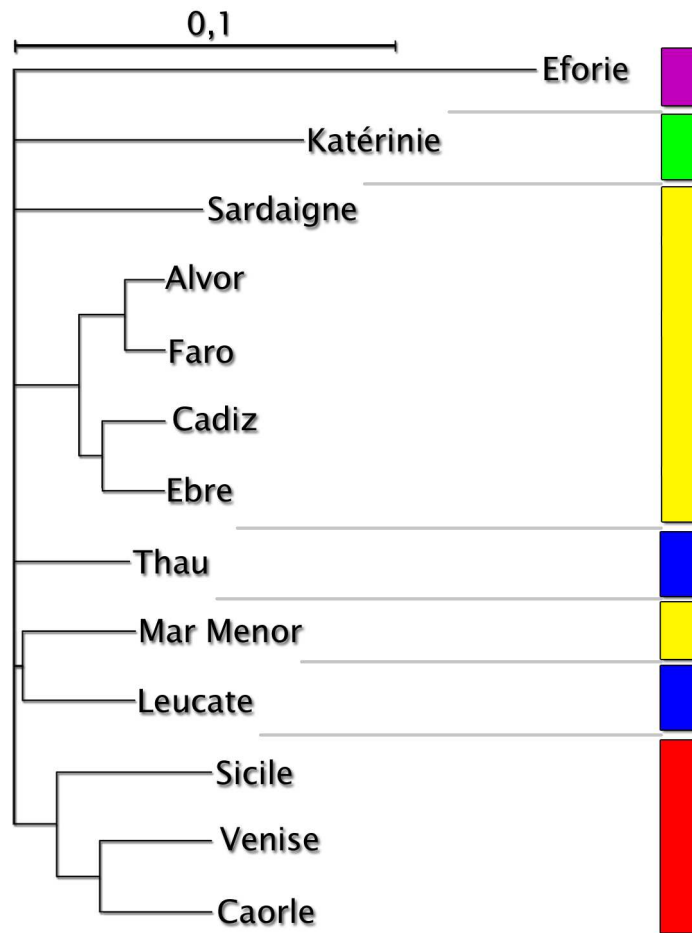


FIG. 7.7: Arbre des populations “naturelles” construit avec les données microsatellites. Les couleurs à droite correspondent aux lignées évolutives majoritaires dans chacune des populations.

7.3 Inférences bayésiennes

7.3.1 Structure génétique inférée avec le logiciel Structure

En utilisant les données nucléaires, les résultats de la recherche objective du nombre de classes (ou groupes de populations) réellement présentes au sein de l'aire d'origine sont présentés dans les figures 7.8 et 7.10.

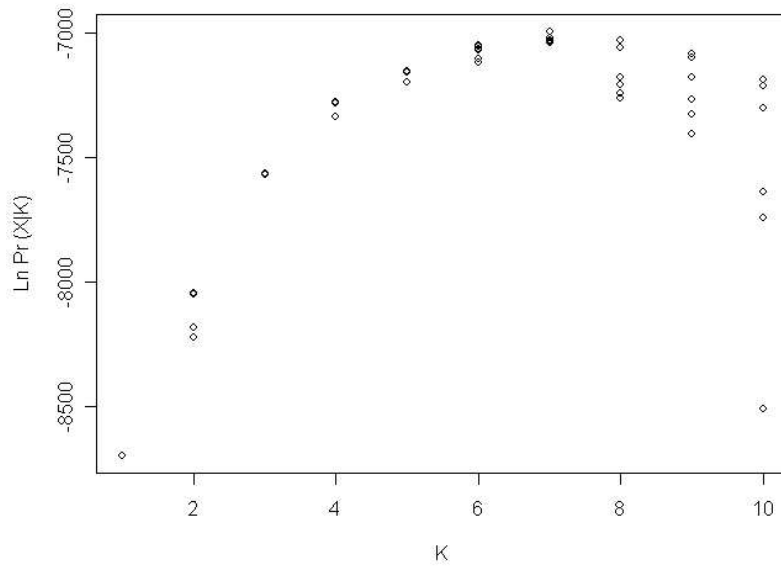
La figure 7.8a montre l'évolution du $\text{Ln } P(D)$ en fonction du K testé (6 simulations ont été réalisées pour chaque valeur de K). Le nombre de classes, d'après ce critère, correspond au K pour lequel le $\text{Ln } P(D)$ est maximal, soit ici, $K = 7$. Toutefois, comme nous l'avons vu plus haut, ce nombre est souvent surestimé en raison du maintien d'une faible augmentation de $\text{Ln } P(D)$ ou de l'apparition d'un plateau quand le K réel est atteint (Evanno *et al.*, 2005). Ainsi, le ΔK a été calculé et est présenté à la figure 7.8b. Le pic franc à $K = 3$ indique sans ambiguïté le nombre de classes inférées à partir des résultats obtenus dans **Structure**.

La figure 7.9 indique pour les deux cas de figure (*i.e.* $K = 3$ et $K = 7$) et pour chaque individu, les probabilités postérieures d'appartenance aux classes identifiées. La figure 7.9a montre que les individus composant les populations d'Alvor, Faro, Cadiz et Ebre sont associés au sein de la même classe. Les populations de la Mar Menor et de la Sardaigne, associées à la région 1 avec les données mitochondriales, et dans une moindre mesure, la population de Sicile, sont ici associées aux populations de Leucate et Thau. La plupart des individus de Venise semblent être issus de la même classe que les individus de Katérinie et Eforie. Enfin, la population de Caorle paraît avoir un profil plus complexe avec un mélange d'individus ayant des probabilités non négligeables d'appartenance à chacune des trois classes.

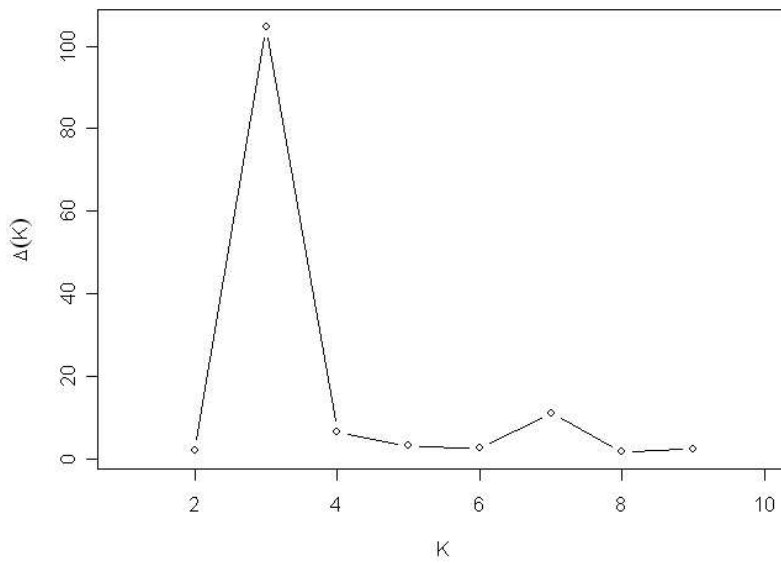
La figure 7.9b permet de retrouver certains de ces résultats, mais la situation de quelques populations est moins tranchée. Par exemple, les individus de la Mar Menor semblent pour la plupart avoir une probabilité partagée d'appartenance à deux classes : la première est caractéristique des individus de Leucate, de Thau et de Sardaigne (classe figurée en bleu foncé), la seconde n'est retrouvée nulle part ailleurs (figurée en vert). La population de Thau paraît abriter des individus issus de plusieurs classes et les populations de Venise, Caorle, Katérinie et Eforie sont maintenant séparées dans 4 classes distinctes.

7.3.2 Structure génétique inférée avec le logiciel Geneland

Le nombre de classes inféré par **Geneland** (figure 7.10) est très différent de celui inféré par la méthode du ΔK et correspond au K inféré avec le $\text{Ln } P(D)$, soit $K = 7$. Toute-



a. Critère préconisé dans **Structure**.



b. Critère défini par Evanno *et al.* (2005).

FIG. 7.8: Inférence du nombre de classes grâce aux simulations réalisées par **Structure**.

fois, des résultats cohérents apparaissent avec les deux méthodes. Ainsi, en intégrant les données géographiques, le regroupement de populations au sein des classes inférées est quasiment identique à celui, basé exclusivement sur les données génétiques, réalisé dans **Structure**.

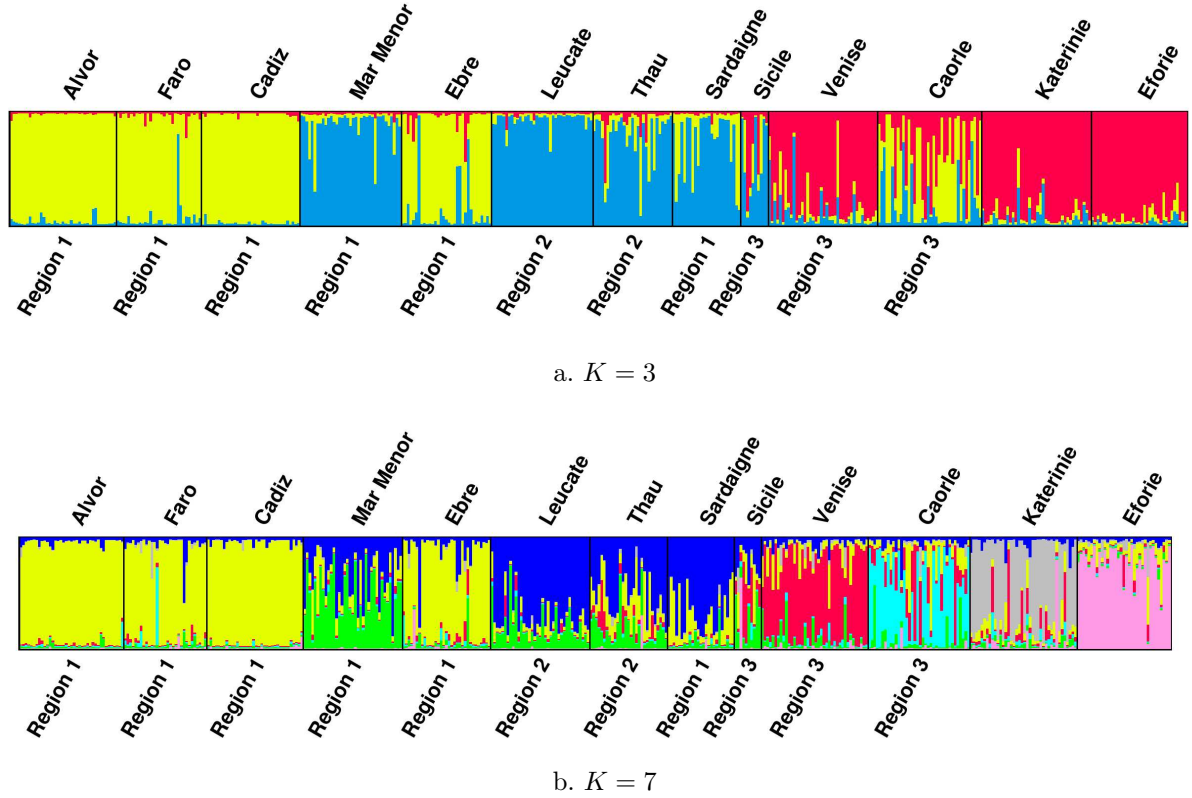


FIG. 7.9: Probabilités postérieures individuelles d'appartenance aux classes identifiées par **Structure**. Chaque ligne verticale correspond à un individu, les couleurs indiquant l'appartenance aux différentes classes. Le nom de chaque population est indiqué ainsi que la région à laquelle les populations ont été rattachées sur la base des analyses mitochondriales.

La figure 7.11 présente les probabilités postérieures d'appartenance de chaque population aux classes inférées par **Geneland**. Étant donné l'échelle de l'aire étudiée, nous avons considéré que tous les individus d'une même population ont été prélevés rigoureusement au même endroit. Ainsi, les résultats obtenus pour une population donnée représentent en fait la somme des résultats individuels de la population considérée.

Une classe regroupant les populations d'Alvor, Faro, Cadiz et Ebre a été identifiée (figure 7.11b). De même, la classe regroupant les populations de la Mar Menor, Leucate Thau et Sardaigne est également présente (figure 7.11c). La troisième classe regroupe les populations de Sicile et de Venise (figure 7.11d). Enfin, les classes 4 à 7 regroupent systématiquement les populations de Caorle, Katérinie et Eforie, ce qui suggère que la sur-représentation du nombre de classes pour ces trois populations est liée à leur grande diversité génétique.

En conclusion, quelle que soit la méthode utilisée, on note une forte convergence entre les regroupements effectués : non seulement les deux types d'approches bayésiennes réalisées sur les données microsatellites fournissent des résultats similaires, mais en outre, les

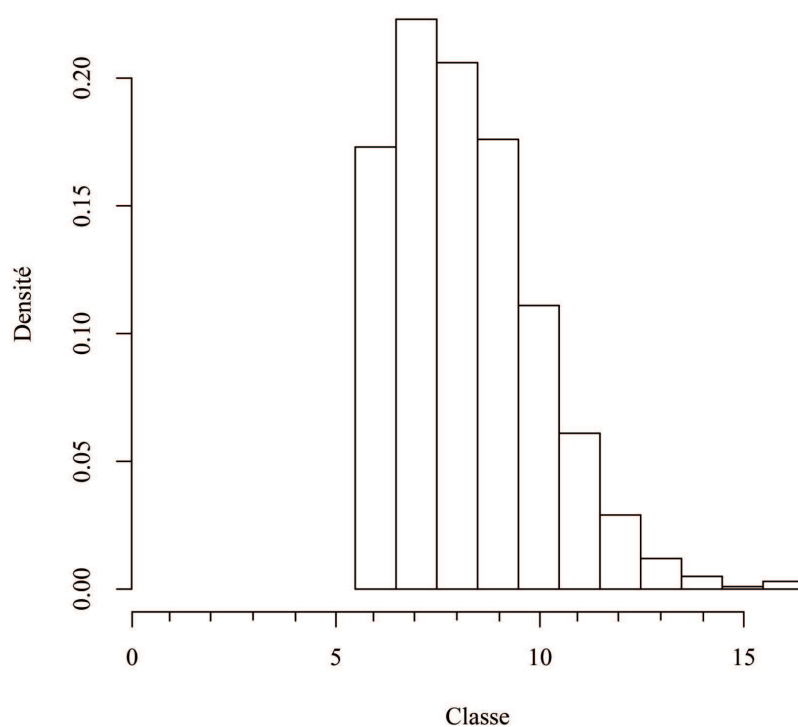
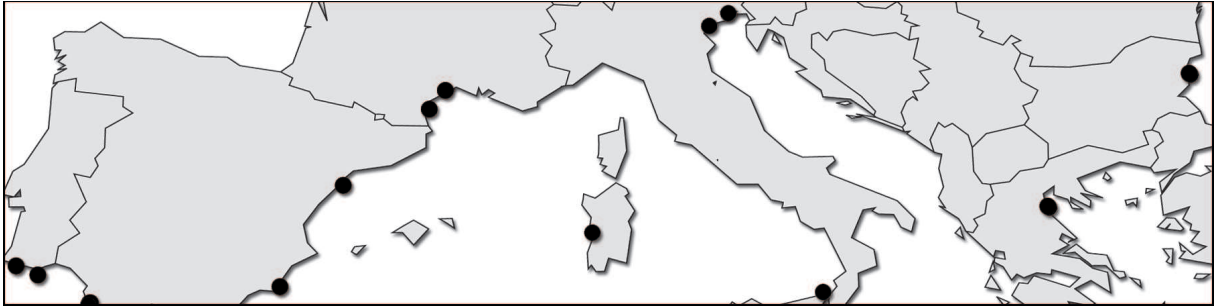
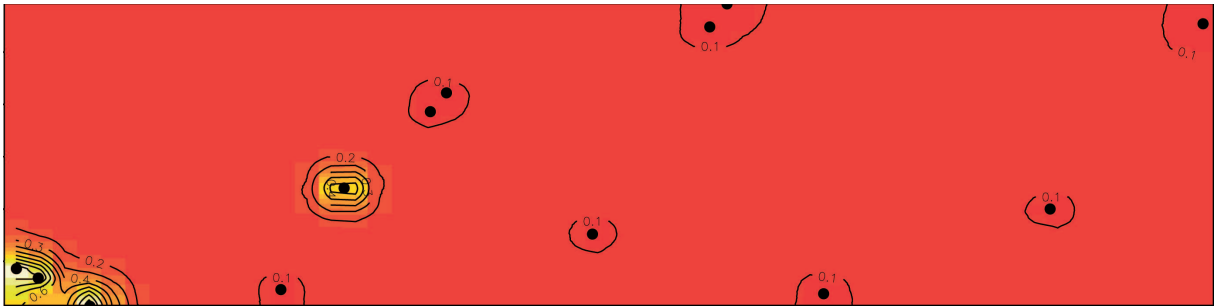


FIG. 7.10: Distribution du nombre de classes inférées par **Geneland**

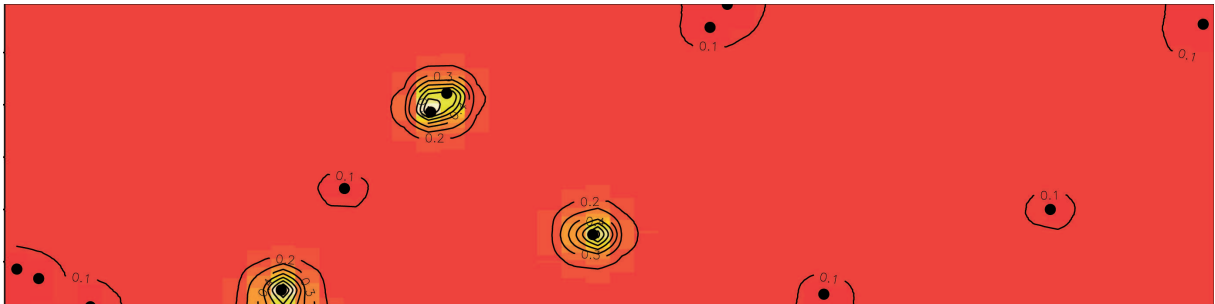
résultats fournis par ces approches sont également cohérents avec les résultats obtenus avec les données mitochondriales.



a. Localisation des sites d'échantillonnage



b. Probabilité postérieure d'appartenance à la classe 1

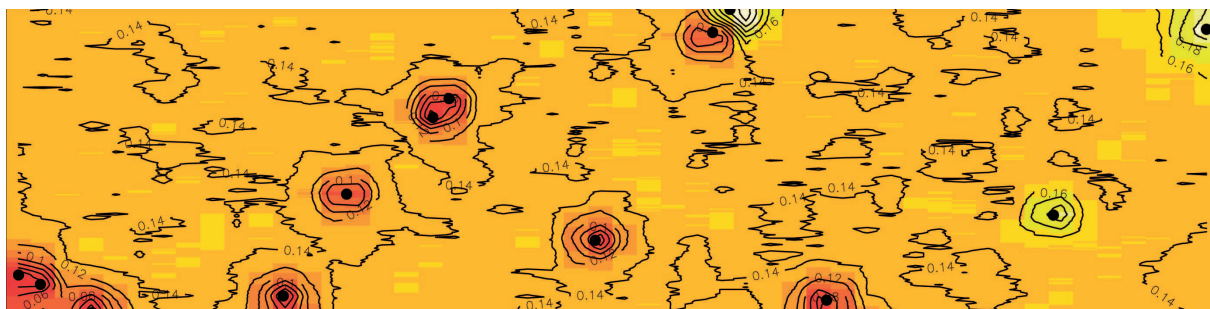


c. Probabilité postérieure d'appartenance à la classe 2

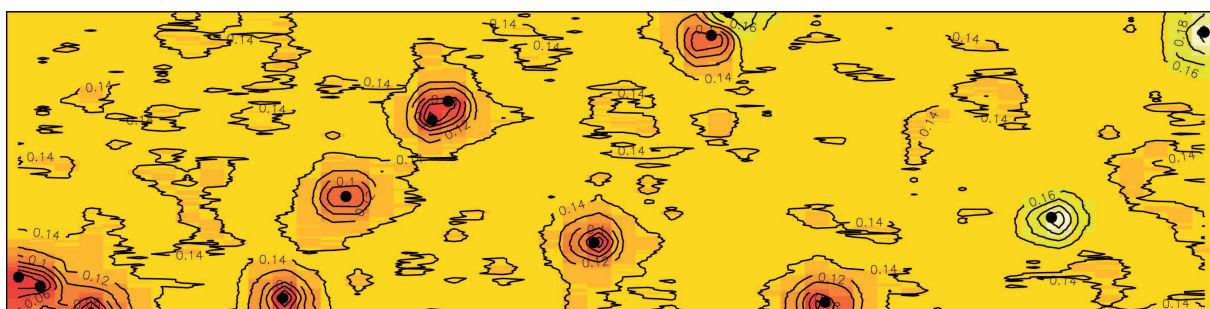


d. Probabilité postérieure d'appartenance à la classe 3

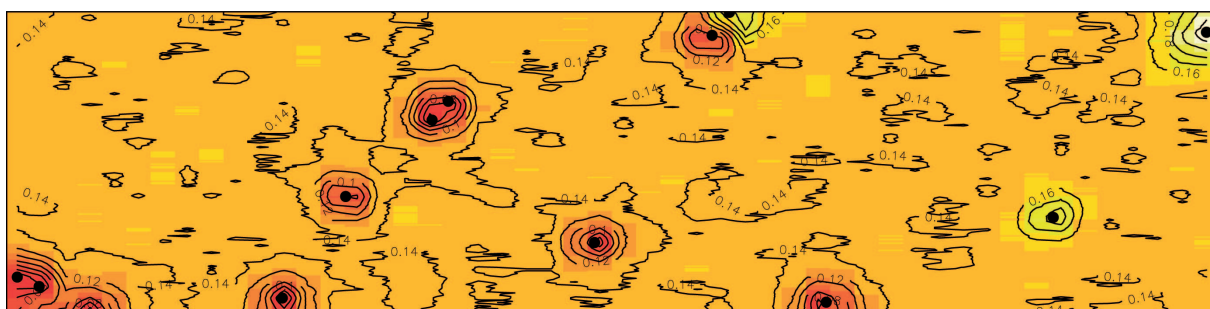
FIG. 7.11: Cartes des probabilités postérieures d'appartenance aux 7 classes identifiées par Geneland. La sous-figure 7.11a rappelle à titre indicatif la position des populations échantillonnées. Les cartes de probabilités postérieures reprennent la position de ces populations sans faire figurer le trait de côte.



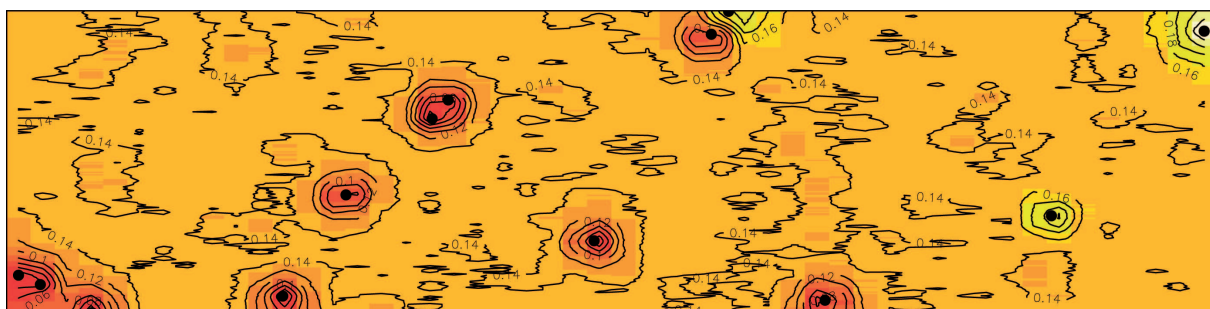
e. Probabilité postérieure d'appartenance à la classe 4



f. Probabilité postérieure d'appartenance à la classe 5



g. Probabilité postérieure d'appartenance à la classe 6



h. Probabilité postérieure d'appartenance à la classe 7

FIG. 7.11: *Suite.* Cartes des probabilités postérieures d'appartenance aux 7 classes identifiées par Geneland.

Discussion

8.1 Profil génétique d'une espèce à faible pouvoir dispersif

À l'échelle de la population, le faible niveau de diversité observé sur les données mitochondriales (2,8 haplotypes et 6,4 sites variables en moyenne, $\hat{H}_{E.mt.moy} = 0,231$ et $\hat{\pi}_{moy} = 0,224.10^{-2}$, tableau 7.1) est comparable à celui observé dans d'autres études portant sur des séquences de COI au niveau infra-spécifique (*e.g.* *Scylla serrata* Gopurenko *et al.*, 1999 ; *Patelloida profunda* Kirkendale et Meyer, 2004 ; *Littorina sp.* Wilding *et al.*, 2000). De même, concernant les locus microsatellites, avec un nombre d'allèles total compris, selon les locus, entre 18 et 66, et une diversité génique variant, selon les populations, entre 0,455 et 0,964 (tableau 7.2), les niveaux de variabilité observés ici sont conformes à ceux observés dans des études similaires, en particulier pour des espèces ayant de faibles capacités dispersives, en raison de leurs traits d'histoire de vie. C'est notamment le cas de *Littorina saxatilis*, espèce ovovivipare (Sokolov *et al.*, 2002) ou de *Buccinum undatum*, gastéropode à développement direct (Weetman *et al.*, 2005). Les valeurs de diversités observées sont en outre inférieures à de nombreuses études portant sur des espèces à fort pouvoir dispersif. C'est par exemple le cas de l'étude réalisée par Huvet *et al.* (2000) sur le bivalve *Crassostrea angulata* : sur un total de 382 individus répartis en 9 populations, les valeurs de $\hat{H}_{E.nuc}$ sont toujours supérieures à 0,934, quels que soient la population et le locus considérés (voir également Dufresne *et al.*, 2002; Launey *et al.*, 2002; Dupont, 2004). Au niveau de la population, les niveaux de diversité génétique observés chez *Cyclope neritea* semblent donc conformes à l'attendu théorique pour une espèce à faible capacité de dispersion, et ce pour le marqueur mitochondrial et les marqueurs nucléaires.

À l'échelle globale, des niveaux de diversité importants sont observés, notamment au

niveau de l'ADN mitochondrial (23 haplotypes et 43 sites variables, $\hat{H}_{E.mt.tot} = 0,744$ et $\hat{\pi}_{tot} = 1,521.10^{-2}$, tableau 7.1). Les différences marquées entre niveaux de diversité moyens et totaux sont en accord avec une forte structure génétique au sein de l'aire “naturelle” de *Cyclope neritea*, cohérente avec les faibles capacités de dispersion de l'espèce. Cette structure est visible non seulement avec les données mitochondriales, mais également avec les données nucléaires et ce malgré une taille efficace quatre fois supérieure. Ainsi, au niveau mitochondrial, cinq lignées évolutives divergentes correspondant à cinq régions distinctes de l'aire “naturelle” de *Cyclope neritea* ont été identifiées (figures 7.2 et 7.3) et au niveau nucléaire, un gradient de diversité est-ouest, ainsi qu'un profil d'isolement par la distance hautement significatif, ont été mis en évidence (figures 7.1 et 7.4). En outre, à l'exception de la population de la Mar Menor, les arbres de populations montrent les mêmes regroupements de populations pour les deux types de marqueurs (figures 7.6 et 7.7), renforçant l'idée d'une forte structure régionale au sein de l'aire “naturelle” de *Cyclope neritea*.

Les profils de diversité et de structure génétique observés ici avec les deux types de marqueurs moléculaires sont donc tout à fait conformes à l'attendu théorique d'une faible diversité locale associée à une forte structure globale, ce qui nous place dans une situation idéale pour retrouver l'origine des populations récemment installées sur les côtes françaises.

8.2 Processus historiques

La structure du réseau d'haplotypes présentée à la figure 7.2 est particulièrement informative sur l'histoire de l'espèce dans son aire “naturelle”. Tout d'abord, l'insertion de l'haplotype de *Cyclope pellucida* sur le centre du réseau semble indiquer que ces haplotypes centraux sont bien les plus anciens, chacune des cinq branches ayant progressivement divergé des haplotypes ancestraux. Le fait que la structure en cinq lignées divergentes soit fortement supportée par la structure géographique dans l'aire “naturelle” (voir la distribution des haplotypes à la figure 7.3) et sur le plan statistique, par l'analyse de variance moléculaire, renforce l'idée d'une divergence historique de ces lignées. L'existence de tels niveaux de divergence traduit un isolement durable (flux de gènes limités, voire absents entre les populations ancestrales dont sont issues les populations actuelles).

En utilisant l'horloge moléculaire la plus rapide calibrée par Marko (2002) sur COI chez plusieurs espèces jumelles (dont des mollusques), les dates de divergence entre les cinq lignées de *Cyclope neritea* et les haplotypes ancestraux (*i.e.* du centre du réseau) sont comprises entre -137 000 et -220 000 ans, soit des dates antérieures à l'ère glaciaire Wis-

consinienne. Toutefois, Taberlet *et al.* (1998) ont montré que les marqueurs moléculaires peuvent être limités pour dater les événements anciens de l'aire Quaternaire. Ainsi, compte tenu de l'imprécision des horloges moléculaires et des difficultés d'extrapolation des taux d'évolutions obtenus pour une espèce à une autre espèce, la datation de la divergence des différentes lignées de *Cyclope neritea* est forcément imprécise. Les profils génétiques observés dans l'aire "naturelle" de *Cyclope neritea*, *i.e.* faible diversité génétique régionale associée à la distribution très large d'un ou de quelques haplotypes, sont en effet tout à fait compatibles avec l'hypothèse d'une recolonisation post-glaciaire relativement rapide (Hewitt, 2001; Marko, 2004). L'hypothèse de populations refuges localisées par exemple au Sud de l'Espagne, de l'Italie et de la Grèce au cours de la dernière période glaciaire (le glaciaire Wisconsinien s'étend de -80 000 à -10 000 ans, le maximum glaciaire étant situé entre -24 000 et -14 000 ans) est compatible avec les profils génétiques observés. En outre, la structure en étoile de quatre des cinq branches du réseau indique des populations ayant récemment subi une phase d'expansion, compatible avec un scénario de recolonisation post-glaciaire des côtes méditerranéennes.

Les informations apportées par les marqueurs microsatellites sont également cohérentes avec une telle hypothèse : le profil marqué d'isolement génétique par la distance géographique indique une colonisation de proche en proche de l'aire occupée actuellement, et le fort gradient de diversité génétique est-ouest suggère que le centre de diversité de l'espèce, situé à l'est de la zone étudiée, pourrait également constituer le centre de dispersion à l'origine de la colonisation de l'ouest du bassin méditerranéen. Ainsi, compte tenu des niveaux de divergence entre lignées haplotypiques d'une part et du gradient de diversité génétique est-ouest d'autre part, il est probable qu'au cours du dernier épisode glaciaire, il y ait eu un plus grand nombre de populations refuges dans l'est que dans l'ouest de l'aire "naturelle", ces populations restant toutefois isolées les unes des autres.

8.3 Processus contemporains

Là encore, les informations apportées par les deux types de marqueurs sont la plupart du temps congruentes.

Sur les données mitochondriales, à l'exception de la population de Katérinie, les seules populations polymorphes sont celles présentant un mélange de lignées évolutives divergentes (voir figure 7.3). Il s'agit des populations de la Mar Menor, de Leucate, de Venise et de Caorle. La zone de Venise et Caorle comme celle de Leucate et Thau sont des bassins conchylicoles très importants et les échanges humains d'espèces cultivées peuvent expliquer les transferts accidentels d'individus d'un bassin de production à un autre. Toutes

ces populations présentent des déficits en hétérozygotes significatifs sur l'ensemble des locus considérés (tableau 7.2). Le fait que les $\hat{F}_{IS}$ soient globalement différents selon les locus permet d'exclure un effet du système reproducteur. En revanche, le déficit en hétérozygotes significatif dans ces populations peut être dû à un effet Wahlund¹⁶ reflétant le processus de “mélange” : si des groupes d'individus, ne possédant pas des fréquences alléliques similaires, ont été mélangés au sein d'une même population, l'observation d'un déficit en hétérozygotes à l'échelle de la population est une conséquence logique. Le fait que ces mélanges de lignées évolutives divergentes soient visibles grâce aux données microsatellites indique que ce processus est récent, puisqu'il a laissé une trace, y compris pour le locus le plus variable du jeu de données (cinq populations présentent des $\hat{F}_{IS}$ significatifs pour le locus 1A10, tableau 7.2). Le mélange d'individus génétiquement divergents au sein de ces populations (probablement en raison d'activités anthropiques), est donc supporté par les deux types de marqueurs utilisés.

Les trois autres populations présentant des $\hat{F}_{IS}$ significatifs (*i.e.* Faro, Thau et Eforie) sont monomorphes au niveau mitochondrial. L'absence de signature de “mélange” au niveau de l'ADN mitochondrial ne permet toutefois pas d'exclure qu'un mélange d'individus issus de lignées évolutives divergentes ait pu avoir lieu. En effet, ces populations sont également des sites conchylicoles importants ; le mélange, au sein de ces populations, d'individus issus de sources différentes en raison de transferts anthropiques ne peut donc pas être exclu. Dans l'hypothèse que les $\hat{F}_{IS}$ significatifs observés dans ces populations soit le reflet d'un mélange de lignées évolutives divergentes, au moins deux cas de figure doivent être envisagés pour expliquer les profils observés :

1. les individus mélangés sont issus de la même région (région au sens des groupes de populations définis plus haut sur la base des ressemblances mitochondriales) : si par exemple, des individus issus du delta de l'Ebre ou de Sardaigne ont été introduits à Faro, ce mélange restera invisible au niveau de l'ADN mitochondrial (ces populations étant monomorphes et présentant le même haplotype), mais les génotypes nucléaires seront vraisemblablement différenciés, dans la mesure où les taux de mutation des microsatellites, très supérieurs à ceux de l'ADN mitochondrial, doivent permettre à ces marqueurs de diverger rapidement au sein de populations n'ayant pas de contact. Dans cette situation, le mélange de lignées, invisible au niveau de l'ADN mitochondrial, devrait être révélé grâce aux génotypes des microsatellites.
2. l'ADN mitochondrial ne permet de tracer que les lignées matriarcales. Si seuls

¹⁶L'effet Wahlund cause la réduction de la proportion d'hétérozygotes attendue à l'équilibre de Hardy-Weinberg au sein d'une population subdivisée : quand chaque sous-population est à l'équilibre de Hardy-Weinberg mais présente des fréquences alléliques différentes, l'hétérozygotie globale sera inférieure à celle attendue en l'absence de méta-populations.

quelques individus mâles, quelle que soit leur population d'origine, ont été introduits dans les populations de Faro, Thau ou Eforie, ou si, dans ces populations, seuls des mâles introduits ont réussi à accéder à la reproduction, la trace du “*mélange*” ne sera visible qu’au niveau nucléaire.

8.4 Apports de l'approche comparative

Sur la base des données mitochondriales, cinq régions principales ont pu être distinguées (péninsule ibérique, côtes méditerranéennes françaises, Adriatique, Mer Égée et Mer Noire, voir figure 7.3). Ces mêmes régions apparaissent clairement, à quelques exceptions près, en réalisant d’une part des arbres de populations (figures 7.6 et 7.7), et d’autre part les simulations bayésiennes sous **Structure** et **Geneland**, quel que soit le nombre de classes inférées par ces logiciels (figures 7.9 et 7.11).

La seule différence notable entre les deux arbres de populations concerne la position de la population de la Mar Menor : associée aux populations typiques de la péninsule ibérique (monomorphes pour l’haplotype H_1) sur la base des données mitochondriales, elle semble plus proche des populations de Leucate et de Thau (populations faisant pourtant partie d’une autre “*région mitochondriale*” : la région “*languedocienne*” des côtes méditerranéennes françaises, figure 7.3) sur la base des données nucléaires. Ce changement de statut de la population de la Mar Menor (*i.e.* population ibérique *vs* population “*languedocienne*”) n’est pas un biais de construction de l’arbre puisqu’il est également observé lors des analyses bayésiennes, quelque soit la méthode employée (**Structure** ou **Geneland**, voir figures 7.9 et 7.11c) ou le nombre de classes inférées (figures 7.9a et 7.9b).

La population de la Mar Menor laissait déjà apparaître un profil particulier au sein des populations de la région ibérique : c’est la seule population polymorphe de cette région (*i.e.* 4 haplotypes). En outre, un unique individu présentant un haplotype caractéristique des populations languedociennes (H_{18}) y a été trouvé, marquant probablement la trace d’un événement de “*mélange*” entre ces populations. Il est donc normal que l’ADN nucléaire garde également la signature de cet événement. En outre, l’association systématique de la population de la Mar Menor avec les populations languedociennes suggère que la majorité des individus échantillonnés dans la Mar Menor doivent avoir gardé la trace de cet événement au niveau de leur ADN nucléaire. Deux hypothèses sont donc susceptibles d’expliquer les différences observées entre les deux types de marqueurs :

1. l’évènement de “*mélange*” a impliqué un nombre inégal de mâles et de femelles, les mâles introduits devant être nettement plus nombreux que les femelles. Cela expliquerait pourquoi si peu d’individus ont gardé la trace de cet événement au

niveau mitochondrial alors que la majorité d'entre eux en garde la trace au niveau nucléaire.

2. si le sex ratio des individus n'est pas en cause, il est alors possible que la fertilité des femelles introduites ait été inférieure à celle des femelles présentes localement.

La population de Sardaigne est également dans cette situation. Bien qu'il n'y ait aucune différence au niveau des arbres, cette population de type ibérique (au niveau mitochondrial) est systématiquement associée à la classe languedocienne par **Structure** et **Geneland**. La figure 7.9a montre toutefois qu'au moins 9 individus de cette population présentent une probabilité non négligeable d'appartenir à la classe ibérique. Là encore, un mélange de populations génétiquement divergentes a probablement eu lieu, mais n'a laissé une trace qu'au niveau nucléaire.

Autre cas de figure, les populations de Venise et Caorle, qui, au niveau mitochondrial, présentent un profil génétique similaire (à Venise, 37 individus sur 40 portent des haplotypes qui sont trouvés à Caorle), sont systématiquement dissociées sur la base des analyses nucléaires. La similitude des profils mitochondriaux de ces populations associée à leur divergence nucléaire suggèrent que, si elles ont très vraisemblablement une origine commune, ces populations doivent être aujourd'hui relativement isolées.

Enfin, les populations les plus à l'est (notamment Caorle, Katérinie et Eforie) sont celles pour lesquelles les résultats diffèrent le plus selon les méthodes employées. Au niveau mitochondrial, ces trois populations appartiennent à des régions et des lignées évolutives distinctes. Au niveau nucléaire, les résultats de **Geneland** semblent résumer les résultats obtenus par **Structure** pour $K = 3$ et pour $K = 7$. Le fait que ces trois populations soient systématiquement associées au sein des mêmes classes indique des similitudes génétiques, conformément aux résultats trouvés par **Structure** pour $K = 3$ (surtout pour Katérinie et Eforie). Par ailleurs, le fait que ces populations soient associées dans quatre classes distinctes (et non au sein d'une seule classe comme c'est le cas pour les autres populations) reflète leur complexité génétique : elles sont similaires mais diversifiées (comme le prouvent les niveaux de diversité génétiques présentés dans le tableau 7.2), ce qui explique pourquoi **Structure** place chacune de ces trois populations dans des classes distinctes pour $K = 7$.

Nous voyons donc ici que l'information apportée par les deux types de marqueurs, bien qu'extrêmement congruente au niveau général, nous permet de mettre en évidence des processus que l'analyse d'un seul type de marqueur n'aurait pas permis de révéler.

Conclusions

Quatre principaux résultats ont été mis en évidence dans cette partie :

1. tout d'abord, les profils génétiques observés dans l'aire "*naturelle*" de l'espèce sont tout à fait conformes à l'attendu théorique concernant les espèces à faible pouvoir dispersif : diversité locale faible comparée aux espèces à fort pouvoir dispersif, et forte structure génétique entre populations ou régions. Ces profils clairs nous placent dans une situation idéale pour aborder la détermination du statut des populations récemment détectées le long des côtes françaises de l'Atlantique et de la Manche.
2. la structure génétique actuelle dans l'aire "*naturelle*" de *Cyclope neritea* semble essentiellement marquée par la trace de processus historiques. En particulier, les profils génétiques observés, tant au niveau mitochondrial que nucléaire, suggèrent une recolonisation post-glaciaire du bassin méditerranéen, depuis (i) des populations refuges durablement isolées les unes des autres, et/ou (ii) le centre de diversité de l'espèce situé à l'est de l'aire occupée par *Cyclope neritea* actuellement. Néanmoins, l'échantillonnage réalisé ici ne rend compte que d'une partie de l'aire de distribution de *Cyclope neritea*. Les scénarios de recolonisation proposés ici restent donc à confirmer.
3. les activités humaines contemporaines semblent avoir des impacts rares mais visibles sur les populations de *Cyclope neritea* (échanges humains probables entre l'Adriatique et la Méditerranée française par exemple). Le mélange d'individus génétiquement divergents au sein de certaines populations, probablement en raison des transferts d'espèces cultivées, a pu être mis en évidence avec les deux types de marqueurs.
4. bien que la convergence des résultats mitochondriaux et nucléaires sur les trois points énoncés ci-dessus soit forte, l'information apportée par les deux types de marqueurs n'est pas redondante. Au contraire, cette approche comparative a permis de mettre en évidence des processus que l'étude d'un seul type de marqueur n'aurait pas permis. En particulier, la situation des populations de la Mar Menor et de Sardaigne

suggère soit un sex ratio biaisé en faveur des mâles dans les groupes d'individus introduits au sein de ces populations, soit un succès reproducteur nettement en faveur des femelles locales.

Les populations de l'aire “naturelle” de *Cyclope neritea* ont donc des signatures génétiques particulièrement claires, qui nous placent dans une situation idéale pour étudier les populations récemment détectées sur les côtes françaises de l'Atlantique et de la Manche.

Troisième partie

Statut des populations de *Cyclope neritea* en limite d'aire et sur la façade atlantique française

Historique de l'expansion récente de l'aire de distribution de *Cyclope neritea*

L'aire de distribution de *Cyclope neritea*, historiquement reconnue comme “naturelle” et présentée dans la partie II, comprend les côtes de la Mer Noire, de la Mer Égée, de l'Adriatique et de la Méditerranée. Dans la partie I, nous avons vu que des populations instables sont observées au niveau de la limite atlantique de cette aire (*e.g.* autour de la frontière franco-espagnole). Jusque dans les années 1950, la population de *Cyclope neritea* la plus septentrionale était enregistrée à Socoa, sur les côtes du Pays Basque français (Kisch, 1950). Or, depuis les années 1970, de nouvelles populations sont apparues graduellement, du sud vers le nord, le long des côtes françaises de l'Atlantique et de la Manche : tout d'abord en 1976 dans le Bassin d'Arcachon (Bachelet *et al.*, 1980), puis, en 1983-84, dans la Baie de Marennes-Oléron (Pigeot, 1988; Sauriau, 1989), sur l'Île de Ré (Tardy *et al.*, 1985) et dans le Golfe du Morbihan (Le Roux *et al.*, 1988), et enfin plus récemment, en 2000, dans la Baie de Morlaix sur la côte nord du Finistère (Bachelet *et al.*, 2004).

Deux hypothèses non exclusives peuvent être formulées pour expliquer cette expansion récente :

1. l'apparition graduelle du sud vers le nord suggère une dispersion naturelle de l'espèce. Cette dispersion pourrait être le fruit d'un relachement des contraintes sélectives ayant par le passé limité l'expansion de *Cyclope neritea* vers le nord : le réchauffement rapide des eaux de surface du Golfe de Gascogne au cours des 30 dernières années (augmentation de 1,8 à 2,4 ° C, IPCC, 2001) pourrait être l'une des causes impliquées.
2. toutefois, les capacités de dispersion limitées de l'espèce semblent contradictoires avec la première hypothèse. Compte tenu de l'absence de stade planctonique au cours du cycle de vie de *Cyclope neritea* (Gomoiu, 1964), une expansion de plus de

1 000 kilomètres en moins de 30 ans semble peu probable. En outre, l'espèce est inféodée aux substrats sablo-vaseux caractéristiques des zones peu battues. Les côtes landaises battues, ou le littoral Breton rocheux sont donc également un frein à la progression de l'espèce vers le nord. Enfin, les populations récemment découvertes de *Cyclope neritea* ont toutes été initialement repérées au sein de bassins conchylicoles importants et à proximité de parcs à huîtres, ce qui semble plutôt accréditer l'hypothèse d'introductions accidentelles en relation avec les pratiques aquacoles humaines.

Les attendus théoriques concernant la diversité et la structure génétique des populations de *Cyclope neritea* le long des côtes françaises de l'Atlantique et de la Manche ne sont pas les mêmes en fonction du processus en jeu (*i.e.* expansion naturelle ou introductions humaines).

- Dans le cas de l'expansion naturelle, un profil d'isolement génétique par la distance géographique est susceptible d'apparaître, à l'instar de celui que nous avons mis en évidence dans l'aire "*naturelle*" de l'espèce (partie II). Les populations étant fondées de proche en proche, les populations proches géographiquement seront théoriquement plus proches sur le plan génétique, que des populations éloignées géographiquement. Les populations les plus éloignées géographiquement doivent en effet diverger sous l'effet de la dérive génétique et des effets de fondation¹⁷ successifs.
- Dans le cas d'introductions humaines, deux situations peuvent être distinguées. Si l'introduction a été ponctuelle (impliquant un nombre limité de fondateurs) et en provenance d'un nombre de sources limitées et peu diversifiées, les populations introduites doivent, là encore, présenter une chute de la diversité génétique en raison d'effets fondateurs marqués (Hellberg *et al.*, 2001; Grosholz, 2002; Dupont *et al.*, 2003). En revanche, si les introductions ont été multiples et récurrentes, du fait de la forte structure génétique observée dans l'aire "*naturelle*", les populations introduites devraient présenter des diversités génétiques plus importantes que les populations sources (Kolbe *et al.*, 2004; Voisin *et al.*, 2005). Dans l'hypothèse d'introductions humaines, les résultats obtenus sur l'ADN mitochondrial à la partie II (populations peu diversifiées mais fortes divergences entre régions) nous placent ainsi dans une situation idéale pour trancher entre introductions multiples et introduction ponctuelle.

Dans le but de préciser l'histoire des populations récemment apparues sur nos côtes et

¹⁷perte de diversité génétique dans les populations récemment créées par rapport aux populations sources, en raison du faible nombre de fondateurs impliqués dans l'évènement de fondation (effet d'échantillonnage du "*pool*" génique). Les effets de dérive génétique associés contribuent de plus à augmenter la structure génétique.

afin que notre étude reflète le mieux possible la situation de l'espèce dans sa nouvelle aire française, nous allons, dans cette partie, étudier cinq populations échantillonnées dans les cinq principales zones récemment colonisées en France. Les trois populations situées en limite de l'aire de distribution historique de l'espèce (*i.e.* San Vicente de la Barquera, Pedreña et Hossegor) seront également étudiées dans cette partie afin de déterminer si leur profil les rapproche des populations naturelles ou si au contraire, il s'agit de populations présentant des caractères similaires à ceux des populations des côtes françaises récemment installées. Pour comparaison avec les données obtenues dans l'aire "*naturelle*", nous utiliserons le marqueur mitochondrial (fragment de 533 paires de bases du gène COI) et les quatre marqueurs microsatellites (locus 2C4, 3C4, 1A10 et 2H10b) utilisés pour les analyses de la partie II.

Dans un second temps, nous nous attacherons à étudier la structure à micro-échelle spatiale au sein d'un site (*e.g.* une baie) dans l'aire récemment colonisée et dans l'aire "*naturelle*" de *Cyclope neritea*. Au sein d'une même baie, deux ou trois populations séparées de quelques centaines de mètres ou de quelques kilomètres seront analysées. Nous avons vu que la structure génétique dans l'aire "*naturelle*" est particulièrement marquée à grande échelle (*i.e.* entre baies séparées de plusieurs dizaines, centaines voire milliers de kilomètres) conformément aux attendus théoriques pour une espèce à développement direct. Néanmoins, ce résultat ne nous permet pas de déterminer si les populations en place de longue date sont également génétiquement différenciées à petite échelle. Autrement dit, ce résultat ne nous permet pas d'appréhender les capacités de dispersion locales de *Cyclope neritea*, élément fondamental pour la réussite du premier stade d'une introduction biologique. La comparaison de la structure locale dans l'aire "*naturelle*" et dans l'aire "*nouvellement colonisée*" nous permettra de mesurer l'impact de la fondation initiale et la capacité d'expansion locale post-introduction de cette espèce. Les baies faisant l'objet de cette comparaison seront (i) Rivedoux, le long des côtes atlantiques françaises et (ii) les sites de l'Ebre, Mar Menor, Caorle et Eforie dans l'aire "*naturelle*".

Une grande partie des résultats présentés dans cette partie a fait l'objet de publications : les résultats préliminaires, ont été publiés dans un article de synthèse paru dans *Marine Ecology Progress Series* (Bachelet *et al.*, 2004, présenté en annexe D.1) et la plupart des données mitochondriales ont fait l'objet d'un article paru dans *Molecular Ecology* (Simon-Bouhet *et al.*, 2006, annexe D.3)

Méthodes statistiques mises en œuvre

9.1 Diversité et structure génétique

Trois groupes de populations ont été considérés :

- le “*groupe 1*” concerne les cinq populations françaises récemment apparues (populations françaises de l’Atlantique et de la Manche, appelées populations FAM dans la suite du manuscrit). Il s’agit (du nord au sud) des populations de Stérec, Séné, Rivedoux, Ronce et Arguin (voir figure 10.2).
- le “*groupe 2*” concerne les trois populations de statut indéterminé, situées à la frontière franco-espagnole : Hossegor, Pedreña et San Vicente de la Barquera (figure 10.2).
- le “*groupe 3*” concerne les populations de l’aire “*naturelle*” de *Cyclope neritea*, déjà analysées dans la partie II. Il s’agit (d’ouest en est) des populations de : Alvor, Faro, Cadiz, Mar Menor, Ebre, Leucate, Thau, Sardaigne, Sicile, Venise, Caorle, Katérinie et Eforie (voir figure 7.3 de la partie II).

L’échantillonnage des côtes françaises a été réalisé afin de refléter au mieux la distribution actuelle de *Cyclope neritea* sur nos côtes. Les huit populations nouvellement étudiées ici (populations au statut indéterminé et populations FAM) ont toutes été échantillonnées, grâce à l’attraction exercée par une proie (voir description de la méthode dans la section 4.2.1).

Pour chaque population, pour chacun des groupes ainsi que pour l’ensemble des huit populations, les mêmes indices de diversité génétique que ceux présentés dans la partie II ont été calculés. Il s’agit, pour les données mitochondriales, du nombre d’haplotypes (N_H), du nombre de sites variables (S), des diversités haplotypiques ($\hat{H}_{E.mt}$) et nucléotidiques ($\hat{\pi}$), et pour les données nucléaires, du nombre d’allèles (N_{all}), de la diversité génique ($\hat{H}_{E.nuc}$) et de la richesse allélique $\hat{R}_{All}$. La méthode de raréfaction de Hurlbert (1971), utilisée pour calculer la richesse allélique, se base sur la plus petite taille de population du jeu de données. L’effectif de la population de Pedreña étant particulièrement réduit ($N = 3$), cette population a été écartée du jeu de données pour réaliser le calcul des richesses alléliques. Après Pedreña, la population présentant

le plus faible effectif est la population de Sicile ($N = 10$). Étant donné qu'un individu de cette population présente des données manquantes pour l'un des locus, il est exclu de l'analyse, les comparaisons de richesse allélique étant donc effectuées, pour toutes les populations, sur la base d'une taille d'échantillon égale à 9.

La structure génétique existant au sein des populations de l'aire considérée a été analysée par le calcul de $\hat{F}_{IS}$ pour chaque locus et chaque population par la méthode de Weir et Cockerham (Weir et Cockerham, 1984, , encadré 6.1 de la partie II). En outre, les valeurs de $\hat{F}_{IS}$, $\hat{F}_{ST}$ et $\hat{F}_{IT}$ ont été calculées dans **Fstat** (Goudet, 1995) pour le groupe 1, le groupe 2, et l'ensemble de ces deux groupes, afin d'analyser la distribution des fréquences alléliques au niveau inter-populations.

9.2 Isolement par la distance

L'hypothèse d'un isolement génétique dû à un isolement géographique (IBD) a été testée pour les données mitochondriales et microsatellites dans le groupe de populations récemment détectées. Pour les deux types de marqueur, le même estimateur de distance génétique a été utilisé, $\hat{F}_{ST}$, calculé pour chaque paire de population et servant à déterminer la distance génétique définie par $\frac{\hat{F}_{ST}}{1-\hat{F}_{ST}}$ (Rousset, 1997). Les distances géographiques entre les différentes paires de populations ont été mesurées en suivant le trait de côte, grâce au logiciel **Google™ earth**. Un test de Mantel a ensuite été réalisé à l'aide du logiciel **R** (**package ade4**, R Development Core Team, 2005), afin de rechercher une corrélation positive entre les matrices de distances géographiques et génétiques (5000 permutations).

9.3 Similitudes entre populations

Afin de mettre en évidence les similitudes entre les 8 populations ciblées ici et les populations de l'aire d'origine de *Cyclope neritea*, plusieurs approches ont été mises en œuvre :

1. Tout d'abord, la composition génétique des huit populations a été comparée avec celle des populations de l'aire "*naturelle*", grâce à l'étude conjointe de la distribution des haplotypes dans ces populations et dans le réseau d'haplotypes construit et présenté dans la partie II (figure 7.2).
2. Ensuite, afin de visualiser les distances génétiques entre toutes les populations étudiées, un arbre de populations a été réalisé avec le logiciel **Mega v3.1** (Kumar *et al.*, 2004) sur les données mitochondriales puis microsatellites. Les mêmes distances et algorithmes de reconstruction que dans la partie II ont été utilisés : distance de "*Kimura 2-paramètres*" (Kimura, 1980) pour les données mitochondriales, et distances minimales de Nei pour les données microsatellites ; méthode de reconstruction phylogénique : "*neighbour-joining*" (Saitou et Nei, 1987).

3. Enfin, nous avons tenté de déterminer l’origine des individus des populations récemment apparues en réalisant, avec le logiciel **GeneClass v2.0.g** (Piry *et al.*, 2004), des tests d’assignation par méthodes bayésiennes, sur la base des données microsatellites. Ce logiciel permet de calculer, pour chaque individu, (i) la population source la plus vraisemblable (méthode de Rannala et Mountain, 1997), grâce à l’attribution d’un “score”, et (ii) les probabilités d’appartenance à chacune des populations sources. Le calcul des scores est basé sur les valeurs de vraisemblances calculées dans le logiciel : avec k populations, le score de l’individu i dans la population l est le suivant :

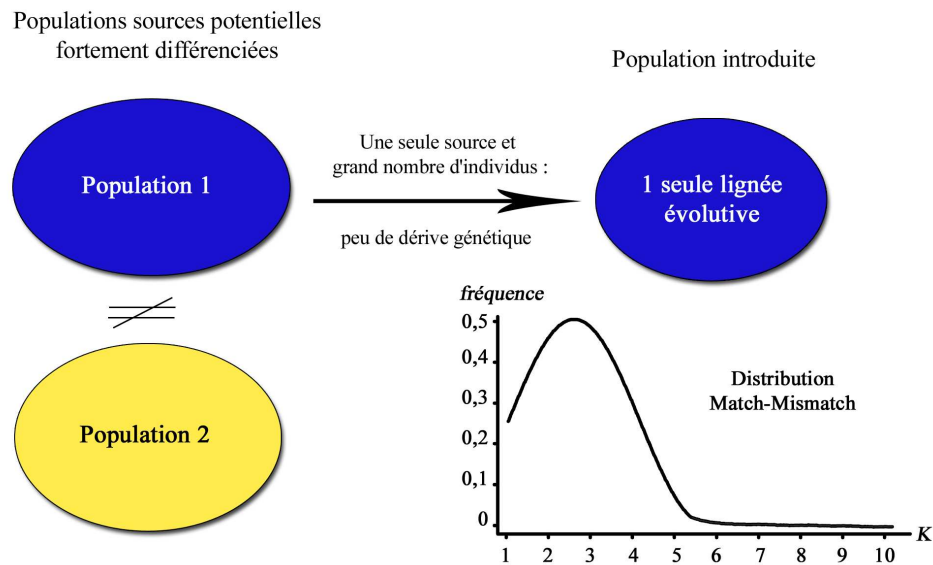
$$Score_{i,l} = \frac{L_{i,l}}{\sum_{j=1}^k L_{i,j}}$$

$L_{i,l}$ étant la vraisemblance de l’appartenance de l’individu i à la population l . Le calcul des probabilités d’appartenance de chaque individu aux populations sources a été réalisé par une procédure de ré-échantillonnage de Monte-Carlo parmi un échantillon simulé de 10 000 individus. Les simulations ont été effectuées en considérant comme populations sources potentielles les populations de l’aire “naturelle” étudiées à la partie II, en incluant les populations situées en limite d’aire de distribution (*i.e.* San Vicente de la Barquera, Pedreña et Hossegor), dans la mesure où, dans le cas d’une expansion naturelle, ces populations pourraient être la source des migrants.

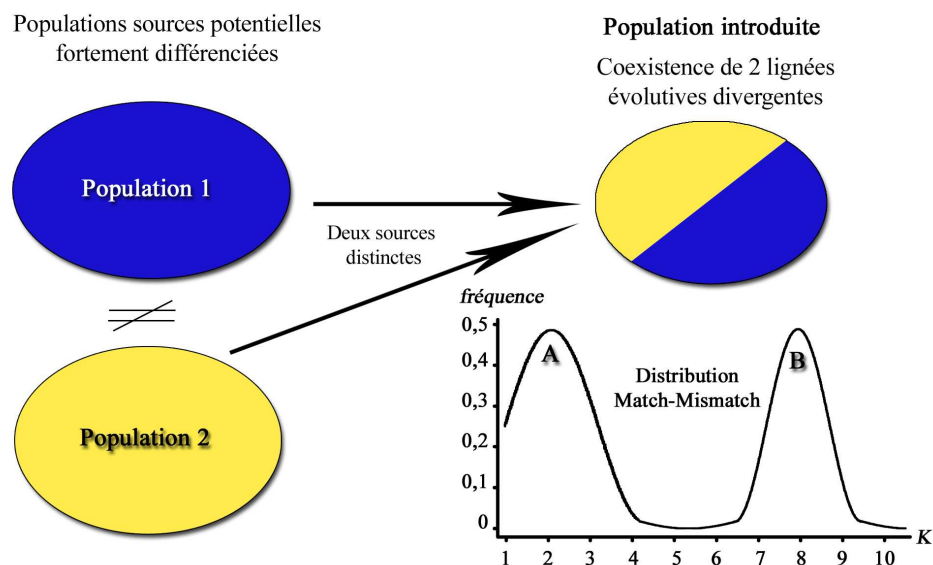
9.4 Profils de “mélanges” et simulations des coalescences

Lorsqu’une population est créée par l’introduction d’individus issus de plusieurs populations sources différenciées, la population introduite présente un profil particulier où deux groupes divergents (ou plus) se trouvent en mélange. Afin de tester l’existence de tels mélanges de lignées évolutives divergentes, nous avons représenté graphiquement la distribution du nombre de différences nucléotidiques entre paires de séquences (K) dans toutes les populations récemment enregistrées (“*distribution match-mismatch*”, voir principe figure 9.1).

En outre, étant donné la faible puissance du test D de Tajima (Galtier *et al.*, 2000, principe du test détaillé dans l’encadré 9.1), nous avons testé, pour chacune des populations, l’écart à l’équilibre mutation-dérive sous un modèle de site infinis en suivant la méthode mise au point par Depaulis et Veuille (1998). Ce test, appelé test K est particulièrement adapté à la détection de situations de “mélange” (Hamblin et Veuille, 1999) : il repose sur la comparaison entre le nombre d’haplotypes réellement observés dans une population (N_H) et le nombre d’haplotypes attendus à l’équilibre mutation-dérive (N_{EH}), étant donné la taille de la population (N) et le nombre de sites variables observés (S).



a. Seule la population 1 fournit des fondateurs à la population introduite. La distribution match-mismatch de la population introduite montre un unique mode pour des valeurs de K faibles : les haplotypes de la population appartenant tous à la même lignée évolutive, ils sont en effet peu divergents.



b. Les deux populations fournissent des fondateurs ; il y a donc au sein de la population introduite, un mélange de lignées évolutives divergentes. Lorsque l'on tire au hasard deux séquences dans cette population, deux cas de figures peuvent être rencontrés : soit les séquences sont issues de la même lignée évolutive, la valeur de K sera alors faible et contribuera à former le pic noté "A" sur le graphique, soit elles proviennent chacune de l'une des deux lignées divergentes ; la valeur de K sera alors grande (du fait de la divergence des lignées évolutives) et contribuera donc à former le pic "B".

FIG. 9.1: Distribution "match-mismatch" théorique d'une population introduite : profils attendus avec et sans mélange de lignées évolutives divergentes.

Définition de l’indice D et exemples de comportement.

La statistique D (Tajima, 1989) s’applique à l’étude du polymorphisme de séquences d’ADN et permet de mettre en évidence des effets sélectifs ou des variations démographiques au sein d’une population. Lorsque l’on s’intéresse au polymorphisme de séquences mitochondriales, les deux indices suivants sont la plupart du temps considérés :

1. le nombre de sites nucléotidiques polymorphes, ou nombre de sites ségrégeants S
2. la diversité nucléotidique moyenne π :

$$\pi = \sum_{ij} p_i p_j \delta_{ij}$$

avec p_i la fréquence du i^{eme} haplotype et δ_{ij} , le nombre de différences nucléotidiques entre les séquences i et j , divisé par la longueur totale de la séquence.

En génétique des populations, le paramètre $\theta = 4N_e\mu$ (N_e : taille efficace de la population, μ : taux de mutation par génération du locus considéré) est couramment utilisé. Sous un modèle en sites infinis (chaque nouvelle mutation touche un nouveau site nucléotidique), on montre que :

$$E(\pi) = \theta \quad \text{et} \quad E(S) = \theta \sum_{i=1}^{n-1} \frac{1}{i}$$

avec n , le nombre d’haplotypes de l’échantillon. Ainsi, deux estimations de θ peuvent être obtenues grâce au polymorphisme de séquences :

$$\hat{\theta}_\pi = \hat{\pi} \quad \text{et} \quad \hat{\theta}_S = \frac{S}{\sum_{i=1}^{n-1} \frac{1}{i}}$$

En l’absence de sélection et à l’équilibre mutation dérive, $\hat{\theta}_\pi$ et $\hat{\theta}_S$ sont identiques. En revanche, ces deux estimateurs ne répondent pas de la même façon à la sélection et aux changements de taille des populations. Ainsi, le calcul de

$$\hat{D} = \hat{\theta}_\pi - \hat{\theta}_S$$

peut être utilisé comme un test statistique afin de vérifier si la population considérée est à l’équilibre mutation-dérive.

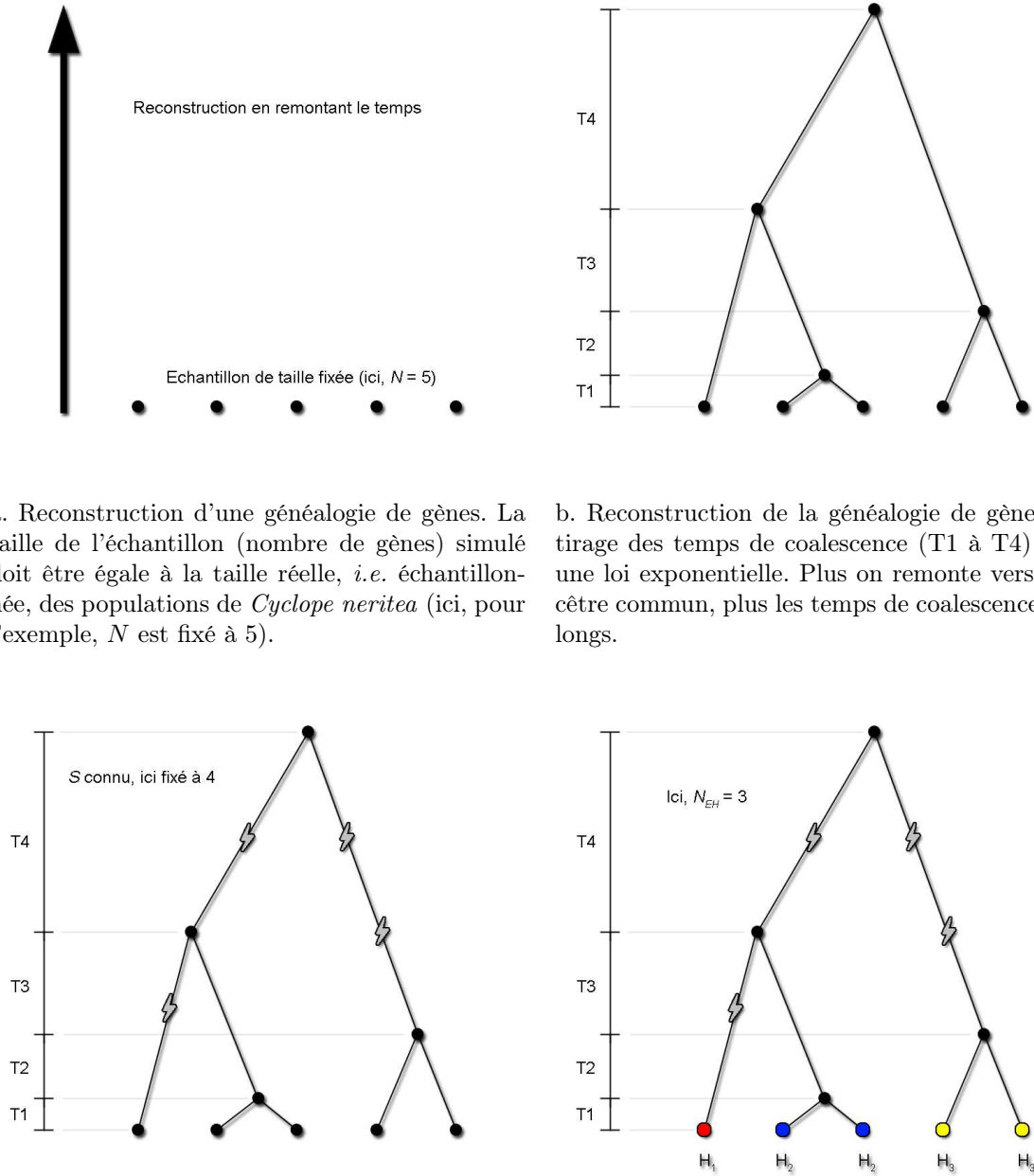
Quand les mutations sont neutres et que la population est à l’équilibre mutation-dérive, $\hat{D}$ est nul. Dans le cas d’une population en expansion démographique, les mutations touchant le locus étudié ont peu de chances d’être perdues et augmentent donc rapidement la valeur de $\hat{\theta}_S$. En revanche, les nouveaux haplotypes seront proches les uns des autres, et la valeur de $\hat{\theta}_\pi$ n’augmentera que plus lentement, au fur et à mesure de la divergence des haplotypes et de l’augmentation de leur fréquence dans la population. $\hat{D}$ prendra donc une valeur négative. Pour des raisons similaires, $\hat{D}$ sera négatif en cas de sélection positive d’un allèle favorable. $\hat{D}$ est en revanche positif en cas d’effet Wahlund ou de réduction démographique.

ENCADRÉ 9.1: L’indice D de Tajima.

Pour réaliser ce test, nous avons modifié un programme initialement développé par Marie-Hélène Müller (INRA Mauguio), qui nous en a fourni le code source. Ce programme, écrit en Delphi, s'appuie sur la théorie de la coalescence¹⁸, en particulier sur l'approche rétrospective. L'approche développée dans le programme utilisé est présentée ci-dessous et illustrée par la figure 9.2 :

1. Tout d'abord, la topologie de la généalogie est reconstituée, en partant des effectifs de l'échantillon réel, jusqu'à l'unique ancêtre commun : en remontant le temps, tous les événements de coalescence ayant eu lieu sont replacés. Selon le modèle de Wright-Fisher, à l'équilibre mutation-dérive, ces temps de coalescence suivent une loi de probabilité exponentielle d'espérance $4N_e(1 - \frac{1}{n})$, avec N_e , la taille efficace de la population et n la taille de l'échantillon étudié. Ainsi, plus on remonte dans le temps, plus il faut attendre longtemps avant d'observer un nouvel événement de coalescence (n diminuant d'une unité à chaque nouvel événement de coalescence), et donc plus la longueur des branches est importante près de l'ancêtre commun.
2. Dans un second temps, les mutations (dont le nombre correspond au nombre de sites ségrégeants S observés dans les populations étudiées) sont placées sur les branches, chaque mutation étant unique (modèle en site infini : absence d'homoplasie). Plus une branche sera longue, plus elle aura de chances de recevoir un nombre de mutations important.
3. Dans un troisième temps, les haplotypes résultant (i) de la topologie de l'arbre et (ii) du placement des mutations, sont identifiés afin de déterminer N_{EH} , le nombre d'haplotypes attendus sous hypothèse d'équilibre mutation-dérive.
4. Enfin, les trois étapes précédentes sont répétées un grand nombre de fois (ici, 10 000), afin de déterminer la distribution nulle de N_{EH} , étant donné (i) la taille de l'échantillon (correspondant ici à l'effectif échantillonné des populations de *Cyclope neritea*), (ii) le nombre de sites variables S observé dans les populations de *Cyclope neritea*, (iii) sous hypothèse d'équilibre mutation-dérive et (iv) avec un modèle en sites infinis. Il est ainsi possible de déterminer d'une part l'intervalle de confiance à 95% de ce paramètre, et d'autre part la proportion de valeurs de N_{EH} supérieures ou inférieures à la valeur de N_H réelle (observé dans les populations de *Cyclope neritea*), afin de déterminer les valeurs de probabilités associées aux valeurs de N_H des populations réelles. Une présentation du principe du programme réalisé sous Delphi ainsi que quelques captures d'écran figurent en annexe C.

¹⁸Un événement de coalescence est la duplication d'un gène ancestral en deux gènes fils.



c. Placement des mutations sur les branches de l'arbre. Le nombre de mutations placées est égal au nombre de sites variables S observé. Le nombre de mutations portées par une branche est d'autant plus grand que la branche est longue.

d. Identification des haplotypes et détermination de la valeur de N_{EH} pour la simulation.

FIG. 9.2: Étapes de la reconstruction d'une généalogie de gènes pour un locus haploïde. Les différentes étapes présentées ici sont répétées un grand nombre de fois afin d'obtenir la distribution théorique de N_{EH} , à laquelle sera comparée le nombre d'haplotypes N_H observé dans la population analysée.

9.5 Étude à micro-échelle

Treize populations ont été échantillonnées en 5 localités grâce à la méthode du dépôt de la proie décrite dans la partie I. Trois populations ont été collectées à Rivedoux (*i.e.* Rivedoux 1, 2 et 3), trois dans la Mar Menor, deux dans le Delta de l'Ebre, trois à Caorle et enfin, deux à Eforie. Pour chaque localité, les populations ont été échantillonnées à proximité les unes des autres. A Rivedoux et Caorle, les trois populations ont été échantillonnées linéairement, 500 mètres séparant 2 populations consécutives. Les deux populations échantillonnées dans le Delta de l'Ebre sont séparées par 3,8 kilomètres. Les deux populations d'Eforie sont distantes de 32,5 kilomètres et enfin, les trois populations de la Mar Menor n'ont pas été échantillonnées de façon linéaire, mais de part et d'autre de la lagune. Les populations Mar Menor 1 et 2 sont distantes de 13,4 kilomètres, 11,7 kilomètres séparent les populations de Mar Menor 1 et 3, et enfin, les populations Mar Menor 2 et 3 sont à 15,7 kilomètres l'une de l'autre (voir figure 9.3).

Ces populations ont été analysées à l'aide des quatres locus microsatellites (2C4, 3C4, 1A10 et 2H10b) et leur structure a été étudiée grâce aux statistiques F :

1. Pour chaque population, les valeurs de $\hat{F}_{IS}$ ont été calculées par la méthode décrite dans la partie II et testée par un test exact implémenté dans **Genepop v3.4** (Raymond et Rousset, 1995) afin de mettre en évidence des écarts aux proportions génotypiques attendues sous hypothèse d'équilibre de Hardy Weinberg.
2. Les valeurs de $\hat{F}_{ST}$ par paires de populations ont été calculées ainsi que les valeurs de probabilité associées à un test exact de différenciation des populations (réalisé avec **Genepop v3.4**), afin d'estimer les niveaux de différenciation entre populations d'une même localité d'une part, et populations de localités distinctes d'autre part.
3. Les valeurs de $\hat{F}_{ST}$ par localité ont également été calculées afin de révéler le niveau de différenciation au sein de chaque groupe de populations proches géographiquement.

Pour visualiser les similitudes entre les populations, un arbre a été réalisé avec les données des treize populations étudiées (la méthode de calcul des distances génétiques ainsi que la méthode de reconstruction de l'arbre sont les mêmes que celles décrites à la partie II et mises en œuvre à la section 9.3). Dans un second temps, une analyse factorielle des correspondances (AFC) utilisant les données alléliques des locus de chaque individu a été réalisée dans **Genetix** (Belkhir *et al.*, 1996-2004). Guinand (1996) a montré que les valeurs d'inertie associées à chaque axe peuvent être assimilées à des combinaisons linéaires des valeurs de F_{st} monolocus. En d'autres termes, cette analyse permet de représenter les divergences génétiques entre individus au sein et entre les populations.



FIG. 9.3: Localisation des sites d'échantillonnage de la Mar Menor en vue de l'étude de la structure génétique à micro-échelle. Source photographique : Google™ Earth.

10

Résultats

10.1 Analyses mitochondriales à l'échelle globale

10.1.1 Diversité moléculaire

Les résultats détaillés concernant les populations de *Cyclope neritea* situées dans l'aire "naturelle" ont été présentés dans la section 7 de la partie II.

La forte diversité génétique observée au niveau global n'est pas répartie de manière égale au sein des trois groupes de populations étudiés (*i.e.* groupe 1 : populations françaises de l'Atlantique et de la Manche (FAM), groupe 2, populations de statut indéterminé, groupe 3 : populations de l'aire "naturelle" de *Cyclope neritea*, voir tableau 10.1). Ainsi, les populations FAM présentent des niveaux de diversité génétique très supérieurs à ceux des groupes 2 et 3. La valeur totale de diversité haplotypique $\hat{H}_{E.mt.tot}$ du groupe 1 (0,833) est supérieure à celle du groupe 3 (0,744). En outre les valeurs de $\hat{H}_{E.mt}$ du groupe 1 sont beaucoup plus homogènes entre elles (comprises entre 0,589 et 0,911, $\hat{H}_{E.mt.moy} = 0,743$) qu'elles ne le sont dans les groupes 2 (comprises entre 0 et 0,514, $\hat{H}_{E.mt.moy} = 0,260$) et 3 (entre 0 et 0,607, $\hat{H}_{E.mt.moy} = 0,231$). La même observation se vérifie pour N_H , S et $\hat{\pi}$. Toutes les populations présentent des niveaux de polymorphismes importants et similaires au sein du groupe 1, alors qu'une population sur deux est monomorphe dans les deux autres groupes (8 sur 16). Cela se traduit, au niveau des différents indices étudiés, par des valeurs moyennes et totales proches pour les populations FAM, mais très différentes pour les groupes 2 et 3. Par exemple, pour les populations de l'aire "naturelle", la valeur de π_{tot} est 6 fois supérieure à la valeur moyenne par population ($1,52.10^{-2}$ et $0,22.10^{-2}$ respectivement) alors que ces deux valeurs sont très proches pour le groupe 1 ($1,52.10^{-2}$ et $1,32.10^{-2}$), ce qui souligne l'existence de structures génétiques différentes au sein des divers groupes.

Parmi les populations FAM, seule Séné semble présenter un profil légèrement différent : cette population présente le plus grand nombre d'haplotypes et la plus forte diversité haplotypique

($N_H = 16$, $\hat{H}_{E.mt} = 0,911$) du groupe 1, mais avec des haplotypes peu divergents les uns des autres, puisque c'est dans cette population que le plus faible nombre de sites variables et la plus faible diversité nucléotidique ($S = 14$, $\hat{\pi} = 0,014.10^{-2}$) sont trouvés.

TAB. 10.1: Diversité génétique des populations de *Cyclope neritea*. Les trois groupes sont présentés ici : populations FAM, populations au statut indéterminé, et populations de l'aire "naturelle". Pour ce dernier groupe (déjà examiné à la partie II), seules les valeurs moyennes et totales sont indiquées. (Ma) : Manche, (A) : Océan Atlantique. SVB : San Vicente de le Barquera, M-O : Marenne-Oléron. Les indices présentés sont : N : nombre d'individus analysés, N_H : nombre d'haplotypes, S : nombre de sites variables, $\hat{H}_{E.mt}$: diversité haplotypique, $\hat{\pi}$: diversité nucléotidique. Pour chaque groupe, la moyenne et l'écart-type (S.D.) de ces indices sont indiqués, ainsi que le total, c'est-à-dire les indices calculés sur l'ensemble des individus du groupe.

Populations	Localisation (Mer-Océan)	N	N_H	S	$\hat{H}_{E.mt}$	$\hat{\pi} \cdot 10^2$
Populations FAM (groupe 1)						
a. Stérec	Baie de Morlaix (Ma)	33	3	18	0,589	1,371
b. Séné	Golfe du Morbihan (A)	32	16	14	0,911	1,014
c. Rivedoux	Île de Ré (A)	39	5	21	0,796	1,503
d. Ronce	Baie de M-O (A)	32	4	19	0,696	1,498
e. Arguin	Bassin d'Arcachon (A)	33	5	19	0,725	1,234
Moyenne		33,8	6,6	18,2	0,743	1,324
(S.D.)		(2,9)	(5,3)	(2,6)	(0,120)	(0,205)
Total		169	21	27	0,833	1,518
Statut indéterminé (groupe 2)						
f. Hossegor	Pays Basque (A)	33	2	1	0,265	0,050
g. Pedreña	Baie de Santander (A)	3	1	0	0,000	0,000
h. SVB	Cantabrica (A)	37	2	12	0,514	1,156
Moyenne		24,3	1,7	4,3	0,260	0,402
(S.D.)		(18,6)	(0,6)	(6,7)	(0,257)	(0,673)
Total		73	3	13	0,484	0,903
Populations de l'aire "naturelle" (groupe 3)						
Moyenne		31,5	2,8	6,4	0,231	0,224
(S.D.)		(8,8)	(2,8)	(6,4)	(0,283)	(0,377)
Total		410	23	43	0,744	1,521
Total		652	39			

10.1.2 Distribution géographique des haplotypes des populations françaises.

Sur les 39 haplotypes identifiés dans l'ensemble du jeu de données ($N = 652$), 15 ont été trouvés exclusivement dans les populations du groupe 1, et un haplotype (H_5) n'a été trouvé que dans les populations du groupe 2. Cependant, les trois haplotypes majoritaires dans l'aire "naturelle" de *Cyclope neritea* (i.e. H_1 , H_{18} et H_{22}) le sont également au sein des groupes 1 et

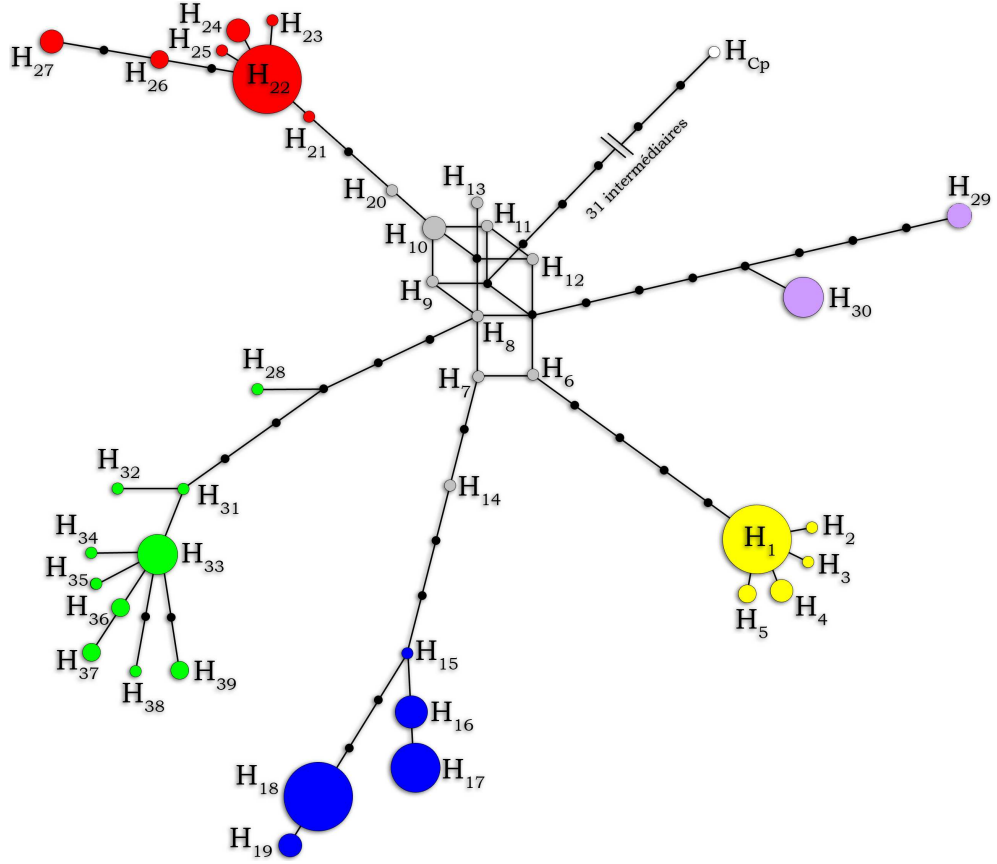


FIG. 10.1: Réseau présentant les relations phylogénétiques entre les haplotypes de *Cyclope neritea*. L'ensemble des données obtenues pour les 21 populations étudiées au cours de cette thèse a été utilisée. Au total, 39 haplotypes ont été identifiés, dont 23 dans l'aire "naturelle" et 16 uniquement trouvés dans les populations FAM et les populations de statut indéterminé. Chaque trait représente une substitution nucléotidique et les points noirs sur les branches représentent des haplotypes intermédiaires manquants. La taille des cercles est proportionnelle à la fréquence des haplotypes. Chaque couleur correspond à une lignée évolutive distincte. L'unique haplotype identifié chez *Cyclope pellucida* est indiqué en blanc.

2. Des haplotypes appartenant à quatre des cinq lignées évolutives divergentes mises en évidence dans l'aire "naturelle" (voir figures 10.1 et 10.2) ont été retrouvés dans les populations FAM et de statut indéterminé, seuls les haplotypes caractéristiques de la Mer Égée n'y figurent pas.

Toutefois, la distribution géographique des haplotypes est très différente dans l'aire "naturelle" et dans l'aire française. En effet, rappelons qu'une forte convergence entre structure génétique et structure géographique a été mise en évidence dans l'aire "naturelle" (partie II) : la diversité locale, très faible, était associée à une très forte structure régionale, avec, au sein de chaque région, très peu de mélanges d'individus issus de lignées évolutives divergentes. La différenciation totale ($\hat{\Phi}_{ST} = 0,868$) était donc principalement due à la différenciation entre régions ($\hat{\Phi}_{CT} = 0,854$), plutôt qu'à la différenciation entre populations d'une même région ($\hat{\Phi}_{SC} = 0,091$).

Au contraire, dans les populations françaises récemment détectées, la différenciation globale

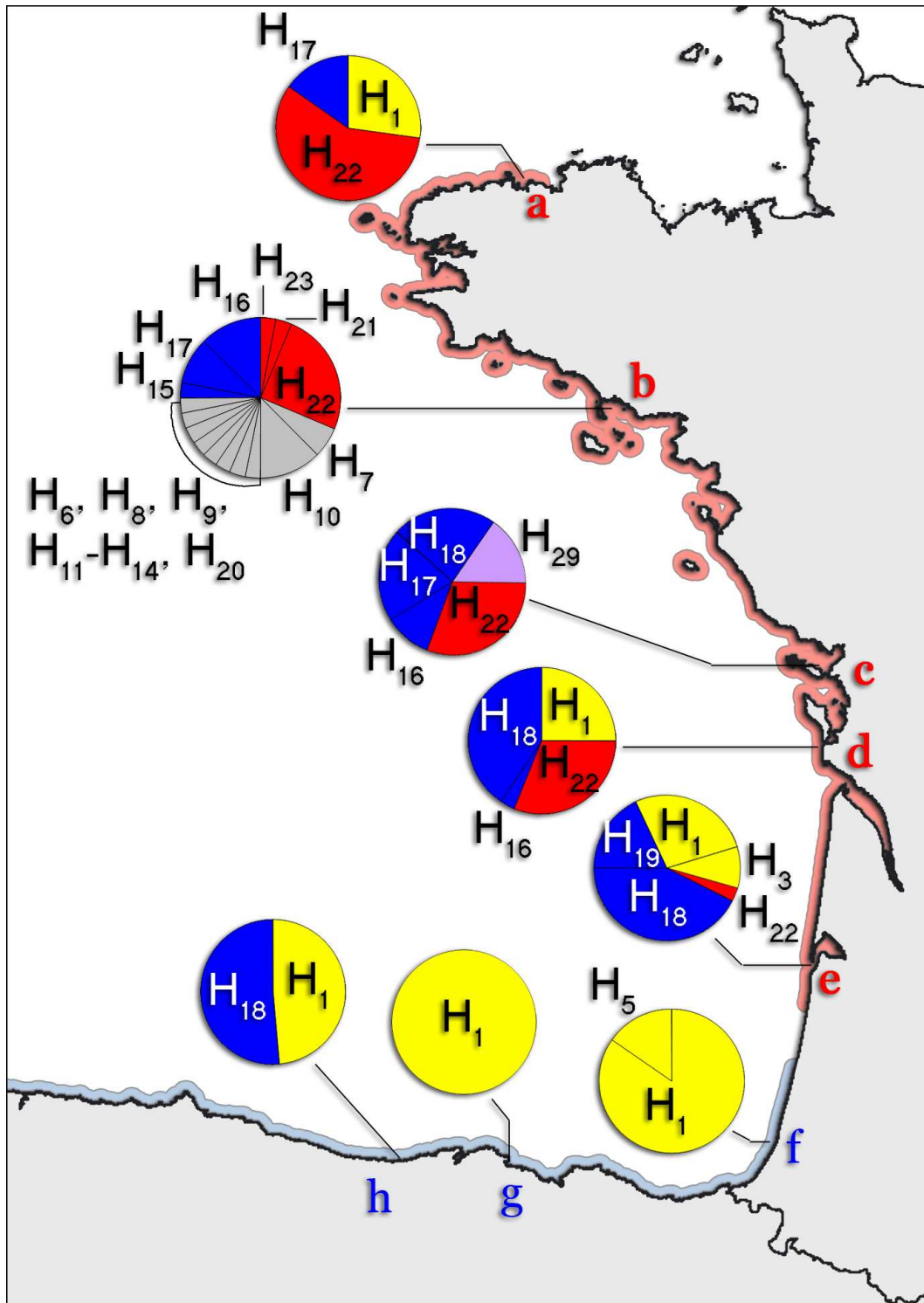


FIG. 10.2: Distribution des haplotypes dans les huit populations étudiées. Le code de couleur fait référence à la position des haplotypes dans les différentes branches du réseau d'haplotypes présenté à la figure 10.1. Les populations FAM sont (en rouge) : a : Stérec, b : Séné, c : Rivedoux, d : Ronce et e : Arguin. Les populations de statut indéterminé sont (en bleu) : f : Hossegor, g : Pedreña et h : San Vicente de la Barquera.

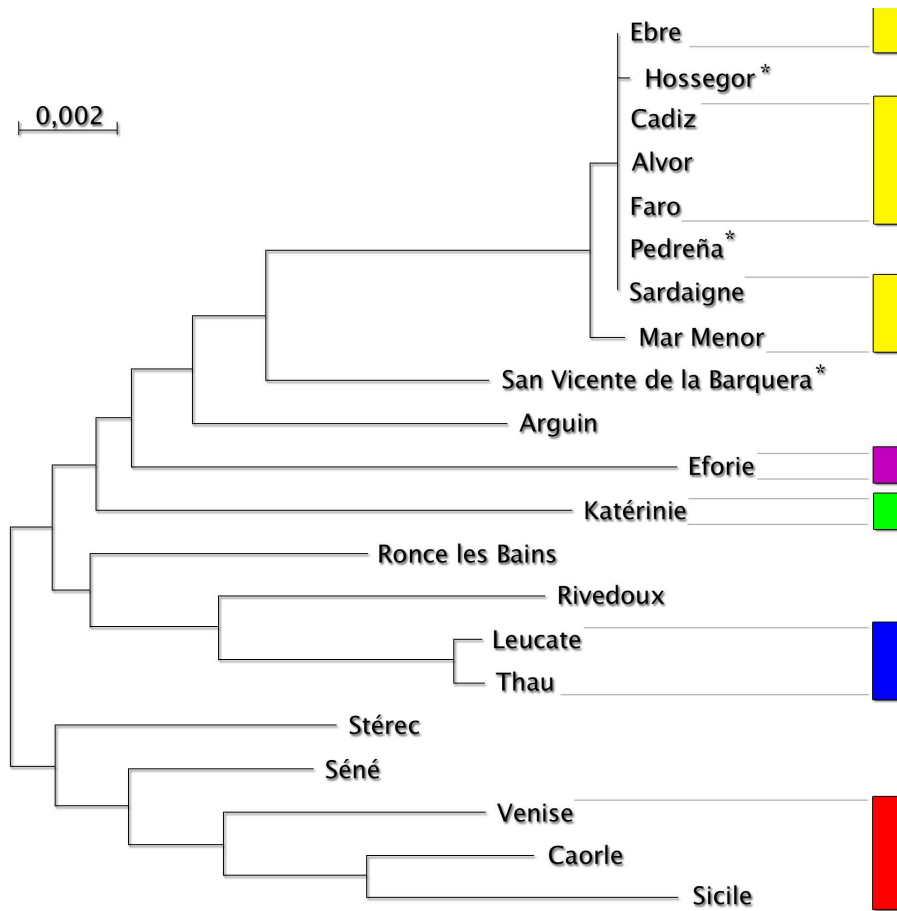


FIG. 10.3: Arbre de populations construit avec les données haplotypiques. Les couleurs à droite correspondent aux lignées évolutives majoritaires dans chacune des populations. Aucune couleur n'a été attribuée aux populations FAM, et les populations de statut indéterminé sont indiquées par une astérisque.

est beaucoup plus faible ($\hat{\Phi}_{ST} = 0,135$, $p < 10^{-4}$) et associée à une forte diversité génétique à l'échelle de la population. Les cinq populations du groupe 1 montrent en effet un mélange d'haplotypes appartenant à deux ou trois lignées évolutives divergentes (figures 10.1 et 10.2) suggérant donc des origines multiples pour les populations FAM.

L'arbre de populations présenté à la figure 10.3 permet de synthétiser ces résultats :

1. les populations de statut indéterminé (*i.e.* Hossegor, Pedreña et San Vicente de la Barquera) sont similaires entre elles et sont associées aux populations du groupe ibérique identifié précédemment dans la partie II. Cette observation est conforme aux indices de diversité génétique observés dans ces populations.
2. les populations FAM sont en revanche dispersées entre les différents groupes de populations de l'aire "naturelle", renforçant l'idée d'origines multiples pour les populations récemment détectées.

L'analyse des distributions "match-mismatch" (figure 10.4) confirme le mélange d'individus

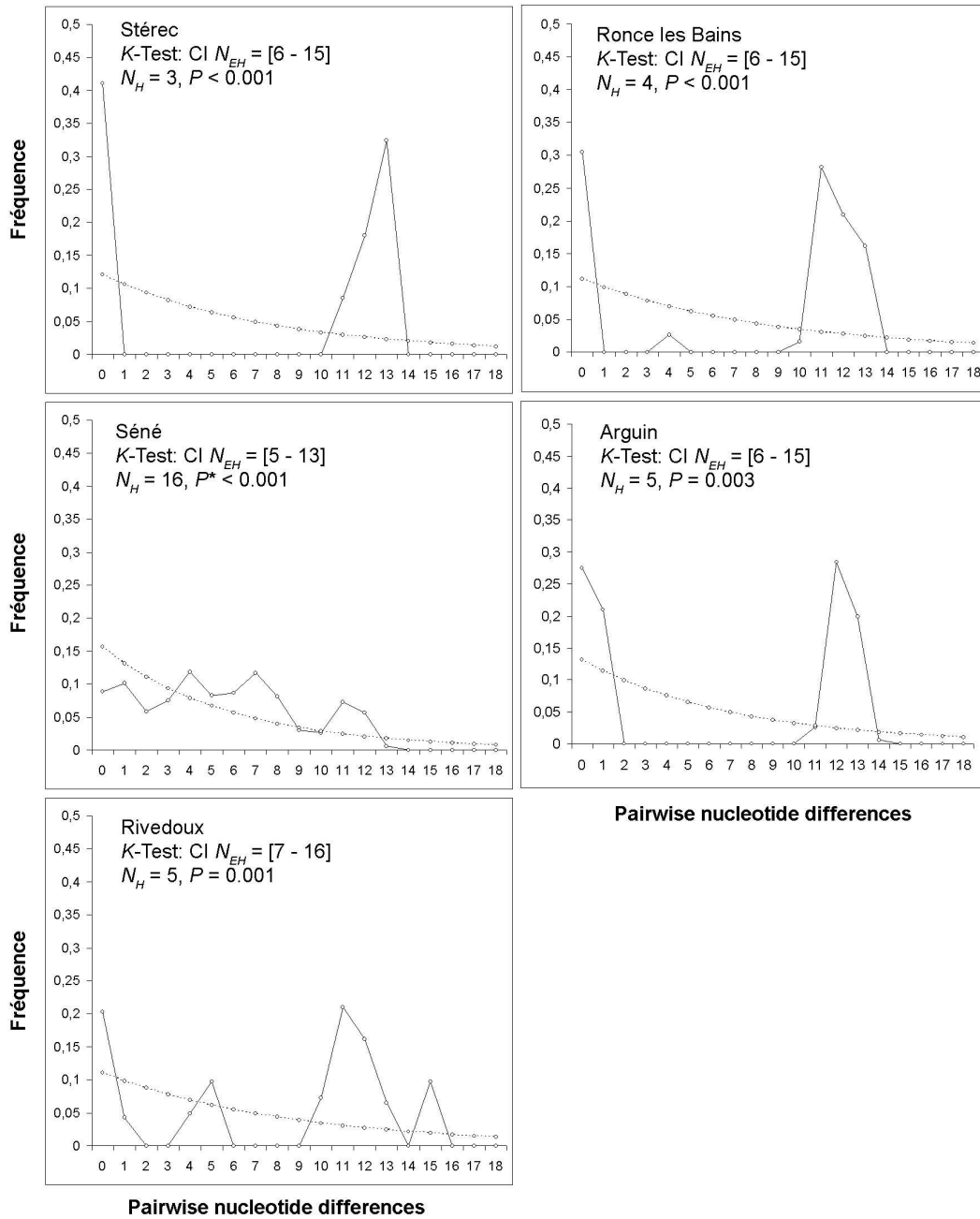


FIG. 10.4: Distribution “match-mismatch” et tests K . Pour chaque population FAM, la distribution observée du nombre de différences nucléotidiques par paire de séquences (lignes continues) ainsi que la distribution attendue sous l’hypothèse nulle d’équilibre démographique (les lignes pointillées) sont représentées. Les résultats des tests K sont indiqués pour chaque population : l’intervalle de confiance à 95% (CI) du nombre d’haplotypes attendu sous hypothèse d’équilibre mutation-dérive (N_{EH}) est présenté. Les valeurs de probabilités (P : probabilité que N_{EH} soit inférieur à N_H) ont été calculées sur la base de 5000 permutations. Pour Séné, P^* est la probabilité que N_{EH} soit supérieur à N_H .

présentant des haplotypes divergents, au sein des populations FAM. Ces distributions, pour Arguin, Rivedoux, Ronce et Stérec, montrent en effet au moins deux pics caractéristiques d'un mélange de lignées divergentes (voir principe de la distribution “*match-mismatch*” figure 9.1). Ce profil est statistiquement confirmé par les résultats du test K puisque, compte tenu du nombre de mutations observées S , le nombre réel d'haplotypes dans ces quatre populations (N_H) est significativement inférieur à celui attendu à l'équilibre mutation-dérive (N_{EH} , figure 10.4). En revanche à Séné, les haplotypes diffèrent les uns des autres sur une large gamme de sites ségrégeants, générant une distribution “*match-mismatch*” plus uniforme (entre 1 et 14 différences selon les paires de séquences considérées). En outre, le nombre d'haplotypes observés dans cette population est significativement plus grand que celui attendu sous hypothèse d'équilibre mutation-dérive (test K , figure 10.4), ce qui suggère une expansion démographique récente ou un polymorphisme balancé¹⁹ ancien (Depaulis et Veuille, 1998).

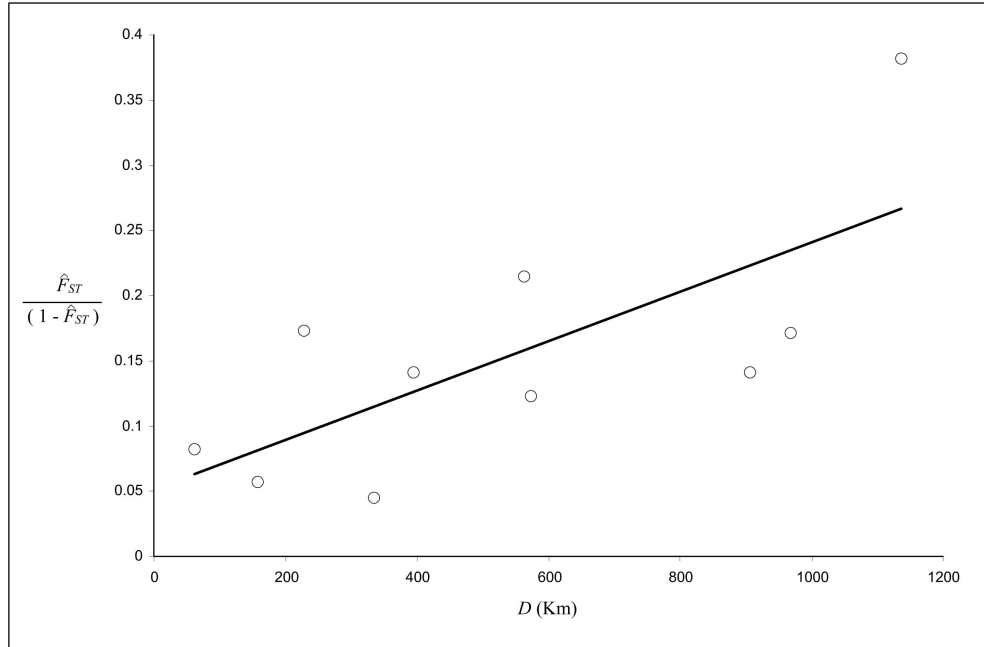


FIG. 10.5: Isolement génétique par la distance géographique entre les cinq populations françaises récemment détectées (données mitochondriales). La distance génétique $\frac{\hat{F}_{ST}}{1 - \hat{F}_{ST}}$ est représentée en fonction de la distance D en kilomètres, mesurée en suivant le trait de côte.

L'ensemble des résultats obtenus suggère donc que les populations FAM sont le résultat d'introductions récentes impliquant de multiples sources. En regard de ce scénario, un dernier résultat surprenant a été observé au sein du groupe 1 : la structure génétique entre paires de populations ($\hat{F}_{ST}$ par paires de populations) augmente avec la distance géographique séparant

¹⁹Polymorphisme apparaissant sous un régime de “*sélection balancée*” qui tend à maintenir stables plusieurs formes alléliques ou haplotypiques au sein d'une population. La sélection balancée s'oppose à la “*sélection directionnelle*”, qui favorise une forme allélique ou haplotypique particulière.

ces populations. Un test de Mantel a révélé un profil d'isolement génétique par la distance géographique significatif ($p = 0,016$, figure 10.5).

10.2 Analyses nucléaires à macro-échelle

10.2.1 Diversité génétique

Comme dans le cas des analyses mitochondriales, les trois groupes de populations ont été étudiés. Les résultats détaillés concernant les populations de *Cyclope neritea* situées dans l'aire "naturelle" ont été présentés dans la section 7 de la partie II.

Sans être aussi tranchée, la comparaison des données microsatellites entre les différents groupes de populations va dans le même sens que celle basée sur les données mitochondriales. Le tableau 10.2 présente les indices de diversité calculés pour les 8 populations des groupes 1 et 2. Les niveaux de diversité observés dans ces deux groupes de populations sont similaires à ceux calculés dans l'aire "naturelle" de *Cyclope neritea*. Par exemple, les valeurs de $\hat{H}_{E.nuc}$ varient entre 0,646 et 0,837 dans les groupes 1 et 2, et entre 0,669 et 0,884 pour les populations de l'aire "naturelle" de l'espèce. Bien qu'il soit difficile de comparer les nombres d'allèles entre groupes d'effectifs différents, on peut toutefois remarquer que les valeurs moyennes multilocus de nombre d'allèles par population sont légèrement plus élevées dans le groupe 1 ($\bar{N}_{all.moy} = 12,6$) que dans les groupes 2 et 3 ($\bar{N}_{all.moy} = 6,7$ et $10,7$ respectivement), alors même que les effectifs du groupe 1 ($N = 169$) sont nettement inférieurs à ceux du groupe 3 ($N = 430$). C'est également le cas des richesses alléliques : pour chaque locus, les valeurs moyennes sont légèrement plus fortes dans les populations du groupe 1 que dans les populations "naturelles" (valeurs moyennes totales : 7,33 et 6,58 respectivement).

Comme avec les données mitochondriales, Séné est la population du groupe 1 présentant le plus grand nombre d'allèles ($\bar{N}_{all} = 14,3$) ainsi que la plus grande richesse allélique ($\hat{R}_{all} = 7,92$). Toutefois, toutes les populations de ce groupe sont similaires du point de vue de l'hétérozygotie attendue.

10.2.2 Structure génétique

Au niveau intra-population, toutes les populations des groupes 1 et 2 (exceptées Ronce et Pedreña) présentent des valeurs de $\hat{F}_{IS}$ positives et significatives (tableau 10.4). Pour toutes ces populations (à l'exception de Hossegor), deux locus sur quatre présentent un déficit en hétérozygotes significatif. Cette observation est en accord avec l'hypothèse d'un effet Wahlund dû au mélange d'individus issus de populations génétiquement différenciées.

Au sein des groupes 1 et 2, et sur l'ensemble de ces deux groupes, les statistiques F révèlent une structure identique à celle observée dans ces populations avec l'ADN mitochondrial (tableau

TAB. 10.2: Diversité des 4 locus microsattellites pour les populations des groupes 1, 2 et 3. Pour ce dernier groupe (déjà examiné à la partie II, figure 7.2), seules les valeurs moyennes et totales sont indiquées. N : nombre d'individus, N_{all} : nombre d'allèles, $\hat{R}_{all}$: richesse allélique, $\hat{H}_{E.nuc}$: diversité génique. SVB : San Vicente de la Barquera.

Populations	N	N_{all}	2C4 $\hat{R}_{all}$	$\hat{H}_{E.nuc}$	N_{all}	3C4 $\hat{R}_{all}$	$\hat{H}_{E.nuc}$	N_{all}	2H10b $\hat{R}_{all}$	$\hat{H}_{E.nuc}$	N_{all}	1A10 $\hat{R}_{all}$	$\hat{H}_{E.nuc}$	$\bar{N}_{all}$	Multilocus $\hat{R}_{all}$	$\hat{H}_{E.nuc}$
Populations FAM (Groupe 1)																
a. Stérec	33	6	4,71	0,730	7	5,82	0,795	19	10,56	0,916	19	10,21	0,905	12,8 (7,2)	7,83 (2,99)	0,837 (0,090)
b. Séné	33	7	5,84	0,789	6	4,49	0,704	12	7,39	0,799	32	13,97	0,967	14,3 (12,1)	7,92 (4,20)	0,815 (0,110)
c. Rivedoux	38	5	7,81	0,783	9	5,84	0,798	20	9,70	0,850	22	10,37	0,909	14 (8,3)	7,68 (2,76)	0,835 (0,057)
d. Ronce	32	7	4,35	0,652	6	5,06	0,776	12	7,57	0,844	15	9,19	0,896	10 (4,2)	6,54 (2,24)	0,792 (0,105)
e. Arguin	33	6	4,21	0,686	7	4,85	0,716	11	7,09	0,814	24	10,58	0,896	12 (8,3)	6,68 (2,87)	0,763 (0,096)
Moyenne (S.D.)	33,8 (2,4)	6,2 (0,8)	4,79 (0,64)	0,728 (0,060)	7,0 (1,2)	5,21 (0,60)	0,758 (0,045)	14,8 (4,3)	8,46 (1,56)	0,845 (0,045)	22,4 (6,4)	10,86 (1,82)	0,915 (0,030)	12,6 (6,8)	7,33 (2,81)	0,811 (0,086)
Total	169	9		0,767	10		0,763	34		0,874	58		0,949	27,8		0,838
Statut indéterminé (Groupe 2)																
f. Hossegor	33	3	2,29	0,488	6	3,83	0,675	12	7,19	0,762	10	6,91	0,835	7,8 (4,0)	5,06 (6,43)	0,690 (0,150)
g. Pedreña	3	2	-	0,667	3	-	0,667	3	-	0,583	2	-	0,667	2,5 (0,6)	-	0,646 (0,042)
h. SVB	37	6	4,13	0,687	8	4,98	0,682	8	5,62	0,799	17	10,98	0,942	9,8 (4,9)	2,39 (3,10)	0,778 (0,122)
Moyenne (S.D.)	24,3 (18,6)	3,7 (2,1)	3,21 (1,30)	0,614 (0,110)	5,7 (2,5)	4,41 (0,82)	0,675 (0,008)	7,7 (4,5)	6,41 (1,11)	0,715 (0,116)	9,8 (7,5)	8,95 (2,88)	0,815 (0,139)	6,7 (4,3)	5,74 (2,66)	0,705 (0,118)
Total	73	6		0,589	10		0,732	15		0,813	18		0,888	12,3		0,756
Populations "naturelles" (Groupe 3)																
Moyenne (S.D.)	33,1 (8,3)	5,6 (1,9)	4,05 (1,05)	0,653 (0,095)	6,0 (2,4)	4,74 (1,19)	0,701 (0,112)	13,2 (4,2)	7,84 (1,54)	0,826 (0,08)9	17,8 (8,8)	9,71 (2,76)	0,887 (0,063)	10,65 (7,14)	6,58 (2,89)	0,767 (0,130)
Total	430	20		0,786	18		0,826	47		0,933	66		0,948	37,75		0,873

10.3). En effet, le déficit en hétérozygotes global, exprimé par $\hat{F}_{IT}$, est principalement dû à un déficit en hétérozygotes entre populations ($\hat{F}_{IS}$ importants) plutôt qu'à une différenciation génétique entre les populations ($\hat{F}_{ST}$ faibles bien que significatifs). Ainsi, il n'est pas surprenant que le test de Mantel, réalisé afin de tester la corrélation entre distance génétique et distance géographique des populations du groupe 1, soit non significatif ($p = 0,813$). Cette situation s'oppose à celle que nous avons observée dans l'aire "naturelle" de *Cyclope neritea* : la différenciation génétique totale se partageait de la même façon au sein des populations, au sein des régions, et entre les régions (voir partie II, tableau 7.4), et un isolement génétique par la distance géographique avait été mis en évidence.

TAB. 10.3: Structure génétique au sein des populations des groupes 1 et 2. Tous les estimateurs sont hautement significatifs ($p < 10^{-5}$, test exact).

	$\hat{F}_{IT}$	$\hat{F}_{IS}$	$\hat{F}_{ST}$
Groupe 1	0,143	0,107	0,040
Groupe 2	0,222	0,150	0,085
Groupe 1 & 2	0,169	0,119	0,056

Cette faible structure au niveau populationnel se retrouve à la figure 10.6 présentant un arbre de populations réalisé avec l'ensemble du jeu de données (populations des groupes 1, 2 et 3). La structure de cet arbre, pour les populations de l'aire "naturelle", est quasiment identique à celle identifiée dans la partie II avec des regroupements de ces populations en fonction de leur position géographique (*e.g.* Alvor, Faro, Cadiz, Ebre et Sardaigne sont associées). En revanche, contrairement à la situation observée dans l'arbre de populations construit avec les données mitochondriales, toutes les populations FAM et de statut indéterminé se retrouvent au sein d'une même branche de l'arbre, soulignant leurs similarités sur le plan nucléaire. Le défaut de cette représentation est lié au fait que les disparités individuelles ne sont pas examinées.

Afin de corroborer l'hypothèse d'un mélange d'individus provenant de sources génétiquement différenciées, nous avons donc réalisé des tests d'assignation individuels avec le logiciel **Geneclass v2.0.g** (Piry *et al.*, 2004). Les résultats sont présentés sous une forme synthétique à la figure 10.7. Les individus du groupe 1 sont assignés aux populations sources potentielles des groupes 2 et 3. Les cinq populations FAM présentent des profils d'assignation similaires. Parmi elles, un nombre important d'individus est en effet assigné aux trois populations du groupe 2. Les assignations vers Pedreña doivent toutefois être considérées avec précaution en raison de ses très faibles effectifs. Caorle et Katérinie sont ensuite les populations qui reçoivent le plus d'assignation. Il est à noter que la population de Séné est celle dont le nombre d'individus assignés à Caorle est le plus grand. Dans une moindre mesure, quelques individus de chaque population française

TAB. 10.4: Valeurs de $\hat{F}_{IS}$ pour chaque population étudiée. Les probabilités associées sont présentées entre parenthèses et les valeurs de $\hat{F}_{IS}$ significatives sont indiquées en gras.

	2C4	3C4	2H10b	1A10	Multilocus
Populations FAM (Groupe 1)					
a. Stérec	0,253 (0,015)	-0,067 (0,903)	0,040 (0,548)	0,130 (0,002)	0,085 (0,005)
b. Séné	0,193 (0,437)	0,139 (0,556)	0,093 (0,008)	0,092 (0,021)	0,127 (0,010)
c. Rivedoux	0,172 (0,452)	-0,178 (0,113)	0,142 (0,039)	0,152 (0,015)	0,160 (0,007)
d. Ronce	0,138 (0,164)	-0,047 (0,584)	0,001 (0,254)	0,107 (0,166)	0,047 (0,200)
e. Arguin	0,160 (0,001)	0,144 (0,436)	-0,030 (0,660)	0,154 (0,028)	0,105 (0,002)
<i>Total</i>	0,184 ($<10^{-4}$)	0,066 (0,603)	0,052 (0,033)	0,127 ($<10^{-4}$)	0,107 ($<10^{-4}$)
Statut indéterminé (Groupe 2)					
f. Hossegor	0,273 (0,119)	0,028 (0,551)	-0,108 (0,805)	0,347 (0,008)	0,130 (0,051)
g. Pedreña	0,500 (1)	-0,500 (1)	-0,143 (1)	0,500 (1)	0,097 (1)
h. SVB	0,056 (0,866)	0,009 (0,041)	0,155 (0,058)	0,369 (0,000)	0,166 ($<10^{-4}$)
<i>Total</i>	0,151 (0,593)	0,000 (0,255)	0,033 (0,418)	0,364 ($<10^{-4}$)	0,150 ($<10^{-4}$)
Total	0,177 (0,001)	0,046 (0,461)	0,046 (0,006)	0,198 ($<10^{-4}$)	0,116 (0,002)

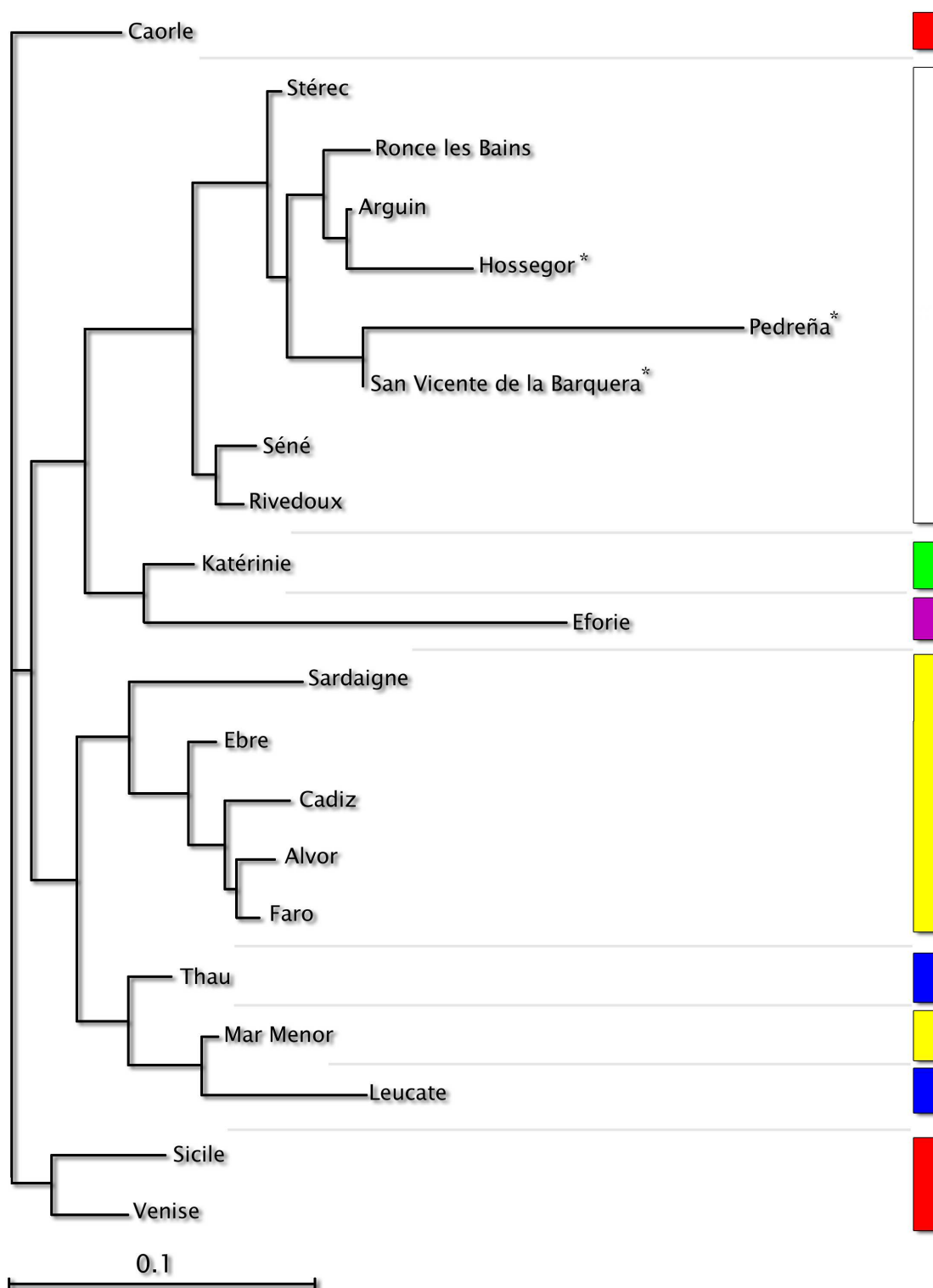


FIG. 10.6: Arbre de populations construit avec les données microsatellites. Les couleurs à droite correspondent aux lignées évolutives majoritaires dans chacune des populations. Comme pour la figure 10.3, les populations du groupe 1 n'ont pas de couleur attitrée (cadre vide) et les populations de statut indéterminé sont signalées par une astérisque.

sont affectés d'une part aux populations de Sicile et de Venise, et d'autre part, aux populations de Leucate, Cadiz et Ebre. Parmi les populations du groupe 1, Séné et Rivedoux présentent des profils d'assignation similaires. Ces deux populations sont les seules n'abritant aucun individu assigné vers San Vicente de la Barquera, et par ailleurs, au sein de ces populations, un nombre d'individus non négligeable est assigné aux populations de Caorle, Katerinie, Ebre et Cadiz.

10.3 Analyses nucléaires à micro-échelle spatiale

Le tableau 10.5 présente les valeurs de $\hat{F}_{IS}$ ainsi que les valeurs de probabilités associées au test d'adéquation à l'équilibre de Hardy-Weinberg. Neuf des treize populations analysées présentent des écarts significatifs aux proportions attendues sous l'équilibre de Hardy-Weinberg, généralement des déficits en hétérozygotes (sauf Rivedoux 2 et Ebre 1). Ces déficits sont la plupart du temps supportés par un seul locus, différent selon les populations. En outre, pour une population donnée, tous les locus ne répondent pas de la même façon, suggérant que le déficit global en hétérozygotes n'est pas dû à des croisements entre individus apparentés.

TAB. 10.5: Valeurs de $\hat{F}_{IS}$ et probabilités associées (p) pour chaque population des cinq sites étudiés. Les valeurs de $\hat{F}_{IS}$ significatives sont présentées en gras. La colonne "*Locus*" indique quels sont les locus présentant des valeurs de $\hat{F}_{IS}$ significatives.

Population	N	$\hat{F}_{IS}$	p	Locus
Rivedoux 1	37	0,166	0,007	2H10, 1A10
Rivedoux 2	40	-0,049	0,050	2H10
Rivedoux 3	40	-0,025	0,097	
Mar Menor 1	37	0,208	$< 10^{-3}$	2C4, 1A10
Mar Menor 2	40	0,146	0,003	3C4
Mar Menor 3	40	0,012	0,369	
Ebre 1	40	-0,034	0,010	3C4, 2H10
Ebre 2	33	0	0,115	
Caorle 1	40	0,072	0,016	2C4
Caorle 2	38	0,141	$< 10^{-5}$	1A10
Caorle 3	40	0,126	0,005	2C4
Eforie 1	35	0,075	$< 10^{-5}$	2C4
Eforie 2	35	-0,097	0,466	

Le tableau 10.6 présente les valeurs de $\hat{F}_{ST}$ par paires de populations. Seuls trois tests exacts sont non significatifs. Ils correspondent aux paires de populations suivantes : Mar Menor 1 et 2,

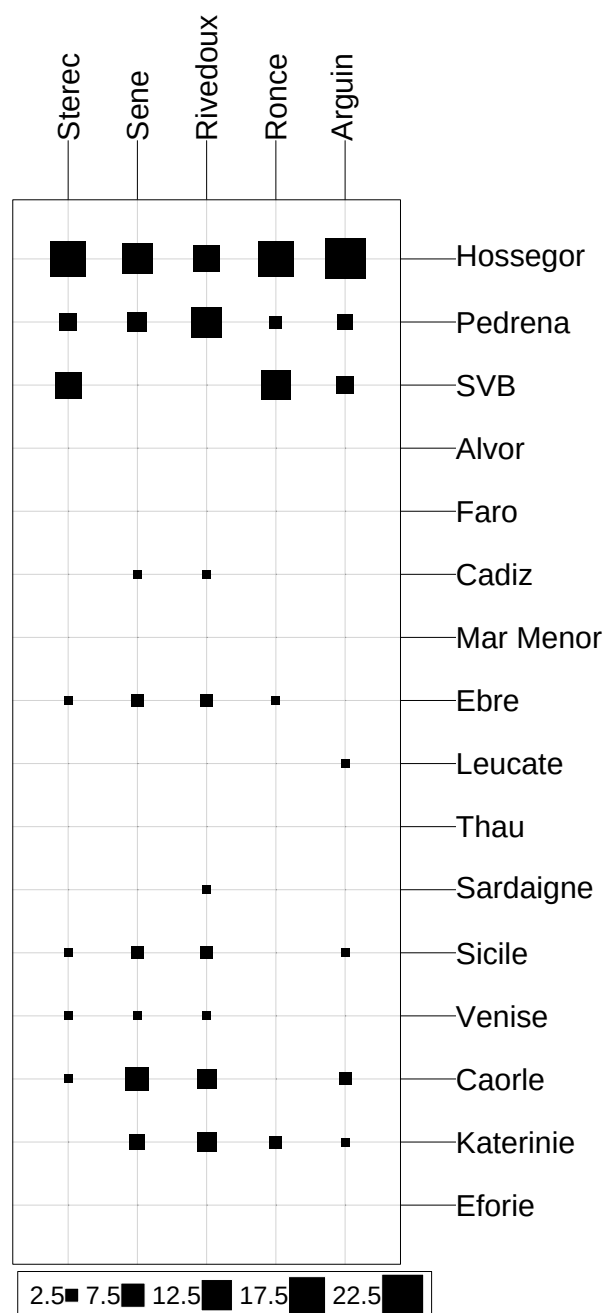


FIG. 10.7: Résultats des tests d'assignation des individus des populations FAM. Les colonnes du diagramme correspondent, de gauche à droite, aux populations FAM ordonnées du nord au sud. Les populations sources potentielles sont ordonnées d'Hossegor à Eforie en longeant le trait de côte. Pour chaque population FAM, le nombre d'individus assignés à chacune des sources potentielles est représenté par des carrés de tailles différentes (échelle au bas de la figure). Par exemple, à Stérec, la majorité des individus est assignée à Hossegor, San Vicente de la Barquera (SVB) puis Pédreña.

TAB. 10.6: $\hat{F}_{ST}$ par paires de populations (au-dessus de la diagonale) et valeurs de probabilités associées (en dessous de la diagonale). Presque toutes les valeurs de probabilités sont inférieures à 10^{-5} : seules les valeurs de probabilités supérieures à ce seuil sont donc indiquées dans le tableau. Les valeurs de Fst **non significatives** sont indiquées en gras. Seules les initiales des populations sont indiquées en en-tête des colonnes. Les cellules grisées permettent de visualiser les valeurs de $\hat{F}_{ST}$ entre populations d'un même site.

	R 1	R 2	R 3	M M 1	M M 2	M M 3	E 1	E 2	C 1	C 2	C 3	E 1	E 2
Rivedoux 1	-	0,002	0,018	0,087	0,100	0,092	0,074	0,067	0,054	0,052	0,040	0,204	0,106
Rivedoux 2	0,044	-	0,026	0,081	0,087	0,081	0,088	0,075	0,038	0,048	0,028	0,203	0,127
Rivedoux 3			-	0,110	0,110	0,099	0,054	0,054	0,070	0,077	0,062	0,197	0,077
Mar Menor 1				-	0,005	0,013	0,154	0,120	0,122	0,107	0,104	0,243	0,169
Mar Menor 2				0,152	-	-0,003	0,170	0,138	0,138	0,134	0,119	0,247	0,171
Mar Menor 3				0,029	0,282	-	0,168	0,133	0,140	0,136	0,120	0,249	0,168
Ebre 1							-	0,021	0,093	0,091	0,098	0,252	0,104
Ebre 2								-	0,076	0,068	0,081	0,244	0,103
Caorle 1									-	0,033	-0,003	0,209	0,145
Caorle 2										-	0,030	0,220	0,142
Caorle 3									0,468		-	0,198	0,144
Eforie 1												-	0,164
Eforie 2													-

Mar Menor 2 et 3, et Caorle 1 et 3. Seules des comparaisons de populations au sein d'un même localité présentent ainsi des valeurs non significatives. A l'exception du site d'Eforie, pour lequel la structure génétique entre populations est importante ($\hat{F}_{ST} = 0,164$, $p < 10^{-5}$, ce qui est cohérent, compte tenu de la distance géographique séparant ces populations), toutes les populations d'un même site sont caractérisées par des valeurs de $\hat{F}_{ST}$ très faibles, que ce soit à Rivedoux (valeurs significatives comprises entre 0,002 et 0,026) ou dans les sites de l'aire "naturelle" (seule une valeur significative pour la Mar Menor : 0,013, $p = 0,029$, et deux pour Caorle : 0,030 et 0,033 $p < 10^{-5}$). Cette faible différenciation intra-localité, y compris dans l'aire "naturelle", suggère qu'à cette échelle, les capacités de dispersion de l'espèce sont fortes. Ce résultat est cohérent avec nos observations sur le terrain : la vitesse des individus de *Cyclope neritea* à la recherche d'une proie est en effet importante pour un gastéropode et a été confirmée en laboratoire (Bachelet *et al.*, 2004, article présenté en annexe D.1).

L'arbre de populations présenté à la figure 10.8 synthétise les informations apportées par les comparaisons entre paires de populations (tableau 10.6). Il montre en effet que les similitudes entre populations d'une même localité sont plus importantes que les similitudes entre populations de sites distincts. Cependant, le site de Rivedoux présente un profil légèrement différent puisque la population de Rivedoux 3 est associée aux populations d'Eforie. Rappelons que l'analyse des données mitochondriales avait montré que Rivedoux était la seule population récemment détectée sur le littoral français à présenter un haplotype caractéristique de la lignée évolutive d'Eforie (voir figures 10.1 et 10.2).

Ce schéma général est supporté par les résultats de l'AFC présentés à la figure 10.9. Les individus appartenant aux populations d'une même localité restent proches les uns des autres en comparaison de la dispersion totale du nuage de points. Toutefois, la dispersion des individus de la Mar Menor est nettement inférieure à celle des deux autres sites, bien que les trois populations de la Mar Menor soient géographiquement plus éloignées les unes des autres que les populations échantillonnées à Rivedoux et Caorle. En outre, les nuages de points individuels des sites de Caorle et de Rivedoux se superposent, indiquant des similarités fortes entre les individus de ces deux sites. Ce résultat est conforme au profil d'assignation de la population de Rivedoux présenté plus haut (voir figure 10.7) : aucun individu de cette population n'a été assigné à la Mar Menor alors qu'un nombre non négligeable a été assigné aux populations de Caorle et Katérinie.

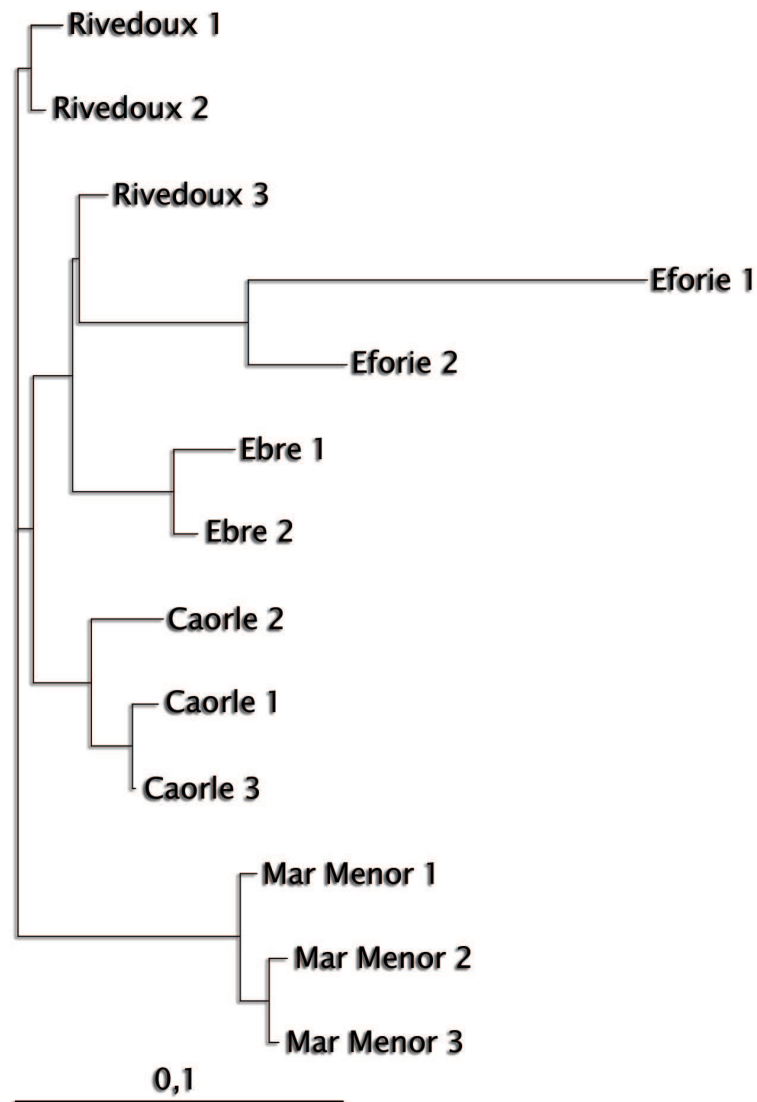
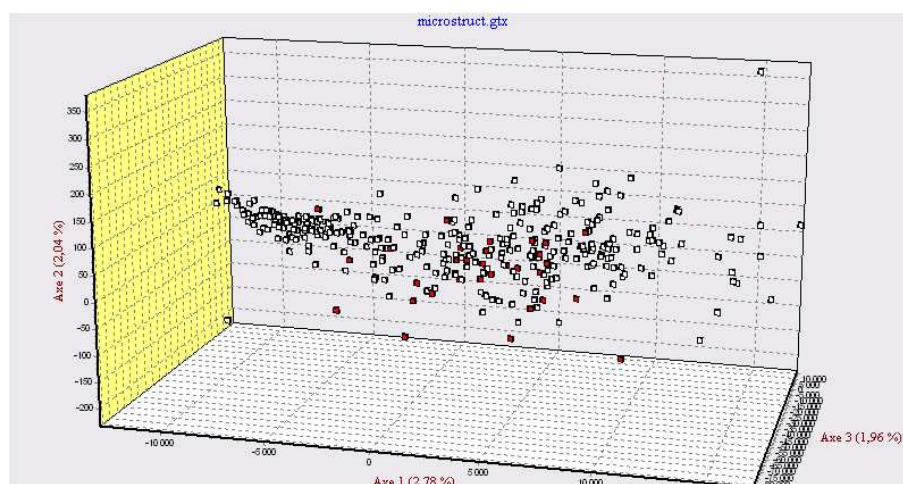
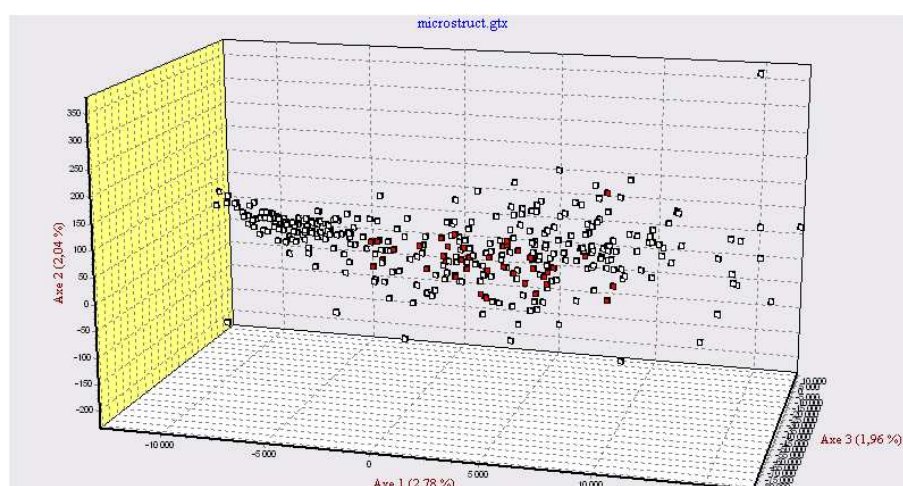


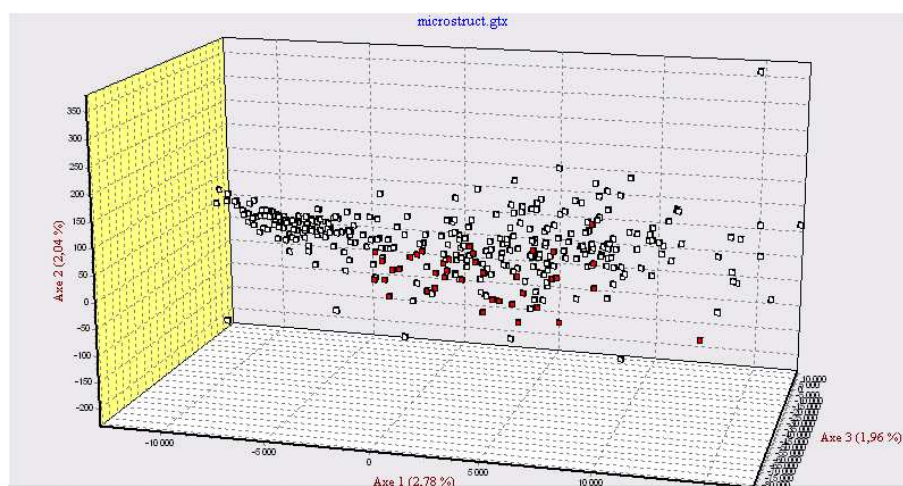
FIG. 10.8: Arbre de populations présentant les similitudes entre les 13 populations utilisées pour étudier la micro-structure.



a. Rivedoux 1

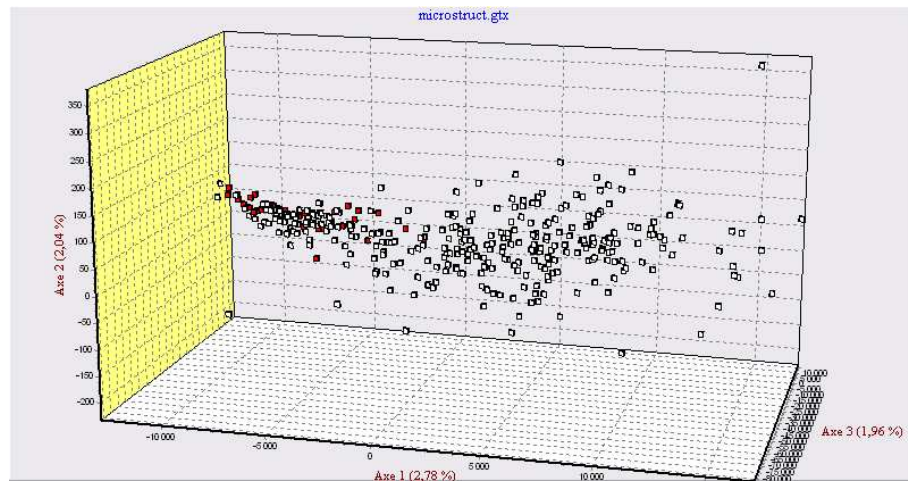


b. Rivedoux 2

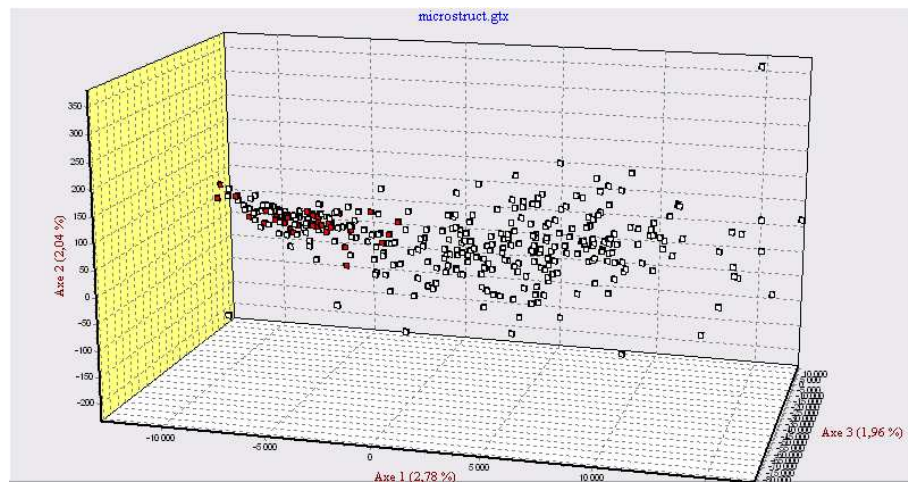


c. Rivedoux 3

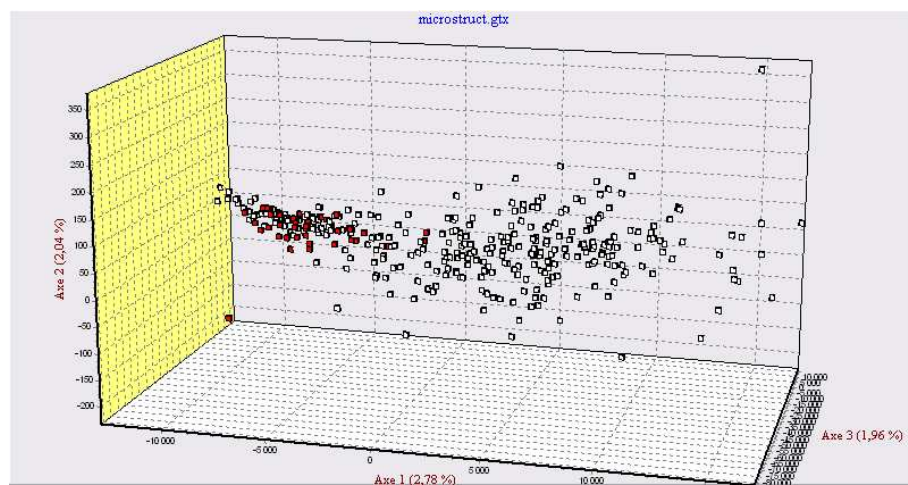
FIG. 10.9: Analyse Factorielle des Correspondances. Chaque point représente un individu dans l'espace factoriel. Chaque graphique permet de repérer les individus d'une population donnée (figurés en rouge).



d. Mar Menor 1

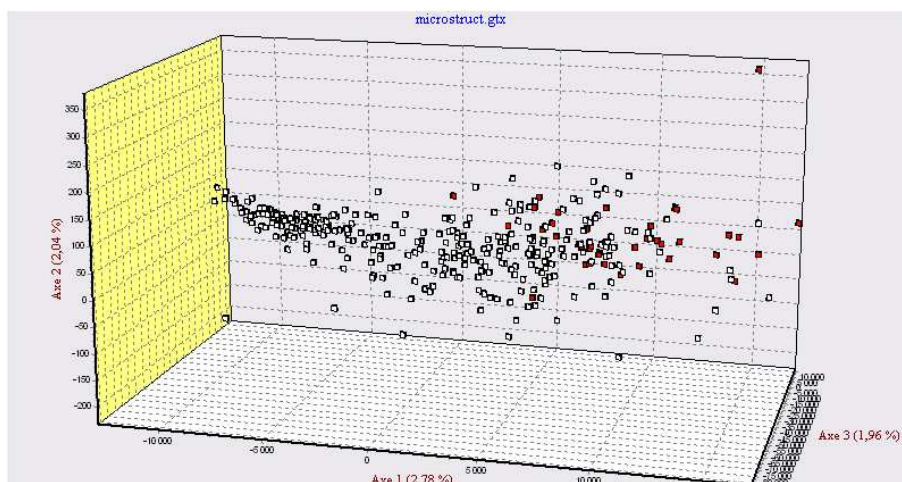


e. Mar Menor 2

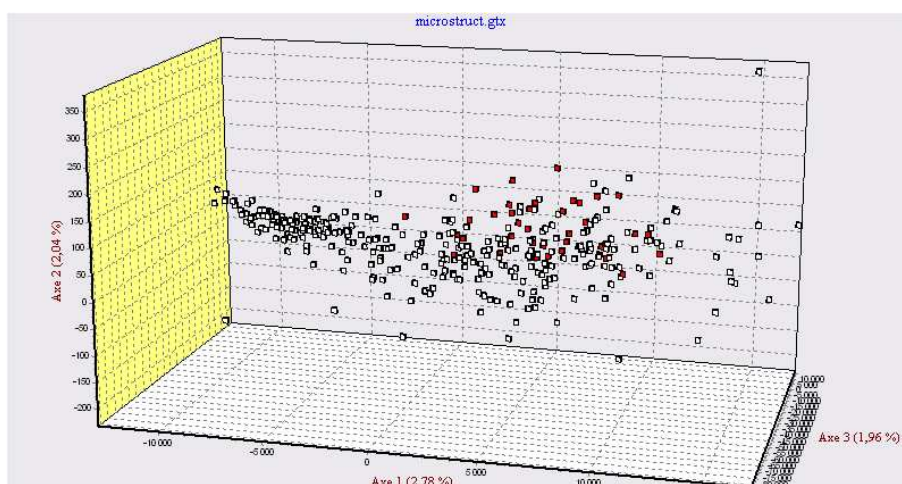


f. Mar Menor 3

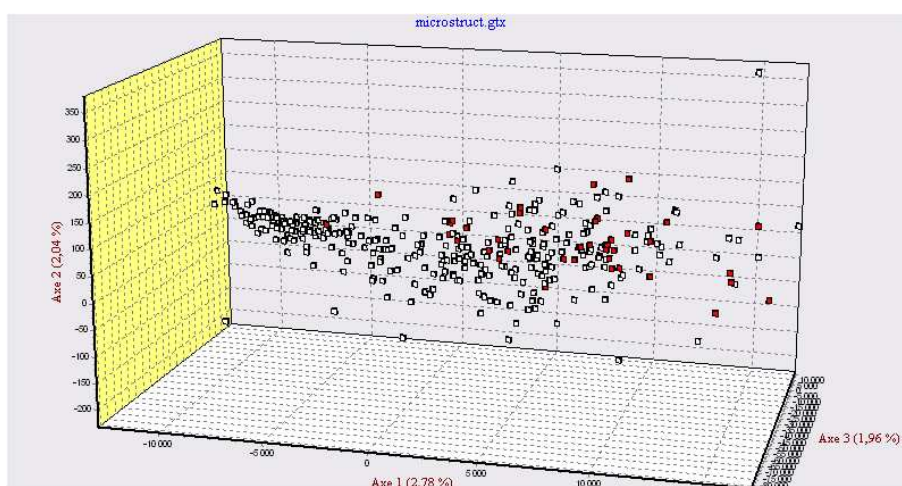
FIG. 10.9: *suite*. Analyse Factorielle des Correspondances.



g. Caorle 1



h. Caorle 2



i. Caorle 3

FIG. 10.9: *suite et fin*. Analyse Factorielle des Correspondances.

Discussion

11.1 Un scénario d'introductions récurrentes depuis de multiples sources

11.1.1 Schéma d'introduction

En comparant les populations récemment détectées sur les côtes françaises de l'Atlantique et de la Manche (populations FAM, groupe 1) et les populations situées dans l'aire "*naturelle*" de *Cyclope neritea* (groupe 3), des différences marquées ont été mises en évidence, grâce à l'analyse des deux types de marqueurs. Dans la partie II, nous avons montré que les populations de l'aire "*naturelle*" de l'espèce étaient caractérisées, au niveau mitochondrial, par une faible diversité génétique intra populations et par une très forte structure au niveau régional. Au contraire, les populations FAM sont caractérisées par une forte diversité génétique au sein des populations et par une faible structure génétique entre populations. De même, les analyses nucléaires ont révélé une diversité génétique dans les populations FAM supérieure à celle observée dans les populations du groupe 3, et une très faible différenciation génétique entre populations alors qu'un isolement génétique par la distance géographique hautement significatif avait pu être mis en évidence dans l'aire "*naturelle*".

Ces différents résultats sont congruents avec un scénario dans lequel les populations FAM ne seraient pas issues d'un événement d'introduction unique depuis une population source fortement diversifiée, mais d'introductions récurrentes en provenance de plusieurs sources. Plus précisément, dans ces populations françaises, nous avons observé, au niveau mitochondrial, de fortes diversités haplotypiques et nucléotidiques, et la présence de deux à trois groupes d'haplotypes fortement divergents. Les distributions "*match-mismatch*", qui présentent (à l'exception de Séné) de multiples pics, ainsi que les tests K significatifs, confirment que ces populations ne sont pas à l'équilibre démographique et que des événements de "*mélange*" (au sens du terme anglophone "*admixture*") d'individus, issus de populations ayant évolué indépendamment, ont

eu lieu. Parmi les treize populations de l'aire "*naturelle*" de *Cyclope neritea*, aucune ne présente un mélange d'haplotypes appartenant aux trois lignées divergentes retrouvées mélangées dans la majorité des populations FAM (figures 10.1 et 7.2). En outre, toutes les populations de l'aire "*naturelle*" sont moins diversifiées que les populations FAM, y compris les populations les plus diversifiées de l'aire "*naturelle*" (Venise et Katérinie), qui restent moins diversifiées que toutes les populations récemment détectées en France. Ainsi, au niveau populationnel, la diversité génétique a augmenté dans la zone où *Cyclope neritea* a été récemment détectée, mais à une échelle plus globale, la forte structure observée dans le bassin méditerranéen a disparu.

Les résultats obtenus grâce aux locus microsatellites viennent confirmer ces observations. Outre la diversité génétique légèrement plus élevée dans les populations FAM que dans l'aire "*naturelle*", les valeurs de $\hat{F}_{IS}$ positives et significatives observées dans quatre des cinq populations FAM suggèrent un effet Wahlund (le système de reproduction n'étant pas en cause, dans la mesure où *Cyclope neritea* est une espèce gonochore ; de plus, les locus ne présentent pas de façon simultanée de déficit en hétérozygotes). Il est probable que cet effet Wahlund est le reflet du mélange de lignées évolutives divergentes à l'échelle des populations. Les résultats des tests d'assignation viennent étayer l'hypothèse d'introductions en provenance de multiples sources. De nombreux individus des populations FAM sont en effet assignés à plusieurs populations de l'aire "*naturelle*". En outre, ces assignations coïncident particulièrement bien avec les similitudes observées sur les données mitochondriales entre populations FAM et populations "*naturelles*". Ainsi :

- la population de Rivedoux est celle présentant le plus grand nombre d'individus assignés à la population de Katérinie. Rivedoux est de plus la seule population à présenter, au niveau mitochondrial, un haplotype de la lignée trouvée principalement en Mer Noire. Rappelons que les populations de Katérinie et d'Eforie sont systématiquement associées au sein des mêmes classes par les analyses bayésiennes réalisées à la partie II sous **Geneland**.
- la population de San Vicente de la Barquera (SVB), qui présente un profil de mélange clair entre individus de "*type ibérique*" et individus de "*types languedociens*" doit avoir gardé la trace de cet événement de mélange au niveau nucléaire. Les assignations d'individus des populations du groupe 1 vers SVB sont donc susceptibles de refléter, dans les populations françaises, le mélange de ces deux lignées divergentes. Cela est d'autant plus probable que les populations de Rivedoux et de Ronce, ne présentant aucun individu assigné vers SVB, sont les populations au sein desquelles les lignées évolutives "*languedociennes*" et "*ibériques*" ne coexistent pas.

Enfin, au niveau nucléaire, la structure génétique interpopulations est également très faible à l'échelle des populations FAM, près de deux fois inférieure à la différenciation entre individus au sein des populations. Cela se traduit notamment par (i) le regroupement de toutes les populations FAM au sein de la même branche de l'arbre de populations de la figure 10.6 et (ii) par les profils d'assignations très ressemblants de toutes les populations FAM. Ainsi, les données nucléaires

permettent également de révéler une augmentation locale de la diversité génétique dans la zone où *Cyclope neritea* a été récemment détectée, cette augmentation s'accompagnant au niveau global d'une diminution de structure par rapport à l'aire "naturelle" de l'espèce.

Ainsi, comme nous l'avions suggéré dans Bachelet *et al.* (2004) au sujet de la population d'Arguin, les résultats obtenus montrent que les populations de *Cyclope neritea* enregistrées le long des côtes françaises ont été introduites depuis plusieurs populations sources de l'aire "naturelle", faiblement diversifiées mais fortement divergentes. Ce résultat est cohérent avec de nombreuses études d'espèces exogènes réalisées en milieu terrestre, dulçaquicole ou marin et pour lesquelles la réduction de polymorphisme attendue suite à l'évènement de fondation n'est pas observé (c'est notamment le cas de la poaceae *Bromus tectorum*, Novak *et al.*, 1993, de la moule brune *Perna perna*, Holland, 2001, de plusieurs gastéropodes thiaridaes, Facon *et al.*, 2003, de la crépidule *Crepidula fornicata*, Dupont *et al.*, 2003, du lézard cubain, Kolbe *et al.*, 2004 ou encore de l'algue brune *Undaria pinnatifida*, Voisin *et al.*, 2005). Comme l'ont montré Kolbe *et al.* (2004), quand des populations naturelles sont monomorphes, l'occurrence d'introductions multiples peut promouvoir l'établissement d'une espèce exogène, en transformant la variation inter-populationnelle de l'aire d'origine en variation intra-populationnelle dans l'aire d'introduction, ce qui permettrait de résoudre en partie le paradoxe des invasions biologiques²⁰.

11.1.2 Vecteurs d'introduction

Les échanges d'huîtres sont connus pour être à l'origine de nombreuses introductions d'espèces exogènes le long des côtes atlantiques européennes (Goulletquer *et al.*, 2002; Wolff et Reise, 2002). Comme le suggérait Sauriau (1991), ces échanges pourraient avoir joué un rôle majeur dans la colonisation des côtes françaises par *Cyclope neritea* puisque toutes les populations FAM recensées à ce jour sont apparues à proximité de parcs à huîtres, où du moins dans des baies à forte activité conchylicole. De tels mécanismes ont déjà été suggérés pour d'autres espèces marines côtières (voir par exemple le cas d'*Ocenebrellus inornatus*, Martel *et al.*, 2004, de la crépidule *Crepidula fornicata*, Dupont *et al.*, 2003, ou encore de l'algue brune *Undaria pinnatifida*, Voisin *et al.*, 2005). Ce vecteur pourrait en outre expliquer l'introduction de nombreux individus issus de populations génétiquement différenciées. Toutefois, les huîtres ne sont pas le seul vecteur d'introduction possible et d'autres espèces cultivées peuvent avoir contribué à l'établissement durable de *Cyclope neritea* sur nos côtes. Le profil génétique particulier de la population de Séné nous en donne un exemple. Cette population montre en particulier :

1. un nombre d'haplotypes observés supérieur à celui attendu sous hypothèse d'équilibre (test

²⁰Le paradoxe de l'invasion biologique est le suivant : comment est-il possible que, dans un milieu donné, une espèce introduite puisse supplanter les espèces locales, théoriquement adaptées à leur milieu de vie, alors même que l'espèce introduite, (i) ayant évolué dans un milieu différent, devrait être moins bien adaptée que les espèces locales dans son aire d'introduction, et (ii) la perte de diversité génétique de l'espèce introduite, en raison de l'effet de fondation, devrait théoriquement réduire ses capacités d'adaptation? (voir Sax et Brown, 2000)

- K, figure 10.4), et le plus grand nombre d'allèles observés parmi les populations FAM
2. une position proche des populations de la Mer Adriatique dans l'arbre de populations réalisé avec les données mitochondriales (figure 10.3)
3. le plus grand nombre d'individus assignés aux populations de Caorle, Venise et Sicile (figure 10.7)
4. sur 65 valeurs de $\hat{\Phi}_{ST}$ calculées entre paires de populations “naturelles” et de populations FAM, la seule valeur de différenciation non significative a été observée entre les populations de Séné et de Venise ($\hat{\Phi}_{ST} = 0,045$, $p = 0,072$, données non présentées).

Pour cette population, l'hypothèse d'un nombre limité de populations sources génétiquement différenciées, l'une d'entre elles ayant récemment subi une expansion démographique, paraît plus parcimonieuse que l'hypothèse d'introductions impliquant de multiples sources plus faiblement différenciées mais fortement divergentes (comme c'est le cas pour les autres populations FAM). Ce scénario est cohérent avec les échanges fréquents de palourdes *Ruditapes philippinarum* entre la lagune de Venise et le Golfe du Morbihan (J.-F. Auvray, Satmar, communication personnelle) où la population de Séné a été échantillonnée. Les échanges de palourdes pourraient ainsi constituer le vecteur principal (voire l'unique vecteur) d'introduction de *Cyclope neritea* dans cette zone.

Cet exemple illustre la difficulté d'identifier un unique vecteur d'introduction pour une espèce exogène sur l'ensemble de sa zone d'introduction. Comme cela a été montré pour d'autres espèces marines introduites (*e.g.* l'algue brune *Undaria pinnatifida*, Voisin *et al.*, 2005), l'étude exhaustive des populations introduites est nécessaire afin d'identifier les facteurs et vecteurs ayant pu promouvoir l'établissement d'une espèce exogène. Toutefois, toutes les populations françaises de *Cyclope neritea* sont caractérisées par au moins un trait commun : leur diversité génétique supérieure à celle des populations “naturelles”, ce qui renforce l'idée de l'importance de la “pression de propagule”²¹ dans le succès d'une invasion (Williamson, 1996; Sax et Brown, 2000).

11.2 Le cas des populations en limite d'aire

Les trois populations situées en bordure de l'aire de distribution “naturelle” historiquement reconnue de *Cyclope neritea* (*i.e.* San Vicente de la Barquera, Pedreña et Hossegor) ont été classées, dans cette partie, “populations de statut indéterminé”. Bien que située à la limite nord de l'aire de distribution “naturelle”, cette zone est généralement considérée comme en faisant partie intégrante, dans la mesure où l'espèce y a été régulièrement recensée par le passé (Sauriau, 1991). Pourtant, les populations trouvées dans cette zone sont toutes situées à proximité de parcs ostréicoles (voir partie I). Par conséquent, ces populations pourraient aussi bien être naturelles

²¹ Terme définissant la pression démographique exercée par l'apport massif de propagules (larves, spores, œufs...) dans un nouvel habitat. Outre le renforcement démographique des populations nouvellement installées, cet apport massif va augmenter la probabilité d'introduction de variants génétiques adaptés aux nouvelles conditions environnementales.

qu'introduites. Puisque cette zone constitue un réservoir potentiel de migrants vers les côtes françaises, il est important de déterminer le statut de ces populations.

Certains résultats supportent l'hypothèse d'un statut naturel de ces populations. En particulier, les trois populations sont génétiquement proches des autres populations de la péninsule ibérique (figures 10.2, 7.3 et 10.3), et leurs niveaux de diversités génétiques sont plus proches des populations de l'aire "*naturelle*" que des populations FAM, à la fois sur la base des données mitochondriales (tableau 10.1) et microsatellites (tableau 10.2).

Toutefois, ces trois populations ne présentent pas exactement le même profil génétique. Hossegor a des caractéristiques que nous avons observées dans le groupe des populations "*naturelles*" (voir tableau 10.1) :

1. les diversités haplotypiques, nucléotidiques et géniques ainsi que la richesse allélique sont plus proches des valeurs du groupe 3 que du groupe 1.
2. les deux haplotypes trouvés dans cette population (*i.e.* H_1 et H_5) appartiennent à la même lignée évolutive (figure 10.1) : seule une mutation les sépare.
3. une analyse bayésienne réalisée dans **Geneland** et incluant, outre les populations "*naturelles*", les trois populations de statut indéterminé, a associé la population d'Hossegor au groupe de populations ibériques (voir les résultats partiels de cette analyse présentés à la figure 11.1b).

A l'opposé, la population de San Vicente de la Barquera présente des caractéristiques que nous avons observées dans la plupart des populations du groupe 1 (populations récemment apparues sur les côtes françaises) :

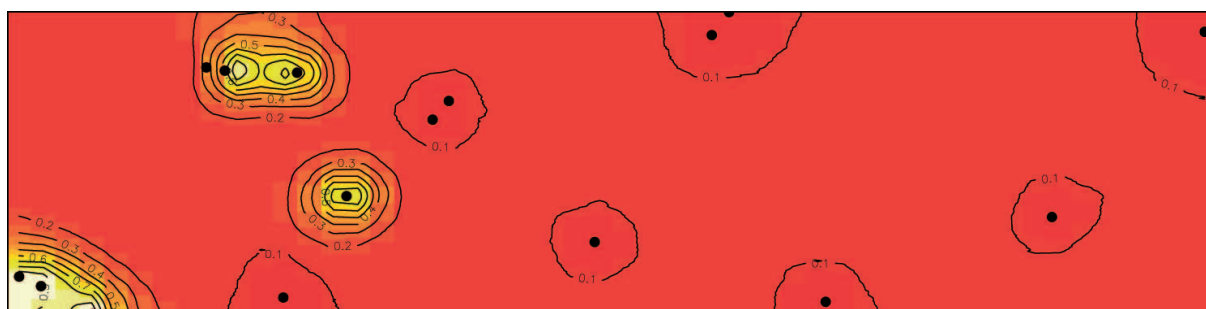
1. une forte diversité génétique
2. la présence d'haplotypes fortement divergents (haplotype "*ibérique*" H_1 et "*languedocien*" H_{18}) n'ayant jamais été trouvés tous les deux en mélange au sein d'une population de l'aire "*naturelle*".
3. l'analyse réalisée dans **Geneland** associe cette population au groupe des populations languedociennes (Thau, Leucate, Sardaigne et Mar Menor, figure 11.1c).

Il est donc très probable qu'une partie des individus de cette population a été transférée par l'homme. Les échanges d'huîtres entre Méditerranée et Atlantique sont en effet fréquents de part et d'autre de la péninsule ibérique (voir par exemple l'introduction probable de *Cyclope neritea* en Galice depuis la Méditerranée ; Advisory Committee on the Marine Environment, 2001; Rolán, 1992).

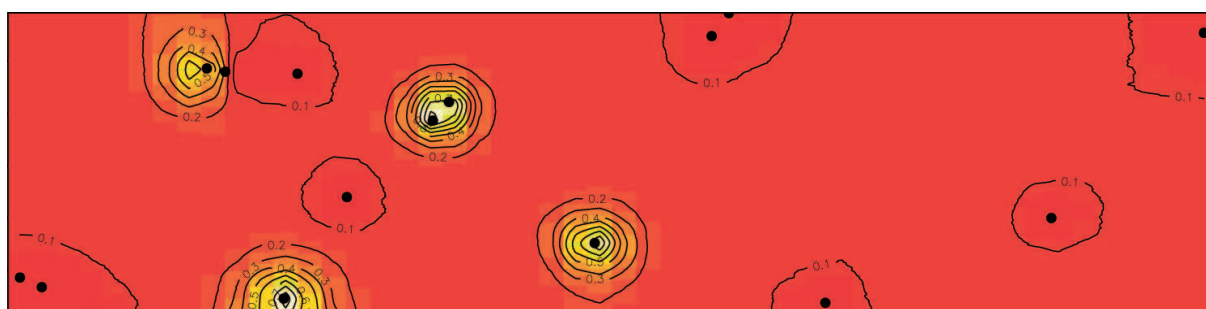
Enfin le faible effectif observé à Pedreña nous empêche de nous prononcer sur le statut de cette population (bien qu'elle soit incluse par **Geneland** dans le groupe ibérique, figure 11.1b). Elle semble en déclin : seuls trois individus vivants ont pu y être échantillonnés alors que de nombreuses coquilles vides (plus que dans aucun autre site d'échantillonnage) ont été trouvées dans les laisses de mer.



a. Localisation des sites d'échantillonnage



b. Association de Pedreña et Hossegor avec les populations ibériques.



c. Association de San Vicente de la Barquera avec les populations "languedociennes".

FIG. 11.1: Résultats partiels produits par **Geneland** en vue de déterminer le statut des populations en limite d'aire.

Ces résultats montrent combien il est difficile d'identifier des processus généraux s'appliquant à une région dans son ensemble. Ici, même des populations proches géographiquement présentent des profils génétiques tranchés : Hossegor a toutes les caractéristiques d'une population "*naturelle*", Pedreña est probablement en déclin, et la population de San Vicente de la Barquera est vraisemblablement le fruit d'un mélange d'individus introduits et d'individus "*locaux*".

11.3 Combinaison de deux facteurs : introductions humaines et dispersion naturelle

Les introductions humaines associées au transport d'espèces cultivées (*i.e.* huîtres et ou palourdes) ont probablement joué un rôle majeur dans le processus d'introduction primaire de *Cyclope neritea* le long des côtes françaises de l'Atlantique et de la Manche. Toutefois, (i) les populations récemment établies sont apparues graduellement sur nos côtes, du sud vers le nord, (ii), un profil significatif d'isolement génétique par la distance géographique a été mis en évidence entre ces populations (figure 10.5) et (iii) la majorité des individus des populations FAM sont assignés à la population de Hossegor, qui présente toutes les caractéristiques d'une population "naturelle".

L'isolement génétique par la distance géographique pourrait être lié à des introductions secondaires en raison d'échanges entre sites conchylicoles, ou à une dispersion naturelle depuis le site d'introduction primaire (*i.e.* Arguin) ou depuis la péninsule ibérique. Dans la mesure où nous ne disposons pas d'informations précises sur les échanges d'espèces cultivées entre bassins de productions, il nous est impossible de rejeter totalement l'hypothèse d'introductions secondaires. Toutefois, la plupart des échanges humains entre les sites conchylicoles français se fait au hasard entre des sites distants, sans continuité particulière. A titre d'illustration, tous les sites ostréicoles français (de la Manche à la Méditerranée) sont dépendants de la production de naissain des bassins d'Arcachon et de Marennes-Oléron (Goulletquer et Heral, 1997). Si la dispersion humaine était le seul processus affectant les populations de *Cyclope neritea*, les haplotypes devraient être mélangés aléatoirement entre les populations, causant l'absence de profil d'isolement par la distance ; il nous faut donc supposer que la dispersion naturelle a joué un rôle dans l'établissement des populations de *Cyclope neritea* sur nos côtes.

L'absence de profil d'isolement par la distance sur les données nucléaires n'exclut pas l'hypothèse de l'expansion naturelle. Dans l'aire "naturelle", la différenciation génétique totale entre populations est moins visible avec les marqueurs nucléaires qu'avec le marqueur mitochondrial. c'est également le cas sur les côtes françaises de l'Atlantique et de la Manche, ce qui explique probablement pourquoi un profil d'isolement par la distance n'est pas visible ici. Toutefois, les profils d'assignation sont sans ambiguïté : pour toutes les populations FAM, la majorité des individus sont assignés vers les populations de Hossegor, Pedreña et San Vicente de la Barquera.

Étant donné les capacités de dispersion limitées de l'espèce et les signatures claires d'introductions récurrentes, il peut paraître surprenant de suggérer que la dispersion naturelle puisse pour partie expliquer le profil d'isolement par la distance observé. Il est de plus peu probable que ce profil particulier ait pu apparaître en seulement vingt à trente ans. Une hypothèse permettant d'expliquer ces contradictions apparentes est celle de la présence, de longue date, de *Cyclope neritea* sur nos côtes, mais en faible densité, l'espèce restant donc indétectable avant les années 1970. Deux observations anciennes réalisées en Baie de Morlaix pourraient accréditer cette thèse :

(i) en 1830, une population dense de *Buccinum neriteum* (synonyme de *Cyclope neritea*) y a été enregistrée par Collard des Cherres (1830)²² et (ii) en 1923, une coquille vide de *Neritula neritea* (autre synonyme de *Cyclope neritea*) y a été observée par Dautzenberg et Fischer (1925). En outre, la distribution discontinue (en “*patches*”) des populations, y compris au sein de l’aire “*naturelle*”, peut rendre l’espèce difficile à détecter quand les densités individuelles sont faibles (voir partie I). Les introductions récentes auraient ensuite contribué à renforcer localement ces populations “*naturelles*” cryptiques.

Cyclope neritea peut ainsi être classifiée comme une “*espèce cryptogénique*” telle que définie par Carlton (1996a). Ces espèces, “ni véritablement natives, ni véritablement introduites”, représentent probablement une large part des espèces dites “*introduites*” nouvellement détectées dans une zone (Carlton, 1996a). Des populations ancestrales de *Cyclope neritea* ont donc pu être présentes sur nos côtes bien avant 1976, soit en raison de migrations naturelles en provenance de la zone la plus au nord de l’aire de distribution “*naturelle*”, soit en raison d’introductions accidentelles anciennes non répertoriées. L’huître *Ostrea edulis* a par exemple été exploitée en France depuis le début du XX^e siècle, avec, dès cette période, des échanges entre la lagune de Thau et le Bassin de Marennes-Oléron (Hinard et Lambert, 1928), susceptibles, comme aujourd’hui, d’avoir servi de vecteur d’introduction pour *Cyclope neritea*. En outre, Sax et Brown (2000) ont démontré que les espèces devenant soudainement abondantes et largement répandues parviennent dans cette situation uniquement après l’échec de nombreuses tentatives de colonisations antérieures, naturelles ou accidentelles. L’établissement récent et pérenne de populations denses de *Cyclope neritea* (de 400 à plus de 1600 individus par mètre carré à Rivedoux, Tardy *et al.*, 1985) pourrait donc avoir été favorisé par la conjugaison des deux processus suivants :

1. une accélération du rythme des introductions accidentelles depuis les années 1970 en raison de l’intensification des transferts d’huîtres entre sites conchyliques atlantiques et méditerranéens.
2. l’augmentation rapide, sur la même période, de la température des eaux de surface du Golfe de Gascogne. Stachowicz *et al.* (2002) ont par exemple montré que le réchauffement global (dont celui des océans), et en particulier les modifications des températures minimales l’hiver et maximales l’été, ont favorisé l’installation de plusieurs espèces d’ascidies introduites.

La combinaison de ces deux processus (Dukes et Mooney, 1999) pourrait avoir favorisé l’établissement des individus introduits et des migrants venant du sud, créant ainsi un renforcement démographique des populations cryptiques de *Cyclope neritea* le long des côtes atlantiques.

En conclusion, les facteurs ayant pu contribuer à l’installation de *Cyclope neritea* sur les côtes françaises de l’Atlantique et de la Manche sont multiples :

²²Il convient toutefois de prendre cet enregistrement avec précaution, l’auteur étant connu, dans ses écrits, pour être responsable de plusieurs erreurs d’identification (S. Gofas, communication personnelle).

- une longue histoire commune avec l’homme à l’origine de renforcements démographiques locaux de populations introduites de façon récurrentes.
- la remontée des températures facilitant l’expansion vers le nord.
- les capacités de dispersion locale de l’espèce. Sur ce dernier point, nous avons mis en évidence, à micro-échelle, qu’en dehors du site d’Eforie ($\hat{F}_{ST} = 0,164$), dont les deux populations sont distantes de plus de 30 kilomètres, les valeurs de $\hat{F}_{ST}$ au sein des baies étudiées sont très faibles, y compris pour des populations “naturelles” en place de longue date et séparées de quelques kilomètres ($\hat{F}_{ST} = 0,016$; 0,005; 0,0021 et 0,020 pour les sites de Rivedoux, Mar Menor, Ebre et Caorle respectivement). Ces résultats suggèrent qu’à l’échelle locale, les capacités de dispersion de *Cyclope neritea* sont relativement importantes. Ainsi, même si les populations introduites sur le littoral français ont plusieurs origines, les échanges fréquents au niveau local vont avoir tendance à homogénéiser les populations. En outre, Bachelet *et al.* (2004) ont montré expérimentalement, que dans certaines conditions (courantologie faible ou nulle, correspondant en partie aux conditions rencontrées sur l’estran à marée basse), cette espèce est capable de se déplacer plus vite sur le substrat que la nasse réticulée *Nassarius reticulatus* (espèce compétitrice potentielle), en particulier pour la recherche de proies.
- enfin, Bachelet *et al.* (2004) ont également montré que *Cyclope neritea* présentait une charge parasitaire nettement inférieure à celle de ses compétiteurs potentiels.

Toutes les conditions sont donc réunies pour que l’installation de populations denses de *Cyclope neritea* le long des côtes françaises de l’Atlantique et de la Manche soit durable.

Conclusions

Tout comme dans l'aire de distribution "*naturelle*" de *Cyclope neritea*, nous venons de voir que dans les zones étudiées ici (Pays Basque espagnol, côtes françaises de l'Atlantique et de la Manche), les informations apportées par l'ADN mitochondrial et les locus nucléaires sont en grande partie congruentes. En particulier, les résultats suivants ont pu être mis en évidence :

1. les trois populations du Pays Basque, dont le statut était auparavant indéterminé, présentent pour certaines des caractéristiques de populations "*naturelles*" (Hossegor) et pour d'autres (San Vicente de la Barquera), un profil de mélange de lignées divergentes, également rencontrées dans les populations FAM.
2. les populations des côtes françaises de l'Atlantique et de la Manche ont certainement une double origine. L'effet des introductions humaines en provenance de multiples sources est indéniablement présent. Il est également probable qu'une expansion naturelle depuis les populations du Pays Basque a pu avoir lieu de manière contemporaine ou plus ancienne.
3. au niveau local, les capacités de dispersion de l'espèce semblent importantes à l'échelle de quelques kilomètres. Cette capacité, associée à (i) la grande diversité génétique des populations françaises, (ii) des capacités intrinsèques (sensibilité des récepteurs sensoriels lui permettant de repérer une proie à distance, et vitesse lui permettant de l'atteindre plus rapidement que d'autres espèces nécrophages) et (iii), une absence de charge parasitaire, font de cette espèce un compétiteur efficace, toutes les conditions étant réunies pour une installation durable, voire une expansion de l'espèce sur nos côtes.

Conclusion générale, perspectives

Expansion et déplacement des aires de distribution : peut-on identifier les processus impliqués ?

Un des objectifs principaux de cette thèse était d'identifier les mécanismes à l'œuvre permettant d'expliquer l'établissement récent du mollusque gastéropode *Cyclope neritea* le long des côtes françaises de l'Atlantique et de la Manche. La particularité de cette espèce tient au fait que les nouvelles populations détectées en France sont apparues graduellement du sud vers le nord et dans le prolongement direct de son aire "*naturelle*" (ou reconnue comme telle), à proximité immédiate de la limite de son aire de distribution. Or, ces régions sont connues pour être des zones particulièrement instables, avec apparitions et disparitions de populations en fonction des fluctuations environnementales (biotiques et abiotiques) et des arrivées de migrants en provenance du cœur de l'aire de distribution de l'espèce. Ces régions en limite d'aire sont donc des zones complexes qui devraient être particulièrement sensibles aux changements environnementaux, notamment climatiques, que ce soit sous un régime de contraction de l'aire de distribution (se traduisant alors par la disparition de ces populations) ou sous un régime d'expansion géographique (les populations précédemment en "*limite d'aire*" se rapprochant du centre de l'aire de distribution). Pourtant, en milieu marin, peu d'études ont permis de documenter avec certitude des déplacements récents de limites d'aires en relation avec des changements climatiques (voir quelques exemples cités dans Harley *et al.*, 2006). En outre, Cox et Moore (2000) rappellent que l'observation des changements de distribution des espèces ne permet généralement pas d'identifier les causes de ces changements en raison de l'interaction de très nombreux facteurs, biotiques et abiotiques. Ces auteurs citent notamment le cas d'oiseaux migrants d'Amérique du nord dont le déclin démographique, amorcé il y a 25 ans, est attribuable autant à des changements climatiques qu'à des altérations de l'habitat de ces oiseaux.

Les mécanismes agissant sur la distribution actuelle des espèces sont donc multiples et en

forte interaction. Dans ce contexte, les activités humaines, et notamment les introductions biologiques, jouent un rôle majeur en interagissant avec les processus écologiques et évolutifs naturels, à l'échelle des communautés (voir figure 12.1) comme à l'échelle des populations d'une espèce donnée (voir figure 12.2). A titre d'illustration, Stachowicz *et al.* (2002) ont montré que la dominance d'espèces d'ascidies exotiques sur les espèces locales était liée à une réaction efficace des espèces introduites en réponse aux modifications des minimas et maximas thermiques (*i.e.* recrutement plus précoce).

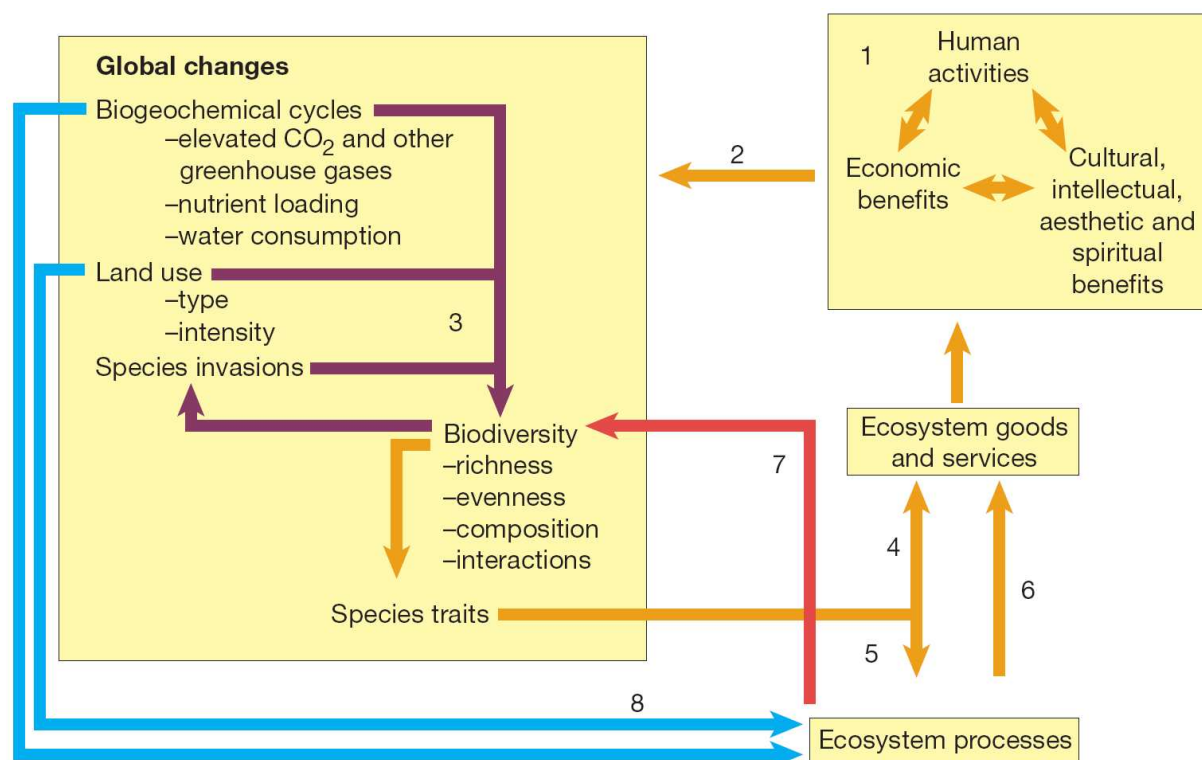
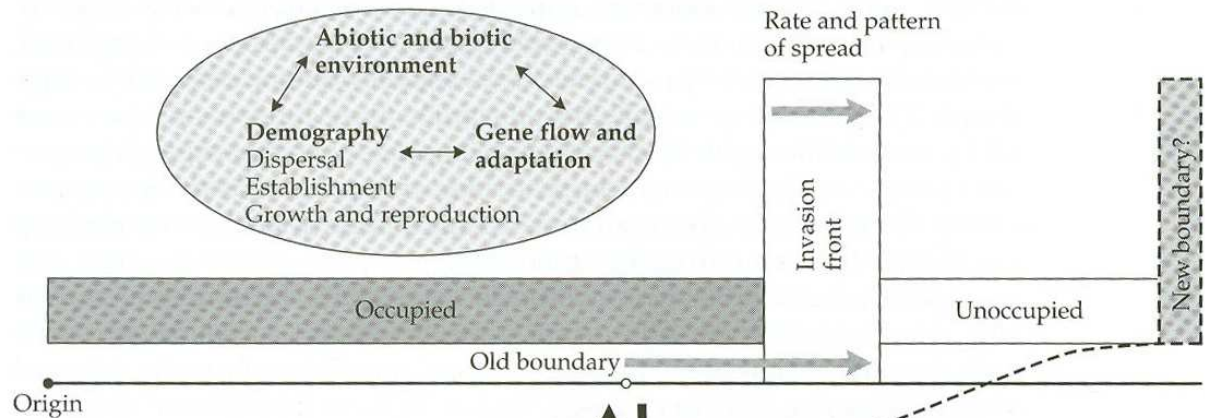


FIG. 12.1: Influences humaines sur la biodiversité et le changement global. Les activités humaines (1) causent des changements globaux significatifs aux niveaux écologique et environnemental (2). Ces changements contribuent à modifier d'une part la biodiversité et d'autre part le contrôle de la susceptibilité aux invasions par cette biodiversité (3, flèches violettes). Les changements de diversité peuvent avoir des conséquences directes sur les utilisations humaines des écosystèmes, et se répercutent donc sur les activités économiques et sociales (4). Les changements de biodiversité peuvent également influencer les processus des écosystèmes (5). S'ils sont altérés, ces processus peuvent se répercuter sur les utilisations humaines des écosystèmes (6) et peuvent en retour, altérer un peu plus la biodiversité (7, flèche rouge). Les changements globaux peuvent également avoir des effets directs sur les processus des écosystèmes (8). Tiré de Chapin *et al.* (2000).

De la même façon, les résultats obtenus sur *Cyclope neritea* au cours de cette thèse permettent d'illustrer (i) l'effet des activités humaines sur l'augmentation de l'aire de distribution d'une espèce dans la continuité de son aire de distribution historique et (ii) la possibilité d'interactions entre processus anthropiques (*e.g.* modification des habitats, introduction d'espèce) et évolution naturelle des populations. L'utilisation de marqueurs moléculaires et la mise en évidence de

(A) Non-equilibrium



(B) Equilibrium

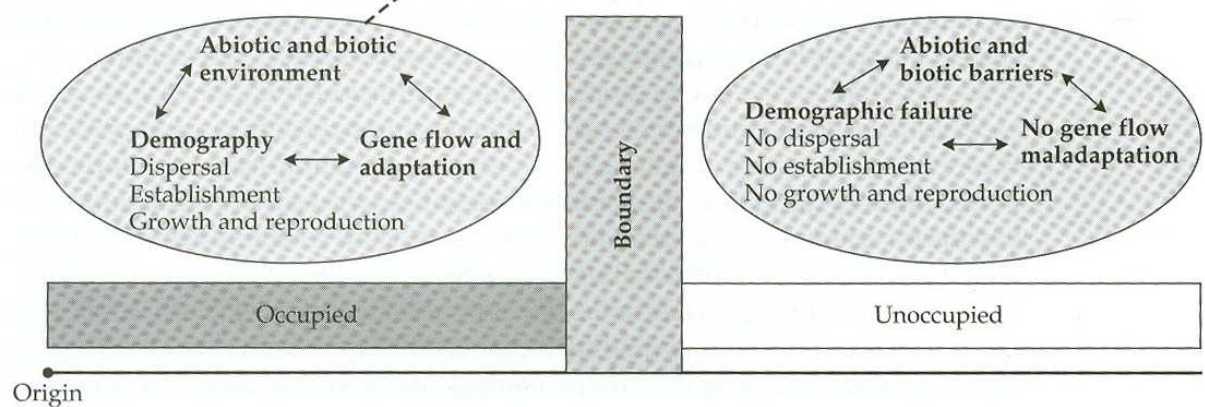


FIG. 12.2: Modèle conceptuel de la dynamique de l'aire de distribution d'une espèce. Les processus influençant la position de la limite d'aire sont présentés en phase de déséquilibre (A) et en phase d'équilibre (B). Pour chaque phase, les flèches pleines connectent les processus interagissant pour déterminer les limites d'aire. Les flèches pointillées indiquent les processus responsables du passage d'une phase à l'autre (*e.g.* rupture d'une barrière géographique). Les espèces exogènes peuvent nous renseigner sur différents processus qui régulent le déplacement des limites d'aire au fil du temps, dont (i) les facteurs pouvant initier une phase de déséquilibre avec expansion géographique (flèche "*expansion*"), (ii) les processus déterminant le rythme et le profil de la dispersion, et (iii) les mécanismes qui permettent à une nouvelle limite d'aire de s'établir de façon stable (flèche "*pinning*"). Tiré de Kinlan et Hastings (2005)

signatures spécifiques dans l'aire de distribution "*naturelle*" de l'espèce nous ont en effet permis de montrer que :

- les populations récemment détectées en France sont issues de processus d'introductions multiples et récurrentes, d'individus en provenance de sources génétiquement divergentes de l'aire "*naturelle*" de l'espèce.
- l'hypothèse d'une expansion naturelle, expliquant pour partie l'établissement des populations françaises, ne pouvait être écartée.
- des disparités peuvent être mises en évidence entre les populations récemment détectées.

La population de Séné présente par exemple un profil génétique particulier qui suggère un lien privilégié avec les populations de l'Adriatique. L'hypothèse d'un vecteur d'introduction différent de l'huître a donc été formulée pour expliquer les caractéristiques génétiques de la population de Séné. Il serait intéressant, à titre de comparaison, d'étudier des espèces ayant une histoire similaire. C'est notamment le cas du gastéropode *Rapana venosa*, qui semble avoir été introduit délibérément en Vendée depuis la Mer Adriatique. Le gastéropode *Gibbula albida* pourrait également être une espèce d'intérêt dans le cadre de cette comparaison : originaire de Mer Noire, il aurait été introduit en Adriatique puis sur les côtes atlantiques françaises, directement ou *via* les côtes méditerranéennes (S. Gofas, communication personnelle).

Les deux processus (*i.e.* introduction et expansion) ont probablement contribué à l'installation durable de *Cyclope neritea* sur les côtes françaises de l'Atlantique et de la Manche. En outre, il n'est pas exclu que des populations cryptiques peu denses, donc non visibles, aient pu être présentes bien avant les premières détectations de l'espèce dans le Bassin d'Arcachon. Le renforcement démographique par migrations naturelles et introductions accidentelles a ensuite pu contribuer à augmenter la densité de ces populations.

Non seulement l'introduction des populations françaises résulte de processus récurrents, mais en outre, ces populations sont issues de sources multiples localisées dans des régions fortement divergentes de l'aire "*naturelle*". L'étude phylogéographique que nous avons menée dans le bassin méditerranéen montre que les échanges entre régions différentes sont très limités et ont pu être à l'origine d'évolutions divergentes des populations. Ces résultats ont des implications différentes concernant, d'une façon générale, les processus d'introductions biologiques et plus spécifiquement la réussite de l'établissement de *Cyclope neritea* le long des côtes françaises.

- Au niveau général, nos résultats nous permettent de compléter la liste croissante des espèces introduites pour lesquelles aucun goulot d'étranglement n'a eu lieu en relation avec le processus d'introduction. En effet, historiquement, des effets de fondation et de goulots d'étranglements étaient considérés comme un attendu théorique des processus d'introductions biologiques, reposant sur l'idée que le nombre de migrants arrivant de façon accidentelle devait être limité (Sax *et al.*, 2005). Or, de nombreuses études tant en milieu marin (*e.g.* l'algue brune *Undaria pinnatifida*, Voisin *et al.*, 2005) qu'en milieu terrestre (*e.g.* le lézard cubain, Kolbe *et al.*, 2004) ont mis en évidence l'absence de tels goulots lors de l'installation de ces espèces introduites dans de nouveaux milieux. Le débat reste néanmoins ouvert concernant la signification de ces résultats en terme de caractéristique associée à une introduction réussie : d'une part, des espèces ont été introduites avec succès bien qu'ayant subi de forts effets de goulots d'étranglements (*e.g.* fourmis d'Argentine, Tsutsui *et al.*, 2000 ; ou moules brunes, Holland, 2001) et d'autre part, les introductions ayant échoué ne fournissent pas de données qui permettraient de tester l'hypothèse d'un effet de dérive génétique suffisamment marqué pour expliquer de tels échecs (Sax *et al.*,

2005).

- A l'échelle de *Cyclope neritea*, l'hypothèse suivante peut être formulée : la présence d'individus fondateurs issus de régions génétiquement diversifiées et évolutivement divergentes pourrait favoriser l'installation durable de l'espèce le long des côtes françaises. Les individus fondateurs trouvent en effet leur source dans des environnements différents et sont ainsi susceptibles de pouvoir s'adapter à des conditions environnementales variables. Cette diversité génétique des populations de *Cyclope neritea* vient s'ajouter à (i) une faible charge parasitaire comparée à d'autres espèces compétitrices potentielles (Bachelet *et al.*, 2004), à (ii) une grande efficacité pour la recherche de proies (Bachelet *et al.*, 2004, observations personnelles) et (iii) une capacité de dispersion à micro-échelle spatiale relativement importante pour une espèce à développement direct (voir l'étude à micro-échelle, partie III). Toutes les conditions sont donc réunies pour que l'établissement de populations denses de *Cyclope neritea* sur nos côtes soit durable.

L'interaction entre facteurs naturels et anthropiques a aussi été mise en évidence dans l'aire de distribution "*naturelle*" de l'espèce. Nous avons vu que les populations récemment détectées en France ne sont pas les seules à présenter des profils de mélange de lignées évolutives divergentes. Ainsi, certaines populations de l'aire "*naturelle*" sont affectées par de tels mélanges visibles aux niveaux mitochondrial et nucléaire (c'est notamment le cas des populations de la Mar Menor et de Caorle). Ces mélanges restent toutefois limités, probablement car au sein de l'aire "*naturelle*" de l'espèce, les populations sont denses, bien installées et adaptées à l'environnement local. Dans ces conditions, il est probable que les individus arrivés de façon ponctuelle en provenance de régions géographiques différentes ne sont pas ou peu compétitifs par rapport aux individus locaux et que par conséquent, leur empreinte dans la population disparaît rapidement. En revanche, pour les populations situées en limite d'aire, les répercussions d'introductions sur le profil génétique de ces populations sont susceptibles d'être bien plus marquées. Ces populations, en raison du phénomène de maladaptation, sont par nature souvent instables. Ainsi, l'introduction de migrants, même en faibles effectifs, peut se traduire par des modifications marquées des fréquences alléliques ou haplotypiques de ces populations, susceptibles de freiner voire d'empêcher totalement la mise en place d'adaptations locales avantageuses. Les transferts humains participent de ce fait au processus de "*migration contre sélection*". La population de San Vicente de la Barquera en est une parfaite illustration : plus de la moitié des individus qui y ont été échantillonnés ne présentent pas le type mitochondrial local, mais le type mitochondrial languedocien. Dans ces conditions, les études phylogéographiques réalisées au sein de l'aire "*naturelle*" des espèces peuvent également être trompeuses : même au sein de l'aire "*naturelle*", il faut envisager que des processus non naturels (*i.e.* anthropiques) puissent être à l'œuvre pour façonner les profils génétiques observés. Cela est particulièrement vrai pour les espèces cultivées, concernées par les échanges d'origine anthropique. Par exemple, les relations phylogéniques au sein du complexe d'espèces du genre *Mytilus* sont particulièrement troubles, et faire abstraction d'éventuels phénomènes de mélanges

de populations au sein même de l'aire "*naturelle*" pourrait conduire à des interprétations erronées de l'histoire de ces espèces.

Concernant *Cyclope neritea*, seules deux populations aux effectifs importants ont pu être analysées dans la zone située en bordure de l'aire "*naturelle*". Afin de vérifier cette hypothèse d'instabilité et de stochasticité de l'effet des introductions en limite d'aire, il conviendrait de renforcer notre échantillonnage dans cette zone et de poursuivre l'étude des populations d'Hossegor et de San Vicente de la Barquera.

Nos résultats soulèvent de nombreuses questions. Notamment, quelles sont les capacités de compétition entre individus issus des différentes "*provinces génétiques*" mises en évidence dans le bassin méditerranéen? Les populations de ces différentes provinces diffèrent-elles par leur capacité d'adaptation vis-à-vis d'un changement climatique? Peut-on proposer des scénarios prédictifs et des attendus théoriques de l'évolution des populations de *Cyclope neritea* en limite d'aire? Plusieurs pistes sont ainsi envisagées en tant que perspectives de ce travail de thèse.

13

Perspectives

13.1 Histoire évolutive de *Cyclope neritea*

Les études phylogéographiques permettent d'analyser les processus de dispersion historiques et contemporains qui affectent la distribution des espèces. Dans notre cas, l'utilisation de cette approche avait pour vocation initiale de déterminer les sources possibles à l'origine des populations de *Cyclope neritea*. Au cours de ce travail, nous avons montré une divergence génétique particulièrement importante entre régions de l'aire "*naturelle*" organisée d'est en ouest. Toutefois, notre échantillonnage répondait avant tout à notre objectif premier d'identification de sources d'introductions potentielles. C'est pourquoi nous avons ciblé des zones d'échantillonnage avec lesquelles des échanges pouvaient avoir eu lieu (en particulier les bassins conchylicoles du nord de la Méditerranée, de l'Adriatique et de la péninsule ibérique).

Par conséquent, nous ne disposons pas d'informations suffisantes sur l'ensemble de l'aire de distribution "*naturelle*" pour pouvoir réaliser une analyse fine des profils observés. Par exemple, au niveau mitochondrial, la plupart des haplotypes identifiés à Séné (et constituant le centre du réseau, voir figures 7.2 et 10.2) n'a été retrouvée dans aucune autre des populations échantillonnées dans l'aire "*naturelle*" de l'espèce (figure 7.3). Deux zones mériteraient d'être plus particulièrement étudiées :

1. les côtes sud du bassin méditerranéen, du Maroc à l'Égypte, probablement beaucoup moins marquées par le dernier épisode glaciaire que les côtes nord de l'aire "*naturelle*". De plus, le Maroc pourrait représenter une autre région propice à l'étude de populations de *Cyclope neritea* en limite d'aire interagissant avec les activités humaines. En effet, les activités aquacoles sur la côte méditerranéenne du Maroc existent depuis 1920 et se sont véritablement développées au cours de ces quinze dernières années. La production marocaine d'huîtres et de palourdes est, en grande partie, destinée à la consommation locale, mais plusieurs espèces de poissons (anguilles, dorade, loup) sont principalement destinées aux exportations vers l'Europe, notamment vers l'Italie, l'Espagne et la France (Abouhala

et al., 1999). Il serait intéressant de suivre l'évolution des populations de ces côtes afin de voir si la détection précoce d'individus introduits est possible.

2. les côtes de la Mer Noire : au cours de chaque épisode glaciaire de l'aire quaternaire, la Mer Noire a été isolée de la Méditerranée en raison de l'abaissement général du niveau des eaux et de la faible profondeur au niveau du détroit du Bosphore. En raison de l'arrivée de nombreux fleuves se jetant dans la Mer Noire (Danube, Dniepr, Dniestr et Don), et de la faible évaporation de ses eaux de surface, le bilan hydrographique de ce bassin est positif. Cela a causé, à chaque période glaciaire une désalinisation importante des eaux de la Mer Noire, ainsi que la disparition de la plupart des espèces marines. Les populations qui y vivent aujourd'hui ont donc recolonisé la zone, depuis la Méditerranée à la suite du dernier maximum glaciaire.

Une étude phylogéographique sur ces deux zones nous permettrait de proposer un scénario d'évolution des populations de *Cyclope neritea* sur son aire de distribution et d'ajuster les dates de divergences estimées entre les différentes lignées évolutives observées dans cette étude, en fonction de données historiques (climatiques, géologiques ...).

Par ailleurs, une étude plus complète de l'espèce *Cyclope neritea*, en regard des autres espèces du genre, permettrait certainement d'apporter des informations sur le plan systématique de ce genre. Par exemple, des chercheurs russes font état de la présence de plusieurs espèces morphologiques de *Cyclope* en Mer Noire (jusqu'à huit selon les sources, D. Micu, communication personnelle). Cela paraît improbable mais suggère toutefois des différences morphologiques marquées au sein de cette zone. Il serait donc intéressant de voir si, contrairement à la population d'Eforie étudiée dans cette thèse, d'autres populations présentent un niveau de polymorphisme conforme à celui des autres populations de l'est de l'aire "naturelle".

13.2 Succès reproducteur et capacité d'adaptation : approches expérimentales

Bien que les données mitochondriales et microsatellites aient révélé des tendances générales similaires, nos résultats ont mis en évidence certaines disparités entre ces deux types de marqueurs. C'est par exemple le cas de la population de Sardaigne, qui, sur la base des données mitochondriales, présentait toutes les caractéristiques d'une population de type ibérique, alors qu'elle était associée aux populations languedociennes sur la base des données microsatellites. Compte tenu des modes d'hérédité de ces deux types de marqueurs, nous avons proposé comme explication un mélange entre individus locaux et migrants venant des populations languedociennes mais qui n'aurait pas été retrouvé dans les lignées maternelles. Les hypothèses permettant d'expliquer pourquoi le mélange de lignées divergentes n'est pas systématiquement visible au niveau de l'ADN mitochondrial sont (i) un biais de représentation des sexes mâles et femelles ou (ii)

l'hypothèse d'un succès reproducteur femelle plus important pour certaines lignées évolutives se retrouvant dans un environnement donné (*i.e.* ici, succès reproducteur des individus femelles de type "ibérique" très largement supérieur au succès des femelles de type "languedocien").

Un moyen de tester ces hypothèses serait de réaliser des croisements entre individus mâles et femelles de différentes origines géographiques dans des conditions contrôlées en aquarium.

Par ailleurs, un autre résultat obtenu au cours de cette thèse est l'impossibilité d'exclure une hypothèse de dispersion naturelle vers le nord, qui pourrait être favorisée par les augmentations de température des eaux du Golfe de Gascogne. Pour affiner notre compréhension de ces processus d'expansion en limite d'aire, il semble intéressant de développer une approche expérimentale, visant dans un premier temps à tester la tolérance à des températures élevées et basses chez des individus provenant de différentes régions géographiques colonisées par *Cyclope neritea*. Nous avons tenté, en cours de thèse, de mettre en place, en aquarium, un test de tolérance à la température en fonction de l'origine des populations étudiées. Les tests, conduits à 13, 20 et 27 °C avec des individus de Thau, Faro et Rivedoux avaient pour objectif d'analyser la valeur sélective en fonction de la température en suivant deux paramètres : la fertilité et la survie. Ces tests ne se sont pas avérés concluants, du fait des très faibles taux de survie observés lors de cette expérience, notamment car la nourriture fournie, se dégradant plus rapidement à haute température, avait causé la mort de nombreux individus. Toutefois, nous restons persuadés que de tels tests sont nécessaires pour évaluer les capacités intrinsèques de dispersion, d'adaptation, de reproduction aux niveaux intra et inter-lignées de l'espèce.

13.3 Développement d'outils théoriques pour tester des scénarios mixtes

En l'absence de références historiques concernant la présence ou l'absence d'une espèce donnée dans une zone particulière, il n'est pas toujours aisé de discriminer des processus de dispersion récents de processus historiques, ou encore des processus d'origine anthropique par rapport à des processus naturels.

Comme nous l'avons vu chez *Cyclope neritea*, la génétique des populations permet de lever un certain nombre d'interrogations en ce qui concerne les populations françaises récemment installées. Toutefois, la présence de certains profils génétiques complexes peut être délicate à interpréter. Les difficultés rencontrées sont en partie dues à l'absence d'attendus théoriques clairs concernant les profils de diversité génétique de populations ayant des histoires démographiques et évolutives complexes (*i.e.* variations de taille efficace des populations, de leur diversité haplotypique ou nucléotidique... Donnelly et Tavaré, 1995; Przeworski *et al.*, 2000). Bien que rarement quantitatifs, les attendus théoriques en terme de profils de diversité dans les populations contemporaines sont en effet parfaitement connus dans le cas de scénarios évolutifs simples mais restent souvent flous dans le cas d'histoires complexes. Ainsi, une population qui subit un goulot d'étran-

glement (*i.e.* réduction drastique de la taille de la population) doit voir sa diversité génétique chuter. Mais qu'en est-il si, par exemple, une population ancestrale se scinde en deux populations qui continuent d'échanger un faible nombre de migrants et dont l'une voit sa taille régulièrement diminuer avant sa remise en contact avec l'autre population ?

L'une des conséquences de ce manque d'attendus théoriques concerne le choix des indices statistiques à utiliser pour tester des hypothèses. Ce choix est parfois arbitraire, tant il est difficile (voire impossible) de savoir à l'avance quel indice sera le plus sensible, et donc le plus à même de révéler un processus particulier (Hudson, 1993; Felsenstein, 1992). Par exemple, avant la formalisation du test K par Depaulis et Veuille (1998), la statistique D mise au point par Tajima (1989) était couramment utilisée pour mettre en évidence des phénomènes de mélanges de populations, alors que Galtier *et al.* (2000) ont montré que cet indice n'était pas toujours le plus puissant.

Dans ce contexte, les approches par modélisation et simulations statistiques sont particulièrement pertinentes. En particulier, la reconstruction de généalogies de gènes suivant les lois de la théorie de la coalescence peut permettre de préciser (i) le comportement des indices statistiques couramment utilisés en génétique des populations (Hudson, 1993), et (ii) les profils génétiques attendus dans les populations ayant une histoire démographique et évolutive complexe.

Le programme réalisé en **Delphi** par Marie-Hélène Muller (2002), que nous avons modifié afin de réaliser des tests K , fonctionne selon ce principe. Son but premier était de tester la sensibilité de plusieurs indices statistiques permettant théoriquement de détecter des variations démographiques ancestrales marquées (augmentation rapide ou réduction drastique de la taille de la population ancestrale). En l'état actuel des choses, il permet à l'utilisateur de simuler l'histoire démographique simple d'une population théorique, et d'observer les effets de cette histoire sur différents indices statistiques utilisés couramment en génétique des populations (*e.g.* nombre d'haplotype, diversité haplotypique, diversité nucléotidique...). La robustesse et la grande simplicité du mode de construction des généalogies permet de répéter un grand nombre de fois la simulation d'un même scénario, 1000 simulations prenant rarement plus de quelques secondes. Travailler sur de multiples répétitions permet d'obtenir la distribution statistique empirique des différents indices calculés, y compris pour les indices dont la distribution statistique est impossible à obtenir analytiquement (Hudson, 1993).

La poursuite du développement de cet outil pourrait notamment permettre :

1. de simuler des scénarios plus complexes, en particulier mettant en situation plusieurs populations susceptibles de subir des événements de mélanges. La modélisation donnera la possibilité de mesurer l'effet de nombreux paramètres (tels que la taille des populations, les proportions d'individus mélangés, la date de divergence ou de remise en contact de populations...) sur les profils de diversité génétique observés dans les populations.
2. de tester la sensibilité d'indices statistiques couramment utilisés en génétique des populations face aux différents scénarios testés. La possibilité d'obtenir facilement les distribu-

tions nulles de nombreux indices permettra de choisir quel test utiliser pour tester tel ou tel scénario.

3. de fournir un outil d'aide à l'interprétation de données génétiques complexes. La comparaison de profils génétiques réels (*e.g.* observés par exemple sur le jeu de données mitochondriales de *Cyclope neritea*) avec les profils génétiques simulés devrait permettre d'estimer si tel ou tel scénario évolutif est compatible avec les données réelles. Si cette comparaison ne constitue pas une preuve en soi, elle donnerait toutefois la possibilité d'orienter l'interprétation de jeux de données complexes.

La mise en œuvre des approches évoquées ci-dessus, *i.e.* suivis de terrain, étude phylogéographique réalisée avec plusieurs types de marqueurs moléculaires, expérimentation et simulations statistiques, devrait permettre de répondre à un grand nombre de questions restées en suspens concernant l'étude de *Cyclope neritea*, un modèle biologique pertinent pour comprendre l'impact des processus historiques ou contemporains, naturels ou anthropiques, sur l'évolution de l'aire de distribution d'une espèce.

Bibliographie

- ABBOTT, R. (1992). Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends in Ecology & Evolution*, **7**(12) : 401–405. 15
- ABOUHALA, A., TALBAOUI, E. et BOUKABOUS, R. (1999). Accès au marché. Le cas du Maroc. *In : Planification de l'aquaculture dans les pays méditerranéens*, pages 61–62. CIHEAM-IAMZ. 157
- ADVISORY COMMITTEE ON THE MARINE ENVIRONMENT (2001). Report of the Working Group on Introductions and Transfers of Marine Organisms. Technical Report, ICES, International Council for the Exploration of the Sea. 56, 141
- ALZIEU, C. (1998). Tributyltin : case study of chronic contaminant in the coastal environment. *Ocean & Coastal Management*, **40**(1) : 23–36. 14
- ANDERSON, R., UNGER, M. et BURRESON, E. (1996). Enhancement of *Perkinsus marinus* disease progression in TBT-exposed oysters (*Crassostrea virginica*). *Marine Environmental Research*, **42** : 177–180. 14
- ARIAS, A. (1976). Contribución al conocimiento de la fauna bentónica de la Bahía de Cádiz. *Investigación pesquera*, **40** : 355–386. 38, 40
- ARNOLD, J., ASMUSSEN, M. et AVISE, J. (1988). An epistatic mating system model can produce permanent cytonuclear disequilibria in a hybrid zone. *Proceedings of the National Academy of Sciences of the United States of America*, **8**(5) : 1893–1896. 59
- ASMUSSEN, M. et BASTEN, C. (1994). Sampling theory for cytonuclear disequilibria. *Genetics*, **138** : 1351–1363. 59
- AVISE, J., NELSON, W., ARNOLD, J., KOEHN, R. et WILLIAMS, G. (1990). The evolutionary genetic status of Icelandic eels. *Evolution*, **44** : 1254–1262. 59
- BACHELET, G., CAZAUX, C., GANTÈS, H. et LABOURG, P. (1980). Contribution à l'étude de la faune marine de la région d'Arcachon, IX. *Bulletin du Centre d'Etude et de Recherche Scientifique de Biarritz*, **13** : 45–64. 21, 27, 103
- BACHELET, G., SIMON-BOUHET, B., DESCLAUX, C., GARCIA-MEUNIER, P., MAIRESSE, G., DE MONTAUDOIN, X., RAIGNÉ, H., RANDRIAMBAO, K., SAURIAU, P.-G. et VIARD, F. (2004). Invasion of the eastern Bay of Biscay by the nassariid gastropod *Cyclope neritea* : origin and effects on the resident fauna. *Marine Ecology Progress Series*, **276** : 147–159. 20, 103, 105, 132, 139, 145, 155

- BALLOUX, F., BRÜNNER, H., LUGON-MOULIN, N., HAUSSE, J. et GOUDET, J. (2000). Microsatellites can be misleading : an empirical and simulation study. *Evolution*, **54**(4) : 1414–1422. 59
- BANDELT, H., FORSTER, P. et RÖHL, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology & Evolution*, **16** : 37–48. 64
- BARRY, J., BAXTER, C., SAGARIN, R. et GILMAN, S. (1995). Climate-related, long-term faunal changes in a California rocky intertidal community. *Science*, **267**(5198) : 672–675. 13
- BELKHIR, K., BORSA, P., CHIKHI, L., RAUFASTE, N. et BONHOMME, F. (1996-2004). GENETIX 4.05, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France). 114
- BERROU, V., MERLE, D., DOMMERGUES, J.-L., CRÔNIER, C. et NÉRAUDEAU, D. (2004). Comparative morphology of pliocene, quaternary and recent shells of *Ocenebra erinaceus* (Linnaeus 1758) and *O. brevirobusta* Houart, 2000 (Mollusca, Muricidae, Ocenebrinae) : reflections of the intra- and interspecific variations. *Geodiversitas*, **26**(2) : 263–295. 39
- BERTNESS, M. et LEONARD, G. (1997). The rôle of positive interactions in communities : lessons from intertidal habitats. *Ecology*, **78**(7) : 1976–1989. 8
- BILLOTE, N., LAGODA, P., RISTERUCCI, A. et BAURENS, F. (1999). Microsatellite-enriched libraries : applied methodology for the tech development of SSR markers in tropical crops. *Fruits*, **54** : 277–288. 195
- BIRKS, H. (1989). Holocene isochrone maps and pattern in tree spreading in the British Isles. *Journal of Biogeography*, **16** : 503–540. 5, 6
- BORJA, A. (1987). Catalogo de los moluscos marinos de la Costa Vasca. *Iberus*, **7** : 211–223. 30, 37, 39, 40
- BOULHIC, M. et TARDY, J. (1986a). *Cyclope neritea* (Linné 1758) : gastéropode prosobanche nouvellement implaté sur les côtes de Charente Maritime ; éthologie, structure de la population. *Halictis*, **15** : 41–49. 20
- BOULHIC, M. et TARDY, J. (1986b). Ponte, développement larvaire et éthologie des juvéniles de *Cyclope neritea* (Linné, 1758). *Halictis*, **15** : 51–58. 20, 29, 30
- BOWEN, B., MEYLAN, A., ROSS, J., LIMPUS, C., BALAZS, G. et AVISE, J. (1992). Global population structure and natural history of the green turtle (*Chelonia mydas*) in terms of matriarchal phylogeny. *Evolution*, **46**(4) : 865–881. 7
- BROWN, J., STEVENS, G. et KAUFMAN, D. (1996). The geographic range : size, shape, boundaries and internal structure. *Annual Review of Ecology & Systematics*, **27** : 597–623. 4, 27
- BRYAN, G., GIBBS, P., HUGGETT, R., CURTIS, L. et BAILEY, D. (1989). Effects of tributyltin pollution on the mud snail, *Ilyanassa obsoleta*, from the York River and Sarah Creek, Chesapeake Bay. *Marine Pollution Bulletin*, **20**(9) : 458–462. 14

-
- BRYAN, G., GIBBS, P., HUMMERSTONE, L. et BURT, G. (1986). The decline of the gastropod *Nucella lapillus* around South-West England : evidence for the effect of tributyltin from antifouling paints. *Journal of the Marine Biological Association of the United Kingdom*, **63**(3) : 611–640. 14
- BUROKER, N. (1983). Population genetics of the American oyster *Crassostrea virginica* along the Atlantic coast and the Gulf of Mexico. *Marine Biology*, **75**(1) : 99–112. 5
- CAEIRO, S., PAINHO, M., COSTA, M. et RAMOS, T. (2002). Sado estuary ecosystem : a management methodology. In : *Proceedings of International Conference on Sustainable Management of Coastal Ecosystems*. Fernando Pessoa University, Porto, Portugal. 40
- CARLTON, J. (1987). Patterns of transoceanic marine biological invasions in the Pacific Ocean. *Bulletin of Marine Science*, **41**(2) : 452–465. 15
- CARLTON, J. (1996a). Biological invasions and cryptogenic species. *Ecology*, **77**(6) : 1653–1655. 17, 22, 144
- CARLTON, J. (1996b). Pattern, process, and prediction in marine invasion ecology. *Biological Conservation*, **78** : 97–106. 15
- CARLTON, J. (1999). Molluscan invasions in marine and estuarine communities. *Malacologia*, **41**(2) : 439–454. 18
- CARLTON, J. et GELLER, J. (1993). Ecological roulette : the global transport of non-indigenous marine organisms. *Science*, **261** : 78–82. 16
- CARR, M. et HIXON, M. (1995). Predation effects on early post-settlement survivorship of coral-reef fishes. *Marine Ecology Progress Series*, **124** : 31–42. 9
- CASSENS, I., VAN WAEREBEEK, K., BEST, P. B., CRESPO, E. A., REYES, J. et MILINKOVITCH, M. C. (2003). The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*) : a critical examination of network methods and rooting procedures. *Molecular Ecology*, **12**(7) : 1781–92. 64
- CHAPIN, F., ZAVALA, E., EVINER, V., NAYLOR, R., VITOUSEK, P., REYNOLDS, H., HOOPER, D., LAVOREL, S., SALA, O., HOBIE, S., MACK, M. et DÍAZ, S. (2000). Consequences of changing biodiversity. *Nature*, **405** : 234–242. 152
- CLARKE, C. (1971). Liberations and dispersal of red deer in northern South Island districts. *New Zealand Journal of Forestry Science*, **1** : 194–207. 9
- CLEMAM (2006). *Check List of European Marine Mollusca*. <http://www.somali.asso.fr/clemam/>. 43, 44
- COBLENTZ, B. E. (1990). Exotic organisms : a dilemma for conservation. *Conservation Biology*, **4**(3) : 261–265. 16
- COHEN, A. N. et CARLTON, J. T. (1998). Accelerating invasion rate in a highly invaded estuary. *Science*, **279** : 555–558. 15, 16

- COHEN, J., SMALL, C., MELLINGER, A., GALLUP, J., SACHS, J., VITOUSEK, P. et MOONEY, H. (1997). Estimates of coastal populations. *Science*, **278**(5341) : 1211–1212. 14
- COLLARD DES CHERRES (1830). Catalogue des Testacés marins du département du Finistère, principalement des côtes de Brest. *Actes de la Société Linnéenne de Bordeaux*, **4** : 3–58. 144
- COLLARES-PEREIRA, M. J., COWX, I., RIBEIRO, F., RODRIGUES, J. et ROGADO, L. (2000). Threats imposed by water resource development schemes on the conservation of endangered fish species in the Guadiana River Basin in Portugal. *Fisheries Management & Ecology*, **7**(1-2) : 167–178. 40
- CONNELL, J. (1975). Some mechanisms producing structure in natural communities : a model and evidences from field experiments. In : *Ecology and evolution of communities*, pages 460–490. Harvard University Press, Cambridge. 4, 9
- COULON, A., GUILLOT, G., COSSON, J.-F., ANGIBAUT, J., AULAGNIER, S., CARGNELUTTI, B., GALAN, M. et HEWISON, A. (2006). Genetic structure is influenced by landscape features : empirical evidence from a roe deer population. *Molecular Ecology*, **15**(6) : 1669. 69
- COX, B. et MOORE, P. (2000). *Biogeography : an ecological and evolutionary approach*. Blackwell Scientific Publications, London. 151
- CROZIER, R. (1990). From population genetics to phylogeny : uses and limits of mitochondrial DNA. *Australian Systematic Botany*, **3** : 111–124. 57
- D'ANTONIO, C. et KARK, S. (2002). Impacts and extent of biotic invasions in terrestrial ecosystems. *Trends in Ecology & Evolution*, **17**(5) : 202–204. 15
- DAUTZENBERG, P. et FISCHER, P. (1925). Les Mollusques marins du Finistère et en particulier de la région de Roscoff. *Travaux de la Station Biologique de Roscoff*, **3** : 1–180. 144
- DAVIES, N., VILLABLANCA, F. et RODERICK, G. (1999). Bioinvasions of the medfly *Ceratitis capitata* : source estimation using DNA sequences at multiple intron loci. *Genetics*, **153** : 351–360. 15
- DAVIS, A., JENKINSON, L., LAWTON, J., SHORROCKS, B. et WOOD, S. (1998). Making mistakes when predicting shifts in species range in response to global warming. *Nature*, **391** : 783–786. 7
- DE BEAUCHAMP, P. (1948). La faune de la zone des marées sur la côte basque. *Compte Rendu Sommaire des Séances. Société de Biogéographie*, **25** : 10–14. 37
- DEITH, M. et SHACKLETON, J. (1988). *Oxygen isotope analyses of marine molluscs from the Franchtchi Cave*. Indiana University Press, Bloomington. 55
- DEMIR, M. (2003). Shells of mollusca collected from the seas of Turkey. *Turkish Journal of Zoology*, **27** : 101–140. 44
- DEPAULIS, F. et VEUILLE, M. (1998). Neutrality tests based on the distribution of haplotypes under an infinite-site model. *Molecular Biology & Evolution*, **15**(12) : 1788–90. 109, 123, 160

-
- DIAZ-ALMELA, E., BOUDRY, P., LAUNEY, S., BONHOMME, F. et LAPEGUE, S. (2004). Reduced female gene flow in the European flat oyster *Ostrea edulis*. *Journal of Heredity*, **95**(6) : 510–516. 59
- DOBZHANSKY, T. (1950). Evolution in the tropics. *American Scientist*, **38** : 209–211. 4
- DONNELLY, P. et TAVARÉ, S. (1995). Coalescents and the genealogical structure under neutrality. *Annual Reviews of Genetics*, **29** : 401–421. 159
- DUFRESNE, F., BOURGET, E. et BERNATCHEZ, L. (2002). Differential patterns of spatial divergence in microsatellite and allozyme alleles : further evidence for locus-specific selection in the acorn barnacle, *Semibalanus balanoides* ? *Molecular Ecology*, **11** : 113–123. 93
- DUKES, J. et MOONEY, H. (1999). Does global change increase the success of biological invaders ? *Trends in Ecology & Evolution*, **14**(4) : 135–139. 144
- DUPONT, L. (2004). *Invasion des côtes françaises par le mollusque exotique Crepidula fornicata : contribution de la dispersion larvaire et du système de reproduction au succès de la colonisation*. These, Université Paris VI. xix, 66, 93
- DUPONT, L., JOLLIVET, D. et VIARD, F. (2003). High genetic diversity and ephemeral drift effects in a successful introduced mollusc (*Crepidula fornicata* : Gastropoda). *Marine Ecology Progress Series*, **235** : 183–195. 104, 139
- EADIE, J. M., BROEKHOVEN, L. et COLGAN, P. (1987). Size ratios and artifacts : Hutchinson's rule revisited. *American Naturalist*, **129**(1) : 1–17. 4
- ESTOUP, A. et ANGERS, B. (1998). Microsatellites and minisatellites for molecular ecology : theoretical and empirical considerations. In : *Advances in molecular ecology*, pages 55–86. NATO press, Amsterdam. 58
- ESTOUP, A., JARNE, P. et CORNUET, J.-M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**(9) : 1591–1604. 59
- ESTOUP, A., LARGIADER, C., PERROT, E. et CHOURROUT, D. (1996). Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology & Biotechnology*, **5**(4) : 295–298. 184
- EVANNO, G., REGNAUT, S. et GOUDET, J. (2005). Detecting the number of clusters of individuals using the software structure : a simulation study. *Molecular Ecology*, **14** : 2611–2620. 70, 87, 88
- EVERETT, R. (2000). Patterns and pathway of biological invasions. *Trends in Ecology & Evolution*, **15**(5) : 177–178. 16
- EXCOFFIER, L., SMOUSE, P. et QUATTRO, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes : application to human mitochondrial DNA restriction data. *Genetics*, **131** : 479–491. 67

- FACON, B., POINTIER, J.-P., GLAUBRECHT, M., POUX, C., JARNE, P. et DAVID, P. (2003). A molecular phylogeography approach to biological invasions of the New World by parthenogenetic Thiarid snails. *Molecular Ecology*, **12**(11) : 3027–3039. 139
- FALUSH, D., STEPHENS, M. et PRITCHARD, J. (2003). Inference of population structure using multilocus genotype data : linked loci and correlated allele frequencies. *Genetics*, **164** : 1567–1587. 69, 72
- FELSENSTEIN, J. (1992). Estimating effective population size from samples of sequences : inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. *Genetical Research*, **56** : 139–147. 160
- FIGUERAS, A. (1956). Moluscos de las playas de Vigo. I. Ecología y distribución. *Investigación pesquera*, **5** : 51–88. 38
- FISCHER, P. (1899). Liste des mollusques recueillis à Guétary et à Saint-Jean-de-Luz. *Bulletin Scientifique de la France et de la Belgique*, **32** : 218–229. 30, 35, 37, 38, 40
- FOLMER, O., BLACK, M., HOEH, W., LUTZ, R. et VRIJENHOEK, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology & Biotechnology*, **3**(5) : 294–299. 47, 48, 188, 189
- FORSYTH, D. et DUNCAN, R. (2001). Propagule size and the relative success of exotic ungulate and bird introductions to New Zealand. *American Naturalist*, **157** : 583–595. 9
- FU, Y. et ARNOLD, J. (1991). On the association of restriction fragment length polymorphisms across species boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, **88** : 3967–3971. 59
- FU, Y. et ARNOLD, J. (1992). Dynamics of cytonuclear disequilibria in finite populations and comparison with a two-locus nuclear system. *Theoretical Population Biology*, **41** : 1–25. 59
- FUTUYAMA, D. (1986). *Evolutionary Biology*. Sunderland, Sinauer Associates Inc. 16
- GALTIER, N., DEPAULIS, F. et BARTON, N. (2000). Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics*, **155** : 981–987. 109, 160
- GALTISOFF, P. (1930). The fecundity of the oyster. *Science*, **72**(1856) : 97–98. 5
- GASTON, K. (1990). Patterns in geographical ranges of species. *Biological Reviews*, **65** : 105–129. 5
- GASTON, K. (1994). Measuring geographic range sizes. *Ecography*, **17** : 198–205. 39
- GASTON, K. (1996). Species-range-size distributions : patterns, mechanisms and implications. *Trends in Ecology & Evolution*, **11**(5) : 197–201. 5
- GASTON, K., BLACKBURN, T. et SPICER, J. (1998). Rapoport's rule : time for an epitaph? *Trends in Ecology & Evolution*, **13**(2) : 70–74. 4
- GASTON, K. et CHOWN, S. (1999). Why Rapoport's rule does not generalise. *Oikos*, **84**(2) : 309–313. 4

-
- GAYLORD, B. et GAINES, S. D. (2000). Temperature or Transport? Range limits in marine species mediated solely by flow. *American Naturalist*, **155**(6) : 769–789. 8
- GELLER, J., CARLTON, J. et POWERS, D. (1994). PCR-based detection of mtDNA haplotypes of native and invading mussels on the northeastern Pacific coast : latitudinal pattern of invasion. *Marine Biology*, **119** : 243–249. 188, 189
- GOMOIU, M. (1964). Biologisches studium der arten *Nassa reticulata* (L.) und *Cyclonassa neritea* (L.) im schwarzen Meer (rumänischer Küstenbereich). *Revue Roumaine de Biologie, Série de Zoologie*, **9** : 39–49. 20, 103
- GOODING, M., GALLARDO, C. et LEBLANC, G. (1999). Imposex in three marine gastropod species in Chile and potential impact on muriciculture. *Marine Pollution Bulletin*, **38**(12) : 1227–1231. 14
- GOPURENKO, D., HUGHES, J. M. et KEENAN, C. P. (1999). Mitochondrial DNA evidence for rapid colonisation of the Indo-West Pacific by the mudcrab *Scylla serrata*. *Marine Biology*, **134**(2) : 227–233. 93
- GOUDET, J. (1995). Fstat version 1.2 : a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**(6) : 485–486. 63, 108
- GOUDET, J. (2005). Hierfstat, a package for R to compute and test variance components and *F*-statistics. *Molecular Ecology Notes*, **5** : 184–186. 65
- GOUDET, J., RAYMOND, M., T., D. et ROUSSET, F. (1996). Testing differentiation in diploid populations. *Genetics*, **144**(4) : 1933–1940. 65
- GOULLETQUER, P., BACHELET, G., SAURIAU, P.-G. et NOEL, P. (2002). Open Atlantic coast of Europe - a century of introduced species into French waters. In : *Invasive aquatic species of Europe. Distribution, impacts and management*, pages 276–290. Kluwer Academic Publishers, Dordrecht. 16, 139
- GOULLETQUER, P. et HERAL, M. (1997). Marine molluscan production trends in France : from fisheries to aquaculture. In : C. Mackenzie Jr., V. Burrell Jr., A. Rosenfield et W. Hobart, éditeurs, *The history, present condition and future of the molluscan fisheries of North and Central America and Europe Volume 3, Europe*, volume 129, pages 137–164. NOAA technical report NMFS, Seattle, WA (USA). 143
- GROSHOLZ, E. (2002). Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology & Evolution*, **17**(1) : 22–27. 16, 104
- GUILLOT, G., ESTOUP, A., MORTIER, A. et COSSON, J. (2005a). A spatial statistical model for landscape genetics. *Genetics*, **170** : 1261–1280. 70, 72
- GUILLOT, G., MORTIER, F. et ESTOUP, F. (2005b). Geneland : a program for landscape genetics. *Molecular Ecology Notes*, **5** : 712–715. 69, 72
- GUINAND, B. (1996). Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. *Biological Journal of the Linnean Society*, **58**(2) : 173–195. 114

- HALL, T. (1999). BioEdit : a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41** : 95–98. 194
- HAMBLIN, M. et VEUILLE, M. (1999). Population structure among African and derived populations of *Drosophila simulans* : Evidence for ancient subdivision and recent admixture. *Genetics*, **153** : 305–317. 109
- HARLEY, C., HUGHES, A., HULTGREN, K., MINER, B., SORTE, C., THORNBUR, C., RODRIGUEZ, L., TOMANEK, L. et WILLIAMS, S. (2006). The impacts of climate change in coastal systems. *Ecology Letters*, **9** : 228–241. 151
- HAWKINS, B. et LAWTON, J. (1995). Latitudinal gradients in butterfly body sizes : is there a general pattern? *Oecologia*, **102**(1) : 31–36. 4
- HAYS, G., RICHARDSON, A. et ROBINSON, C. (2005). Climate change and marine plankton. *Trends in Ecology & Evolution*, **20**(6) : 337–344. 13
- HEINER, C., HUNKAPILLER, K., CHEN, S., GLASS, J. et CHEN, E. (1998). Sequencing multimegabase-template DNA with BigDye terminator chemistry. *Genome Research*, **8**(5) : 557–561. 192
- HELLBERG, M. E., BALCH, D. P. et ROY, K. (2001). Climate-driven range expansion and morphological evolution in a marine gastropod. *Science*, **292**(5522) : 1707–1710. 104
- HERNANDEZ OTERO, J. et JIMENEZ MILLAN, F. (1972). Distribución de los moluscos : gasterópodos y pelecipodos, marinos, de las costas de Galicia. *Cuad. Biol.*, **1** : 79–93. 38
- HEWITT, G. (2001). Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**(3) : 537–549. 95
- HIDALGO, J. (1917). Fauna Malacológica de España, Portugal y las Baleares. Moluscos testáceo marinos. *Trab. Mus. nac. Cienc. nat., Madrid (Ser. zool.)*, **30** : 1–752. 30, 37, 38, 39, 40
- HINARD, G. et LAMBERT, L. (1928). Tableau de l'ostréiculture française (deuxième partie). *Revue des Travaux de l'Office des Pêches Maritimes*, **1**(4) : 61–127. 144
- HOLLAND, B. S. (2001). Invasion without a bottleneck : Microsatellite variation in natural and invasive populations of the brown mussel *Perna perna* (L). *Marine Biotechnology*, **3**(5) : 407–415. 139, 154
- HOLT, R. (2003). On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, **5** : 159–178. 5, 29
- HUDSON, R. (1993). The how and why of generating gene genealogies. *In : Mechanisms of molecular evolution*, pages 23–26. Sinauer, Sunderland, MA. 160
- HUGHES, L. (2000). Biological consequences of global warming : is the signal already apparent? *Trends in Ecology & Evolution*, **15**(2) : 56–61. 13
- HURLBERT, S. (1971). The nonconcept of species diversity : a critique and alternative parameters. *Ecology*, **52** : 577–586. 63, 107

-
- HURRELL, J. (1995). Decadal trends in the North Atlantic Oscillation : regional temperatures and precipitation. *Science*, **269**(5224) : 376–679. 12, 13
- HUVET, A., LAPÈGUE, S., MAGOULAS, A. et BOUDRY, P. (2000). Mitochondrial and nuclear DNA phylogeography of *Crassostrea angulata*, the Portuguese oyster endangered in Europe. *Conservation Genetics*, **1** : 251–262. 93
- IPCC (2001). *Climate change 2001 : the scientific basis*. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,. 11, 12, 22, 30, 103
- JABLONSKI, D. (1997). Body-size evolution in Cretaceous molluscs and the status of Cope's Rule. *Nature*, **385** : 250–252. 4
- JARNES, P. et LAGODA, P. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, **11**(10) : 424–429. 56, 58
- JEWETT, S., DEAN, T., SMITH, R. et BLANCHARD, A. (1999). Exxon Valdez oil spill : impacts and recovery in the soft-bottom benthic community in and adjacent to eelgrass beds. *Marine Ecology Progress Series*, **185** : 59–83. 14
- KEANE, R. M. et CRAWLEY, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, **17**(4) : 164–169. 15
- KIJAS, J., FOWLER, J., GARBETT, C. et THOMAS, M. (1994). Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Bio Techniques*, **16** : 657–662. 195
- KIMURA, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16** : 111–120. 64, 108
- KINLAN, B. et HASTINGS, A. (2005). Rates of population spread and geographic range expansion : what exotic species tell us. In : *Species invasions - Insights into ecology, evolution and biogeography*, pages 381–419. Sax, D.F. & Stachowicz, J.J. & Gaines, S.D. 153
- KIRKENDALE, L. A. et MEYER, C. P. (2004). Phylogeography of the *Patelloida profunda* group (Gastropoda : Lottidae) : diversification in a dispersal-driven marine system. *Molecular Ecology*, **13**(9) : 2749–2762. 93
- KIRKPATRICK, M. et BARTON, N. (1997). Evolution of a species' range. *American Naturalist*, **150**(1) : 1–23. 10, 29
- KISCH, B. S. (1950). Growth stages of *Cyclope neritea* (L.). *Journal of Conchology*, **23**(4) : 108. 21, 37, 40, 103
- KOCHER, T., THOMAS, W., MEYER, A., EDWARDS, S., PAABO, S., VILLABLANCA, F. et WILSON, A. (1989). Dynamics of mitochondrial DNA evolution in animals : amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America*, **86** : 6196–62000. 188, 189

- KOLBE, J. J., GLOR, R. E., RODRIGUEZ SCHETTINO, L., LARA, A. C., LARSON, A. et LOSOS, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431** : 177–81. 57, 104, 139, 154
- KUHN, S. et STINER, M. (1998). The earliest aurignacian of Riparo Mochi (Liguria, Italy). *Current Anthropology*, **39**(special issue : The Neanderthal Problem and the Evolution of Human Behavior) : S175–S189. 55
- KUMAR, S. et KURTZ, S. (1993). Properties of a two-dimensional Poisson-Voronoi tessellation : a Monte Carlo study. *Materials characterization*, **31**(1) : 55–68. 72
- KUMAR, S., TAMURA, K. et NEI, M. (2004). MEGA3 : integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5** : 150–163. 64, 108
- LAGARDÈRE, J.-P. (1966). Recherches sur la biologie et l'écologie de la macrofaune des substrats meubles de la côte des Landes et de la côte Basque. *Bulletin du Centre d'Études et de Recherches scientifiques de Biarritz*, **6** : 143–209. 37, 39, 40
- LAUNEY, S., LEDU, C., BOUDRY, P., BONHOMME, F. et NACIRI-GRAVEN, Y. (2002). Geographic structure in the european flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. *Journal of Heredity*, **93**(5) : 331–351. 93
- LE ROUX, A. (1994). Sexualité et reproduction de *Cyclope neritea* (mollusque gastéropode) dans le golfe du Morbihan (Bretagne-Sud). *Haliotis*, **23** : 51–59. 20
- LE ROUX, A., MAHEO, R. et DAGUSAN, J. (1988). Présence et particularités écologiques de *Cyclope neritea* (Linné 1758) (gastéropode prosobranch) dans le Golfe du Morbihan. *Haliotis*, **18** : 310. 21, 103
- LEVINSON, G. et GUTMAN, G. (1987). Slipped and mispairing : a major mechanism for DNA sequence evolution. *Molecular Biology & Evolution*, **4** : 203–221. 58
- LINNÉ, C. v. (1758). *Systema Naturae. Editio decima. 1. Regnum Animale Holmiae, Laurentii Salvii iv.* 43
- LOMOLINO, M., RIDDLE, B. et BROWN, J. (2005). *Biogeography*. Sinauer Associates, Inc, Sunderland, Massachusetts, third edition. 3, 8, 9
- LOPEZ-COTELO, I., VIEITEZ, J. et DIAZ-PINEDA, F. (1982). Tipos de comunidades bentónicas de la playa del Puntal (Bahía de Santander). *Cahiers de Biologie Marine*, **23** : 53–69. 37, 40
- MAC ARTHUR, R. (1972). *Geographical ecology : patterns in the distribution of species*. Harper & Row, New York. 4
- MARKO, P. (2002). Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology & Evolution*, **19** : 2005–2021. 94
- MARKO, P. (2004). What's larvae got to do with it ? Disparate patterns of post-glacial population structure in two benthic marine gastropods with identical dispersal potential. *Molecular Ecology*, **13**(3) : 597–611. 95

-
- MARRINER, A., MORHANGE, C., BOUDAGHER-FADEL, M., BOURCIER, M. et CARBONEL, P. (2005). Geoarchaeology of Tyre's ancient northern harbour, Phoenicia. *Journal of Archaeological Science*, **32** : 1302–1327. 56
- MARTEL, C., VIARD, F., BOURGUET, D. et GARCIA-MEUNIER, P. (2004). Invasion of the marine gastropod *Ocenebrellus inornatus* on the French Atlantic Coast. I. Scenario for the source of introduction. *Journal of Experimental Biology & Ecology*, **305** : 155–170. 139
- MIESZKOWSKA, N., KENDALL, M., HAWKINS, S., LEAPER, R., WILLIAMSON, P., HARDMAN-MOUNTFORD, N. et SOUTHWARD, A. (2006). Changes in the range of some common rocky shore species in Britain - a response to climate change? *Hydrobiologia*, **555** : 241–251. 13
- MOLINARI, R., OLSON, D. et REVERDIN, G. (1990). Surface current distributions in the tropical Indian Ocean derived from compilations of surface buoy trajectories. *Journal of Geophysical Research*, **95**(C5) : 7217–7238. 8
- MOONEY, H. et CLELAND, E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(10) : 5446–5451. 16
- MORTON, J. (1960). The habits of *Cyclope neritea*, a style-bearing stenoglossan gastropod. *Proceedings of the Malacological Society of London*, **34** : 96–105. 37, 40
- MULLER, M.-H. (2002). *Flux de gènes et domestication : l'histoire évolutive d'un complexe sauvage-cultivé. Approches expérimentale et théorique chez la luzerne pérenne (Medicago sativa L.)*. These, Université de Montpellier 2. xv, 221
- MULLIS, K., FALOONA, F., SCHARF, S., SAIKI, R., HORN, G. et ERLICH, H. (1986). Specific enzymatic amplification of DNA in vitro : the polymerase chain reaction. *In : Cold Spring Harbor Symposia on Quantitative Biology*, volume 51, pages 263–273. 185
- MUN, J., BOHONAK, A. et RODERICK, G. (2003). Population structure of pumpkin fruit fly *Bactrocera depressa* (Tephritidae) in Korea and Japan : pliocene allopatry or ancient invasion? *Molecular Ecology*, **12** : 2941–2951. 18
- MYERS, J., SIMBERLOFF, D., KURIS, A. et CAREY, J. (2000). Eradication revisited : dealing with exotic species. *Trends in Ecology & Evolution*, **15**(8) : 316–320. 15
- NEI, M. (1972). Genetic distance between populations. *American Naturalist*, **106** : 283–291. 64
- NEI, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New-York. 62
- NIXON, S. (1988). Physical energy inputs and the comparative ecology of lake and marine ecosystems. *Limnology & Oceanography*, **33**(4, part 2) : 1005–1025. 7, 8
- NIXON, S. (1995). Coastal marine eutrophication : a definition, social causes and future concerns. *Ophelia*, **41** : 199–219. 14
- NOBRE, A. (1932). *Moluscos marinhos de Portugal*. Instituto de Zoologica de Universidade do Pôrto, Imprensa Portuguesa, Pôrto. 35, 38, 40
- NOVAK, S., MACK, R. et SOLTIS, P. (1993). Genetic variation in *Bromus tectorum* (Poaceae) : introduction dynamics in North America. *Canadian Journal of Botany*, **71** : 1441–1448. 139

- O'FOIGHIL, D., GAFFNER, P., WILBUR, A. et HILBISH, T. (1998). Mitochondrial cytochrome oxydase I gene sequences support an Asian origin for the Portuguese oyster *Crassostrea angulata*. *Marine Biology*, **131** : 497–503. 188, 189
- PALUMBI, S. (1994). Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology & Systematics*, **25** : 547–572. 5, 7
- PALUMBI, S. (2001). Humans as the world's greatest evolutionary force. *Science*, **293**(5536) : 1786–1790. 10, 16
- PARDINI, A., JONES, C., NOBLE, L., KREISER, B., MALCOLM, H., BRUCE, B., STEVENS, J., CLIFF, G., SCHOLL, M., FRANCIS, M., DUFFY, C. et MARTIN, A. (2001). Sex-biased dispersal of great white sharks. *Nature*, **412** : 139–140. 7
- PARMESAN, C. et YOHE, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421** : 37–42. 13, 22
- PENAS, E. et GONZALEZ, G. (1983). Relationships between benthic infauna and environmental factors in three beaches of the Ría de Arosa embayment (Spain using canonical correlation analysis. *Journal of Experimental Biology & Ecology*, **68** : 245–256. 38
- PETIT, R., EL MOUSADIK, A. et PONS, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**(4) : 844–855. 63
- PIELOU, E. (1979). *Biogeography*. John Wiley and Sons, New York. 5
- PIGEOT, J. (1988). Présence du gastéropode nassaridé *Cyclope neritea* (Linné, 1788) dans le bassin de Marennes-Oléron. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**(6) : 717–719. 21, 22, 103
- PIRY, S., ALAPETITE, A., CORNUET, J.-M., PAETKAU, D., BAUDOUIN, L. et ESTOUP, A. (2004). GeneClass2 : a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95** : 536–539. 109, 126
- POR, F. (1971). One hundred years of Suez Canal - A century of Lessepsian Migration : retrospect and viewpoints. *Systematic Zoology*, **20**(2) : 138–159. 15
- POSADA, D. et CRANDALL, K. A. (2001). Intraspecific gene genealogies : trees grafting into networks. *Trends in Ecology & Evolution*, **16**(1) : 37–45. 64
- PRITCHARD, J., STEPHENS, M. et DONNELLY, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155** : 945–959. 69, 70
- PRZEWORSKI, M., HUDSON, R. et DI RIENZO, A. (2000). Adjusting the focus on human variation. *Trends in Genetics*, **16**(7) : 296–302. 159
- R DEVELOPMENT CORE TEAM (2005). *R : a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 65, 67, 108
- RALLO, A. (1981). Aportaciones al conocimiento de la fauna bentónica litoral vizcaína. *Reun. bien. R. Soc. esp. Hist. nat. Oviedo*, **5** : 51–60. 37, 39

-
- RANNALA, B. et MOUNTAIN, J. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, **94** : 9197–9201. 109
- RAYMOND, M. et ROUSSET, F. (1995). GENEPOP (version 1.2) : population genetics software for exact tests and ecumenicism. *The Journal of Heredity*, **86** : 248–249. 63, 65, 114
- REISE, K., GOLLASCH, S. et WOLFF, W. (1999). Introduced marine species of the North Sea coasts. *Helgoländer Meeresuntersuchungen*, **52** : 219–234. 15, 16
- REVERDIN, G., FRANKIGNOUL, C., KESTENARE, E. et MCPHADEN, M. (1994). Seasonal variability in the surface currents of the equatorial Pacific. *Journal of Geophysical Research*, **99**(C10) : 20 323–20 344. 8
- RICCIARDI, A. et MACISAAC, H. (2000). Recent mass invasion of the north american great lakes by ponto-caspian species. *Trends in Ecology & Evolution*, **15**(2) : 62–65. 15
- RISSO, A. (1826). *Histoire naturelle des principales productions de l’Europe Méridionale et particulièrement de celles des environs de Nice et des Alpes Maritimes*. Levrault, Paris. 44
- RODHOUSE, P., McDONALD, J., NEWELL, R. et KOEHN, R. (1986). Gamete production, somatic growth and multiple-locus enzyme heterozygosity in *Mytilus edulis*. *Marine Biology*, **90**(2) : 209–214. 5
- RODRIGUES, A. et QUINTINO, V. (1991). Malacofauna da Ría de Alvor (costa sul de Portugal). Estudo faunístico e cenótico. *Iberus*, **10**. 35, 38
- ROHDE, K., HEAP, M. et HEAP, D. (1993). Rapoport’s rule does not apply to marine teleosts and cannot explain latitudinal gradients in species richness. *American Naturalist*, **142**(1) : 1–16. 4
- ROLÁN, E. (1984). *Moluscos de la Ría de Vigo. I. Gasterópodos*. Collego Universidade de Vigo, Vellgraf S.A., Santiago de Compostela. 38
- ROLÁN, E. (1992). Dos especies más de moluscos mediterráneos introducidas en la bahía de O Grove (Galicia, Spain). *Thalassas*, **10** : 135. 39, 56, 141
- ROOT, T., PRICE, J., HALL, K., SCHNEIDER, S., ROSENZWEIG, C. et POUNDS, J. (2003). Fingerprints of global warming on wild animals and plants. *Nature*, **421** : 57–60. 13
- ROUSSET, F. (1997). Genetic differentiation and estimation of gene flow from F -statistics under isolation by distance. *Genetics*, **145** : 1219–1228. 69, 108
- ROY, K., JABLONSKI, D. et VALENTINE, W. (2001). Climate change, species range limits and body size in marine bivalves. *Ecology Letters*, **4** : 366–370. 6
- ROZAS, J. et ROZAS, R. (1999). DnaSP version 3 : an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15** : 174–175. 62
- ROZEN, S. et SKALETSKY, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In : *Bioinformatics methods and protocols*, Methods in Molecular Ecology, pages 365–386. Humana Press, Totowa. 47, 197

- RUIZ, G. M., CARLTON, J., GROSHOLZ, E. D. et HINES, A. H. (1997). Global invasions of marine and estuarine habitats by non-indigenous species : mechanisms, extent, and consequences. *American Zoologist*, **37** : 621–632. 15, 16
- RUSSELL, M. et LINDBERG, D. (1988). Real and random patterns associated with molluscan spatial and temporal distributions. *Paleobiology*, **14**(4) : 322–330. 39
- SAITOU, N. et NEI, M. (1987). The neighbor-joining method : a new method of reconstructing phylogenetic trees. *Molecular Biology & Evolution*, **4** : 406–425. 64, 108
- SAKAI, A. K., ALLENDORF, F. W., HOLT, J. S., LODGE, D. M., MOLOFSKY, J., WITH, K. A., BAUGHMAN, S., CABIN, R. J., COHEN, J. E., ELLSTRAND, N. C., MCCAULEY, D. E., O'NEIL, P., PARKER, I. M., THOMPSON, J. N. et WELLER, S. G. (2001). The population biology of invasive species. *Annual Review of Ecology & Systematics*, **32** : 305–332. 15, 23
- SAURIAU, P.-G. (1989). *Cyclope neritea* (Linné, 1758) dans le bassin de Marennes-Oléron, bilan des observations de 1984 à 1986. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**(8) : 995–1007. 21, 22, 103
- SAURIAU, P.-G. (1991). Spread of *Cyclope neritea* (Mollusca : Gastropoda) along the north-eastern Atlantic coasts in relation to oyster culture and to climatic fluctuations. *Marine Biology*, **109** : 299–309. 21, 27, 30, 37, 38, 39, 139, 140
- SAX, D. et BROWN, J. (2000). The paradox of invasion. *Global Ecology & Biogeography*, **9**(5) : 363–371. 139, 140, 144
- SAX, D., STACHOWICZ, J. et GAINES, S. (2005). *Species invasions - Insights into ecology, evolution and biogeography*. Sax, D.F., Stachowicz, J.J. & Gaines, S.D. 154
- SCHLÖTTERER, C. et TAUTZ, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, **20** : 211–215. 58
- SCHNEIDER, S., ROESSLI, D. et EXCOFFIER, L. (2001). Arlequin : a software for population data analysis (Version 2.000). 67
- SCOTLAND, R. (1992). Cladistic theory. In : *Cladistics : a practical course in systematics*, pages 3–13. Clarendon Press, Oxford. 59
- SHACKLETON, J. (1988). *Marine molluscan remains from Franchtchi Cave*. Indiana University Press, Bloomington. 55
- SHEA, K. et CHESSON, P. (2002). Community ecology theory as a framework for biological invasions. *Trends in Ecology & Evolution*, **17**(4) : 170–176. 16
- SHIGESADA, N. et KAWASAKI, K. (1997). *Biological Invasions : Theory and Practice*. Oxford University Press., Oxford. 15
- SIMBERLOFF, D. et BOECKLEN, W. (1981). Santa Rosalina reconsidered : size ratios and competition. *Evolution*, **35**(6) : 1206–1228. 4
- SIMON-BOUHET, B., DAGUIN, C., GARCIA-MEUNIER, P. et VIARD, F. (2005). Polymorphic microsatellites for the study of newly established populations of the gastropod *Cyclope neritea*. *Molecular Ecology Notes*, **5**(1) : 121–123. 195, 199

-
- SIMON-BOUHET, B., GARCIA-MEUNIER, P. et VIARD, F. (2006). Mitochondrial gene genealogies provide evidence for multiple geographical origins for *Cyclope neritea* (Nassaridae) populations recently recorded in France. *Molecular Ecology*, **15**(6) : 1699–1711. 21, 105
- SMITH, K. et BROWN, J. (2002). Patterns of diversity, depth range and body size among pelagic fishes along a gradient of depth. *Global Ecology & Biogeography*, **11**(4) : 313–322. 4
- SOKOLOV, E., SOKOLOVA, I. et PÖRTNER, H.-O. (2002). Polymorphic microsatellite DNA markers from the marine gastropod *Littorina saxatilis*. *Molecular Ecology Notes*, **2**(1) : 27–29. 93
- STACHOWICZ, J., TERWIN, J., WHITLATCH, R. et OSMAN, R. (2002). Linking climate change and biological invasions : ocean warming facilitates nonindigenous species invasions. *Proceedings of the National Academy of Sciences of the United States of America*, **99** : 15 497–15 500. 144, 152
- STEVENS, G. (1992). The elevational gradient in altitudinal range : an extension of rapoport's latitudinal rule to altitude. *American Naturalist*, **140**(6) : 893–911. 4
- STRATHMANN, R. et STRATHMANN, M. (1982). The relationship between adult size and brooding in marine invertebrates. *American Naturalist*, **119**(1) : 91–101. 5
- TABERLET, P., FUMAGALLI, L., WUST-SAUCY, A.-G. et COSSON, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**(4) : 453–464. 95
- TAJIMA, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105** : 437–460. 62
- TAJIMA, F. (1989). The effect of change in population size on DNA polymorphism. *Genetics*, **123** : 597–601. 111, 160
- TARDY, J., GAILLARD, J. et PORTÈRES, G. (1985). *Cyclope neritea* (Linné, 1788), une espèce de gastéropode nassaridé nouvelle pour les côtes du centre-ouest Atlantique de France. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**(3) : 391–396. 18, 21, 30, 103, 144
- TAYLOR, A., ALLEN, J. et CLARK, P. (2002). Extraction of a weak climatic signal by an ecosystem. *Nature*, **416** : 629–632. 12
- THOMPSON, J. D., HIGGINS, D. G. et GIBSON, T. J. (1994). CLUSTAL W : improving the sensitivity of progressive sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22** : 4673–4680. 194
- TILMAN, D., MAY, R., LEHMAN, C. et NOWAK, M. (1994). Habitat destruction and the extinction debt. *Nature*, **371** : 65–66. 14
- TIMMERMAN, A., OBERHUBER, J., BACHER, A., ESCH, M., LATIF, M. et ROECKNER, E. (1999). Increased El Niño frequency in a climate model forced by future greenhouse warming. *Nature*, **398** : 694–697. 11

- TINDALL, K. et KUNKEL, T. (1988). Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry*, **27**(16) : 6008–6013. 185
- TROUVE, S., DEGEN, L. et GOUDET, J. (2005). Ecological components and evolution of selfing in the freshwater snail *Galba truncatula*. *Journal of Evolutionary Biology*, **18** : 358–370. 65
- TSUTSUI, N. D., SUAREZ, A. V., HOLWAY, D. A. et CASE, T. J. (2000). Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(11) : 5948–5953. 15, 154
- VALIELA, I. (1995). *Marine ecological processes*. Springer, second edition. 7
- VANHAEREN, M. et D'ERRICO, F. (2001). La parure de l'enfant de La Madeleine (fouilles Peyrony). Un nouveau regard sur l'enfance au paléolithique supérieur. *Paleo*, **13** : 201–240. 55
- VERMEIJ, G. (1982). Environmental change and the evolutionary history of the periwinkle *Littorina littorea* in North America. *Evolution*, **36**(3) : 561–580. 18
- VERNESI, C., CRESTANELLO, B., PECCHIOLI, E., TARTARI, D., CARAMELLI, D., HAUFFE, H. et BERTORELLE, G. (2003). The genetic impact of demographic decline and reintroduction in the wild boar (*Sus scrofa*) : a microsatellite analysis. *Molecular Ecology*, **12** : 585–595. 70
- VIEITEZ, J. (1981). Estudio de la comunidades bentónicas de dos playas de las Rías de Pontevedra y Vigo (Galicia, España). *Boln Inst. esp. Oceanogr.*, **6** : 242–258. 38
- VITOUSEK, P., MOONEY, H., LUBCHENCO, J. et MELILLO, J. (1997). Human domination of Earth's ecosystems. *Science*, **277** : 494–499. 10, 14
- VOISIN, M., ENGEL, C. et VIARD, F. (2005). Differential suffling of native genetic diversity across introduced region in a brown alga : aquaculture vs. maritime traffic effects. *Proceedings of the National Academy of Sciences of the United States of America*, **102** : 5432–5437. 57, 104, 139, 140, 154
- WALKER, D. et KENDRICK, G. (1998). Threats to macroalgal diversity : marine habitat destruction and fragmentation, pollution and introduced species. *Botanica Marina*, **41**(1) : 105–112. 14
- WARES, J., GOLDWATER, D., KONG, B. et CUNNINGHAM, C. (2002). Refuting a controversial case of human-mediated marine species introduction. *Ecology Letters*, **5**(4) : 577–584. 18
- WASSON, K., ZABIN, C., BEDINGER, L., DIAZ, M. et PEARSE, J. (2001). Biological invasions of estuaries without international shipping : the importance of intraregional transport. *Biological Conservation*, **102** : 143–153. 16
- WEETMAN, D., HAUSER, L., SHAW, P. et BAYES, M. (2005). Microsatellite markers for the whelk *Buccinum undatum*. *Molecular Ecology Notes*, **5**(2) : 361–362. 93
- WEIR, B. et COCKERHAM, C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**(6) : 1358–1370. 65, 66, 108

-
- WILDING, C., GRAHAME, J. et MILL, P. (2000). Mitochondrial DNA CoI haplotype variation in sibling species of rough periwinkles. *Heredity*, **85**(1) : 62–74. 93
- WILLIAMSON, M. (1996). *Biological invasions*. Chapman & Hall, London. 9, 15, 16, 140
- WILLIG, M., KAUFMAN, D. et STEVENS, R. (2003). Latitudinal gradients of biodiversity : pattern, process, scale and synthesis. *Annual Review of Ecology, Evolution & Systematics*, **34** : 273–309. 3
- WILLIS, E. (1974). Populations and local extinctions of birds on Barro Colorado Island, Panama. *Ecological Monographs*, **44** : 153–169. 9
- WOLFF, W. et REISE, K. (2002). Oyster imports as a vector for introduction of alien species into northern and western coastal waters. In : *Invasive aquatic species of Europe. Distribution, impact and management*. Kluwer Academic Publishers, Dordrecht/Boston/London. 139
- WOLSTENHOLME, D. (1992). Animal mitochondrial DNA : structure and evolution. *International Review of Cytology*, **141** : 173–216. 57, 58
- WRIGHT, S. (1931). Evolution in Mendelian populations. *Genetics*, **16** : 97–159. 57
- WRIGHT, S. (1938). Size of population and breeding structure in relation to evolution. *Science*, **87** : 430–431. 57
- WRIGHT, S. (1951). The genetical structure of populations. *Annals of Eugenics*, **15** : 323–354. 66, 67
- ZEISSET, I. et BEEBEE, T. (2001). Determination of biogeographical range : an application of molecular phylogeography to the European pool frog *Rana lessonae*. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268** : 933–938. 70
- ZOUROS, E., BALL, A., SAAVEDRA, C. et FREEMAN, K. (1994). An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. *Proceedings of the National Academy of Sciences of the United States of America*, **91** : 7463–7467. 57
- ZOUROS, E., FREEMAN, K., BALL, A. et POGSON, G. (1992). Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature*, **359** : 421–414. 57

Annexes

A

Compléments méthodologiques

Préambule

Cette thèse faisant largement appel à la biologie moléculaire, nous avons choisi de consacrer une annexe distincte à la description des techniques de laboratoire mises en œuvre pour la création et l'utilisation des marqueurs moléculaires indispensables à notre étude. Cette annexe répond à un double objectif : (i) entrer dans le détail des protocoles mis en place afin de rappeler au non-spécialiste la démarche suivie, et (ii) éviter les redites qui pourraient alourdir notre propos dans les trois parties principales de ce document. Il sera donc fréquemment fait référence à ces compléments méthodologiques dans le corps du texte, les sections “matériels et méthode” des parties I, II et III se consacrant plutôt à la description des méthodes statistiques mises en œuvre spécifiquement dans chacune de ces parties.

A.1 Méthodes non spécifiques aux marqueurs mis en œuvre

A.1.1 Extraction de l'ADN génomique total

Les individus collectés sur le terrain ont été conservés dans l'éthanol à 95% en attendant les extractions d'ADN. L'extraction et la purification de l'ADN génomique total des individus est la première étape de la grande majorité des protocoles de biologie moléculaire. La méthode consiste à :

1. **prélever quelques mg de tissus** : après avoir cassé les coquilles de *Cyclope neritea*, moins de 10 mg de tissus musculaires ont été prélevés dans le pied de chaque individu.
2. **faire digérer le tissu par une action mécanique et enzymatique** : après découpe du tissu musculaire, les échantillons sont mis à digérer, dans 200 à 500 μL de tampon (dont la nature dépend de la suite du protocole), en présence de 10 μL de protéinase K à 20 mg/mL. La digestion s'effectue à 55 °C avec agitation permanente, pendant une durée minimum de 3 heures.
3. **séparer les acides nucléiques des autres composants cellulaires**. Parmi les nombreuses techniques existantes, nous en avons utilisé deux :
 - (a) la **“pseudo-extraction” au chelex** : avec cette méthode (mise au point par Estoup *et al.*, 1996), l'étape de digestion est réalisée en présence de billes de Chelex à 5% ayant une forte affinité pour la plupart des composés cellulaires à l'exception des acides nucléiques. Après l'incubation à 55 °C, les échantillons sont placés 20 min à 100 °C afin de faire éclater les cellules. Une étape de centrifugation (5 min à 5200 g_n) fait chuter les billes de chelex dans le culot, la récupération du surnageant permettant d'isoler les acides nucléiques. Cette méthode est appelée “pseudo-extraction” car elle ne permet pas d'obtenir des purifications de bonne qualité, de nombreux éléments susceptibles de parasiter les étapes ultérieures des protocoles étant souvent récupérés dans le surnageant. En outre, la conservation à long terme des ADN extraits par cette méthode est délicate. C'est pourquoi, bien que peu coûteuses, seules quelques extractions chelex ont été réalisées en début de thèse, notamment au cours des phases de mises au point de marqueurs mitochondriaux.
 - (b) **l'extraction sur colonne de résine** : ce protocole plus coûteux mais aussi plus performant a été utilisé pour la très grande majorité des échantillons traités. Nous avons utilisé les kits prêts à l'emploi proposés par deux fournisseurs : Nalgen® et

Qiagen® (DNeasy Tissue Kit™). Ces kits comprennent des tampons de composition inconnue et des colonnes individuelles au fond desquelles se trouve une résine à forte affinité pour l'ADN. Le protocole précis dépend du fabricant, mais le principe est toujours le même : une digestion classique des tissus est réalisée à 55 °C. Cette digestion est suivie d'un passage de 10 min à 70 °C afin d'achever la lyse cellulaire. Les échantillons, placés dans les colonnes de résine subissent deux à trois étapes de rinçages par ajout de tampons alcooliques puis centrifugation à 13000 g_n (1 à 3 min). La centrifugation assure l'élimination de la quasi-totalité des composés cellulaires potentiellement gênants pour les étapes ultérieures des protocoles. Enfin, la récupération de l'ADN est réalisée par ajout d'un tampon d'élution qui détache l'ADN de la membrane de résine et une dernière étape de centrifugation. Cette méthode fournit de bons rendements d'extraction (jusqu'à 20 μ g d'ADN pour 10 mg de tissus) et de purification et permet en outre une très bonne conservation des ADN extraits à -20 °C.

A.1.2 La Réaction de Polymérisation en Chaîne : principe

L'étape d'amplification de l'ADN a pour objectifs (i) de sélectionner, chez tous les individus étudiés, la même zone précise du génome (ou "*locus*") et (ii) de dupliquer en un grand nombre de copies le fragment d'ADN choisi, afin de permettre par exemple son séquençage ou son génotypage. L'amplification fait appel à la technique de Réaction de Polymérisation en Chaîne ("*PCR*", pour "*Polymerase Chain Reaction*") mise au point par Mullis *et al.* (1986, Prix Nobel de Chimie en 1993) qui consiste en une succession de réactions de réplication d'une matrice d'ADN.

Le milieu réactionnel contient tous les éléments nécessaires à la synthèse de nouveaux brins d'ADN : l'ADN matrice de chaque individu à analyser, la *Taq*-polymérase²³ et son co-facteur le $MgCl_2$, les quatre désoxyribonucléotides (dNTP) libres en large excès (dATP, dTTP, dCTP et dGTP), qui seront assemblés par l'enzyme pour synthétiser les nouveaux brins d'ADN, et un couple d'oligonucléotides (amorces, ou "*primer*"), défini spécifiquement pour encadrer une zone cible. Différents additifs sont parfois ajoutés dans le milieu réactionnel, (*e.g.* albumine de sérum bovin ou BSA) afin de neutraliser d'éventuels inhibiteurs de réplication.

Le mélange réactionnel est ensuite introduit dans un thermocycleur qui va placer les échantillons à 3 températures successives et qui va recommencer en effectuant de 25 à plus de 35 cycles. Chaque cycle comprend :

1. **une phase de dénaturation** à une température comprise entre 93 °C et 95 °C. La rupture des liaisons hydrogènes liant les deux brins de la double hélice permet d'obtenir des ADN simples brins.
2. **une phase d'hybridation** (ou annealing) dont la température dépend du couple d'amorces utilisé. Au cours de cette étape, la température des échantillons est redescendue progressivement. Seules les amorces, parce qu'elles sont de petite taille, vont pouvoir s'accrocher rapidement à l'ADN matrice qui reste simple brin. La température d'hybridation est généralement comprise entre 45 et 60 °C.
3. **une phase d'élongation** à 72 °C, au cours de laquelle la *Taq*-polymérase va synthétiser le brin complémentaire au brin matrice, en se servant des amorces pour point de départ et

²³Enzyme de réplication thermorésistante, elle a été isolée pour la première fois chez la bactérie thermophile *Thermus aquaticus*, et a permis une automatisation de la technique avec des taux d'erreurs extrêmement faibles (moins d'une substitution pour 9 000 nucléotides assemblés et moins d'un déplacement du cadre de lecture sur 41 000 bases polymérisées Tindall et Kunkel, 1988).

des nucléotides libres présents dans le milieu réactionnel.

Les conditions spécifiques aux différentes PCR que nous avons réalisé au cours de cette thèse (composition des mélanges réactionnels et détail des cycles) seront précisées dans les sections suivantes.

Les produits de chaque cycle de synthèse sont utilisés comme matrices pour les cycles suivants. Comme le montre la figure A.1, les premières copies de l'ADN cible n'apparaissent dans le milieu réactionnel qu'à la fin du troisième cycle. Chaque copie initiale de l'ADN matrice est en effet plus longue que la séquence cible souhaitée. Le nombre de ces fragments de taille inappropriée augmente de façon linéaire dans le mélange réactionnel (leur nombre vaut $2c$ avec c = nombre de cycles). Par contre, les copies des brins nouvellement synthétisés ont toutes la longueur recherchée : à partir du troisième cycle, leur nombre augmente exponentiellement ($2^c - 2c$).

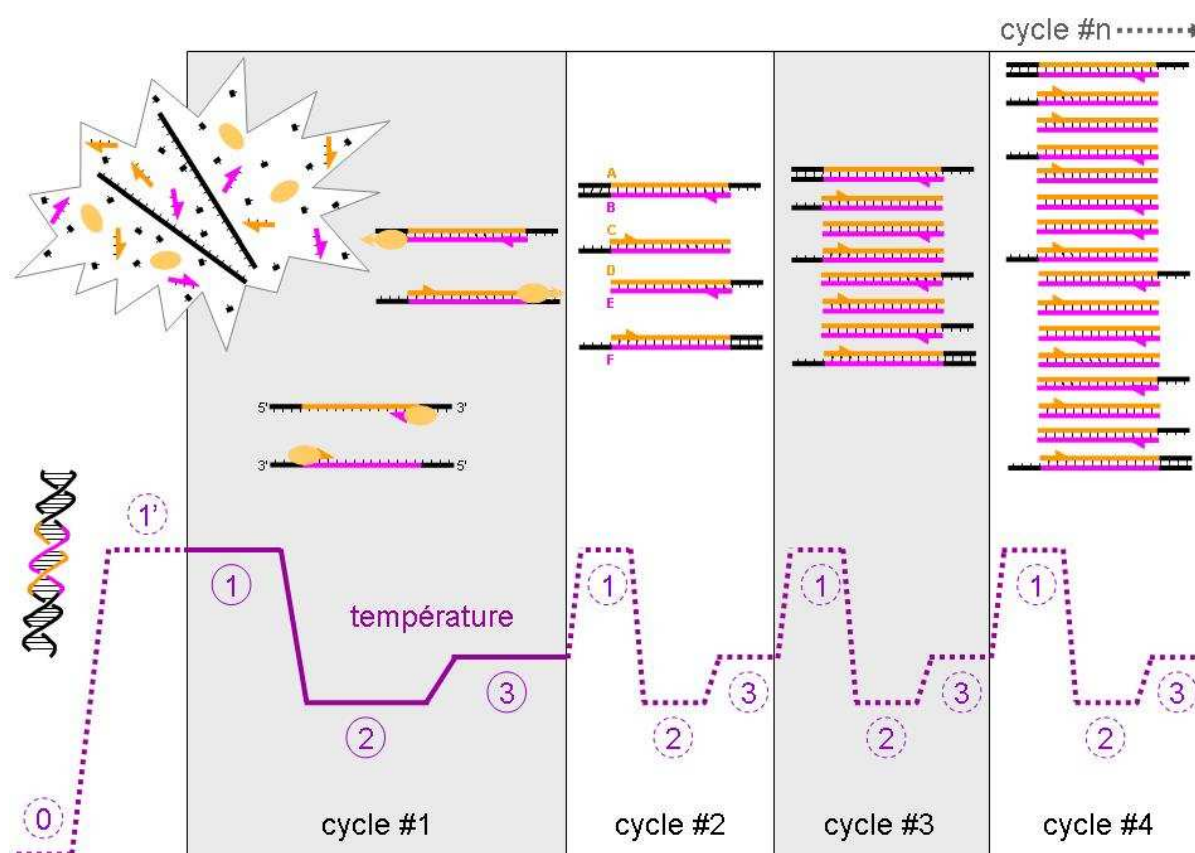


FIG. A.1: Les étapes de la Réaction de Polymérisation en Chaîne. En haut à gauche, les flèches roses et oranges représentent les amorces, les petits carrés noirs les dNTP et les ovales oranges la *Taq*-polymérase. Le locus cible est indiqué en couleur au milieu de la matrice d'ADN. En bas, les numéros correspondent aux étapes de : dénaturation (1), d'hybridation (2) et d'élongation (3). Dès le quatrième cycle, le nombre de copies de la taille souhaitée représente 50% de la totalité des molécules du mélange réactionnel. Source : <http://fr.wikipedia.org/wiki/Utilisateur:Ygonaar>.

Ainsi, le nombre total de copies d'ADN étant doublé à chaque cycle, 2^c copies sont théoriquement synthétisées dans le milieu réactionnel. Après 25 cycles, plus de 33 millions de copies sont donc présentes (ce qui est généralement suffisant pour une utilisation ultérieure), dont plus de 99,999% ont la taille recherchée (voir figure A.2).

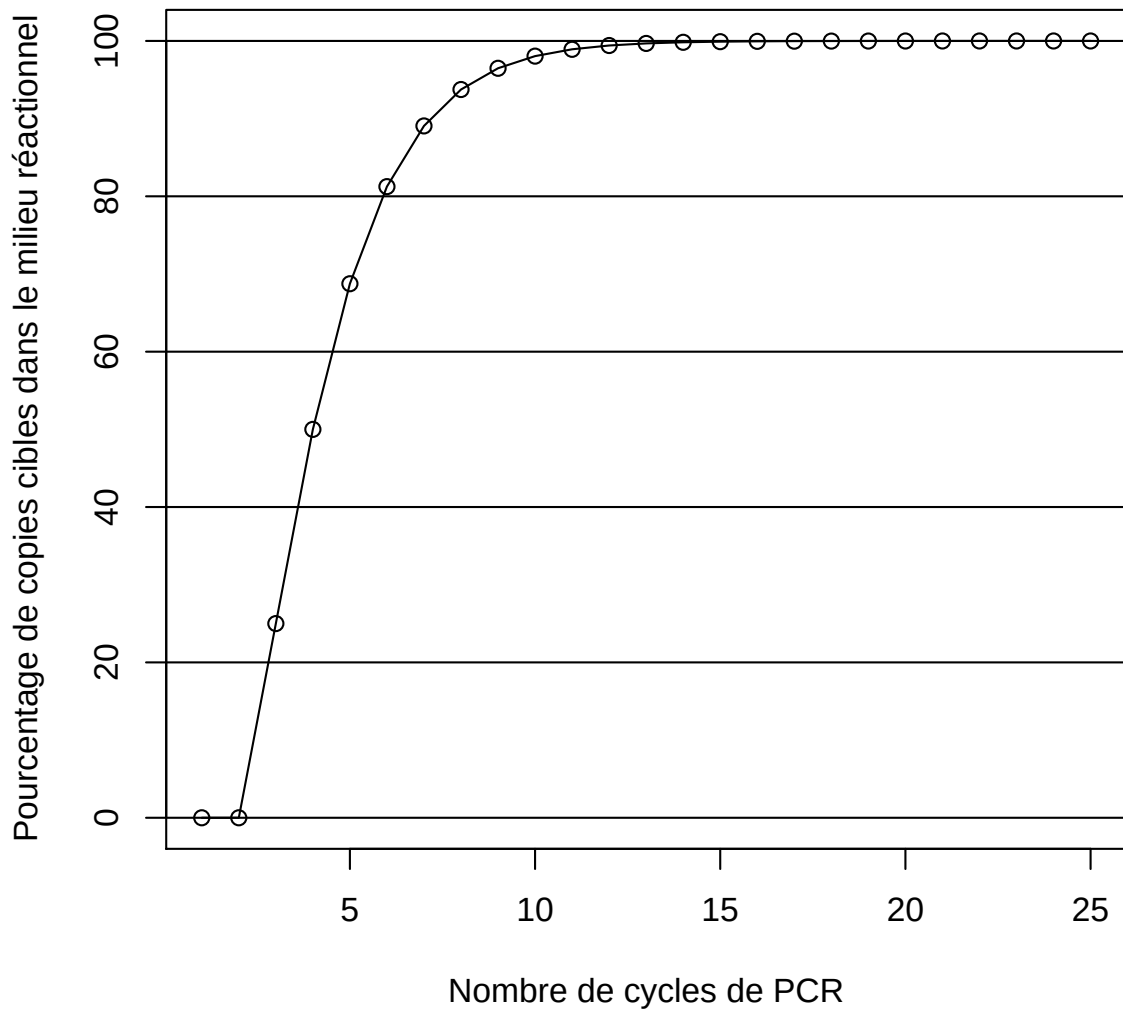


FIG. A.2: Évolution théorique du pourcentage de copies d'ADN de taille souhaitée dans le milieu réactionnel au cours d'une réaction de polymérisation en chaîne.

Cette technique de réplique ciblée *in vitro* permet donc d'obtenir, à partir d'un échantillon complexe et peu abondant, d'importantes quantités d'un fragment d'ADN spécifique et de longueur définie.

A.2 Mise au point d'un marqueur mitochondrial

A.2.1 Amorces universelles

L'une des difficultés des études phylogéographiques est d'utiliser des marqueurs ayant des taux d'évolutions compatibles avec l'échelle spatiale et temporelle des processus que nous cherchons à révéler. En pratique, il faut donc pouvoir définir des amorces encadrant une zone variable du génome. Mais celles-ci doivent pouvoir se fixer rigoureusement au même endroit chez tous les individus étudiés, la zone d'accroche de l'amorce doit donc être la plus conservée possible.

Dans le cas de l'ADN mitochondrial (ADN_{mt}), plusieurs amorces, dites "*universelles*", remplissent ces conditions : définies dans des régions phylogénétiquement conservées du génome, elles sont susceptibles d'amplifier des régions polymorphes précises chez un large spectre d'espèces (Kocher *et al.*, 1989; Folmer *et al.*, 1994; Geller *et al.*, 1994; O'Foighil *et al.*, 1998). Ces amorces permettent d'éviter la recherche souvent longue et difficile de nouveaux marqueurs moléculaires spécifiques de l'espèce étudiée. Les couples d'amorces universelles présentés dans le tableau A.1 ont été testés sur plusieurs individus de *Cyclope neritea* afin de savoir quel locus était le plus à même de convenir pour une étude à grande échelle. Ces amorces flanquent notamment les gènes codant pour les ARN_r 12 S et 16 S, la région de contrôle de l'ADN_{mt} ainsi que les sous-unités I et III de la cytochrome c oxydase.

Pour chacun des couples testés, la solution réactionnelle, d'un volume total de 50 μ L contient, dans un tampon 1X, environ 200 ng d'ADN matrice, 1,5 mM de MgCl₂, 0,6 mM de dNTP, 50 pmole de chacune des deux amorces (sens et anti-sens), et une unité de *Taq*-polymérase Thermoprime Plus Hi-Fidelity (ABgene ®). Les réactions de PCR ont été menées dans un thermocycleur "DNA Thermal Cycler PTC 100" (MJResearch) qui nous a permis de tester différentes températures d'appariement pour chaque couple d'amorces. Le détail des cycles est présenté dans le tableau A.2.

Les résultats de l'amplification sont ensuite visualisés sous lumière UV après électrophorèse sur gel d'agarose à 2% (Ultra-pure DNA-Grade agarose, Eurogenetec) contenant du bromure d'éthidium, un agent intercalant qui va s'insérer entre les brins d'ADN au cours de leur migration. Excité entre 370 et 530 nanomètres, il émet dans le visible à 622 nm (voir un exemple de photo de gel à la figure A.3).

L'obtention de produits de PCR ne garantit pas toujours que l'amplification soit réussie : de très faibles variations des conditions expérimentales peuvent suffire à perturber la spécificité de l'amplification. Cela se traduit, sur les gels d'agarose (voir figure A.4), par l'apparition de bandes surnuméraires et/ou d'un bruit de fond important (si des fragments "*parasites*" sont amplifiés), ou par l'absence de toute bande si les conditions expérimentales ne sont pas réunies pour réaliser l'amplification. En outre, la présence d'une seule bande sur gel d'agarose ne garantit pas non plus l'accès à une séquence propre : les gels d'agarose étant peu résolutifs, plusieurs fragments d'ADN de tailles similaires peuvent en effet se superposer et ainsi rendre le séquençage difficile, voire impossible. Les résultats qualitatifs des amplifications réalisées avec les différentes amorces universelles sont présentés dans le tableau A.3.

Les amorces universelles permettent donc assez rapidement d'obtenir un amplifiat séquençable. Toutefois, comme l'attestent les résultats qualitatifs du tableau A.3, les produits de PCR ne sont pas toujours d'une qualité suffisante pour les étapes ultérieures des protocoles. Si l'espèce étudiée est phylogénétiquement trop éloignée des espèces ayant servi à la mise au point de ces amorces universelles, la spécificité n'est pas toujours excellente. C'est pourquoi, quand un amplifiat d'une qualité suffisante peut être obtenu, le séquençage et la redéfinition d'amorces internes spécifiques de l'espèce étudiée sont indispensables à la réalisation ultérieure de PCR et

TAB. A.1: Description des différentes amorces testées

Locus	Nom de l'amorce	Séquence	Taille du fragment cible (bp)	Référence
12S	12S H	5'-GAGGGTGACGGGGGGTGTGT-3'	386	Kocher <i>et al.</i> (1989)
	12S L	5'-AAACTGGGATTAGATACCCCACTAC-3'		
16S	16S H	5'-CCGGTCTGAACCTCAGATCAGATCAGT-3'	555-570	Geller <i>et al.</i> (1994)
	16S L	5'-CGCCTGTTTATCAAAAACAT -3'		
COI	HCO2198	5'-TAAACTTCAGGGTGACCAAAAAATCA-3'	708	Folmer <i>et al.</i> (1994)
	LCO1490	5'-GGTCAACAAATCATAAAGATATTGG-3'		
COIII	CO III	5'-ACCACATCTACAAAATGCCAATATC-3'	1000	O'Foighil <i>et al.</i> (1998)
	H O III L	5'-ATTAGTTGATCCTAGGCCTTGACC-3'		
Région de Contrôle	Control Region H	5'-TAACTGCAGAGGCTAGGACCAAAACCT-3'	1000	Kocher <i>et al.</i> (1989)
	Control Region L	5'-TCAAAGCTACACCAGTCTTGTAAACC-3'		

TAB. A.2: Conditions de PCR testées avec les différents couples d'amorces universelles (RC : Région de Contrôle). Ces PCR étaient toutes précédées d'une étape de dénaturation à 94 °C pendant 5 min, puis suivies par une phase d'élongation terminale de 7 min à 72 °C.

Locus	Dénaturation	Appariement	Élongation	Nombre de cycles
12S	30 s à 94 °C	30 s à 48-63 °C	1 min à 72 °C	40
16S	30 s à 94 °C	30 s à 45-64 °C	1 min à 72 °C	40
COI	30 s à 94 °C	30 s à 40-60 °C	1 min à 72 °C	40
COIII	30 s à 94 °C	30 s à 40-62 °C	1 min à 72 °C	40
RC	30 s à 94 °C	30 s à 45-55 °C	1 min à 72 °C	40

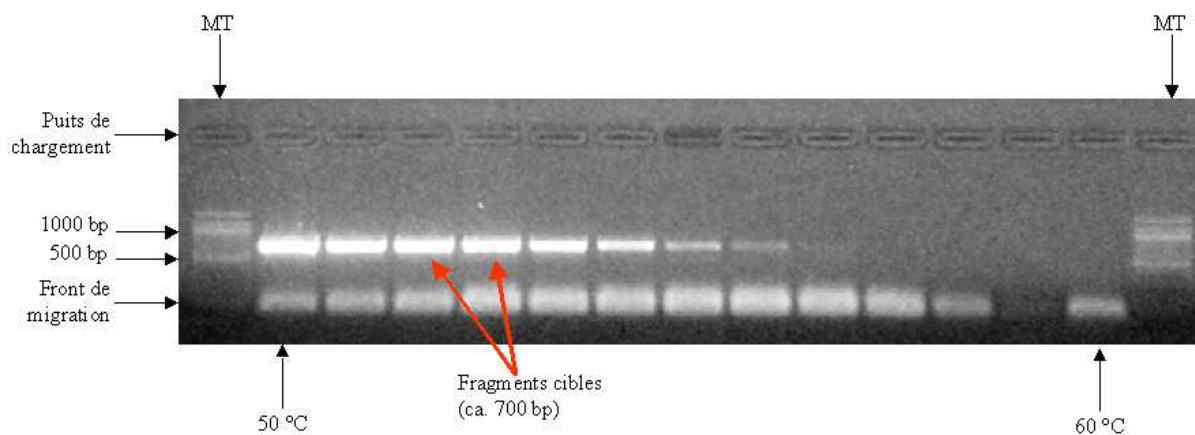


FIG. A.3: Exemple de résultat d'amplifications obtenus pour COI avec un gradient de température. L'amplification a été testée sur le même individu à 12 températures d'hybridation différentes (l'avant dernier puits est vide), comprises entre 50 (premier puits du gel à gauche) et 60 °C (dernier puits à droite). Les flèches marquées "MT" correspondent aux puits contenant le marqueur de taille. Au niveau du front de migration se retrouvent toutes les amorces et les nucléotides libres n'ayant pas été utilisés au cours de la PCR. Les températures permettant une amplification efficace s'étendent de 50 à 54 °C.

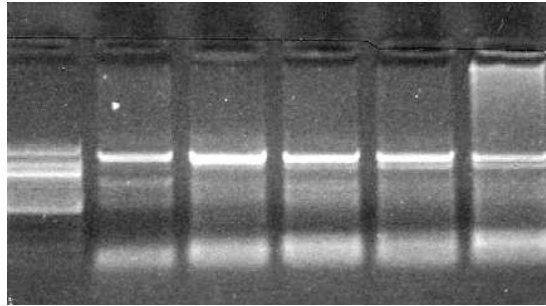


FIG. A.4: Exemple de PCR ne donnant pas les résultats escomptés. Cette PCR, réalisée avec les amorces universelles de la Région de Contrôle, dans les mêmes conditions pour les 5 individus testés (le premier puits à gauche étant occupé par le marqueur de taille), se caractérise par un bruit de fond important et par la présence de bandes surnuméraires qui n'étaient pas ciblées par l'amplification, probablement en raison d'un manque de spécificité de la réaction.

TAB. A.3: Résultats qualitatifs obtenus pour chaque locus testé.

Locus	Obtention d'un produit d'amplification	Amplification toujours non spécifique	Séquences obtenues exploitables	Polymorphisme
12S	non	-	-	-
16S	oui	non	non	-
COI	oui	non	oui	oui
COIII	oui	oui	-	-
RC	oui	oui	-	-

de séquençage en série, ce qu'exige la génétique des populations. Nous allons donc maintenant détailler le protocole que nous avons utilisé pour séquencer le seul produit de PCR utilisable à l'issue des tests d'amorces universelles : un fragment du gène codant pour la sous-unité I de la cytochrome oxydase.

A.2.2 Séquençage

Purification des produits de PCR

À l'issue de la réaction de PCR, le milieu réactionnel contient non seulement le fragment d'ADN cible, mais également les restes d'amorces et de nucléotides n'ayant pas été incorporés dans de nouvelles copies ainsi que divers éléments pouvant interférer avec le séquençage des fragments amplifiés (*e.g.* restes d'enzymes, BSA...). Il convient donc de purifier ces produits de PCR afin d'obtenir une solution d'ADN cible propre au séquençage. La purification a ici été réalisée sur plaques MultiScreen-PCR MANU03010 (Millipore). Les échantillons sont simplement déposés sur la membrane, puis filtrés par application d'un vide ou par centrifugation. La membrane ne retenant que les grosses molécules, seul l'ADN cible (plusieurs centaines de paires de bases) est retenu. Les échantillons sont ensuite repris dans 50 μ L d'eau milliQ.

Réaction de séquence et séquençage

La réaction de séquence consiste à réaliser une nouvelle PCR, en utilisant l'ADN purifié comme matrice, des nucléotides (dNTP) classiques et des dNTP modifiés par l'incorporation de marqueurs fluorescents : les "*BigDye™ Terminator*" (Heiner *et al.*, 1998). Chaque type de dNTP modifié (dATP*, dTTP*, dGTP* et dCTP*) possède un marquage propre (ou Dye) et a la particularité d'interdire la poursuite de la répllication après son incorporation dans le fragment d'ADN néo-synthétisé. dNTP classiques et marqués sont présents dans un rapport particulier, de sorte que les fragments d'ADN de toutes les longueurs possibles (de 1 à 700 paires de bases (bp) si l'ADN matrice fait 700 bp) seront présents dans la solution, chaque fragment étant terminé par un unique nucléotide marqué. En outre, une seule amorce (et non les deux amorces d'un même couple) sera utilisée. Ainsi, les ADN à séquencer seront simples brins, la composition du brin complémentaire étant déduite de celle du brin séquencé. Toutefois, nous avons systématiquement réalisé deux réactions de séquence ainsi que deux séquençages pour chaque individu étudié, en utilisant l'amorce sens d'une part puis l'amorce anti-sens d'autre part. Le séquençage n'étant pas toujours infallible, avoir deux séquences du même fragment nous a permis de corriger certaines erreurs de lecture du séquenceur automatique.

Les réactions de séquence sont réalisées dans un volume total de 5 μ L et contiennent environ 100 ng d'ADN purifié, 0,75 μ L d'un tampon 5X fourni (400 mM Tris HCl pH 9,0 et 10 mM MgCl₂), 5 pmole de l'une des deux amorces et 0,5 μ L du mélange de Big Dye V3.1 (Applied Biosystem) contenant les dNTP marqués et non marqués. La réaction de PCR comprend, outre une étape de dénaturation initiale (5 min à 96 °C), 50 cycles de dénaturation (30 s à 96 °C), hybridation (30 s à la température d'appariement, pour COI, 50 °C) et élongation (4 min à 60 °C). A l'issue de cette réaction, une autre étape de purification sur plaques MAHVN4510 (Millipore) est réalisée afin d'éliminer les dNTP* non incorporés.

Enfin, le séquençage a été réalisé par un séquenceur à 16 capillaires ABI prism® 3100 (Perkin-Elmer Applied Biosystems, Foster City, CA). Pour chaque échantillon, les fragments migrent dans un capillaire en fonction de leur taille. Lorsque les fragments arrivent au niveau du laser du séquenceur, les Dye Terminator sont excités, et chaque type de dNTP* émet dans une

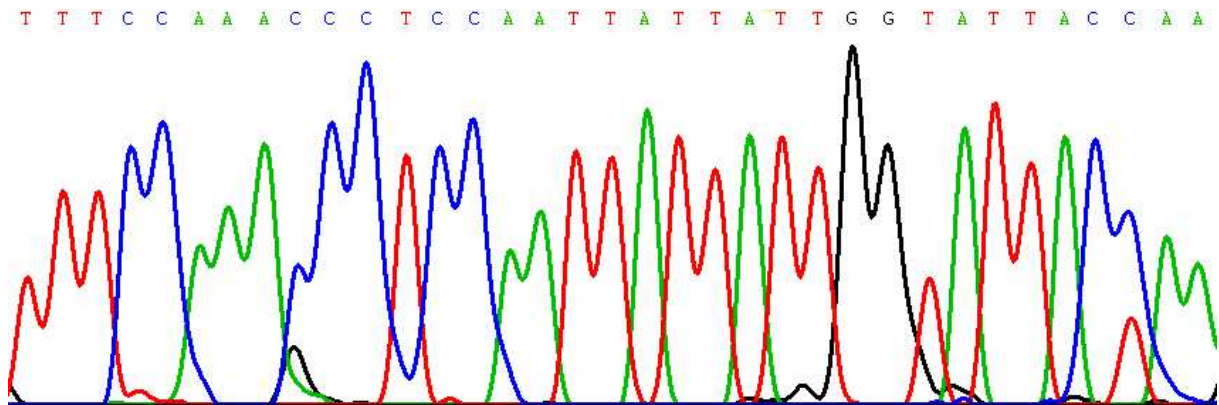


FIG. A.5: Exemple de chromatogramme.

longueur d'onde spécifique. Chaque nouveau fragment passant devant le laser permet donc la lecture d'une base supplémentaire de la séquence nucléotidique du fragment cible. Les résultats sont délivrés à la fois sous forme de séquence texte et de chromatogramme (voir figure A.5), chaque pic de couleur correspondant à l'émission lumineuse d'un Dye Terminator excité par le laser du séquenceur.

A.2.3 Définition d'un couple d'amorces spécifiques

Lorsqu'une séquence a été obtenue grâce aux amorces universelles, il ne reste plus qu'à redéfinir des amorces internes complémentaires de cette séquence (encadrant la zone la plus large possible afin d'obtenir un fragment cible contenant le plus d'informations possible) et à tester leur efficacité.

La définition d'amorces doit suivre un certain nombre de règles. Par exemple, afin de faciliter l'accroche sur son ADN matrice, une amorce dont l'extrémité 3' se termine par un G ou un C sera recherchée; les triples liaisons hydrogènes se créant spontanément entre G et C au cours de la phase d'hybridation sont en effet plus stables que les doubles liaisons entre A et T. Pour la même raison, le contenu total de l'amorce en GC va influencer la température d'hybridation. Un autre critère de choix est la taille de l'amorce. Plus elle sera longue, plus sa température d'hybridation sera élevée, garantissant une bonne spécificité de l'accroche. Toutefois, augmenter inconsidérément la longueur des amorces augmente parallèlement les chances de voir apparaître des structures secondaires stables au sein ou entre les amorces d'un même couple, ce qui diminue évidemment l'efficacité de l'hybridation. La taille est donc généralement comprise entre 20 et 25 bases. Ainsi, le choix d'une amorce résulte toujours d'un compromis entre différents critères de choix.

Concernant *Cyclope neritea*, les amorces que nous avons définies sont les suivantes :

- Amorce sens : 5'-GTAAAATTTTCGATCTGTTA-3' (Cy2)
- Amorce antisens : 5'-GGATTAGTTGGTACAGC-3' (CyR)

Avec ces nouvelles amorces, les PCR ont été réalisées dans les conditions suivantes : volume total : 50 μ L, ADN : 5 à 50 ng, amorces : 0,5 μ M, dNTP : 0,2 mM, MgCl₂ : 1,5 mM, Taq-polymérase : 1,25 Unité et Tampon 10X (Tris HCl pH 8,8 à 50 mM à 25 ° C et du (NH₄)₂SO₄, ABgene) : 5 μ L.

Les cycles de PCR comprennent une dénaturation initiale de 3 min à 93 ° C suivie de 40 cycles de 30 s à 93 ° C, 30 s à 50 ° C et 60 s à 72 ° C. Une élongation finale de 10 min à 72 ° C termine

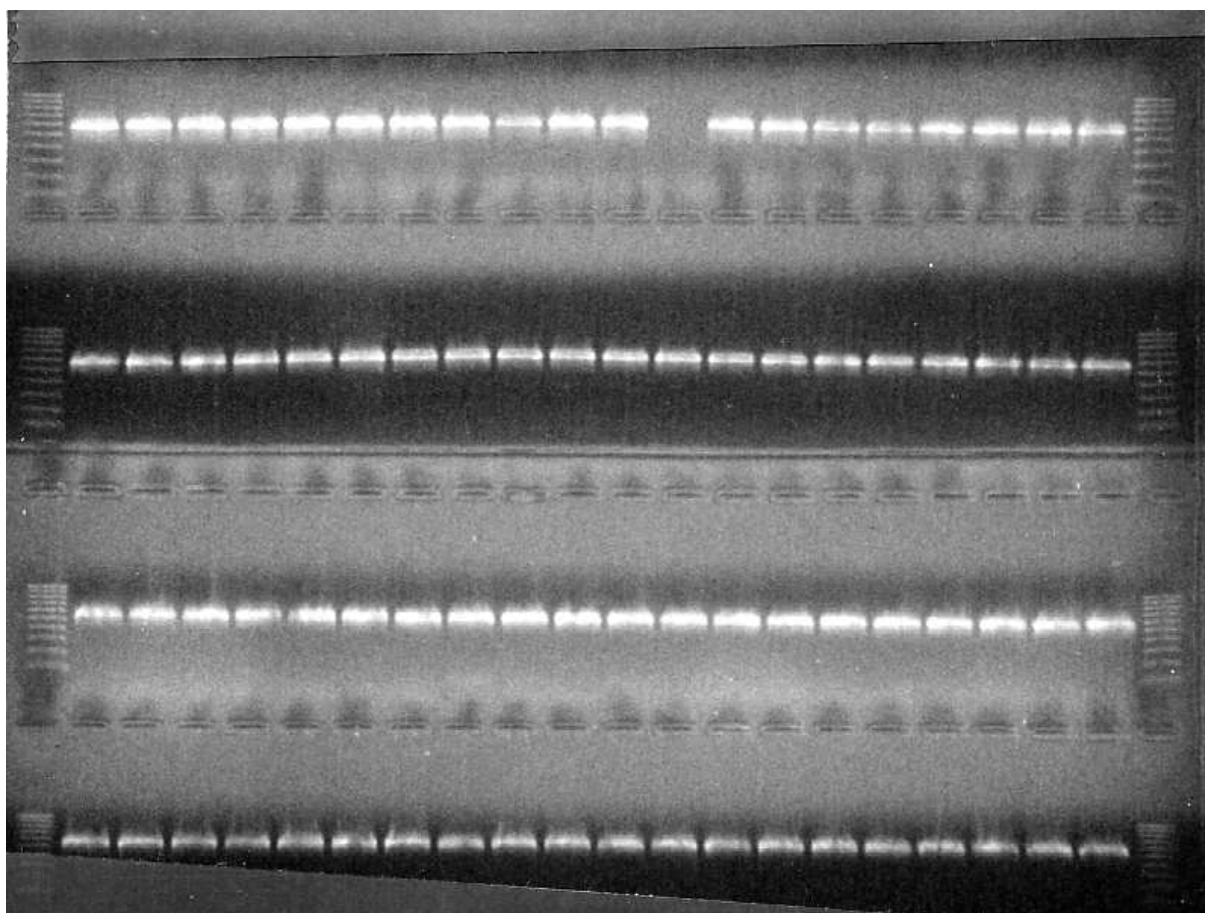


FIG. A.6: Résultats de PCR réalisée avec les amorces Cy2 et CyR. Sur 80 individus testés (issus de deux populations naturelles de *Cyclope neritea*), 79 amplifications ont réussi.

la PCR. De la même façon qu'avec le couple d'amorces universelles (HCO2198 et LCO1490), les mêmes étapes de purification sur plaque millipore, de réaction de séquence et de séquençage ont été réalisées. Ces amorces ont donné de très bons résultats (voir par exemple la figure A.6 montrant les résultats d'une PCR réalisée sur 80 individus issus de deux populations), un fragment de 533 paires de bases du locus COI ayant ainsi pu être séquencé chez plus de 650 individus. Les séquences obtenues ont été alignées grâce au logiciel *Clustal W* (Thompson *et al.*, 1994) puis vérifiées grâce à la comparaison des séquences sens et anti-sens dans le logiciel *Bioedit* (Hall, 1999), afin de permettre la détermination des haplotypes (séquences nucléotidiques uniques), données de base des analyses statistiques ultérieures.

A.3 Mise au point de marqueurs microsatellites

La mise au point de marqueurs microsatellites pour une nouvelle espèce est souvent plus compliquée que la mise au point de marqueurs mitochondriaux. Il est en effet difficile de savoir *a priori* à quel endroit précis du génome les microsatellites d'intérêt vont être localisés. Ces marqueurs se trouvent généralement dans des zones très variables du génome, il n'existe donc pas d'amorces universelles permettant l'amplification rapide de tels marqueurs. Le protocole expérimental suivi pour la mise au point des marqueurs microsatellites spécifiques à *Cyclope neritea* est décrit succinctement dans la "primer note" présentée en annexe D.2, (Simon-Bouhet *et al.*, 2005). Les différentes étapes sont détaillées ci-dessous.

A.3.1 Construction d'une banque enrichie

La banque réalisée par Claire Daguin (EGPM, Station Biologique de Roscoff), à partir d'ADN génomiques extraits sur colonne de résine, suit la méthode de Billote *et al.* (1999) basée sur Kijas *et al.* (1994). Afin de disposer d'une quantité de matériel suffisante, les ADN issus de trois individus ont été mélangés.

Digestion et ligation

Les ADN sont tout d'abord digérés par l'enzyme de restriction RsaI (Promega) coupant les sites 5'-GTAC-3' entre le T et le A et produisant des extrémités franches. Environ 1,5 µg d'ADN est mis en présence de spermidine (4 mM) et de 30 unités d'enzyme pendant 7 heures à 37 °C. Les ADN digérés sont ensuite purifiés sur colonne Nucleospin® (Macherey-Nagel) afin d'éliminer l'enzyme et la spermidine (volume initial : 370 µL, volume final : 80 µL).

Une étape de ligation à deux adaptateurs (Rsa21 et Rsa25) est ensuite réalisée :

- Rsa21 : 5'-CTCTTGCTTACGCGTGGACTA-3'
- Rsa25 : 5'-TAGTCCACGCGTAAGCAAGAGCACA-3'

Le milieu réactionnel (100 µL) contient environ 1 µg d'ADN, les deux adaptateurs (à 10 µM chacun) et 9 unités d'enzyme T4-ligase (Promega). La ligation s'effectue à 18 °C pendant 12 heures. Puis, une nouvelle étape de purification sur colonne Nucleospin® (Macherey-Nagel) est réalisée (volume initial : 100 µL, volume final : 50 µL).

Une étape de PCR post-ligation est ensuite effectuée (sur un thermocycleur MJresearch PTC200) afin d'obtenir assez d'ADN pour les étapes ultérieures du protocole. Le volume réactionnel total de 50 µL contient 7 µL d'ADN purifié, 0,5 mM de dNTP, 300 ng d'amorces (complémentaires des adaptateurs ayant servi pour la ligation), 3 mM de MgCl₂ et 5 unités de *Taq*-polymérase (Thermoprime ABgene). Les cycles de PCR sont les suivants : dénaturation initiale de 4 min à 94 °C, puis 20 cycles de 30 s à 94 °C, 60 s à 60 °C et 90 s à 72 °C, et enfin, une élongation terminale de 5 min à 72 °C.

Une dernière purification sur colonne Nucleospin® est effectuée avant l'enrichissement (volume initial : 45 µL, volume final : 50 µL).

Enrichissement

L'enrichissement consiste à éliminer, parmi la multitude de fragments présents dans le milieu réactionnel, ceux ne contenant vraisemblablement pas le motif microsatellite recherché. Pour cela, 50 µL d'ADN purifié sont dilués dans 400 µL d'eau milliQ et dénaturés à 95 °C pendant 15 min. 150 pmole d'oligonucléotides biotinylés en 5' et 13 µL de tampon SSC sont ajoutés. Les

oligonucléotides ont pour séquence (AC)₁₀, et vont aller s'accrocher sur des fragments possédant un motif microsatellite complémentaire au cours des 10 min d'incubation à température ambiante. 100 μ L de tampon SSC à 0,5X contenant des microbilles Magnesphere® paramagnétique (Promega) recouvertes de streptavidine sont ensuite ajoutés. Grâce à la streptavidine, ces billes vont s'accrocher à l'extrémité 5' des oligonucléotides biotinylés (au cours d'une nouvelle incubation de 10 min à température ambiante), puis elles seront récupérées grâce à leurs propriétés magnétiques. Ces billes sont ensuite lavées plusieurs fois avec un tampon SSC (0,5X), puis reprises dans un volume final de 250 μ L ne contenant plus, théoriquement, que des fragments d'ADN possédant un motif (AC) répété.

L'enrichissement est suivi d'une étape de purification sur colonne Nucleospin® (volume initial : 250 μ L, volume final : 100 μ L), d'une PCR post-enrichissement (volume total : 100 μ L, ADN : 10 μ L, dNTP : 0,5 mM, amorce complémentaire des adaptateurs : 560 ng, MgCl₂ : 3 mM et *Taq*-polymérase thermoprime ABgene : 10 μ L — dénaturation initiale de 1 min à 95 °C, puis 20 cycles de 40 s à 94 °C, 60 s à 60 °C et 2 min à 72 °C, et enfin, élongation terminale de 5 min à 72 °C) et d'une autre purification sur colonne Nucleospin® (volume initial : 90 μ L, volume final : 100 μ L).

Ligation, transformation et Dot Blot

Nous disposons donc maintenant de nombreux fragments d'ADN amplifiés et purifiés, chacun d'eux étant censés contenir un motif d'intérêt. Il s'agit maintenant de séparer ces différents fragments et de les répliquer afin de permettre leur séquençage. Pour cela, la ligation au plasmide pGEM®-T easy (Promega) est réalisée : dans un volume total de 10 μ L, 50 ng de plasmide, 3 unités de ligase (Promega) et 3 μ L d'ADN enrichi, amplifié et purifié sont mis à incuber à 4 °C pendant 12 heures.

Ces plasmides, contenant maintenant chacun un fragment d'ADN différent, sont insérés par transformation (par choc thermique, passage de 0 à 42 °C puis retour à 0 °C) dans des bactéries compétentes (*Escherichia coli* souche JM109). Les clones positifs, repérés par une réaction colorée, sont repiqués 24 heures plus tard. Trois plaques contenant chacune 96 clones positifs ont ainsi été préparées.

Une réaction de PCR est réalisée pour amplifier chacun des inserts contenus dans les plasmides bactériens. Pour cela, les cultures liquides sont mises à incuber pendant 5 heures, puis diluées au cinquième dans l'eau et dénaturées pendant 15 min à 99 °C. 2,5 μ L de cette solution contenant les plasmides bactériens sont prélevés pour chaque individu et serviront de matrice pour la PCR. La réaction s'effectue dans un volume total de 20 μ L et contient, outre l'ADN, 0,5 mM de dNTP, 3 mM de MgCl₂, 0,5 unités de *Taq*-polymérase Upitherm (Uptima) et enfin, 100 ng de deux amorces complémentaires des séquences plasmidiques flquant les inserts :

- pucF 5'-GTTTTCCAGTCACGACGTTGTA-3'
- pucR 5'-GGAAACAGCTATGACCATG-3'

Le déroulement de la PCR est le suivant : la dénaturation initiale de 4 min à 94 °C est suivie de 30 cycles de 45 s à 94 °C, 45 s à 55 °C et 45 s à 72 °C, et se termine par une phase d'élongation de 5 min à 72 °C.

Enfin, un dot blot des produits de PCR sur une membrane de nylon chargée positivement (Roche) permet de déterminer quelles sont les ligations/transformations ayant fonctionné : l'hybridation (pendant 12 heures à 56 °C) d'une sonde (TG)₁₀ marquée à la digoxygénine (DIG-(TG)₁₀) permet en effet la révélation par chémoluminescence des fragments d'ADN contenant le motif recherché (Dig luminescent detection kit (Roche) et exposition contre un film Hyperfilm ECL (Amersham Pharmacia Biotech) pendant 1 heure).

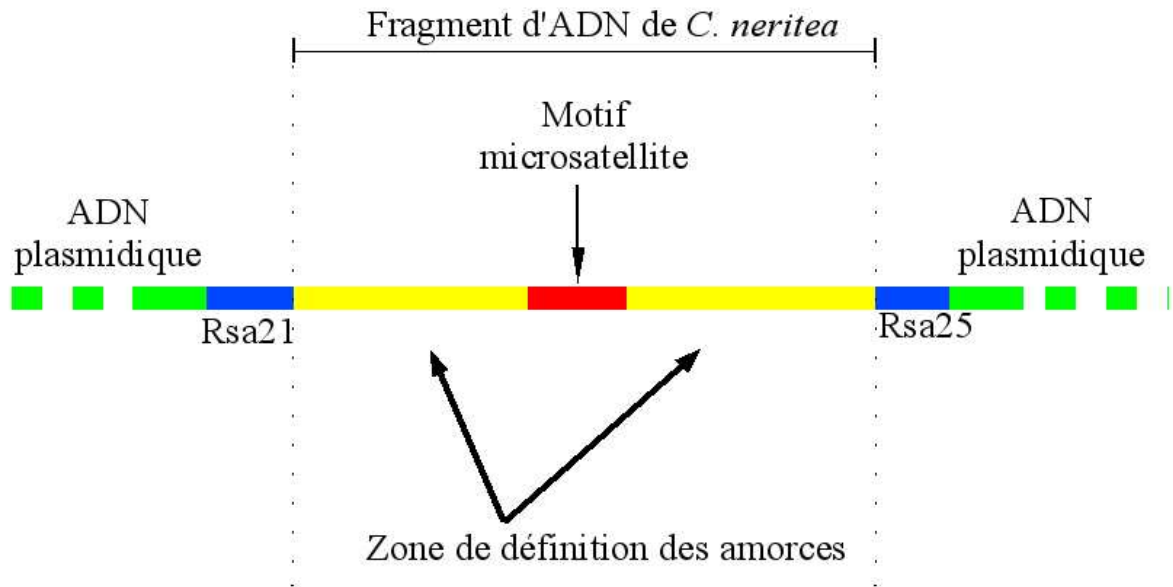


FIG. A.7: Représentation schématique d'un fragment d'ADN de *Cyclope neritea* obtenu par séquençage d'un clone positif. Ce cas de figure est idéal dans la mesure où un motif microsatellite est présent et il est situé au milieu de la séquence nucléotidique, ce qui laisse un grand choix pour la définition d'amorces de part et d'autre de ce fragment.

Au total, 144 clones positifs (clones présentant une luminescence forte et dont la taille de l'insert est comprise entre 300 et 900 paires de bases) ont été séquencés grâce à la méthode décrite à la section A.2.2 (page 192).

A.3.2 Séquençage et sélection des locus utilisables

Le séquençage des clones positifs est réalisé en utilisant, pour la réaction de séquence, les amorces complémentaires des séquences plasmidiques flanquant le fragment d'ADN de *Cyclope neritea*. La description schématique d'une séquence idéale de clone positif est présentée à la figure A.7. Cette séquence présente, de l'extérieur vers l'intérieur, un fragment d'ADN plasmidique correspondant aux amorces utilisées au cours de la réaction de séquence, les adaptateurs Rsa21 et Rsa25 utilisés au début du protocole, et enfin, au centre, un fragment d'ADN de *Cyclope neritea* comprenant un motif microsatellite.

Cette séquence théorique est dite "idéale" car le motif microsatellite présent est idéalement placé au sein de la séquence de *Cyclope neritea*. Dans de nombreux cas, le motif est soit collé à l'un des deux adaptateurs, soit trop long et complexe rendant la définition d'amorces encadrant le microsatellite impossible. Ainsi, sur les 144 clones positifs séquencés, 36 paires d'amorces ont pu être définies grâce au logiciel Primer 3 (Rozen et Skaletsky, 2000). L'efficacité et la spécificité de ces amorces ont été testées par PCR, les résultats d'amplification étant chargés sur gels d'agarose. A l'issue de ces tests, seuls 24 couples d'amorces ont donné des résultats satisfaisants : PCR spécifique (une seule bande visible sur agarose), peu de bruit de fond, taille du fragment correcte (entre 150 et 500 paires de bases) et répétabilité de l'amplification sur plusieurs individus.

Les 24 amorces "sens" des couples choisis ont ensuite été marquées par deux types de Dye fluorescents : après excitation par un laser, les amorces émettent dans l'infrarouge à 700 (IRD700)

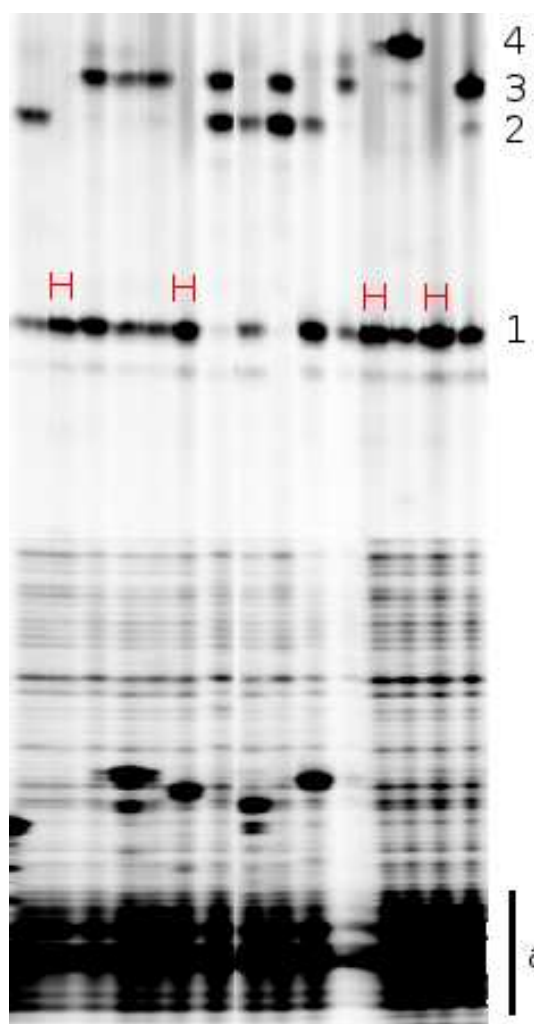


FIG. A.8: Exemple d'image obtenue après migration d'un locus microsatellite sur gel de polyacrylamide (15 individus). Le "a" indique le patch d'amorces marquées n'ayant pas été incorporées au cours de la PCR. Quatre allèles sont visibles pour ce locus (numérotés de 1 à 4 par ordre de taille croissant), 4 individus sont homozygotes ("H" rouge), les autres sont hétérozygotes.

ou à 800 nm (IRD800). Au lieu de marquer les nucléotides comme dans le cas du séquençage de l'ADN mitochondrial, seules les amorces sont ici marquées. Dans la mesure où seule la longueur (donc les profils électrophorétiques) des fragments amplifiés en PCR est importante, ce marquage permettra de visualiser les fragments après migration sur gel de polyacrylamide dénaturant. Contrairement à l'agarose, ces gels sont extrêmement résolutifs et permettent de séparer des fragments dont la taille ne varie que d'une simple paire de base. La figure A.8 présente un exemple d'image générée par l'un des séquenceurs à plaques (par opposition au séquenceur à capillaire) ayant servi à génotyper les ADN (Li-Cor NEN global IR2 DNA sequencer).

La visualisation sur gel d'acrylamide des résultats de nouveaux tests a souvent révélé des problèmes qu'il était impossible de détecter sur gel d'agarose : plusieurs locus produisant des fragments de taille similaire peuvent par exemple être amplifiés par un seul couple d'amorces. L'amplification n'étant pas toujours spécifique, plus de deux bandes apparaissent chez certains individus (les bandes étant relativement proches, elles apparaissaient superposées sur gel agarose).

Certains locus se sont révélés strictement monomorphes, donc inutilisables. D'autres étaient tellement polymorphes que la détermination précise de la taille des allèles s'avérait délicate (plus de 100 allèles avec plus de 250 paires de bases d'écart entre les plus petits et les plus grands). D'autres encore n'ont pas pu être amplifiés proprement, la présence systématique de bandes parasites empêchant la détermination claire des génotypes. Ainsi, sur les 24 couples d'amorces testés sur gel d'acrylamide, seuls 8 se sont avérés exploitables. Leurs principales caractéristiques sont présentées dans la figure A.9, tiré de Simon-Bouhet *et al.* (2005).

Ces 8 locus ont été utilisés pour réaliser le génotypage en série des populations de *Cyclope neritea* sur un séquenceur Li-Cor (Li-Cor 4300 DNA Analyser), à l'aide du logiciel Saga génération 2 v3.2.1 (voir la capture d'écran présentée à la figure A.10). Mais là encore, l'utilisation de ces marqueurs à grande échelle a révélé la présence probable d'allèles nuls²⁴, faussant grandement les résultats. Au final, seuls 4 locus microsatellites se sont avérés non ambigus et ont été utilisés en routine au cours de ce travail de thèse. Il s'agit des locus Cn-2C4, Cn-3C4, Cn-1A10 et Cn-2H10b.

Enfin, tout comme l'haplotype est la donnée de base pour les analyses mitochondriales, le génotype est la source de toutes les analyses microsatellites. Chaque individu sera donc caractérisé par la taille de ses deux allèles (taille identique s'il est homozygote) à chacun des locus analysés.

La liste de tous les haplotypes obtenus et les fréquences alléliques aux locus microsatellites sont présentés en annexe B.

²⁴Des déficits en hétérozygotes très marqués nous ont forcés à considérer la présence d'allèles nuls : certaines mutations du motif microsatellite peuvent entraîner des modifications de la zone d'accroche de l'une ou des deux amorces. Quand cela se produit, l'allèle en question n'est pas amplifié et des individus en réalité hétérozygotes sont détectés homozygotes.

Table 1 Characteristics of the eight microsatellite loci isolated from *Cyclope neritæ* (GenBank Accession nos: AY711365–AY711372). T_i and T_a stand for the initial and optimal annealing temperature of the touch-down PCR procedure; Cycle 1 refers to the number of cycles used to reach the optimal annealing temperature whereas Cycle 2 refers to the number of cycles used at the optimal annealing temperature. N_{all} and H_T : number of alleles and total heterozygosity over the whole sample ($N = 48$), N_{all} and H_E : number of alleles and expected heterozygosity at the population level, F_{IS} : estimator for deviation to Hardy–Weinberg proportions and associated P -values (exact test)

Locus	Core sequence	Primer sequences (5'–3')	T_i – T_a (°C) (Cycle 1)	Cycle 2	MgCl ₂ (mM)	Allele size range (bp)	N_{all}	H_T	Thau ($N = 24$)			Faro ($N = 24$)		
									N_{all}	H_E	F_{IS}	N_{all}	H_E	F_{IS}
Chr-2C4	(GT) ₁₅	F: ATCGAGCAGAAAGACATGAC R: GAAAGTGTGACCCGTGAG	65–58 (8)	24	1.5	151–172	5	0.66	5	0.61	0.106 ($P = 0.305$)	2	0.51	0.269 ($P = 0.237$)
Chr-2H10b	(GA) ₁₆ GG(GA) ₅ G GAGAGA(ACG) ₄	F: CCACGGTGTCCCTACTT R: AATGAGCGGTAAATCTGACAA	60–52 (8)	24	1.5	118–193	9	0.83	6	0.73	0.033 ($P = 0.096$)	6	0.80	0.134 ($P = 0.449$)
Chr-3C4	(TGTG) ₉ (TGGC) ₂ (TGTC) ₆	F: TCTCCCAATCCAGTATAAG R: AACCCAACTTGCCCACTATG	58–50 (8)	32	2.5	206–239	6	0.72	5	0.69	0.091 ($P = 0.323$)	4	0.58	–0.247 ($P = 0.787$)
Chr-1H8	(GT) ₂₈	F: CCTTCTTGGTGTGCTGTTCTG R: GGAGTAGGTAGAGTTAGTGAATGTC	62–52 (10)	25	1.5	130–142	5	0.47	4	0.43	0.609 ($P = 0.000$)	3	0.52	0.666 ($P = 0.000$)
Chr-1A10	(TG) ₁₀ (AG) ₁₅ GGG (AG) ₂₁ (TG) ₃	F: GCGTTACAGCAAAATGC R: CAGATCCGTCTTCAGC	60–52 (8)	24	1.5	152–224	15	0.89	13	0.90	0.246 ($P = 0.079$)	6	0.74	0.084 ($P = 0.451$)
Chr-1A8	(GA) ₃₆ N ₄₀ (GT) ₃ A (TG) ₆ N ₁₅ (GT) ₄	F: GACACTGCCAGAAACAGTC R: GAAGTAAACAGTCGAAATTAATATAC	64–54 (10)	25	2.5	237–324	18	0.93	16	0.93	0.460 ($P = 0.000$)	3	0.61	0.702 ($P = 0.006$)
Chr-3B2	(GT) ₁₂ GC(GT) ₂ GC(GT) ₃	F: CTACCAACACACCGATGTCAG R: CACCTATGTTGGTGCAGC	65–55 (10)	24	1.5	172–190	4	0.71	3	0.60	0.665 ($P = 0.001$)	2	0.26	0.324 ($P = 0.288$)
Chr-2B7	(CA) ₇	F: CAGTGGCAAGCTGCTTTAGATC R: CAACAACGACAAAGGCAAGC	63–55 (8)	22	2.5	176–178	2	0.42	2	0.23	0.254 ($P = 0.310$)	2	0.51	–0.154 ($P = 0.680$)
Overall (SD)										0.64 (0.23)	0.290 ($P = 0.000$)		0.57 (0.17)	0.171 ($P = 0.005$)

© 2005 Blackwell Publishing Ltd, *Molecular Ecology Notes*, 10.1111/j.1471-8286.2005.00857.x

FIG. A.9: Principales caractéristiques des microsatellites testés au niveau de la population (nom des locus, conditions d'amplifications et indices de diversité sommaires).

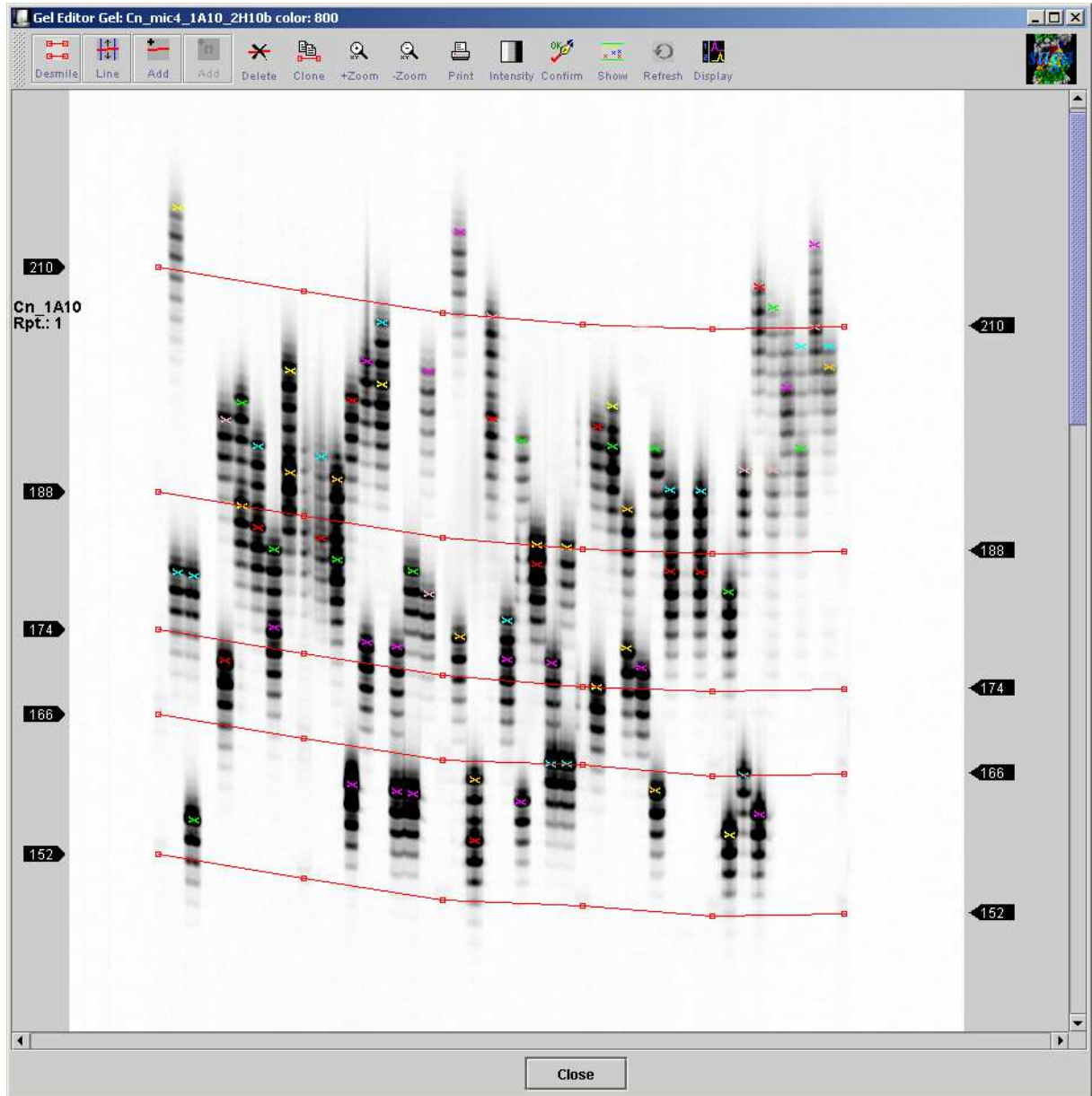


FIG. A.10: Capture d'écran du logiciel Saga. Un marqueur de taille est intercalé tous les huit puits (peu visible ici par manque de contraste) afin de visualiser "l'effet sourire" ou "smiling" apparaissant au cours de la migration en raison d'une température non homogène sur toute la largeur du gel. Chaque ligne rouge indique donc une taille d'allèle donnée et permet de visualiser les distorsions. Pour chaque individu, les croix de couleur placées automatiquement par le logiciel (édition/correction à la main grandement recommandées!) indiquent la taille des allèles. La *Taq*-polymérase réalisant des erreurs au cours de la PCR (suppression d'un ou plusieurs motifs par glissement), plusieurs échos ou bandes parasites, généralement moins intenses que les allèles et de plus petite taille, apparaissent sous les bandes des allèles. Les glissements d'un motif sont plus fréquents que les glissements de deux motifs, eux-même plus fréquents que les glissements de trois motifs, etc. Au sein des échos, les allèles sont donc généralement situés plus haut, et leur bande est la plus intense.

B

Données moléculaires

B.1 Fréquences alléliques par locus et par Population

TAB. B.1: Fréquences alléliques du locus 2C4 pour chaque populations étudiée. L'ordre des populations (présentées en colonnes) est le suivant : 1. Stérec, 2. Séné, 3. Rivedoux, 4. Ronce, 5. Arguin, 6. Hossegor, 7. Pedreña, 8. San Vicente de la Barqueire, 9. Alvor, 10. Faro, 11. Cadiz, 12. Mar Menor, 13. Ebre, 14. Leucate, 15. Thau, 16. Sardaigne 17. Sicile, 18. Caorle, 19. Venise, 20. Katérinie, 21. Eforie. Le nombre d'allèles observé N_{all} , le nombre de génotypes observés N_{gen} ainsi que l'hétérozygotie observée H_O et attendue H_E sont présentés.

2C4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
N_{gen}	66	66	74	64	66	62	6	74	78	54	72	74	66	74	54	50	20	76	80	74	58	
N_{all}	6	7	5	7	6	3	2	6	4	4	5	4	5	6	6	4	5	7	11	6	5.48	
1	0.01	0.00	0.00	0.02	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01
2	0.38	0.32	0.23	0.31	0.24	0.35	0.50	0.41	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13
3	0.00	0.06	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01
4	0.04	0.30	0.31	0.02	0.03	0.00	0.00	0.00	0.41	0.44	0.53	0.15	0.65	0.03	0.04	0.26	0.30	0.53	0.45	0.18	0.15	0.23
5	0.33	0.09	0.23	0.50	0.47	0.63	0.50	0.36	0.01	0.00	0.01	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.28	0.00	0.17
6	0.14	0.11	0.10	0.09	0.21	0.00	0.00	0.15	0.00	0.04	0.00	0.19	0.04	0.07	0.13	0.08	0.40	0.08	0.24	0.16	0.00	0.11
7	0.09	0.09	0.14	0.02	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02
8	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.54	0.50	0.43	0.32	0.26	0.19	0.37	0.56	0.00	0.29	0.03	0.00	0.00	0.17
9	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.05	0.04	0.00	0.01	0.34	0.03	0.53	0.41	0.02	0.20	0.07	0.17	0.00	0.28	0.10
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.10	0.04	0.09	0.03	0.50	0.04
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.04	0.06	0.00	0.00	0.00	0.00	0.00	0.01
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.01
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00
H_O	0.55	0.64	0.65	0.56	0.58	0.35	0.33	0.65	0.67	0.56	0.56	0.41	0.48	0.54	0.56	0.56	0.30	0.47	0.57	0.81	0.28	0.53
H_E	0.73	0.79	0.78	0.65	0.68	0.49	0.60	0.69	0.55	0.56	0.54	0.73	0.51	0.66	0.69	0.62	0.74	0.64	0.71	0.83	0.66	0.66

TAB. B.2: Fréquences alléliques du locus 2C3 pour chaque populations étudiée.

3C4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
N_{gen}	66	66	64	64	62	64	6	74	78	60	72	74	66	74	58	50	18	76	78	80	70	
N_{all}	7	6	9	6	7	6	3	8	4	4	5	5	5	6	6	4	7	7	6	13	6	6.19
1	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.01
4	0.03	0.00	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.05	0.07	0.00	0.04	0.00	0.02	0.00	0.00	0.01	0.00	0.11	0.73	0.05
5	0.30	0.42	0.31	0.36	0.40	0.39	0.17	0.19	0.24	0.43	0.56	0.04	0.35	0.03	0.24	0.26	0.06	0.01	0.05	0.25	0.11	0.25
6	0.29	0.32	0.27	0.17	0.34	0.39	0.33	0.51	0.49	0.40	0.28	0.53	0.39	0.58	0.31	0.40	0.33	0.13	0.09	0.04	0.04	0.32
7	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
8	0.14	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.02
9	0.00	0.00	0.02	0.00	0.00	0.00	0.50	0.12	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.11	0.00	0.31	0.00	0.00	0.05
10	0.04	0.15	0.11	0.19	0.10	0.17	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.04
11	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
12	0.09	0.04	0.17	0.19	0.03	0.00	0.00	0.11	0.15	0.00	0.08	0.23	0.14	0.04	0.00	0.00	0.00	0.17	0.00	0.01	0.00	0.07
13	0.11	0.04	0.08	0.08	0.10	0.02	0.00	0.01	0.00	0.12	0.01	0.16	0.08	0.16	0.10	0.30	0.06	0.25	0.08	0.19	0.00	0.09
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.04	0.00	0.18	0.15	0.04	0.17	0.17	0.32	0.14	0.03	0.06
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.22	0.00	0.15	0.01	0.00	0.03
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
H_O	0.85	0.61	0.66	0.81	0.61	0.66	1.00	0.68	0.72	0.57	0.72	0.70	0.82	0.62	0.93	0.60	0.67	0.87	0.80	0.88	0.46	0.72
H_E	0.80	0.70	0.80	0.78	0.71	0.68	0.73	0.68	0.68	0.65	0.61	0.65	0.71	0.61	0.79	0.69	0.84	0.81	0.77	0.85	0.46	0.71

TAB. B.3: Fréquences alléliques du locus 2H10b pour chaque populations étudiée.

2H10b	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
N_{gen}	66	58	74	64	62	64	6	74	74	62	72	74	66	74	58	50	20	74	80	80	70	
N_{all}	19	12	20	12	11	12	3	8	9	16	8	12	14	11	13	9	9	20	20	18	13	12.81
1	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.00	0.00	0.00	0.00	0.00	0.02
2	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
3	0.03	0.00	0.00	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.06	0.00	0.00	0.01
4	0.04	0.19	0.07	0.25	0.26	0.03	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.02	0.00	0.00	0.00	0.00	0.04	0.01	0.00	0.04
5	0.00	0.07	0.04	0.00	0.03	0.03	0.00	0.00	0.01	0.02	0.00	0.07	0.03	0.12	0.15	0.02	0.10	0.00	0.04	0.00	0.00	0.04
6	0.20	0.40	0.36	0.06	0.32	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.09
7	0.03	0.00	0.04	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.01	0.05	0.00	0.00	0.00	0.00	0.03	0.00	0.01
8	0.06	0.00	0.07	0.03	0.00	0.00	0.00	0.00	0.23	0.27	0.17	0.01	0.27	0.04	0.17	0.02	0.10	0.27	0.23	0.06	0.00	0.10
9	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00
10	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.14	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.15	0.15	0.16	0.00	0.04
11	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.20	0.63	0.04
12	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.01
13	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.00	0.00
14	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.00	0.51	0.14	0.10	0.40	0.01	0.01	0.05	0.00	0.07
15	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
16	0.03	0.02	0.04	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.04	0.00	0.01	0.03	0.06	0.03	0.02
17	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.01
18	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.01	0.00	0.00
19	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
20	0.00	0.07	0.04	0.00	0.07	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.10	0.06	0.02	0.02
21	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
22	0.17	0.05	0.08	0.08	0.08	0.11	0.67	0.27	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.07	0.16	0.03	0.10	0.10
23	0.08	0.09	0.05	0.27	0.11	0.14	0.17	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06
24	0.03	0.00	0.01	0.02	0.00	0.06	0.00	0.04	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.03	0.00	0.02
25	0.06	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.01
26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
27	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.01

TAB. B.4: Fréquences alléliques du locus 2H10b pour chaque populations étudiée (*suite et fin*).

2H10b	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
28	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.03	0.00	0.04	0.00	0.00	0.05	0.00	0.00	0.00	0.01	0.03	0.01	0.01
29	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.04	0.01	0.01	0.02	0.04	0.00	0.00	0.00	0.04	0.00	0.01
30	0.00	0.00	0.01	0.00	0.03	0.00	0.00	0.03	0.00	0.03	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.04	0.01	0.01	0.01
31	0.00	0.00	0.03	0.02	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
32	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.07	0.08	0.19	0.04	0.17	0.16	0.19	0.26	0.05	0.00	0.00	0.00	0.00	0.06
33	0.09	0.00	0.00	0.12	0.03	0.02	0.17	0.26	0.24	0.18	0.26	0.07	0.17	0.03	0.14	0.04	0.00	0.01	0.01	0.00	0.04	0.09
34	0.00	0.00	0.00	0.08	0.00	0.05	0.00	0.03	0.00	0.05	0.01	0.03	0.11	0.07	0.02	0.12	0.00	0.00	0.00	0.00	0.04	0.03
35	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.08	0.07	0.04	0.11	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.02
36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00
40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00
43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.05	0.01	0.00	0.00	0.01
44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.08	0.10	0.00	0.01	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.02
45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.02	0.18	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01
47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
H_O	0.88	0.72	0.73	0.84	0.84	0.84	0.67	0.68	0.89	0.77	0.61	0.68	0.82	0.62	0.76	0.84	0.70	0.73	0.82	0.95	0.74	0.77
H_E	0.92	0.80	0.85	0.84	0.81	0.76	0.60	0.80	0.83	0.88	0.83	0.85	0.86	0.70	0.88	0.79	0.83	0.89	0.89	0.91	0.59	0.81

TAB. B.5: Fréquences alléliques du locus 1A10 pour chaque populations étudiée.

1A10	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
N_{gen}	66	66	70	60	66	66	6	74	78	58	72	72	66	72	54	50	20	74	80	74	66	
N_{all}	19	32	22	15	24	10	2	17	11	6	10	11	18	17	20	10	12	32	25	31	28	17.71
1	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
2	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
3	0.00	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
5	0.06	0.11	0.21	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02
6	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00
7	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.09	0.00	0.20	0.16	0.26	0.04	0.00	0.04
8	0.06	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01
9	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00
10	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.07	0.01	0.01
11	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.02	0.00	0.15	0.01	0.07	0.01	0.01	0.02
12	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.08	0.01
13	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.00
14	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
15	0.12	0.04	0.13	0.02	0.06	0.08	0.50	0.04	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.08	0.04	0.06
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00
17	0.00	0.11	0.00	0.00	0.01	0.00	0.00	0.00	0.04	0.15	0.17	0.00	0.06	0.00	0.00	0.08	0.05	0.05	0.09	0.04	0.00	0.04
18	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
19	0.00	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.01	0.04	0.15	0.00	0.01
20	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
21	0.24	0.03	0.01	0.03	0.03	0.06	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.04	0.04	0.01	0.01	0.03
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	0.01	0.03	0.00	0.15	0.03	0.21	0.50	0.01	0.24	0.29	0.06	0.00	0.12	0.00	0.00	0.02	0.10	0.05	0.04	0.08	0.01	0.09
24	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.01	0.00	0.01	0.00	0.01
25	0.03	0.06	0.01	0.00	0.04	0.01	0.00	0.00	0.04	0.00	0.18	0.00	0.14	0.00	0.02	0.00	0.05	0.07	0.03	0.05	0.09	0.04
26	0.04	0.03	0.16	0.02	0.01	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.01	0.00	0.02
27	0.12	0.01	0.00	0.23	0.15	0.30	0.00	0.11	0.01	0.00	0.00	0.00	0.01	0.03	0.11	0.00	0.05	0.00	0.04	0.05	0.03	0.06

TAB. B.6: Fréquences alléliques du locus 1A10 pour chaque populations étudiée (*suite*).

1A10	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
28	0.01	0.03	0.07	0.08	0.27	0.03	0.00	0.08	0.00	0.00	0.00	0.29	0.00	0.10	0.00	0.00	0.00	0.08	0.01	0.01	0.00	0.05
29	0.03	0.00	0.03	0.00	0.03	0.14	0.00	0.08	0.31	0.24	0.24	0.06	0.03	0.04	0.07	0.08	0.05	0.01	0.10	0.05	0.04	0.08
30	0.01	0.01	0.00	0.07	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.15	0.00	0.01	0.06	0.22	0.05	0.05	0.03	0.00	0.00	0.03
31	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.06	0.09	0.01	0.03	0.03	0.07	0.09	0.00	0.05	0.01	0.00	0.01	0.01	0.02
32	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.07	0.00	0.07	0.11	0.02	0.00	0.03	0.00	0.00	0.00	0.02
33	0.06	0.01	0.03	0.10	0.00	0.12	0.00	0.08	0.00	0.00	0.03	0.03	0.01	0.01	0.02	0.00	0.00	0.04	0.01	0.04	0.06	0.03
34	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01
35	0.06	0.00	0.04	0.10	0.09	0.01	0.00	0.11	0.01	0.00	0.06	0.08	0.11	0.43	0.04	0.32	0.00	0.00	0.05	0.01	0.08	0.08
36	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
37	0.01	0.00	0.00	0.02	0.01	0.03	0.00	0.05	0.19	0.17	0.18	0.12	0.09	0.07	0.07	0.06	0.15	0.01	0.00	0.01	0.04	0.06
38	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.04	0.05	0.07	0.11	0.15	0.00	0.13	0.00	0.00	0.00	0.00	0.01	0.00	0.03
39	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.03	0.00	0.00	0.00	0.01
40	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.01	0.00	0.04	0.01	0.02	0.04	0.00	0.00	0.03	0.03	0.04	0.01
41	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.06	0.00	0.02	0.02	0.00	0.00	0.00	0.03	0.06	0.01
42	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.00
43	0.00	0.01	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.06	0.01
44	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00
45	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
46	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.04	0.01
47	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
48	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.00
49	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
50	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
51	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.00	0.00	0.00
52	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
53	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00
54	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
55	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
56	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00

TAB. B.7: Fréquences alléliques du locus 1A10 pour chaque populations étudiée (*suite et fin*).

1A10	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Moy
57	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
58	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
59	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
60	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
H_O	0.79	0.88	0.77	0.80	0.76	0.55	0.33	0.59	0.90	0.59	0.78	0.67	0.88	0.44	0.74	0.76	0.90	0.76	0.95	0.89	1.00	0.75
H_E	0.90	0.97	0.91	0.89	0.89	0.83	0.60	0.94	0.81	0.81	0.85	0.86	0.92	0.80	0.94	0.83	0.93	0.95	0.91	0.95	0.96	0.88

TAB. B.8: Moyennes et écart-types de N_{all} , H_O et H_E pour chaque population.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
$\bar{N}_{all}$	12.75	14.25	14.00	10.00	12.00	7.75	2.50	9.75	7.00	7.50	7.00	8.00	10.50	10.00	11.25	7.25	8.00	16.00	14.50	18.25	13.25
S.D.	7.23	12.12	8.29	4.24	8.29	4.03	0.58	4.92	3.56	5.75	2.45	4.08	6.56	5.23	6.70	2.75	3.37	12.57	9.47	8.99	10.37
$\bar{H}_O$	0.77	0.71	0.70	0.76	0.70	0.60	0.58	0.65	0.79	0.62	0.67	0.61	0.75	0.56	0.75	0.69	0.64	0.71	0.79	0.88	0.62
S.D.	0.15	0.12	0.06	0.13	0.12	0.20	0.32	0.04	0.12	0.10	0.10	0.14	0.18	0.08	0.15	0.13	0.25	0.17	0.16	0.06	0.32
$\bar{H}_E$	0.84	0.81	0.83	0.79	0.78	0.69	0.63	0.78	0.71	0.72	0.71	0.77	0.75	0.69	0.83	0.73	0.83	0.82	0.82	0.88	0.67
S.D.	0.09	0.11	0.06	0.10	0.10	0.15	0.07	0.12	0.13	0.14	0.15	0.10	0.18	0.08	0.11	0.09	0.08	0.14	0.10	0.06	0.21

B.2 Données mitochondriales brutes

Séquence complète de l'hapotype H_1 , majoritaire dans notre jeu de données.

```

1 AAAGTATTGT AATTGCTCCA GCCAATACCG GTAAAGATAG AAGCAATAAA ATAGCTGTAA
61 TTTTACAGA TCATACAAAT AAAGGAAGAC GCTCAAATTG TATTCCTCGT CATCGTATAT
121 TTACAATAGT TGTAATAAAA TTTACAGCTC CTAAAATTGA AGACACACCT GCAAGGTGTA
181 AAGAAAAAAT AGCTAAATCA ACTGAACCAC CGGCGTGAGC TAAATTAGCG GATAGAGGAG
241 GATAAACTGT TCACCCTGTT CCTGCCCCCTC TTTCAACTGC AGCTGAAGAT AGTAAAAGGA
301 GTAAAGCAGG AGGAAGCAAT CAAAATCTTA TATTATTTAA TCGAGGAAAA GCTATATCAG
361 GAGCTCCTAA TATTAAAGGA ACTAATCAGT TTCCAAAGCC TCCAATTATT ATTGGTATTA
421 CTAAGAAAAA AATTATTACA AAAGCATGAG CTGTTACAAT TACATTGTAA AGTTGGTCAT
481 CACCAAGAAG TGCCCCAGGC TGTCCAAGTT CAGCACGAAT AAGAAGTCTT AAA

```

TAB. B.9: Position des sites variables des 39 haplotypes de *Cyclope neritea*. Les tirets correspondent au même état de caractère que l'haplotype H_1 . L'effectif total par haplotype est indiqué (N).

[illegible]

TAB. B.10: Nombre d'haplotypes par populations. Les populations sont : St : Stérec, Sé : Séné, Riv : Rivedoux Ro : Ronce, Ar :Arguin, Hos : Hossegor, Ped : Pedreña, SVB, San Vicente de la Barquera, Alv : Alvor, Far : Faro, Cad : Cadiz, MM : Mar Menor, Eb : Ebre, Leu, Leucate, Th : Thau, Sar : Sardaigne, Sic : Sicile, Cao : Caorle, Ven : Venise, Kat : Katerinie, Efo : Eforie

	St	Sé	Riv	Ro	Ar	Hos	Ped	SVB	Alv	Far	Cad	MM	Eb	Leu	Th	Sar	Sic	Cao	Ven	Kat	Efo
H_1	9			8	9	28	3	18	39	31	36	24	33			25					
H_2												3									
H_3					3																
H_4												9									
H_5						5															
H_6		1																			
H_7		2																			
H_8		1																			
H_9		1																			
H_{10}		4																			
H_{11}		1																			
H_{12}		1																			
H_{13}		1																			
H_{14}		1																			
H_{15}		1																			
H_{16}		4		1										1					3		
H_{17}	5	3	8															16	3		
H_{18}			9	13	14			19				1		34	29						
H_{19}					6																
H_{20}		1																			
H_{21}		1																			
H_{22}	19	8	12	10	1									2				20	25		
H_{23}		1																			
H_{24}																		1	6		
H_{25}																		1			
H_{26}																	3				
H_{27}																	4				
H_{28}																		1			
H_{29}			6															1	1		
H_{30}																					26
H_{31}																				1	
H_{32}																				1	
H_{33}																				20	
H_{34}																				1	
H_{35}																				1	
H_{36}																				2	
H_{37}																				2	
H_{38}																				1	
H_{39}																				3	

C

Présentation du programme réalisé sous
Delphi.

Le mode de fonctionnement du programme de coalescence est présenté à la figure C.5. Ce programme s'articule autour d'unités comprenant des procédures et fonctions spécifiques, chaque unité effectuant une tâche particulière. L'affichage de la boîte de dialogue de saisie (voir les exemples présentés aux figures C.1 et C.2) est par exemple géré par UNITE 1, et permet à l'utilisateur de n'avoir accès qu'aux paramètres pertinents pour le type de simulation choisi.

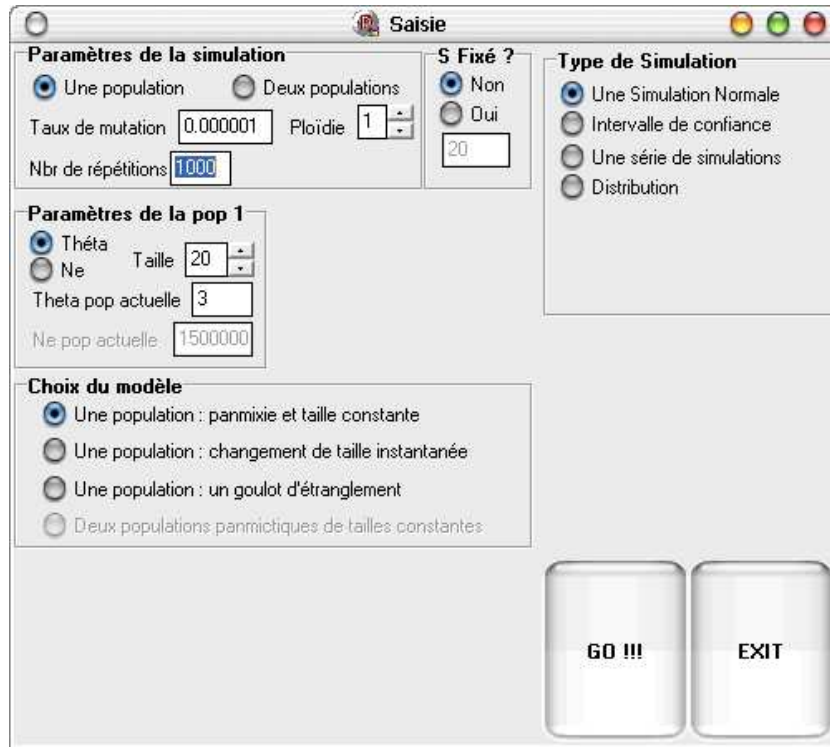


FIG. C.1: Exemple d'affichage de la fenêtre de saisie dans le cas d'une simulation impliquant une population panmictique de taille constante.

Dans le cas le plus simple (*i.e.* une seule population simulée, panmictique et de taille constante), l'utilisateur fixe tout d'abord le taux de mutation μ , le degré de ploïdie et le nombre de répétitions (plus ce nombre sera grand, plus le temps de calcul sera long, mais plus l'estimation des valeurs d'indices statistiques sera précise). L'utilisateur fixe ensuite les paramètres propres à la population simulée : la taille de l'échantillon N et la taille efficace de la population N_e (θ et N_e étant reliés par la relation $\theta = xN_e\mu$, x étant le degré de ploïdie, l'utilisateur peut choisir de fixer un paramètre ou l'autre).

Lorsque la simulation est lancée, une procédure détermine tout d'abord la marche à suivre (*i.e.* le type de résultats à collecter) en fonction du type de simulation choisi (cadre "Type de simulation" de la figure C.1). Une autre procédure permet de collecter les paramètres entrés par l'utilisateur dans la boîte de dialogue. La reconstruction des

FIG. C.2: Exemple d'affichage de la fenêtre de saisie dans le cas d'une simulation impliquant une populations subissant un goulot d'étranglement. L'utilisateur souhaite obtenir l'intervalle de confiance du nombre d'haplotypes observés dans l'échantillon de 20 gènes issus d'une population ayant subi ce goulot.

généalogies de gènes commence alors et suit les étapes décrites à la figure 9.2, page 113 :

1. Construction de l'arbre par tirage des temps de coalescence dans une loi exponentielle (unité ARBO)
2. Tirage, pour chaque branche, du nombre de mutations dans une loi de Poisson (unité MUTE)
3. Calcul des différents indices pour chaque généalogie de gènes et stockage en vue de calculs de moyennes, de distributions, d'intervalles de confiance... (unité CALCUL).

Enfin, à l'issue du nombre de répétitions, les indices calculés sont affichés dans une boîte de dialogue de sortie (voir les exemples présentés aux figures C.3 et C.4).

Au cours de cette thèse, nous avons essentiellement modifié les unités ARBO et CALCUL. Le calcul du test K a par exemple nécessité la programmation du calcul de l'intervalle de confiance du nombre d'haplotypes. Pour réaliser ce test, l'unité MUTE fonctionne différemment : les mutations ne sont pas tirées au hasard dans une loi de probabilité, car leur nombre total est connu : il est fixé à la valeur réelle observée dans la population réelle étudiée. L'unité se contente alors de répartir ces mutations dans l'arbre en tenant compte

The 'Resultat' window displays the following data:

Segragating Sites	Nombre d'allèles	Diversité Haplotypique	Pairwise Difference
Moyenne: 10.516	Observé Moyenne: 6.51401	Observé Moyenne: 0.74699	Moyenne: 2.97744
Variance: 23.91185	Variance: 3.4635	Variance: 0.01864	Variance: 3.39803
D de Tajima Valeur: -0.06229	Attendu Moyenne: 6.57244	Attendu Moyenne: 0.75	Moy des var de chaque simulation: 8.57829
Variance: 0.8711			

Buttons: Init, Fermer, Exit

FIG. C.3: Indices statistiques calculés dans le cas d'une simulation simple.

The 'Resul.' window shows the 'Intervalle de Confiance' section with the following inputs:

Intervalle de Confiance	
β	10
P Value	

FIG. C.4: Exemple de résultats obtenus lorsque l'utilisateur souhaite calculer l'intervalle de confiance à 95% d'un indice statistique.

de la longueur respective des différentes branches.

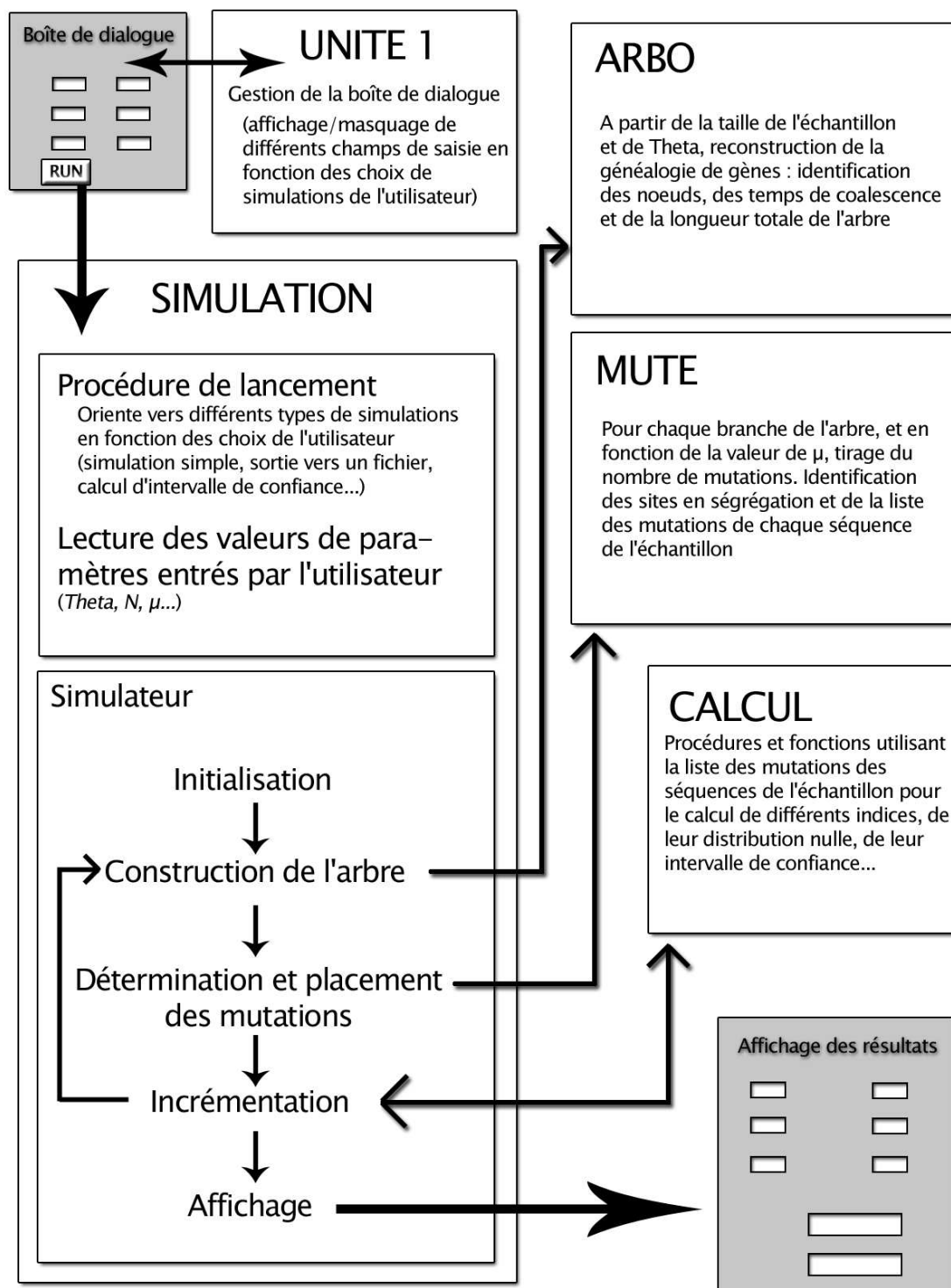


FIG. C.5: Diagramme schématisant le mode de fonctionnement du programme de coalescence. D'après Muller (2002).

D

Articles publiés

D.1 Marine Ecology Progress Series

Invasion of the eastern Bay of Biscay by the nassariid gastropod *Cyclope neritea*: origin and effects on resident fauna

Guy Bachelet^{1,*}, Benoît Simon-Bouhet^{2,3}, Céline Desclaux¹,
Pascale Garcia-Meunier², Guillaume Mairesse¹, Xavier de Montaudouin¹,
Hélène Raigné¹, Karine Randriambao¹, Pierre-Guy Sauriau⁴, Frédérique Viard³

¹Laboratoire d'Océanographie Biologique, UMR 5805 CNRS-Université Bordeaux 1, Station Marine d'Arcachon, 2 rue du Professeur Jolyet, 33120 Arcachon, France

²Laboratoire de Biologie et Environnement Marins, EA 3168, Université de La Rochelle, 17000 La Rochelle, France

³Evolution et Génétique des Populations Marines, UMR 7127 CNRS, Station Biologique de Roscoff, BP 74, 29682 Roscoff cedex, France

⁴Centre de Recherche sur les Ecosystèmes Marins et Aquacoles, UMR 10 CNRS-IFREMER, Place du Séminaire, BP 5, 17137 L'Houmeau, France

ABSTRACT: The distribution area of the nassariid gastropod *Cyclope neritea* (L.) includes the Mediterranean and the Black Sea, as well as the Atlantic coasts of the southern Iberian Peninsula. The species has spread north to the eastern Bay of Biscay (Arcachon Bay in 1976, Marennnes-Oléron area and Morbihan Gulf in 1983–84). This spread might be explained either by (1) a natural spread favoured by environmental changes (e.g. an increase of temperature) or (2) a sudden range expansion due to the introduction of individuals from distant native populations. Molecular genetic analyses based on mitochondrial markers suggest that the present *C. neritea* population in Arcachon Bay has been introduced, probably unintentionally with oyster transfers, from several source populations, genetically similar to the populations analysed in this study, i.e. those in the western Mediterranean and in south Portugal. Within its new distribution area, *C. neritea* could potentially compete with the autochthonous nassariid *Nassarius reticulatus*, both species being scavengers. Although *C. neritea* tends to occur mainly in relatively clean sands in the intertidal and *N. reticulatus* in subtidal, organic rich sediments, the habitats of both species partially overlap. Laboratory experiments showed that in still water conditions *C. neritea* was more active and reached carrion faster than *N. reticulatus*, thereby having a competitive advantage over the latter; flow conditions ($\sim 1 \text{ cm s}^{-1}$) appeared to stimulate the activity of *N. reticulatus*. Analysis of parasite load in both species in Arcachon Bay indicated that *N. reticulatus* was more heavily parasitized by digenean trematodes than *C. neritea*. This suggests that the spread and population dynamics of *C. neritea* along the French Atlantic coast has been favoured by the lack of parasites. Altogether, recurrent introduction, competitive ability and lack of heavy parasitic load might explain the successful settlement of *C. neritea* along the French Atlantic coast.

KEY WORDS: Introduced species · Scavenging gastropods · Population genetics · Interspecific competition · Behaviour · Parasites

— Resale or republication not permitted without written consent of the publisher —

INTRODUCTION

The introduction of alien species has been recognized as a major threat to marine ecosystems, with a set of ecological and evolutionary consequences from species

to ecosystem levels (Grosholz 2002). In this respect, introduced species are considered one of the most serious threats to the conservation of natural biodiversity (Coblentz 1990, Cohen & Carlton 1998, Mooney & Cleland 2001, Palumbi 2001, Shea & Chesson 2002)

*Email: g.bachelet@epoc.u-bordeaux1.fr

and are identified as the second cause (after habitat destruction by human activities) of the loss of biological diversity (Vitousek et al. 1997, Wilcove et al. 1998). As recently emphasized by Ruiz et al. (2000), multiple approaches, although rare, are required to understand the patterns and processes of marine biological invasions. For example, to identify the proximate factors for the success or the failure of invasion, studies of life history traits, species interactions (including the consequences on resident species) and the genetic architecture of invasive populations are complementary (Parker et al. 1999, Ruiz et al. 2000, Sakai et al. 2001, Grosholz 2002, Lee 2002).

Dukes & Mooney (1999), Carlton (2000) and Stachowicz et al. (2002) have pointed out the likely relationships between global and climate change and the success of biological invasions. In this context, the nassariid gastropod *Cyclope neritea* (L., 1758) is an interesting model species. This gastropod is 1 of the 87 marine and brackish water species that are known or suspected to have been accidentally introduced to the Atlantic and Channel coasts of France (Gouilletquer et al. 2002). *C. neritea* is native to the Mediterranean and Black Seas and to the Atlantic coasts of southern Spain and Portugal as far north as Setubal (Hidalgo 1917, Nobre 1931). Up to the early 1970s, discontinuous populations were also recorded in the north of Spain and on the French Basque coast (Kisch 1950, Morton 1960; see also Sauriau 1991 for a thorough analysis of occurrence records of *C. neritea* on the Atlantic coasts). New *C. neritea* populations appeared on the French Atlantic coasts during the 1970s and 1980s: (1) in 1976 in Arcachon Bay (Bachelet et al. 1980), (2) in 1983–84 further north, in the Marennes-Oléron Bay (Pigeot 1988, Sauriau 1989), at the Isle of Ré (Tardy et al. 1985), and in the Gulf of Morbihan (Le Roux et al. 1988). Arcachon Bay and the Gulf of Morbihan are located approximately 150 km and >500 km, respectively, north of the early 1970s range of the species. It should be noted that, in all these newly colonized areas, *C. neritea* populations are now self-reproducing and abundant, with densities >100 ind m⁻² (Tardy et al. 1985, Bachelet et al. 1990, Le Roux 1994, Afli & Chenier 2002). A large *C. neritea* population was also observed in 2003 in the Bay of Morlaix, English Channel (B. Simon-Bouhet & F. Viard, pers. obs.).

Two non-exclusive hypotheses can be put forward to explain the present distribution of *Cyclope neritea* along the French Atlantic coasts: (1) a natural range expansion towards the north due to environmental (e.g. climatic) changes, and/or (2) accidental introductions related to shellfish culture (e.g. oyster transfers between the Mediterranean and the Bay of Biscay). As suggested by Sauriau (1991), both processes may have been responsible, since the aforementioned new

populations of *C. neritea* have potentially benefited from a slight warming of coastal waters during the period 1970 to 1990. However, when a species is introduced close to its natural distribution area, it is difficult to discriminate both processes using field survey only. We addressed this issue by sequencing a mitochondrial DNA gene fragment (the Cytochrome C Oxidase I, COI), which is known to be a powerful tool in determining the origin and history of new populations (Sakai et al. 2001) and can trace back historical processes at within-species level (Avise et al. 1987, Avise 1991).

Competitive interactions between native and introduced species, leading to fitness reduction in native populations, have been experimentally demonstrated in marine species and found to be a major determinant of the success of an invasion (see e.g. reviews by Carlton 1992, Ruiz et al. 1999, Grosholz 2002). In European coastal waters an autochthonous species with which *Cyclope neritea* might compete is the netted dogwhelk *Nassarius reticulatus* (L.), as both species belong to the same family and feed as scavengers. We made field observations to determine if both species when living in sympatry occupy a similar habitat, and performed laboratory experiments to determine if there was competition for food between the 2 nassariids.

It has also been suggested that parasite release is important in allowing an introduced species to become invasive (Calvo-Ugarteburu & McQuaid 1998b). Parasitic infection in introduced species is mainly limited by 2 fundamental principles (Torchin et al. 2002): (1) the intensity and prevalence of infection tend to increase with host population density, which suggests that in a recently introduced host-parasite system, host density will often be too low for the parasite to establish and (2) parasites, in particular digenean trematodes that exhibit high prevalence in molluscs (Lauckner 1980), often have a complex life-cycle, including 2 or more host species. Consequently, many introduced species lack most or all of their native natural parasites (Calvo-Ugarteburu & McQuaid 1998a, Torchin et al. 2001), which is advantageous when in potential competition with native species. In the present work we compared the parasite prevalence and richness between *Cyclope neritea* and *Nassarius reticulatus*. The disadvantage in being parasitized was assessed through a survival experiment.

Based on an integrative and pluridisciplinary approach, the aims of this study are therefore to give insights about the history of the introduction, together with the identification of biological factors that may have promoted the settlement and then the maintenance of an exotic marine species, recently introduced along the European North Atlantic coast.

MATERIALS AND METHODS

Study species and area. *Cyclope neritea* is a small mollusc (~15 mm maximum shell width) which lives in shallow and intertidal habitats, predominantly in calm waters, where it usually lies buried a few mm below, or crawls at, the sediment surface. Like most members of the family Nassariidae, it is predominantly a scavenger, feeding on carrion, especially dead bivalves (Morton 1960, Britton & Morton 1994), though it may also behave as a deposit feeder (Southward et al. 1997). *C. neritea* has no dispersive larval stage: the embryos undergo direct development during 2 to 5 wk within egg capsules (each capsule usually containing a single embryo) deposited on hard substrata such as cockle shells, then hatch as benthic juveniles (Kisch 1950, Gomoiu 1964, Massé et al. 1978, Boulhic & Tardy 1986, D'Asaro 1993). The first *C. neritea* population observed outside its common native range (Mediterranean, Black Sea and Atlantic coasts of the Iberian Peninsula) was in Arcachon Bay, where it has been monitored since 1976. We focused on this particular population to develop several parallel studies (population genetics, habitat, competition and parasitic load).

Sampling, DNA extraction, amplification and sequencing. Adult specimens of *Cyclope neritea* were collected from 3 intertidal locations (Table 1): (1) Faro (S Portugal) and (2) Thau Lagoon (French Mediterranean), which are both located within the natural range of the species; these populations may have contributed to the settlement of *C. neritea* in France; and (3) Arcachon Bay (site 'Arguin'), where the first new population of *C. neritea* on the French Atlantic to the north of the Basque coast was recorded. The specimens were attracted with a dead fish, collected, removed from their shell and stored individually in 95% ethanol before DNA extraction. Total DNA was extracted from <15 mg of foot muscle using DNeasyTM Tissue Kit according to the manufacturer's protocol (Quiagen). From preliminary sequences carried out with the universal primers originally defined by Folmer et al. (1994) (i.e. LCO1491 and HCO2198) to amplify the COI gene, we designed a

new set of internal primers more specific to *C. neritea*: Cy2 5'-GTAAAATTTTCGATCTGTTA-3' (forward) and CyR 5'-GGATTAGTTGGTACAGC-3' (reverse). The amplifications were performed using these primers in 50 µl reactions containing 5 to 50 ng DNA, 0.5 µM of each primer, 0.2 mM dNTPs, 1.5 mM MgCl₂, 1× Reaction Buffer (containing 50 mM Tris HCl pH 8.8 at 25°C, 200 mM (NH₄)₂SO₄, ABgene) and 1.25 U Red Hot[®] DNA Polymerase (ABgene). A MJResearch PTC 100 Thermal Cycler was used with a cycling profile as follows: initial denaturation step at 93°C (3 min) followed by 40 cycles at 93°C (30 s), 53°C (30 s), 72°C (60 s) and a final extension step at 72°C (10 min). Double stranded PCR products were cleaned using MultiScreen-PCR MANU03010 plates (Millipore) and cycle-sequenced using ABI PRISM[®] BigDyeTM Terminators v3.0 Cycle Sequencing Kit (Heiner et al. 1998) following the manufacturer's protocol (Applied Biosystems). Unincorporated BigDye terminators were removed using MultiScreen MAHVN4510 plates (Millipore). In order to minimize sequencing errors, both strands were sequenced for each individual using an ABI PRISM[®] 3100 Automated DNA Sequencer (Perkin-Elmer Applied Biosystems). Sequence data were aligned using ClustalW (Thompson et al. 1994) and ambiguities were checked comparing each sequence with its complementary fragment using BioEdit (Hall 1999).

Sequence analysis. A haplotype is defined by a unique stretch of nucleotides. Haplotype diversity H_e (Nei 1987) and average pairwise difference K (i.e. average number of nucleotide differences between pairs of sequences) (Tajima 1983) were calculated for each population using the software DnaSP (Rozas & Rozas 1999). Haplotype frequencies and pairwise F_{ST} values (calculated from frequency values; Weir & Cockerham 1984) were estimated using Arlequin v.2.0 (Schneider et al. 2001). Using a permutation procedure as implemented in Arlequin v.2.0, we assessed the statistical significance of pairwise F_{ST} values under the null hypothesis of no difference between populations.

For polymorphic populations, the mismatch distribution (i.e. the distribution of the observed number of pairwise nucleotide site differences) was computed in DnaSP, in order to examine the occurrence of mixing of evolutionary divergent haplotype groups at the population level. This analysis was used to put forward hypotheses about the number of genetically differentiated populations at the origin of the establishment of a given population.

Finally, phylogenetic analyses were performed with the software Network

Table 1. *Cyclope neritea*. Sampling sites (geographic co-ordinates) and molecular diversity of populations. The number of individuals sequenced (N), the number of haplotypes (H), the haplotype diversity (H_e) and the average pairwise differences (K) are given for each studied population and for the entire data set

	Latitude	Longitude	N	H	H_e	K
Thau Lagoon	43° 27' N	3° 38' E	29	1	0.000	0.000
Faro (Rio Formosa)	37° 00' N	7° 58' W	31	1	0.000	0.000
Arcachon Bay (Arguin)	44° 35' N	1° 14' W	33	5	0.725	6.576
Total	-	-	93	5	0.602	6.344

which builds haplotypic networks from free-recombinant data, based on the median-joining algorithm (Bandelt et al. 1999). This process combines both minimum spanning trees and maximum parsimony approaches. Such networks show the evolutionary relationships (based on mutational events) between the different haplotypes observed over the whole data set.

Habitat preferences. The habitat, in terms of bathymetric distribution and sediment properties, of *Cyclope neritea* and *Nassarius reticulatus* was assessed in Arcachon Bay in March 2002. In this macrotidal lagoon, 5 bathymetric transects were performed, with 4 to 10 sampling stations on each transect. Stations were distributed over a tidal range from -8.8 to $+3.1$ m. At each station, a trap (diameter 26 cm), made of a plastic bag with mesh openings of 4 mm on the bottom (to prevent loss of nassariids when raised to the surface) and 15 mm on the sides (to allow nassariids to enter the trap), was immersed at high tide and deposited on the sea bottom. A sampling unit was defined as a trap containing 20 g fresh weight of bait (dead *Crepidula fornicata*) immersed for 30 min. Seven replicates were taken at each station. Nassariids of both species were counted in each sampling unit. At each station, surface sediment was collected with an Ekman grab and analysed for grain size distribution (wet sieving) and organic matter content (loss of weight of dry sediment at 550°C during 2 h).

Experiments on interspecific competition for food. Laboratory experiments were designed to investigate the behaviour and the chemoreception ability of *Cyclope neritea* and *Nassarius reticulatus*. These experiments were performed in a long straight PVC channel ($160 \times 7 \times 3.5$ cm) without sediment and filled to a depth of 3 cm with seawater. No attempt was made to control environmental parameters; however, the experiments were run during the day under natural lighting, and within a short period (16 to 20 May) during which conditions remained largely constant. During the experiments, water temperature and salinity ranges were 17.7 to 18.9°C and 24 to 28‰, respectively. Nassariids were collected in Arcachon Bay and maintained without food in aquaria with running oxygenated seawater for 2 months prior to the experiments. Similar-sized adult *C. neritea* (shell width = 11 to 15 mm) and *N. reticulatus* (shell height = 18 to 27 mm) were used. All experiments were carried under 2 different hydrodynamic conditions: (1) in still water, and (2) in flow conditions, where sea-water flowed unidirectionally at a surface velocity of 0.81 to 0.94 cm s^{-1} .

The first set of experiments was conducted without any prey in order to observe the spontaneous behaviour of both nassariid species. Individuals of

the 2 species (4 *Cyclope neritea* and 4 *Nassarius reticulatus*; 8 replicates) were arranged in a line at an equal distance from both ends of the channel. The number of individuals moving spontaneously (in still water and flow conditions) over a distance of at least 10 cm in 30 min was recorded.

In a second set of experiments we measured the ability of both nassariid species to locate and reach a prey (a dead cockle *Cerastoderma edule*). Individuals of the 2 species (4 *Cyclope neritea* and 4 *Nassarius reticulatus*; 8 replicates in still water, 4 replicates in running seawater) were placed at one end of the channel and allowed to feed upon the carrion which was placed at distances of 20, 40 or 60 cm. In flow conditions, seawater flowed unidirectionally towards both the carrion and the nassariids. The times at which each individual reached the prey were recorded and their speed was calculated. Individuals that did not reach the carrion within 30 min were assigned zero.

Survey of parasite abundance. From October 2000 to July 2001, sympatric populations of *Nassarius reticulatus* and *Cyclope neritea* were collected monthly by hand on a sandflat ('Arguin') in Arcachon Bay. In the laboratory, 100 snails of each species were gently crushed and dissected under a stereomicroscope. Mean prevalence (i.e. percentage of infected individuals; Bush et al. 1997) was calculated for each parasite species or family.

Survival experiment in relation to parasites. Due to the low parasite infection in *Cyclope neritea* (see 'Results'), the survival experiment was performed with *Nassarius reticulatus* only. Survival of infected and non-infected *N. reticulatus* individuals was compared to assess the disadvantage of being parasitized for the native species. A large quantity of snails was collected at 'Arguin', and subsequently isolated in dishes with seawater at $\sim 24^{\circ}\text{C}$. Each dish was observed under a stereomicroscope to detect digenean larvae (cercariae). Nassariids that did not shed parasites were not always free of infection (Curtis & Hubbard 1990). Therefore, a higher number of control snails ($\times 3$) were kept in order to check non-infection when individuals died. Parasitized and control snails had the same size structure. A total of 158 snails (36 parasitized and 122 control) were gathered in a large aquarium, in running seawater at ambient temperature. Each individual was marked to calculate its lifespan during the experiment (total duration: 280 d). Snails were not fed in order to weaken them and exacerbate parasite detrimental effects. Although unrealistic, being starved equalized all gastropods at the same 'diet' and prevented them from trophic infestation (de Montaudouin et al. 2003).

RESULTS

Mitochondrial DNA (mtDNA) polymorphism within and among native and introduced *Cyclope neritea* populations

PCR amplification yielded a 553 base pair (bp) long fragment of the COI gene. Nineteen polymorphic sites were found in the 3 populations, and out of the entire sample ($N = 93$), 5 different mtDNA haplotypes were identified (Table 2). Over the whole data set, 2 haplotypes were found at very high frequency (f), namely A ($f = 0.462$) and B ($f = 0.430$). Haplotypes C, D and E were found more rarely, with frequencies of 0.065, 0.032 and 0.011, respectively. At the population level, Thau and Faro populations revealed only 1 haplotype (A and B, respectively), which was in contrast to the Arcachon population where the 5 haplotypes (A to E) were found, with A and B representing about 70% of the total number of sequences for this population (see haplotype distribution in Fig. 1). Consequently, at Arcachon, the haplotype diversity was very high ($H_e = 0.725$, Table 1) as well as the average pairwise difference ($K = 6.576$, Table 1). This high value of K was in agreement with the mismatch distribution from which 2 main peaks were observed, 1 of them centred on values of 13 to 14 nucleotide differences (Fig. 2). Such values are indicative of mixing of groups of individuals characterized by evolutionary divergent haplotypes within the Arcachon population. This pattern can be related to the evolutionary relationships between the observed haplotypes, as shown in the haplotypic network (Fig. 3). From this network, 3 major clusters were identified: (1) Haplotypes A and C, (2) Haplotypes B and D, and (3) Haplotype E. Haplotypes were highly divergent among these groups (13 mutations between A

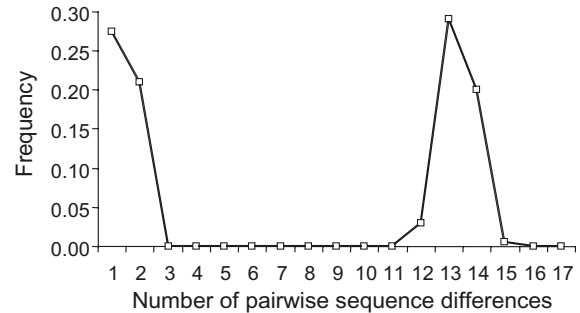


Fig. 2. *Cyclope neritea*. Mismatch distribution for the population of Arcachon Bay. The frequency of each observed value of pairwise difference is plotted

and B, 11 between A and E, and 13 between B and E; Fig. 3) but haplotypes were very close within each group (2 mutations steps between A and C, 1 between B and D; Fig. 3).

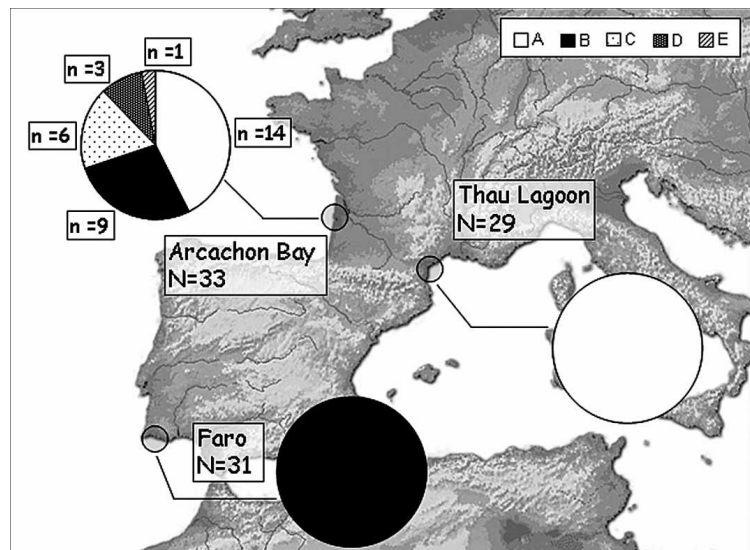


Fig. 1. *Cyclope neritea*. Haplotype distribution within the 3 study populations. Total number of individuals sequenced in each population (N) and number of each haplotype in the population of Arcachon Bay (n) are indicated. Haplotypes A to E are distinguished by different patterns

Table 2. *Cyclope neritea*. Definition of Haplotypes A to E found within the 3 populations of Thau, Faro and Arcachon. Positions of the polymorphic sites are given. '-' indicates that the same character is present in Haplotype A

	Positions (bp)																		
	1	44	50	53	104	218	224	254	266	292	317	347	377	382	398	422	425	479	503
A	A	T	G	G	T	A	A	T	T	G	T	G	G	C	C	C	G	G	C
B	-	C	A	A	-	-	-	C	C	-	C	A	A	-	G	T	-	A	T
C	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-
D	-	C	A	A	-	-	-	C	C	-	C	A	A	A	G	T	-	A	T
E	G	-	-	A	C	G	G	C	C	A	C	-	-	-	-	-	A	A	-

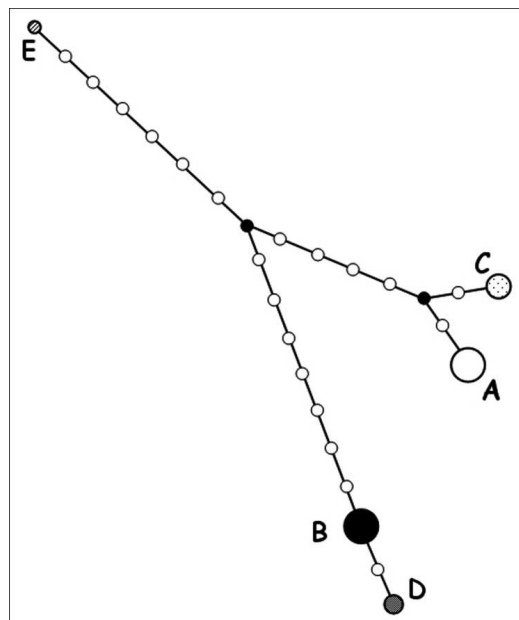


Fig. 3. *Cyclope neritea*. Haplotypic network showing the phylogenetic relationships between the haplotypes. (○) Single mutational events. (●) Missing haplotypes between observed haplotypes in the data set

Pairwise F_{ST} values were all significant ($p < 10^{-5}$) indicating a large and significant genetic structure among the 3 populations. Because of the lack of common haplotypes, the highest (and maximum) genetic differentiation was found when comparing the populations of Thau and Faro ($F_{ST} = 1$). High genetic differentiation levels were also found when comparing the populations of Arcachon and Faro ($F_{ST} = 0.58$) and the populations of Arcachon and Thau ($F_{ST} = 0.32$).

Habitat of *Cyclope neritea* and *Nassarius reticulatus*

In Arcachon Bay *Cyclope neritea* occurred at tidal levels between -0.8 and $+3.1$ m, while *Nassarius reticulatus* was found between -5.2 m and $+1.9$ m (observations at the 5 study transects). As an example, Fig. 4 shows the zonation pattern along the bathymetric transect at 'Arguin', where the abundance of *C. neritea* was the highest at intertidal levels and that of *N. reticulatus* increased with depth down to -4.0 m, with an overlap of both species distribution between -1.0 and $+1.3$ m. *C. neritea* did not occur in fine, organic-rich sediments, whereas *N. reticulatus* were found in all sediment types sampled at Arcachon (Fig. 5). *C. neritea* did not tolerate sediment organic content $> 1\%$, whereas *N. reticulatus* were still abundant in sediment with organic content as high as 14% .

Competition for food between *Cyclope neritea* and *Nassarius reticulatus*

When placed without any dead prey in the seawater channel, the individuals of both nassariid species either moved erratically or did not move at all, regardless of hydrodynamic conditions. Nevertheless, a 2-way ANOVA indicated that *Cyclope neritea* moved spontaneously in a higher proportion than *Nassarius reticulatus* in both hydrodynamic conditions ($p < 0.05$) and that this spontaneous moving was greater in flow conditions than in still water for both species ($p < 0.05$) (Table 3; Fig. 6). Moreover, while the proportion of *C. neritea* moving without the presence of carrion was almost the same in still water ($53.1 \pm 9.9\%$) and in flow ($59.4 \pm 11.5\%$), a higher proportion of *N. reticulatus* moved spontaneously in flow than in still water conditions ($53.1 \pm 11.0\%$ and $18.7 \pm 12.5\%$, respectively)

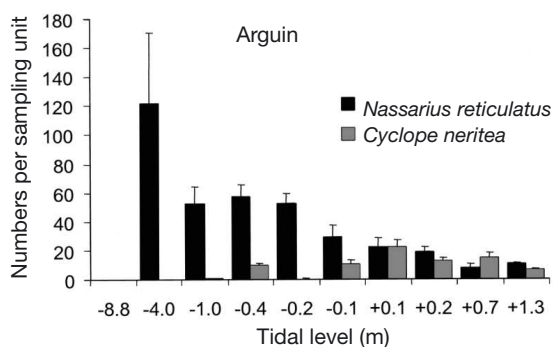


Fig. 4. *Cyclope neritea* and *Nassarius reticulatus*. Mean ($\pm$ SE) number of individuals of both species per sampling unit (see text for explanations), at various tidal levels along a transect at 'Arguin' (Arcachon Bay)

(Fig. 6) In the presence of carrion, the speed at which individuals of both species moved towards, and reached, food was higher in flow than in still water conditions (Fig. 7), due to chemical stimulus conveyed by water. Due to heteroscedasticity that prevented us from performing a 3-way ANOVA between hydrody-

Table 3. *Cyclope neritea* and *Nassarius reticulatus*. Results of a 2-way ANOVA, testing for the effect of species and hydrodynamic conditions on the spontaneous moving of individuals of both species when no carrion was available for 30 min. Data were previously arcsin-transformed and homogeneity of variances was checked by Cochran test. ns: $p > 0.05$; *: $p < 0.05$

Source of variation	df	MS	F
Species	1	0.848	4.330*
Hydrodynamism	1	1.027	5.244*
Species $\times$ Hydrodynamism	1	0.303	1.548 ns
Error	28	0.196	

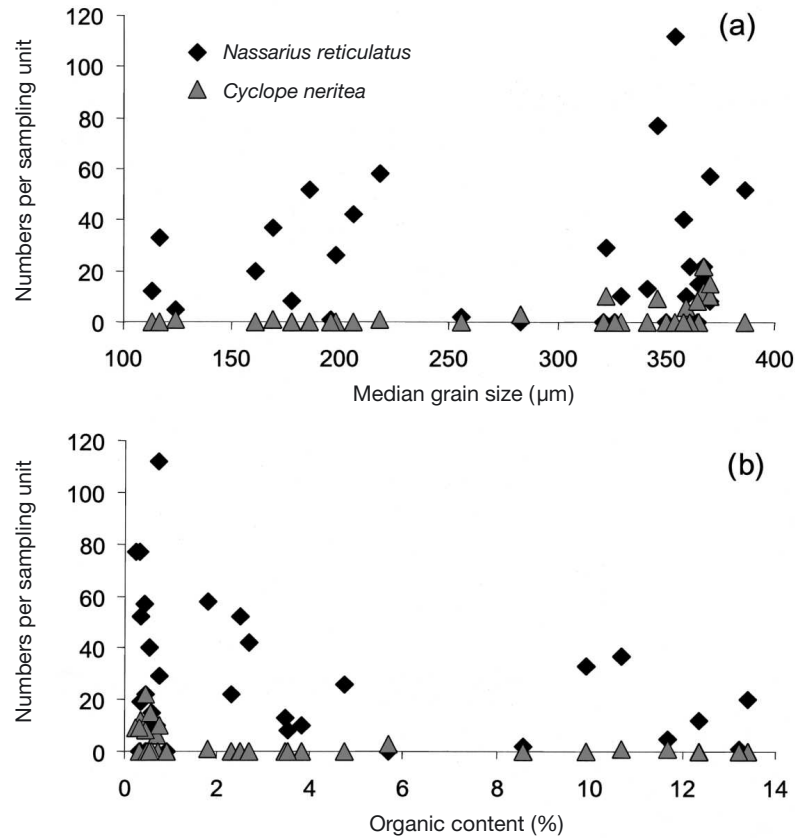


Fig. 5. *Cyclope neritea* and *Nassarius reticulatus*. Numbers of individuals of both species per sampling unit (see text for explanations) in relation to (a) median grain size and (b) organic content of the sediment at 35 stations located along 5 transects in Arcachon Bay

namic conditions, species and distance to the prey, the results were analysed separately for each hydro-dynamic condition by means of a 2-way ANOVA

(Table 4). In still water conditions, *N. reticulatus* took longer to arrive at carrion (i.e. speed was slower) than *C. neritea* ($p < 0.001$) and the speeds of both species

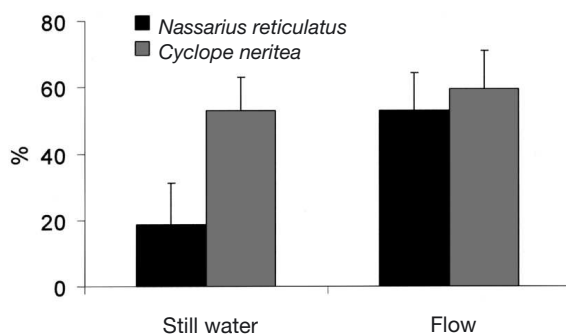


Fig. 6. *Cyclope neritea* and *Nassarius reticulatus*. Mean (+SE) percentage of individuals of both species moving spontaneously (i.e. without bait) over a minimum distance of 10 cm during 30 min, in laboratory experiments with still water or flow conditions

Table 4. *Cyclope neritea* and *Nassarius reticulatus*. Results of 2-way ANOVAs, testing for the effect of species and distance to the prey on the speed of individuals of both species to reach their prey (a) in still water and (b) in flow conditions. Data were previously $\log(x+1)$ transformed and homogeneity of variances was checked by Cochran test. ns: $p > 0.05$; **: $p < 0.01$; ***: $p < 0.001$

Source of variation	df	MS	F
(a) in still water conditions:			
Species	1	1.656	18.202***
Distance	2	3.022	33.214***
Species $\times$ Distance	2	0.043	0.471 ns
Error	281	0.091	
(b) in flow conditions:			
Species	1	36.741	9.949**
Distance	2	74.702	20.228***
Species $\times$ Distance	2	30.783	8.335***
Error	90	3.693	

decreased the further away the carrion was placed from the nassariids ($p < 0.001$) (Table 4a; Fig. 7). In flow conditions, the speed of both species was similar when nassariids were at distances of 20 and 40 cm from the carrion, but at 60 cm *N. reticulatus* arrived at prey more quickly than *C. neritea* (Fig. 7); as a consequence, the species $\times$ distance interaction was highly significant ($p < 0.001$) (Table 4b).

Parasite load in *Cyclope neritea* and *Nassarius reticulatus*

Five digenean families were found in *Nassarius reticulatus* from Arcachon Bay (site 'Arguin'): Schistosomatidae, Zoogonidae, Lepocreadiidae, Echinostomatidae (mainly *Himasthla quissetensis*), Microphallidae, and unidentified sporocysts. Total prevalence fluctuated between 6 and 18% (Fig. 8). In contrast, *Cyclope neritea* was hardly parasitized, with only 1% infection by Microphallidae in November, February and April (Fig. 8). At experiment completion (280 d), the percentage of parasitized *N. reticulatus* among surviving individuals was identical to that of the dead ones, i.e. 50%. However, the peak of mortality was reached earlier in parasitized snails, i.e. 100–120 d as opposed to 140–160 d in uninfected snails (Fig. 9). From the 60th day to the mortality peak, parasitized snails had a higher mortality rate.

DISCUSSION

Vacant habitat and/or high competitive ability, low parasitic load, high genetic variance and climate change are among the factors shown to be most critical in the establishment of an exotic species (Williamson 1996, Carlton 2000, Sakai et al. 2001, Grosholz 2002, Stachowicz et al. 2002). From field, empirical and experimental studies, we analyzed patterns and processes associated with the recent settlement (less than 30 yr) of *Cyclope neritea* populations along the Atlantic French coast. We particularly studied the first established population (1976) in Arcachon Bay.

We first tackled the genetic architecture of the Arcachon population compared with populations from the original range of the species. Following accidental introductions of alien species, significant decreases of population genetic diversity compared to the native range have been regularly reported, e.g. in plants (Amsellem et al. 2000), in insects (Villablanca et al. 1998, Tsutsui et al. 2000), and in marine invertebrates (Grosholz 2002). In contrast with these studies, the recently settled French Atlantic population of *Cyclope neritea* (Arcachon) exhibited a far greater genetic

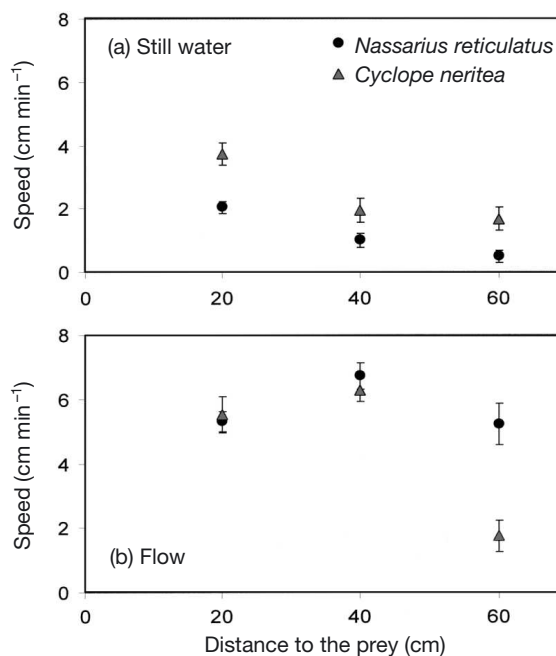


Fig. 7. *Cyclope neritea* and *Nassarius reticulatus*. Mean ($\pm$ SE) speed by which individuals of both species reached carrion placed at 20, 40 or 60 cm in laboratory experiments with (a) still water or (b) flow conditions

diversity than the populations located within the natural range of the species (Thau and Faro). In contrast to the populations of Thau and Faro, for which a total lack of polymorphism was found (only 1 haplotype per population, $H_e = 0$), the recently settled population of Arcachon showed a very high level of polymorphism (5 haplotypes, $H_e = 0.725$).

The high genetic diversity observed at Arcachon could either be due to an introduction from a genetically diverse source population or to recurrent introductions from genetically differentiated sources (Bohonak et al. 2001). Our data support the latter hypothesis. Indeed, the 5 haplotypes found in this population were highly divergent with a mean number of pairwise differences (K) equal to 6.576. Together with the mismatch analysis, this is indicative of a genetic admixture within the Arcachon population and suggests the establishment of individuals originating from genetically distant populations. Moreover, haplotypes found most frequently in this population were fixed in the population of Thau (Haplotype A) and in the population of Faro (Haplotype B). This suggests that independent introductions from either these 2 sites or from genetically related sites occurred. In addition, Haplotype E, genetically divergent from the other haplotypes, was not found in the 2 populations sampled within the natural range of the species. This could indi-

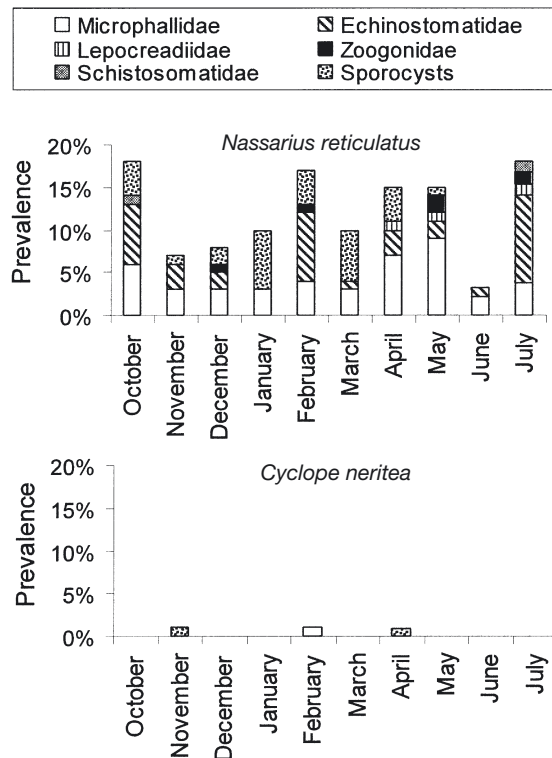


Fig. 8. *Cyclope neritea* and *Nassarius reticulatus*. Digenean prevalence (% of infected snails) in *N. reticulatus* (top) and *C. neritea* (bottom), from October 2000 to July 2001, in Arcachon Bay ('Arguin')

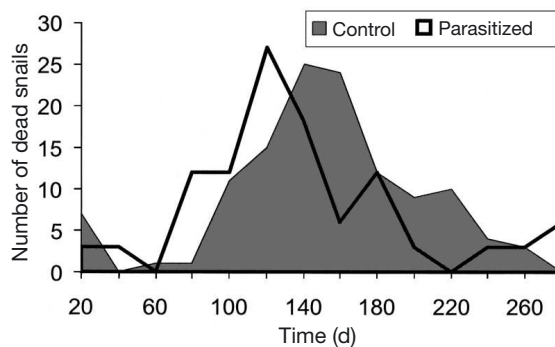


Fig. 9. *Nassarius reticulatus*. Mortality of parasitized and healthy (control) individuals during a 280 d experiment in running seawater

cate that other sources (e.g. eastern Mediterranean or Adriatic Sea) may have contributed to the settlement of the Arcachon population. The Adriatic Sea, where there are high densities of *Cyclope neritea*, is suspected as being at the origin of several gastropod introductions on the French Atlantic coasts because of

the Manila clam *Tapes philippinarum* shellfish trade (Gouletquer et al. 2002). Examples of possible introductions are *Gibbula albida*, noticed in the Morbihan Gulf in 1973 (Deleamarre & Le Neuthiec 1995) and in Arcachon Bay in 1986 (Bachelet et al. 1990), and *Rapana venosa*, first introduced in the Black Sea during the 1940s and the Adriatic Sea during the 1970s (Zibrowius 1991), then in southern Brittany in the late 1990s (Gouletquer et al. 2002).

The recently settled *Cyclope neritea* population at Arcachon probably has multiple origins: a Mediterranean source with the transfer of Haplotype A (and probably C) and an Atlantic source with the transfer of Haplotype B (and probably D). However, because we lack other populations from the native range of the species, it is difficult to specify the true number of sources for the French populations as well as the number of independent introduction events (the origin of Haplotype E is yet unknown). It would be interesting to examine more closely the reasons for the observed divergence between haplotypes A and B which could be indicative of a phylogeographic break across Gibraltar and signal the occurrence of morphologically identical cryptic species. The presence at Arcachon of the haplotype found in Thau Lagoon implies human mediated dispersal mechanisms related to shellfish culture (e.g. transfers of *Crassostrea gigas* since the 1960s between French oyster farms; Sauriau 1991). However, the presence at Arcachon of the haplotype found in Faro could be due either to a natural range expansion of the species towards the north or to introduction events related to human activities (e.g. imports of *Crassostrea angulata* from Portugal since the 1860s; Gruet & Baudet 1997). These 2 processes (natural expansion and introduction) are characterized by different evolutionary time scales (Ribera & Boudouresque 1995). Because of the low dispersal ability of *C. neritea* (i.e. direct development), a natural range expansion towards the north would probably be associated with a step by step process contributing to a cline of the frequency of the haplotype found in Faro. Conversely, recurrent introductions are likely to leave jump dispersal genetic signatures with random fluctuations of frequencies between populations. It is therefore critical to analyze other populations located along the Atlantic coasts of Portugal, Spain and France to discriminate these 2 hypotheses (i.e. natural expansion or human-mediated dispersal from Portugal). The analysis of other recently settled French populations of *C. neritea* would also be useful to check for the consistency of the pattern observed in the Arcachon population.

Recent studies documented the importance of the genetic architecture (genetic variance, response to natural selection, etc.) of invading populations as being a crucial mechanism of biological invasions

(Reznick & Ghalambor 2001, Lee 2002). In several cases, bottleneck effects have been correlated with rapid rates of evolution (Tsutsui et al. 2000, Lee 2002). The large genetic diversity in neutral markers of *Cyclope neritea* from Arcachon cannot be directly correlated with adaptative evolution. However, the studied population is likely to originate from individuals coming from different geographic origins, and may have evolved under various environmental conditions; this would favour its adaptation to a wide range of environmental conditions and promote its long-term establishment and spread.

Mechanisms responsible for biological invasions are not only linked to intrinsic characteristics of the invaders, but also to interspecific interactions at the community/ecosystem level. Both competitive interaction and parasitism are considered as critical factors for the success of an invasion (Mooney & Cleland 2001, Keane & Crawley 2002). Non-indigenous species introduced in a new environment are expected to interfere with native species that exploit common food resources. We therefore hypothesized that *Cyclope neritea*, as a scavenger, would interact with the autochthonous dogwhelk *Nassarius reticulatus* which is probably the commonest scavenger gastropod in shallow marine waters of western Europe (Britton & Morton 1994). The *N. reticulatus* complex actually includes 2 closely related species, *N. reticulatus* (L.) and *N. nitidus* (Jeffreys) (Sanjuan et al. 1997); however, we found only *N. reticulatus* in field sampling in Arcachon Bay and used this species in competition experiments. A prerequisite to any study on interspecific competition is that species encounter each other so they must occur in the same habitat. Both *C. neritea* and *N. reticulatus* are extremely eurythermic and euryhaline but, according to Mars (1950), *N. reticulatus* is slightly less euryhaline than *C. neritea* (15 to 39‰ vs. 6 to 42‰, respectively). Actually, *C. neritea* occurs everywhere in Arcachon Bay, whereas *N. reticulatus* does not inhabit the most eastern parts of the bay where salinity may fall well below 15‰ in winter, due to large inputs of freshwater from rivers and drainage trenches (Bachelet & Dauvin 1993). The zonation pattern found in Arcachon Bay, i.e. –0.8 to +3.1 m for *C. neritea* and –5.2 to +1.9 m for *N. reticulatus*, is in agreement with literature data which indicate that *C. neritea* occurs in the Atlantic mainly in the intertidal (Morton 1960, Sauriau 1989) while *N. reticulatus* extends from low tide level to 36 m depth (Fretter & Graham 1962). Although both species may occur on a wide range of sediments, *C. neritea* in Arcachon Bay seems to be restricted to sands with low organic content, whereas *N. reticulatus* generally thrives best in substrates with high organic content (Tallmark 1980). Therefore, it appears that the habitats of both gastropods partially

overlap in Arcachon Bay, with populations of both species being in sympatry in relatively clean medium and coarse sands located at tidal levels between ca. –1 and +1 m, where some interspecific competition may be expected.

Competition for food between benthic scavengers has seldom been demonstrated experimentally. To our knowledge, this has been attempted only by Morton & Yuen (2000) who used a combination of methods such as the assessment of food preferences, speed to bait, consumption rates, and the observation of individual behaviour when species were feeding together. We tested neither different bait species nor consumption rates by the 2 nassariid species. Behavioural observations (not reported here) in aquaria showed that up to 20 individuals of each species could be found aggregated on a single dead cockle without interfering with each other, probably because both species possess an extendible proboscis that allows feeding at a distance of 2 to 3 cm; similar observations were made by Morton (1960) on *Cyclope neritea* and by Morton & Yuen (2000) on *Nassarius festivus*. In the present study, we chose the speed of individuals to reach carrion and begin feeding as a method to assess competition. Our laboratory experiments clearly showed that, in static water, *C. neritea* moved towards food faster than *N. reticulatus*, whatever the distances (20, 40, or 60 cm) between the prey and the nassariids were. This trend changed in flow conditions, at least when the nassariids were at distances of 20 and 40 cm from the prey, when both species moved with a similar speed. These results were consistent with the spontaneous moving observed without prey, *C. neritea* being active in both static and moving water, and *N. reticulatus* moving spontaneously mainly in flow conditions. It appears that: (1) in still water experimental conditions, similar to those prevailing on intertidal sand flats with slowly running water at low tide (i.e. the habitat preferred by *C. neritea*), *C. neritea* was more active and reached its prey faster than *N. reticulatus*, thus having a competitive advantage over the latter; (2) in flow conditions, similar to those prevailing in subtidal levels (i.e. the habitat preferred by *N. reticulatus*), the activity of *N. reticulatus* increased compared to static conditions and its efficiency to reach a prey became similar to that of *C. neritea*. Nevertheless, even in low hydrodynamic conditions, competitive exclusion of *N. reticulatus* by *C. neritea* is unlikely, due to the abundance of bivalve carrion in the field, especially in shellfish areas such as Arcachon Bay. The fact that the 2 species basically have different substrate preferences would be an additional enhancing factor for the establishment of *C. neritea* because it can occupy an open niche in its new distribution area.

The present study demonstrated that the native *Nassarius reticulatus* was much more heavily parasitized than the introduced *Cyclope neritea*. A similar result was found in South Africa with the native mussel *Perna perna* and the introduced mussel *Mytilus galloprovincialis* (Calvo-Ugarteburu & McQuaid 1998a). The lack of infection in *C. neritea* could be due to a natural resistance to digenean parasites, as observed in the gastropod *Crepidula fornicata* in its native area (Pechenik et al. 2001). However, this is not the case because *C. neritea* in the Mediterranean can be heavily parasitized with the same digenean species as those observed at Arcachon in *N. reticulatus*, e.g. *Himasthla quissetensis* (Prévot 1974). Therefore, it is difficult to explain why *C. neritea* was not parasitized in Arcachon Bay where both parasite and host species (final host: lariid birds; second intermediate host: cockles) were all present. Two hypotheses can be proposed. The first one is a misidentification of the parasites. For example, is *Himasthla quissetensis* really the same species that parasitizes *C. neritea* (Prévot 1974), *N. reticulatus* (Russell-Pinto 1993, Desclaux et al. 2002) and *Ilyanassa obsoleta* (Stunkard 1938)? If not, the lack of infection of *C. neritea* in the introduction site could be due to the absence of the different hosts of the parasite cycle (the lariid birds and the cockles in the Mediterranean belong to the same genus but not to the same species). An alternative hypothesis in explaining the lack of parasites in *C. neritea* is that populations of the introduced nassariid have acquired a higher resistance to digenean infections.

When an introduced species does not bring its natural suite of enemies/parasites with it, it should experience a type of release or 'freedom' in its new environment. This release will be of a magnitude proportional to the ecological importance of the natural enemies left behind (Torchin et al. 2002). In the case of a competition between introduced and indigenous species, parasite release can be an important advantage (Calvo-Ugarteburu & McQuaid 1998b). Indeed, parasitic infection is recognized as a stressful factor which may lower the resistance of the host and its ability to challenge an unparasitized species. The experiment comparing mortality rates of infected and non-infected *Nassarius reticulatus* emphasized the disadvantage of being parasitized, as already suggested in a similar experiment with the mudsnail *Hydrobia ulvae* (de Montaudouin et al. 2003).

By using a multidisciplinary approach, this paper investigated several features expected to promote the long-term establishment of a recent invader that colonized Arcachon Bay 30 yr ago. Our results suggest that the combination of various factors, i.e. recurrent introduction of individuals originating from different environments, a significant competitive ability and the lack

of heavy parasitic load, explains the settlement of *Cyclope neritea* in Arcachon Bay. Whether these factors can promote a greater invasion success in the near future and the long-term establishment of this species along the Atlantic French coast, although likely, is still an opened issue. In particular, the role played by climate change is still to be investigated since, as already reported for marine benthic communities (Glémarec 1979) and invasive species such as *C. neritea* (Sauriau 1991) and *Crassostrea gigas* (Drinkwaard 1999), climate warming is expected to favour population dynamics and further northwards expansion of warm-temperate species. Experimental studies and genetic analysis of other French populations of *C. neritea* would be useful to check for this effect as well as to verify the consistency of the patterns observed in the Arcachon population.

Acknowledgements. This research was supported by the 'Invasions Biologiques (INVABIO)' programme (project No. D4E/SRP/01115) of the Ministère de l'Ecologie et du Développement Durable, France. F.V. received support for the mtDNA analyses from the CNRS (ATIP programme) and the Brittany Region (Programme 'Accueil et Emergence' and Sequencing Platform–GenoMer/Genopole Ouest). C.D. and B.S.-B. were supported by a PhD grant from the Ministère de la Jeunesse, de l'Education Nationale et de la Recherche. We acknowledge helpful comments from 2 anonymous referees.

LITERATURE CITED

- Afli A, Chenier F (2002) Etat de santé de la macrofaune benthique et rôle des espèces invasives dans le golfe du Morbihan (Bretagne, France). *Vie Milieu* 52:43–57
- Amsellem L, Noyer JL, Le Bourgeois T, Hossaert-McKey M (2000) Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. *Mol Ecol* 9:443–455
- Avisé JC (1991) Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annu Rev Genet* 25:45–69
- Avisé JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JT, Reeb CA, Saunders NC (1987) Intraspecific phylogeny: the mitochondrial DNA gap between population genetics and systematics. *Annu Rev Ecol Syst* 18:489–522
- Bachelet G, Dauvin JC (1993) Distribution quantitative de la macrofaune benthique des sables intertidaux du bassin d'Arcachon. *Oceanol Acta* 16:83–97
- Bachelet G, Cazaux C, Gantès H, Labourg PJ (1980) Contribution à l'étude de la faune marine de la région d'Arcachon, IX. *Bull Cent Etud Rech Sci Biarritz* 13:45–64
- Bachelet G, Labourg PJ, Madani I (1990) Nouvelles signalisations de Mollusques (Prosobranches et Bivalves) dans le Bassin d'Arcachon. *Cah Biol Mar* 31:87–92
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Bohonak AJ, Davies N, Villablanca FX, Roderick GK (2001) Invasion genetics of New World medflies: testing alternative colonization scenarios. *Biol Invasions* 3:103–111

- Boulhic M, Tardy J (1986) Ponte, développement larvaire et éthologie des juvéniles de *Cyclope neritea* (Linné, 1758). *Haliotis* 15:51–58
- Britton JC, Morton B (1994) Marine carrion and scavengers. *Oceanogr Mar Biol Annu Rev* 32:369–434
- Bush AO, Lafferty KD, Lotz JM, Shostak AW (1997) Parasitology meets ecology on its own terms: Margolis et al. revisited. *J Parasitol* 83:575–583
- Calvo-Ugarteburu G, McQuaid CD (1998a) Parasitism and introduced species: epidemiology of trematodes in the intertidal mussels *Perna perna* and *Mytilus galloprovincialis*. *J Exp Mar Biol Ecol* 220:47–65
- Calvo-Ugarteburu G, McQuaid CD (1998b) Parasitism and invasive species: effects of digenetic trematodes on mussels. *Mar Ecol Prog Ser* 169:149–163
- Carlton JT (1992) Introduced marine and estuarine mollusks of North America: an end-of-the-20th-century perspective. *J Shellfish Res* 11:489–505
- Carlton JT (2000) Global change and biological invasions in the oceans. In: Mooney HA, Hobbs RJ (eds) *Invasive species in a changing world*. Island Press, Washington, DC, p 31–53
- Coblentz BE (1990) Exotic organisms: a dilemma for conservation. *Conserv Biol* 4:261–265
- Cohen AN, Carlton JT (1998) Accelerating invasion rate in a highly invaded estuary. *Science* 279:555–558
- Curtis LA, Hubbard KM (1990) Trematode infections in a gastropod host misrepresented by observing shed cercariae. *J Exp Mar Biol Ecol* 143:131–137
- D'Asaro CN (1993) Gunnar Thorson's world-wide collection of prosobranch egg capsules: Nassariidae. *Ophelia* 38:149–215
- Deleamarre JL, Le Neuthiec R (1995) Présence de *Gibbula albida* dans le Golfe du Morbihan (Morbihan) et dans la région de Paimpol (Côtes d'Armor). *Bull Soc Sci Nat Ouest Fr* 17:54–61
- de Montaudouin X, Blanchet H, Kisielewski I, Desclaux C, Bachelet G (2003) Digenetic trematodes moderately alter *Hydrobia ulvae* population size structure. *J Mar Biol Assoc UK* 83:297–305
- Desclaux C, de Montaudouin X, Bachelet G (2002) Cockle emergence at the sediment surface: 'favourization' mechanism by digenetic parasites? *Dis Aquat Org* 52:137–149
- Drinkwaard AC (1999) Introductions and developments of oysters in the North Sea area: a review. *Helgol Meeresunters* 52:301–308
- Dukes JS, Mooney HA (1999) Does global change increase the success of biological invaders? *Trends Ecol Evol* 14:135–139
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3:294–299
- Fretter V, Graham A (1962) *British prosobranch molluscs*. Ray Society, London
- Glémarec M (1979) Les fluctuations temporelles des peuplements benthiques liées aux fluctuations climatiques. *Oceanol Acta* 2:365–371
- Gomoiu MT (1964) Biologisches Studium der Arten *Nassa reticulata* L. und *Cyclonassa neritea* (L.) im Schwarzen Meer (rumänischer Küstenbereich). *Rev Roum Biol Ser Zool* 9:39–49
- Gouletquer P, Bachelet G, Sauriau PG, Noel P (2002) Open Atlantic coast of Europe—a century of introduced species into French waters. In: Leppäkoski E, Gollasch S, Olenin S (eds) *Invasive aquatic species of Europe*. Distribution, impacts and management. Kluwer Academic Publishers, Dordrecht, p 276–290
- Grosholz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends Ecol Evol* 17:22–27
- Gruet Y, Baudet J (1997) Les introductions d'espèces d'invertébrés marins. In: Dauvin JC (ed) *Les biocénoses marines et littorales françaises des côtes Atlantique, Manche et Mer du Nord: synthèse, menaces et perspectives*. Laboratoire de Biologie des Invertébrés Marins et Malacologie—Service du Patrimoine Naturel/Institut d'Ecologie et de Gestion de la biodiversité/Muséum National d'Histoire Naturelle IEGB/MNHN, Paris, p 242–250
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Heiner CR, Hunkapiller KL, Chen SM, Glass JI, Chen EY (1998) Sequencing multimegabase-template DNA with BigDye terminator chemistry. *Genome Res* 8:557–561
- Hidalgo JG (1917) Fauna malacológica de España, Portugal y las Baleares. *Moluscos testáceos marinos*. *Trab Mus Nac Cienc Nat Madrid Zool* 30:1–752
- Keane RM, Crawley MJ (2002) Exotic plant invasions and the enemy release hypothesis. *Trends Ecol Evol* 17:164–170
- Kisch BS (1950) Growth stages of *Cyclope neritea* (L.). *J Conchol* 23:108
- Lauckner G (1980) Diseases of Mollusca: Gastropoda. In: Kinne O (ed) *Diseases of marine animals*. Wiley, Chichester, p 311–424
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends Ecol Evol* 17:386–391
- Le Roux A (1994) Sexualité et reproduction de *Cyclope neritea* (Mollusque gastéropode) dans le golfe du Morbihan (Bretagne-Sud). *Haliotis* 23:51–59
- Le Roux A, Maheo R, Daguzan J (1988) Présence et particularités écologiques de *Cyclope neritea* (Linné, 1758) (Gastéropode Prosobranchie) dans le Golfe du Morbihan. *Haliotis* 18:310
- Mars P (1950) Euryhalinité de quelques Mollusques méditerranéens. *Vie Milieu* 1:441–448
- Massé H, Nodot C, Macé AM (1978) Influence de la température sur la reproduction et la survie de quelques Nassariidae (Mollusca, Gastropoda). In: McLusky DS, Berry AJ (eds) *Physiology and behaviour of marine organisms*. Pergamon Press, Oxford, p 367–374
- Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. *Proc Natl Acad Sci USA* 98:5446–5451
- Morton B, Yuen WY (2000) The feeding behaviour and competition for carrion between two sympatric scavengers on a sandy shore in Hong Kong: the gastropod, *Nassarius festivus* (Powys) and the hermit crab, *Diogenes edwardsii* (De Haan). *J Exp Mar Biol Ecol* 246:1–29
- Morton JE (1960) The habits of *Cyclope neritea*, a style-bearing stenoglossan gastropod. *Proc Malac Soc Lond* 34:96–105
- Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York
- Nobre A (1931) *Moluscos marinhos de Portugal*. Imprensa Portuguesa, Porto
- Palumbi SR (2001) Humans as the world's greatest evolutionary force. *Science* 293:1786–1790
- Parker IM, Simberloff D, Lonsdale WM, Goodell K and 7 others (1999) Impact: toward a framework for understanding the ecological effects of invaders. *Biol Invasions* 1:3–19
- Pechenik JA, Fried B, Simpkins HL (2001) *Crepidula fornicata* is not a first intermediate host for trematodes: who is? *J Exp Mar Biol Ecol* 261:211–224
- Pigeot J (1988) Présence du gastéropode nassaridé *Cyclope neritea* (Linné, 1758) dans le bassin de Marennes-Oléron. *Ann Soc Sci Nat Charente-Marit* 7:717–719

- Prévot G (1974) Recherches sur le cycle biologique et l'écologie de quelques trématodes nouveaux parasites de *Larus argentatus michaellis* Naumann dans le midi de la France. PhD thesis, University of Aix-Marseille
- Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112–113:183–198
- Ribera MA, Boudouresque CF (1995) Introduced marine plants, with special reference to macroalgae: mechanisms and impact. In: Round FE, Chapman DJ (eds) *Progress in phycological research*, Vol 11. Biopress, p 187–268
- Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15:174–175
- Ruiz GM, Fofonoff P, Hines AH, Grosholz ED (1999) Non-indigenous species as stressors in estuarine and marine communities: assessing invasion impacts and interactions. *Limnol Oceanogr* 44:950–972
- Ruiz GM, Fofonoff PW, Carlton JT, Wonham MJ, Hines AH (2000) Invasion of coastal marine communities in North America: apparent patterns, processes, and biases. *Annu Rev Ecol Syst* 31:481–531
- Russell-Pinto F (1993) Espécies de digenea que infectam *Cerastoderma edule* (n.v. berbigão) em Portugal. Caracterização da resposta do hospedeiro à infestação. PhD thesis, University of Porto
- Sakai AK, Allendorf FW, Holt JS, Lodge DM and 11 others (2001) The population biology of invasive species. *Annu Rev Ecol Syst* 32:305–332
- Sanjuan A, Pérez-Losada M, Rolán E (1997) Allozyme evidence for cryptic speciation in sympatric populations of *Nassarius* spp. (Mollusca: Gastropoda). *J Mar Biol Assoc UK* 77:773–784
- Sauriau PG (1989) *Cyclope neritea* (Linné, 1758) dans le bassin de Marennes-Oléron. Bilan des observations de 1984 à 1986. *Ann Soc Sci Nat Charente-Marit* 7:995–1004
- Sauriau PG (1991) Spread of *Cyclope neritea* (Mollusca: Gastropoda) along the north-eastern Atlantic coasts in relation to oyster culture and to climatic fluctuations. *Mar Biol* 109:299–309
- Schneider S, Roessli D, Excoffier L (2001) Arlequin: a software for population data analysis (Version 2.000). Genetics and Biometry Laboratory, University of Geneva
- Shea K, Chesson P (2002) Community ecology theory as a framework for biological invasions. *Trends Ecol Evol* 17:170–176
- Southward AJ, Southward EC, Dando PR, Hughes JA, Kennicutt II MC, Alcalá-Herrera J, Leahy Y (1997) Behaviour and feeding of the nassariid gastropod *Cyclope neritea*, abundant at hydrothermal brine seeps off Milos (Aegean Sea). *J Mar Biol Assoc UK* 77:753–771
- Stachowicz JJ, Terwin JR, Whitlatch RB, Osman RW (2002) Linking climate change and biological invasions: ocean warming facilitates nonindigenous species invasions. *Proc Natl Acad Sci USA* 99:15497–15500
- Stunkard HW (1938) The morphology and life cycle of the trematode *Himasthla quissetensis* (Miller and Northup, 1926). *Biol Bull (Woods Hole)* 75:145–164
- Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105:437–460
- Tallmark B (1980) Population dynamics of *Nassarius reticulatus* (Gastropoda, Prosobranchia) in Gullmar Fjord, Sweden. *Mar Ecol Prog Ser* 3:51–62
- Tardy J, Gaillard J, Portères G (1985) *Cyclope neritea* (Linné 1788) une espèce de gastéropode nassariidé nouvelle pour les côtes du centre-ouest Atlantique de France. *Ann Soc Sci Nat Charente-Marit* 7:391–396
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Torchin ME, Lafferty KD, Kuris AM (2001) Release from parasites as natural enemies: increased performance of a globally introduced marine crab. *Biol Invasions* 3:333–345
- Torchin ME, Lafferty KD, Kuris AM (2002) Parasites and marine invasions. *Parasitology* 124:137–151
- Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species. *Proc Natl Acad Sci USA* 97:5948–5953
- Villablanca FX, Roderick GK, Palumbi SR (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Mol Ecol* 7:547–560
- Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of Earth's ecosystems. *Science* 277:494–499
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Assessing the relative importance of habitat destruction, alien species, pollution, over-exploitation, and disease. *BioScience* 48:607–616
- Williamson M (1996) *Biological invasions*. Chapman & Hall, London
- Zibrowius H (1991) Ongoing modification of the Mediterranean marine fauna and flora by the establishment of exotic species. *Mésogée* 51:83–107

Editorial responsibility: Otto Kinne (Editor),
Oldendorf/Luhe, Germany

Submitted: November 21, 2003; Accepted: April 22, 2004
Proofs received from author(s): July 14, 2004

D.2 Molecular Ecology Notes

TECHNICAL NOTE

Polymorphic microsatellites for the study of newly established populations of the gastropod *Cyclope neritea*

BENOÎT SIMON-BOUHET,*† CLAIRE DAGUIN,* PASCALE GARCIA-MEUNIER† and FRÉDÉRIQUE VIARD*

**Evolution et Génétique des Populations Marines, UMR 7127 CNRS, Station Biologique de Roscoff, B.P. 74, 29682 Roscoff cedex, France, †Laboratoire de Biologie et Environnement Marins, Institut de la Mer et du Littoral, Avenue Lazaret, Port des Minimes, 17000 La Rochelle, France*

Abstract

Both human-mediated introductions and climatic changes may promote the settlement of species in new areas outside of their natural geographical range. To investigate the settlement of recently established populations of the neogastropod *Cyclope neritea*, we developed eight microsatellite markers. Their usefulness was studied in two native populations previously found to be monomorphic with mitochondrial markers. The eight loci were found to be polymorphic in both populations, with two to 18 alleles per locus. This result shows promise for these loci in studies of recently founded populations of *C. neritea*.

Keywords: biological invasion, *Cyclope neritea*, marine gastropod, microsatellite

Received 30 September 2004; revision accepted 15 October 2004

Over the last few decades, the release of alien marine species has greatly increased (Occhipinti-Ambrogi & Savini 2003). For species found in new areas close to their native range, both human-mediated introductions and natural spread may explain the range expansion (Dukes & Mooney 1999). The origin of recently settled populations of the marine neogastropod *Cyclope neritea* along the French Atlantic coasts is questionable (for a detailed description of its colonization pattern, see Bachelet *et al.* 2004). Two processes could be involved in the settlement of these new populations: (i) a natural spread towards north favoured by environmental changes and / or (ii) a human introduction from distant native populations, related to the release of marine organisms for aquaculture. We aim to investigate fine scale mechanisms involved in the early stages of the settlement of the new *C. neritea* populations in contrast with populations from the recognized native range. We developed eight microsatellite markers and tested their usefulness by analysing two native populations.

Genomic DNA from three individuals from the Thau lagoon (French Mediterranean Sea) was isolated from < 15 mg of foot muscle using DNeasy Tissue Kit (QIAGEN). A genomic library enriched for CA repeated microsatellites was constructed following the protocol described by

Billote *et al.* (1999) based on Kijas *et al.* (1994). Briefly, total DNA was digested with *RsaI* (Promega). After purification on Nucleospin Column (Macherey-Nagel), 1 µg of digested DNA was ligated to *RsaI* adapters (Rsa21F: 5'-CTCTTGCTTACGCGTGGACTA-3' and Rsa25R: 5'-TAGTCCACGCGTAAGCAAGAGCACA-3'). For the enrichment procedure, hybridization of DNA fragments to biotinylated (AC)₁₀ probes attached to streptavidin-coated magnetic beads (Streptavidin MagneSphere® Paramagnetic Particles; Promega) was carried out. The purified enriched fraction was ligated into pGEM-T easy (Promega) and transformed into *Escherichia coli* JM109 competent cells. Recombinants with appropriate insert sizes were determined by polymerase chain reaction (PCR) (350–800 base pairs; 288 clones), transferred to positively charged nylon membranes and screened by hybridization with digoxigenin-labelled DIG-(TG)₁₀ probes. From a total of 144 positive clones sequenced on a 3100 Genetic Analyser using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems), 36 primer pairs were designed using PRIMER 3 software (Rozen & Skaletsky 2000). After preliminary tests for amplification on agarose gels, 24 forward primers were fluorescently labelled with infra-red fluorescent dye IRD700 or IRD800 for screening on a Li-Cor NEN Global IR2 DNA sequencer. The genomic DNA for genotyping was obtained using DNeasy Tissue Kit (QIAGEN). Amplification conditions were optimized for eight loci (Table 1) for

Correspondence: Benoît Simon-Bouhet, Fax: +335 4650 0294; E-mail: bsimonbo@univ-ir.fr

Table 1 Characteristics of the eight microsatellite loci isolated from *Cyclope neritina* (GenBank Accession nos: AY771365–AY771372). T_i and T_a stand for the initial and optimal annealing temperature of the touch-down PCR procedure; Cycle 1 refers to the number of cycles used to reach the optimal annealing temperature whereas Cycle 2 refers to the number of cycles used at the optimal annealing temperature. N_{Ta} and H_T : number of alleles and total heterozygosity over the whole sample ($N = 48$), N_{all} and H_E : number of alleles and expected heterozygosity at the population level, F_{is} : estimator for deviation to Hardy–Weinberg proportions and associated P -values (exact test)

Locus	Core sequence	Primer sequences (5'–3')	T_i – T_a (°C) (Cycle 1)	Cycle 2	MgCl ₂ (mM)	Allele size range (bp)	Thau ($N = 24$)				Faro ($N = 24$)			
							N_{Ta}	H_T	N_{all}	H_E	F_{is}	N_{all}	H_E	F_{is}
Cn-2C4	(GT) ₁₅	F: ATCGAGCAGAAAGACATGAC R: GAAAGTGTGACCCCGTGAG	65–58 (8)	24	1.5	151–172	5	0.66	5	0.61	0.106 ($P = 0.305$)	2	0.51	0.269 ($P = 0.237$)
Cn-2H10b	(GA) ₁₆ GG(GA) ₅ G GAGAGA(AC)G) ₄	F: CCACGGTGTGCTCCCTACTTT R: AATGAGCCGTTTAACTGACAA	60–52 (8)	24	1.5	118–193	9	0.83	6	0.73	0.033 ($P = 0.096$)	6	0.80	0.134 ($P = 0.449$)
Cn-3C4	(TGTC) ₉ (TGGC) ₂ (TGTC) ₆	F: TCTCCCAATCCAGTATATAAG R: AACCCAACTTGCCCAACTATG	58–50 (8)	32	2.5	206–239	6	0.72	5	0.69	0.091 ($P = 0.323$)	4	0.58	–0.247 ($P = 0.787$)
Cn-1H8	(GT) ₂₈	F: CCTTCTTGCTGCTGTCTTCTG R: GGAGTAGGTAGAGTTAGTGAATGTC	62–52 (10)	25	1.5	130–142	5	0.47	4	0.43	0.609 ($P = 0.000$)	3	0.52	0.666 ($P = 0.000$)
Cn-1A10	(TG) ₁₀ (AG) ₁₅ GGG (AG) ₂₁ (TG) ₃	F: GCGTTACAGGCAATGC R: CAGAAATCCGTGCTTCAGC	60–52 (8)	24	1.5	152–224	15	0.89	13	0.90	0.246 ($P = 0.079$)	6	0.74	0.084 ($P = 0.451$)
Cn-1A8	(GA) ₃₆ N ₄₀ (GT) ₃ A (TG) ₆ (N) ₁₅ (GT) ₄	F: GACACTGCCCGAAGACAGTC R: GAAGTAAACAGTGGAAATTAATATAC	64–54 (10)	25	2.5	237–324	18	0.93	16	0.93	0.460 ($P = 0.000$)	3	0.61	0.702 ($P = 0.006$)
Cn-3B2	(GT) ₁₂ GC(GT) ₂ GC(GT) ₃	F: CTACACACACCCAGTCTCCAG R: CACCTATGTTCCGTGCGC	65–55 (10)	24	1.5	172–190	4	0.71	3	0.60	0.665 ($P = 0.001$)	2	0.26	0.324 ($P = 0.288$)
Cn-2B7	(CA) ₇	F: CAGTGGCAAGTCTTTAGATC R: CAACAACGACAAAAGGCAAGC	63–55 (8)	22	2.5	176–178	2	0.42	2	0.23	0.254 ($P = 0.310$)	2	0.51	–0.154 ($P = 0.680$)
Overall (SD)										0.64 (0.23)	0.290 ($P = 0.000$)		0.57 (0.17)	0.171 ($P = 0.005$)

which preliminary tests showed unambiguous patterns. Each reaction was performed in 20 µL and contained approximately 50 ng of DNA, 0.05 µM of fluorescently labelled forward primer, 0.075 µM of unlabelled forward primer, 0.125 µM of unlabelled reverse primer, 1.5 or 2.5 µM of MgCl₂ depending on the locus (Table 1), 0.25 mM of each dNTP, 0.1 g/L of bovine serum albumin, 1X PCR buffer [75 mM Tris-HCl pH 8.8 at 25 °C, 20 mM (NH₄)₂SO₄ and 0.01% Tween 20] and 0.50 U Thermoprime Plus DNA Polymerase (ABgene), using a MJ Research PTC-200 thermocycler (MJ Research Inc.). Thermal cycling conditions followed a locus-specific touch-down PCR procedure: 3 min of initial denaturation step at 95 °C followed by cycles (Cycle 1 in Table 1) of denaturation at 95 °C for 45 s, annealing at temperature decreasing by 1 °C per cycle from T_i to T_a (Table 1) for 45 s and extension at 72 °C for 45 s, then followed by a locus-specific number of cycles (Cycle 2 in Table 1) at 95 °C for 45 s, annealing at final T_a (Table 1) for 45 s and extension at 72 °C for 45 s. A final extension step was carried out for 7 min at 72 °C. PCR products were screened on a 6.5% polyacrylamide gel using a Li-Cor NEN Global IR2 DNA sequencer system. Allele sizes were determined using a known DNA sequence.

Polymorphism at the eight loci was tested in two populations of *C. neritea* from its native range [Thau (French Mediterranean Sea), $N = 24$ and Faro (south Portugal, Atlantic), $N = 24$] that were found to be monomorphic when sequencing a mitochondrial gene fragment (Cytochrome Oxidase I, Bachelet *et al.* 2004). No linkage disequilibria across loci were detected using GENEPop version 3.4 (Raymond & Rousset 1995). This software was also used to estimate the number of alleles, the expected heterozygosities and adequacy of genotypic proportions to Hardy–Weinberg expectations by an approximate exact test (detailed in Rousset & Raymond 1995; Table 1). Over the 48 individuals analysed, the number of alleles ranged from two to 18 and expected heterozygosities from 0.42 to 0.93, respectively. At the population level, deviations from Hardy–Weinberg equilibrium were variable across loci and populations. For instance, Cn-1A8 showed a large heterozygote deficiency only in Faro that might indicate the presence of null alleles as a result of a phylogeographical break in the species' native range (Bachelet *et al.* 2004). A larger number of individuals and populations should thus be examined to better assess the reliability of this locus. Nonetheless, both

populations were found to be polymorphic [e.g. mean number of 6.75 (SE 5.01) and 3.50 (SE 1.69) alleles in Thau and Faro, respectively] and exhibited similar levels of expected heterozygosity. The microsatellite loci developed in the present study are thus considerably more polymorphic than the mitochondrial marker previously used and are of interest to study fine-scale processes involved in the settlement of new populations of *C. neritea*.

Acknowledgements

The authors are grateful to Carolyn Engel for improvements of the manuscript. This project is part of the INVABIO program of the Ministère de l'Ecologie et du Développement Durable (project n°D4E/SRP/01115). F.V. acknowledges support from the European Network of Excellence 'Marine Genomics' (contract n°505403) and the CNRS (ATIP program). This project also benefited from the facilities of the Genomer and Ouest-Genopole sequencing platform. B.S.-B. was supported by a PhD grant from the Ministère de l'Education et de la Recherche.

References

- Bachelet G, Simon-Bouhet B, Desclaux C *et al.* (2004) Invasion of the eastern Bay of Biscay by the nassariid gastropod *Cyclope neritea*: origin and effects on the resident fauna. *Marine Ecology Progress Series*, **276**, 147–159.
- Billote N, Lagoda P, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the tech development of SSR markers in tropical crops. *Fruits*, **54**, 277–288.
- Dukes JS, Mooney HA (1999) Does global change increase the success of biological invaders? *Trends in Ecology and Evolution*, **14**, 135–139.
- Kijas JMH, Fowler JCS, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the Citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Biotechniques*, **16**, 657–662.
- Occhipinti-Ambrogi A, Savini D (2003) Biological invasions as a component of global change in stressed marine ecosystems. *Marine Pollution Bulletin*, **46**, 542–551.
- Raymond M, Rousset F (1995) GENEPop (version 1.2): a population genetics software for exact tests and ecumenicism. *The Journal of Heredity*, **86**, 248–249.
- Rousset F, Raymond M (1995) Testing heterozygote excess and deficiency. *Genetics*, **140**, 1413–1419.
- Rozen S, Skaletsky H (2000) PRIMER 3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, New Jersey.

D.3 Molecular Ecology

Multiple introductions promote range expansion of the mollusc *Cyclope neritea* (Nassariidae) in France: evidence from mitochondrial sequence data

B. SIMON-BOUHET,*† P. GARCIA-MEUNIER† and F. VIARD*

**Evolution et Génétique des Populations Marines, UMR 7144 CNRS-UPMC, Station Biologique de Roscoff, B.P. 74, 29682 Roscoff cedex, France, †Laboratoire de Biologie et Environnement Marins, FRE 2727 CNRS-ULR, Avenue Michel Crépeau, 17042 La Rochelle, France*

Abstract

Since the 1970s, the nassariid gastropod *Cyclope neritea* has been extending its range north along the French Atlantic coasts from the Iberian Peninsula. This may be due to natural spread because of the recent warming of the northeastern Atlantic. However, human-mediated introductions related to shellfish culture may also be a probable explanation for this sudden range expansion. To examine these two hypotheses, we carried out a comprehensive study based on mitochondrial gene sequences (cytochrome oxidase I) of the five recently colonized French bays as well as 14 populations located in the recognized native range of the species. From a total of 594 individuals, we observed 29 haplotypes to split into three divergent clades. In the native range, we observed a low molecular diversity, strong genetic structure and agreement between geography and gene genealogies. Along the French coasts, we observed the opposite: high genetic diversity and low genetic structure. Our results show that recurrent human-mediated introductions from several geographical areas in the native range may be a source for the French Atlantic populations. However, despite the low dispersal ability of *C. neritea*, the isolation-by-distance pattern in France suggested that this gastropod may have been present (although unnoticed) on the French Atlantic coasts before the 1970s. As *C. neritea* shows characteristics of a cryptogenic species, the classification of Atlantic populations as either native or introduced is not straightforward. Cryptogenic species should be studied further to determine the status of new populations close to their recognized native range.

Keywords: biological introductions, cryptogenic species, *Cyclope neritea*, cytochrome oxidase I, marine gastropod, range limit

Received 16 September 2005; revision accepted 14 December 2005

Introduction

In marine habitats and especially coastal areas, the rate of biological invasions has increased considerably in the last few decades, altering community dynamics (Dukes & Mooney 1999; Grosholz 2002). Nevertheless, the processes underlying the sustainable settlement of nonindigenous species (NIS) in coastal habitats are still not fully known (Carlton 1996; Grosholz 2002). Introduction processes are obviously highly related to the increase in transport in ship

ballast waters as well as to intentional or accidental releases for aquaculture and fisheries (Carlton & Geller 1993; Cohen & Carlton 1998). The sustainable settlement of NIS in coastal habitats may also benefit from other facets of global change. Fast environmental changes such as habitat or climatic changes may indeed alter native communities and weaken the competitive ability of indigenous species, thus facilitating the settlement of introduced species (Dukes & Mooney 1999; Carlton 2000; Sax & Brown 2000). For example, Stachowicz et al. (2002) recently demonstrated that changes in the maximum and minimum environmental temperatures may have promoted the establishment of two invasive ascidians at the expense of native species. Temperature is known to be a critical factor influencing the

Correspondence: Benoît Simon-Bouhet, Fax: +33 546500294; E-mail: bsimonbo@univ-lr.fr, bsimonbo@sb-roscoff.fr

1700 B. SIMON-BOUHET, P. GARCIA-MEUNIER and F. VIARD

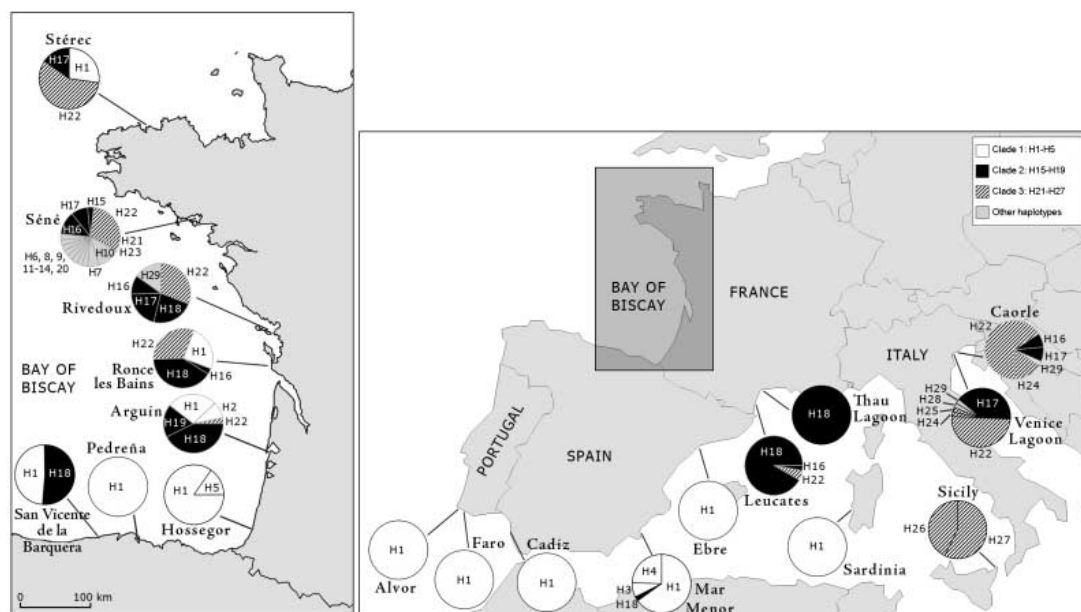


Fig. 1 Location of the study populations and haplotype frequencies at the population level. Native populations are shown on the right whereas FAEC populations and populations of unknown status are shown on the detailed map of the Bay of Biscay (left side). For each population, haplotype names are indicated and colour/figures (see box on the upper-right corner) in pie charts refer to the defined clades in Fig. 2.

natural distribution of animal and plant species: a wide array of species colonized western and northern Europe from the Iberian Peninsula after the ice ages (Comes & Kadereit 1998; Taberlet *et al.* 1998; or Gomez & Lunt, *in press*) and species located on both sides of a biogeographic boundary can extend or restrict their natural range in response to local changes in environmental conditions (Gaylord & Gaines 2000). Natural processes might thus interact with human-mediated introductions.

The Bay of Biscay (northeastern Atlantic, Fig. 1) is an interesting area to study this phenomenon because the mean temperature of this geographical area has increased by 0.6–0.8 °C per decade during the last 30 years (IPCC 2001). This rapid and recent warming has promoted the displacement of mollusc species within this bay (e.g. *Macoma balthica*, Hummel *et al.* 2000). Moreover, there are numerous and important shellfish culture sites present along the coasts of the Bay of Biscay and the exchange of commercial species between distant shell-farming sites have been shown to be an efficient way of primary and secondary introductions of NIS, including molluscs (e.g. *Crepidula fornicata*, see References in Dupont *et al.* 2003), in new biotas (Wolff & Reise 2002). Therefore, changes in temperature and the transport of oysters may both induce community changes in the Bay of Biscay. The recent expansion of the distribution range of the mollusc *Cyclope neritea* makes this species an appropriate model to examine the role played by

these two processes. This nassariid gastropod is native to the Mediterranean and Black seas, and to the Atlantic coasts of Morocco, southern Spain and Portugal. Until the 1950s, only discontinuous and ephemeral populations were recorded along the northern Atlantic coasts of Spain and the southern French Basque coasts, with the Basque country being defined as the northern edge of its natural range (for a detailed description, see Sauriau 1991). Since 1976, new populations have been recorded gradually moving from south to north along the French Atlantic coasts: first in Arcachon Bay, then in 1983–1984, in the Marennes-Oléron Bay, the Île de Ré, and the Gulf of Morbihan, and more recently from 2000, along English Channel coasts (see References in Bachelet *et al.* 2004). *C. neritea* is rather uncommon because new populations of this species appeared just at the edge of its previously recognized natural range. On the one hand, the gradual appearance of *C. neritea* populations from south to north suggests a natural spread towards north during the past 30 years due to environmental changes. On the other hand, a natural spread over such a short period of time seems unlikely given the limited dispersal ability of this species. As a direct developer, *C. neritea* has no planktonic larval stage (Gomoiu 1964) and is restricted to a very particular substrate (i.e. sheltered spots, characterized by sandy-muddy substrates). Therefore, the discontinuous habitat due to the rocky shores along the southern French Atlantic coasts may slow down natural migrations (Sauriau

RANGE EXPANSION OF *C. NERITEA*: AN mtDNA SURVEY 1701

1991). Moreover, all the newly established French populations of *C. neritea* were found near oysters beds. Therefore, the appearance of *C. neritea* along the French Atlantic and English Channel (FAEC) coasts may be due to human-mediated introductions via oyster exchanges between Mediterranean and Atlantic shellfish culture sites (Boulhic & Tardy 1986; Pigeot 1988; Le Roux *et al.* 1988; Sauriau 1991).

In a previous study, Bachelet *et al.* (2004) focused on Arcachon Bay, which was the first site in which *C. neritea* was observed outside of its recognized native range. Mitochondrial analyses revealed the coexistence of three highly divergent lineages in this population, reinforcing the idea of accidental introductions due to human-mediated activities. However, based on this single population, the authors could not differentiate between a single introduction from a highly diversified source population or multiple and recurrent introductions from more poorly diversified but highly divergent source populations. Moreover, by analysing only one population, Bachelet *et al.* (2004) could not determine the colonization history of *C. neritea* along the French coasts. Here, we consider: (i) whether the population inhabiting Arcachon Bay exhibits a pattern congruent with a single-source or a multiple-source introduction scenario. As accidental NIS introductions along the FAEC coasts are mainly related to oyster exchanges (Goulletquer *et al.* 2002), we focused our native range sampling on areas where important shellfish cultures occur (i.e. the Iberian Peninsula, the French Mediterranean coasts and the Northern Adriatic Sea); (ii) whether the recently established FAEC populations all share a common mechanism for their establishment and whether the FAEC populations are genetically related. In particular, we tested for a genetic isolation pattern between the known primary site of introduction (i.e. Arcachon Bay, colonized in 1976) and the more recently and gradually colonized sites along the French coasts; and (iii) whether we can identify the status (either native or introduced) of the populations located at the edge of the recognized natural range of *C. neritea* (i.e. the northern part of the Iberian Peninsula and French Basque Country).

Mitochondrial gene studies were shown to be reliable for phylogeography studies of native species as well as for analysing the mechanisms involved in the settlement of introduced species (Kolbe *et al.* 2004; Voisin *et al.* 2005). Moreover, because the effective population size for mitochondrial genes is four-times lower than for nuclear markers, any genetic drift effects are emphasized and isolation by distance patterns should be easier to detect (Diaz-Almela *et al.* 2004). Therefore, we carried out a comprehensive sampling of *C. neritea* along the FAEC coasts (five populations) as well as 11 populations from the native range and three populations located at the edge of the previously recognized range of *C. neritea*. Classical population genetics analyses were combined with intraspecific phylogenetic approaches by analysing a fragment of the

cytochrome oxidase I mitochondrial gene (*COI*, 533 bp) in 594 individuals.

Materials and methods

Sampling

A survey and exhaustive sampling of French Atlantic and English Channel (FAEC) coasts was carried out between March 2002 and June 2003 in each bay in which *Cyclope neritea* has been recorded since the 1970s (Table 1, Fig. 1). We believe this sampling fits exactly the real distribution of the species in its current FAEC range. We collected five FAEC populations: (i) Arguin, the primary site of introduction in 1983 in Arcachon Bay (which was previously studied by Bachelet *et al.* 2004), (ii) Ronces les Bains (Marennes-Oléron Bay, Sauriau 1989), Rivedoux (Île de Ré, Tardy *et al.* 1985) and Séné (Gulf of Morbihan, Le Roux *et al.* 1988), three bays in which *C. neritea* was first observed in 1983–84; and (iii) Stérec (Morlaix Bay, English Channel), the most northerly population, known on French coasts since the end of the 1990s. We also sampled three populations located at the northern edge of the recognized natural range (French Basque Country) between December 2003 and March 2004: (i) Hossegor (French Basque Country), where *C. neritea* went unnoticed until our sampling and located at 50 km North of Socoa where the species was first recorded in 1950 (Kisch 1950) and (ii) two northern Spanish populations, Pedreña and San Vicente de la Barquera. We also sampled 11 populations between March and December 2004 in the recognized native range of the species to allow comparisons with the FAEC populations (Table 1; Fig. 1): (i) along the Atlantic coasts of Portugal and Spain (3 populations), (ii) along the Mediterranean coasts of Spain, France and Italy (6 populations) and (iii) along the coasts of the Adriatic Sea (Italy, 2 populations). Thau Lagoon (France) and Faro (Portugal) had been previously studied by Bachelet *et al.* (2004). For each population, we collected adult specimens, which were stored in 95% ethanol before DNA extraction.

DNA extraction, amplification and sequencing

We followed the protocol detailed in Bachelet *et al.* (2004). Briefly, total DNA was extracted from less than 15 mg of foot muscle using the DNeasy Tissue Kit according to the manufacturer's protocol (QIAGEN). Amplification of the *COI* gene was carried out using specific primers and the polymerase chain reaction (PCR) products were directly sequenced using ABI PRISM BigDye Terminators version 3.0 Cycle Sequencing Kit following the manufacturer's protocol (Applied Biosystems). Both strands were sequenced for each individual using an ABI PRISM-3100 Automated DNA Sequencer (PerkinElmer Applied Biosystems) and sequence data were aligned using CLUSTAL W (Thompson *et al.* 1994).

1991). Moreover, all the newly established French populations of *C. neritea* were found near oysters beds. Therefore, the appearance of *C. neritea* along the French Atlantic and English Channel (FAEC) coasts may be due to human-mediated introductions via oyster exchanges between Mediterranean and Atlantic shellfish culture sites (Boulhic & Tardy 1986; Pigeot 1988; Le Roux *et al.* 1988; Sauriau 1991).

In a previous study, Bachelet *et al.* (2004) focused on Arcachon Bay, which was the first site in which *C. neritea* was observed outside of its recognized native range. Mitochondrial analyses revealed the coexistence of three highly divergent lineages in this population, reinforcing the idea of accidental introductions due to human-mediated activities. However, based on this single population, the authors could not differentiate between a single introduction from a highly diversified source population or multiple and recurrent introductions from more poorly diversified but highly divergent source populations. Moreover, by analysing only one population, Bachelet *et al.* (2004) could not determine the colonization history of *C. neritea* along the French coasts. Here, we consider: (i) whether the population inhabiting Arcachon Bay exhibits a pattern congruent with a single-source or a multiple-source introduction scenario. As accidental NIS introductions along the FAEC coasts are mainly related to oyster exchanges (Goulletquer *et al.* 2002), we focused our native range sampling on areas where important shellfish cultures occur (i.e. the Iberian Peninsula, the French Mediterranean coasts and the Northern Adriatic Sea); (ii) whether the recently established FAEC populations all share a common mechanism for their establishment and whether the FAEC populations are genetically related. In particular, we tested for a genetic isolation pattern between the known primary site of introduction (i.e. Arcachon Bay, colonized in 1976) and the more recently and gradually colonized sites along the French coasts; and (iii) whether we can identify the status (either native or introduced) of the populations located at the edge of the recognized natural range of *C. neritea* (i.e. the northern part of the Iberian Peninsula and French Basque Country).

Mitochondrial gene studies were shown to be reliable for phylogeography studies of native species as well as for analysing the mechanisms involved in the settlement of introduced species (Kolbe *et al.* 2004; Voisin *et al.* 2005). Moreover, because the effective population size for mitochondrial genes is four-times lower than for nuclear markers, any genetic drift effects are emphasized and isolation by distance patterns should be easier to detect (Diaz-Almela *et al.* 2004). Therefore, we carried out a comprehensive sampling of *C. neritea* along the FAEC coasts (five populations) as well as 11 populations from the native range and three populations located at the edge of the previously recognized range of *C. neritea*. Classical population genetics analyses were combined with intraspecific phylogenetic approaches by analysing a fragment of the

cytochrome oxidase I mitochondrial gene (*COI*, 533 bp) in 594 individuals.

Materials and methods

Sampling

A survey and exhaustive sampling of French Atlantic and English Channel (FAEC) coasts was carried out between March 2002 and June 2003 in each bay in which *Cyclope neritea* has been recorded since the 1970s (Table 1, Fig. 1). We believe this sampling fits exactly the real distribution of the species in its current FAEC range. We collected five FAEC populations: (i) Arguin, the primary site of introduction in 1983 in Arcachon Bay (which was previously studied by Bachelet *et al.* 2004), (ii) Ronce les Bains (Marennes-Oléron Bay, Sauriau 1989), Rivedoux (Île de Ré, Tardy *et al.* 1985) and Séné (Gulf of Morbihan, Le Roux *et al.* 1988), three bays in which *C. neritea* was first observed in 1983–84; and (iii) Stérec (Morlaix Bay, English Channel), the most northerly population, known on French coasts since the end of the 1990s. We also sampled three populations located at the northern edge of the recognized natural range (French Basque Country) between December 2003 and March 2004: (i) Hossegor (French Basque Country), where *C. neritea* went unnoticed until our sampling and located at 50 km North of Socoa where the species was first recorded in 1950 (Kisch 1950) and (ii) two northern Spanish populations, Pedreña and San Vicente de la Barquera. We also sampled 11 populations between March and December 2004 in the recognized native range of the species to allow comparisons with the FAEC populations (Table 1; Fig. 1): (i) along the Atlantic coasts of Portugal and Spain (3 populations), (ii) along the Mediterranean coasts of Spain, France and Italy (6 populations) and (iii) along the coasts of the Adriatic Sea (Italy, 2 populations). Thau Lagoon (France) and Faro (Portugal) had been previously studied by Bachelet *et al.* (2004). For each population, we collected adult specimens, which were stored in 95% ethanol before DNA extraction.

DNA extraction, amplification and sequencing

We followed the protocol detailed in Bachelet *et al.* (2004). Briefly, total DNA was extracted from less than 15 mg of foot muscle using the DNeasy Tissue Kit according to the manufacturer's protocol (QIAGEN). Amplification of the *COI* gene was carried out using specific primers and the polymerase chain reaction (PCR) products were directly sequenced using ABI PRISM BigDye Terminators version 3.0 Cycle Sequencing Kit following the manufacturer's protocol (Applied Biosystems). Both strands were sequenced for each individual using an ABI PRISM-3100 Automated DNA Sequencer (PerkinElmer Applied Biosystems) and sequence data were aligned using CLUSTAL W (Thompson *et al.* 1994).

1702 B. SIMON-BOUHET, P. GARCIA-MEUNIER and F. VIARD

Table 1 Geographic locations of the 19 study populations. Numbers of individuals (N) are indicated together with genetic diversity estimates (N_H , number of haplotypes; S , number of segregating sites; H_E , haplotypic diversity; π , nucleotide diversity)

Population	Location	Country (Sea-Ocean)	N	N_H	S	H_E	$\pi \times 10^2$
Recently established populations (set 1)							
Stérec	Morlaix Bay, N. Brittany	France (EC)	33	3	18	0.589	1.371
Séné	Gulf of Morbihan, S. Brittany	France (A)	32	16	14	0.911	1.014
Rivedoux	Île de Ré	France (A)	39	5	21	0.796	1.503
Ronce les bains	Marenne-Oléron Bay	France (A)	32	4	19	0.696	1.498
Arguin	Arcachon Bay	France (A)	33	5	19	0.725	1.234
Mean for recently established populations			33.8	6.6	18.2	0.743	1.324
(Standard deviation)			(2.9)	(5.3)	(2.6)	(0.120)	(0.205)
Total over all recently established populations			169	21	27	0.833	1.518
Populations of unknown status (set 2)							
Hossegor	Basque Country	France (A)	33	2	1	0.265	0.050
Pedreña	Santander Bay, Cantabrica	Spain (A)	3	1	0	0.000	0.000
San Vicente de la Barquera	Cantabrica	Spain (A)	37	2	12	0.514	1.156
Mean for populations with unknown status			24.3	1.7	4.3	0.260	0.402
(Standard deviation)			(18.6)	(0.6)	(6.7)	(0.257)	(0.653)
Total over all populations with unknown status			73	3	13	0.484	0.903
Native populations (set 3)							
Alvor	Algarve	Portugal (A)	39	1	0	0.000	0.000
Faro	Ria Formosa, Algarve	Portugal (A)	31	1	0	0.000	0.000
Cadiz	Andalucia	Spain (A)	36	1	0	0.000	0.000
Mar Menor	Murcia	Spain (MS)	37	4	13	0.527	0.216
Ebre	Cataluña	Spain (MS)	33	1	0	0.000	0.000
Leucates	Languedoc-Roussillon	France (MS)	37	3	13	0.156	0.255
Thau Lagoon	Languedoc-Roussillon	France (MS)	29	1	0	0.000	0.000
Sardinia	Golfo di Oristano	Italy (MS)	25	1	0	0.000	0.000
Sicily	Punta del Faro (strait of Messina)	Italy (MS)	7	2	2	0.571	0.214
Venice Lagoon	Veneto	Italy (AS)	40	6	25	0.603	1.251
Caorle	Veneto	Italy (AS)	38	5	21	0.543	0.769
Mean for native populations			32	2.4	6.7	0.218	0.246
(Standard deviation)			(9.5)	(1.9)	(9.5)	(0.276)	(0.406)
Total over all native populations			352	13	35	0.661	1.378
Total			594	29	36	0.734	1.479

EC, English Channel; A, Atlantic; MS, Mediterranean Sea; AS, Adriatic Sea.

Note: populations were sampled by the authors except for Sardinia, Sicily and Venice Lagoon for which samples were provided by Jeroen Jansen, Francesco Patti and Davide Tagliapietra, respectively.

Sequence analysis

For each population, we examined the genetic diversity by calculating the number of mitochondrial haplotypes, N_H , and polymorphic sites, S , the nucleotide diversity, θ (i.e. average number of nucleotide differences between pairs of sequences), and the gene diversity, H_E (Nei 1987), using DNASP version 3.53 software (Rozas & Rozas 1999). We defined three population sets (Table 1): recently established populations (i.e. populations recorded since the 1970s along the FAEC coasts, set 1), populations of unknown status (i.e. populations located at the edge of the recognized native range of the species, set 2) and native populations (i.e. populations from the recognized native range of the species, set

3). For each set, we calculated the mean (and associated standard deviations) and overall N_H , S , H_E and θ -values.

We investigated the overall or regional population genetic structure by performing analyses of molecular variance (AMOVA, Excoffier *et al.* 1992) as implemented in ARLEQUIN version 2.0. software (Schneider *et al.* 2001). Fixation indices (Wright 1951) analogous to F_{ST} , F_{CT} and F_{SC} parameters, were calculated to analyse the genetic differentiation between populations over the whole native range (Φ_{ST}), among groups of populations (regions) within the native range (Φ_{CT}) and among populations within groups (regions; Φ_{SC}) as defined in Excoffier *et al.* (1992). Φ -statistics take into account both the haplotype frequencies and the molecular distances between haplotypes. We assessed the statistical

significance of the Φ indices using a nonparametric permutation procedure implemented in ARLEQUIN version 2.0. We used the isolation by distance (IBD) model to examine the genetic relationships between FAEC populations. IBD was determined by plotting genetic distance (i.e. $\hat{F}_{ST}/(1 - \hat{F}_{ST})$, Rousset 1997) against linear geographic distance (in kilometres) measured following the coastline. These distances were calculated for each FAEC population pair and we used a Mantel test to check a positive correlation between matrices of pairwise geographic and genetic distances using IBD 1.5 software (Bohonak 2002).

Given the pattern that was observed by Bachelet *et al.* (2004) in Arguin, we had to discriminate between two introduction processes: (i) a single introduction event from a highly diversified source population or (ii) multiple introductions from slightly diversified but highly divergent source populations. Therefore, as well as molecular diversity, we also analysed the haplotype distribution across populations after determining their genealogical relationships by a haplotypic network (Cassens *et al.* 2003). The network was constructed using a median-joining algorithm implemented in NETWORK 4.0.0.0 software (Bandelt *et al.* 1999). We then applied a maximum parsimony algorithm to simplify the complex branching pattern and to represent all the most parsimonious intraspecific phylogenies. We illustrated similarities between FAEC and native populations by building a population tree using MEGA version 3.0 software (Kumar *et al.* 2004): pairwise genetic distances matrices were calculated using the Kimura 2-parameters model (Kimura 1980) and phylogenetic reconstructions were carried out using a neighbour-joining algorithm (Saitou & Nei 1987). Finally, for FAEC populations, we calculated the distribution of the observed number of pairwise nucleotide site differences (i.e. mismatch distribution) in DNASP version 3.53 to examine the mixing of evolutionary divergent lineages at the population level. Given the low resolving power of Tajima's *D*-test (Galtier *et al.* 2000), we carried out tests for departure from mutation-drift equilibrium under an infinite-site model following the method of Depaulis & Veuille (1998). This test (*K*-test) is based on the comparison between the observed number of haplotypes (N_H) and the expected number under equilibrium (N_{EH}) given the sample size (N) and the observed number of segregating sites (S). It is particularly relevant to our study because it can detect recent population admixture (Hamblin & Veuille 1999). We used a computer program provided by M. H. Muller (2002) to compute confidence intervals and one-tailed *P* values for N_H .

Results

Molecular diversity over the dataset

A high molecular diversity was found over the whole study

although not evenly distributed across population sets. We analysed a 533-bp fragment of the *COI* gene and observed only synonymous substitutions (Appendix). Over the whole data set ($N = 594$ individuals), we identified 36 polymorphic sites defining 29 haplotypes (Table 1, Appendix; GenBank Accession nos: AY789970–AY789991 and DQ087210–DQ087216). The haplotype distribution at the population level is shown in Fig. 1. Over the whole data set, we found three haplotypes (H1, H18 and H22) at a high frequency, $f = 0.443$, 0.200 and 0.163, respectively (sum = 0.806) compared to the others ($0.002 < f < 0.059$). We observed large differences in gene and molecular diversities for the three population sets (Table 1): gene and molecular diversity indices were higher in recently established populations (set 1) than in populations of unknown status (set 2) or from the native range (set 3). For example, H_E ranged from 0.589 to 0.911 (mean = 0.743) in set 1, whereas it ranged from 0 to 0.514 (mean = 0.260) in set 2 and from 0 to 0.603 (mean = 0.218) in set 3. We observed the same for N_H , S and π . All populations presented high and similar polymorphism levels in set 1, whereas the other two sets showed monomorphic populations (7 out of 14). Molecular and genetic diversities were not distributed similarly across populations according to the set: overall and mean values for diversity indices were closer in set 1 than in sets 2 and 3. For example, overall value for π for set 3 was six times higher than mean value (1.38×10^2 and 0.25×10^2 respectively) whereas overall and mean values were almost identical for set 1 (1.52×10^2 and 1.32×10^2 respectively). Finally, in the recently established populations, only Séné differed in the nucleotidic and haplotypic diversity estimates: this population showed the highest number of haplotypes and highest haplotypic diversity ($N_H = 16$, $H_E = 0.911$) whereas the number of segregating sites and nucleotide diversity ($S = 14$, $\pi = 0.01024$) were the lowest of the FAEC populations.

Genealogical relationships and geographic distribution of haplotypes

Gene lineages were clearly split across regions within the native range. The most likely (i.e. parsimonious) genealogies between haplotypes are shown in Fig. 2. We found that the three most frequent haplotypes (i.e. H1, H18 and H22) belonged to three main clades (clade 1: H1–H5, clade 2: H15–H19 and clade 3: H21–H27). The genetic distances between these clades were large (11–20 mutational steps) compared to the genetic distances within clades (1–6 mutational steps). We found that only two haplotypes, H28 and H29, were distant from these three clades (9–15 mutational steps according to the cluster). The 10 other haplotypes, found at low frequencies, were close to each other and represented an unresolved torso in the middle of the network. The overall structure of the network was

1704 B. SIMON-BOUHET, P. GARCIA-MEUNIER and F. VIARD

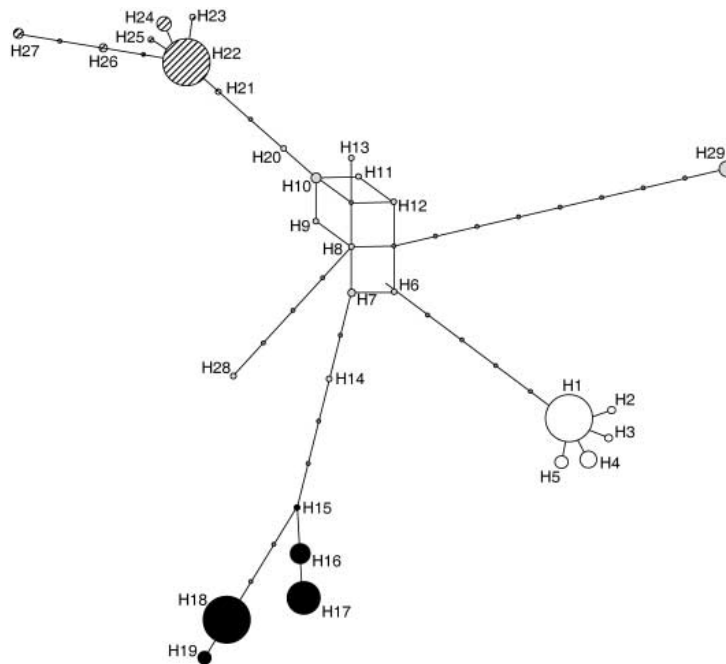


Fig. 2 Haplotype network showing the phylogenetic relationships between haplotypes. Circle sizes are proportional to haplotype frequencies over the whole data set. The three main clades are identified by different colours (plain black, plain white and black and white hatched circles). (●): inferred mutational step.

consistent with the geographic location of the native populations (Fig. 1). Within the native range, we could clearly identify three subgroups of populations corresponding to three geographic regions: (i) Southwestern populations (i.e. Iberian Peninsula plus Sardinia) were characterized almost exclusively by haplotypes belonging to clade 1; (ii) French Mediterranean populations were mainly characterized by haplotypes belonging to clade 2; and (iii) the Eastern populations (Sicily and Adriatic Sea populations) were mainly characterized by haplotypes belonging to clade 3. This apparent genetic clustering within the natural range was statistically significant according to AMOVA: the overall high and significant genetic structure ($\Phi_{ST} = 0.869$, $P < 10^{-3}$) was due to a large genetic differentiation between these three regions ($\Phi_{CT} = 0.846$, $P < 10^{-3}$) compared to the low genetic differentiation among the populations within these regions ($\Phi_{SC} = 0.152$, $P < 10^{-3}$). The correspondence between genetic and geographic structures is well illustrated by a neighbour-joining tree (Fig. 3), in which the three subgroups are clearly separated. We found that the populations of unknown status were closely related to the Iberian Peninsula populations. Conversely, the recently established populations were distributed over the whole tree suggesting several origins for the FAEC populations.

Genetic pattern along the French Atlantic and English Channel coasts

Contrary to the pattern observed among native populations, we found the genetic differentiation between the FAEC populations to be low ($\Phi_{ST} = 0.135$) but significant and associated with a high population genetic diversity. The five recently established populations exhibited haplotypes belonging to two or three of the clades as shown by the network analysis (Figs 1 and 2). Mismatch distribution analyses corroborated this co-occurrence of individuals having evolutionary divergent haplotypes (Fig. 4). Arguin, Ronces, Rivedoux and Stérec all showed at least two major peaks characterizing two or more groups of genetically divergent haplotypes co-occurring at the population level. In these four populations, the observed number of haplotypes (N_H) was significantly lower than expected under mutation-drift equilibrium (N_{EH} , K -tests in Fig. 4), suggesting a recent admixture (Hamblin & Veuille 1999). However, in Séné we observed a larger number of haplotypes that differed from each other by a large range of mutational steps (from 1 to 14), generating a more uniform mismatch distribution. Moreover, in this population was observed a larger number of haplotypes than expected (K -test, Fig. 4), suggesting a recent population expansion

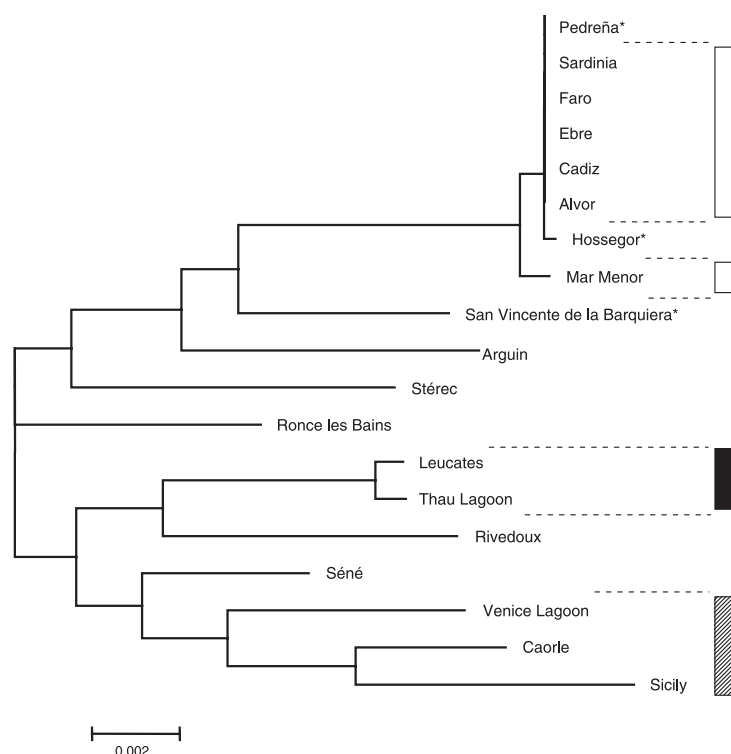
RANGE EXPANSION OF *C. NERITEA*: AN mtDNA SURVEY 1705

Fig. 3 Population neighbour-joining tree. Genetic distances between populations were estimated using the Kimura 2-parameters model. Coloured rectangles on the right of the tree show the main clade (as defined in Fig. 2) found in each native population. **sp.*: populations of unknown status.

or ancient balanced polymorphism (Depaulis & Veuille 1998). Finally, within the FAEC populations, we observed an increase in the pairwise F_{ST} estimates when we considered more distant populations, which resulted in a significant pattern of isolation by distance ($P = 0.016$, Fig. 5).

Discussion

Recurrent introductions from multiple sources

Although the overall level of mitochondrial polymorphism observed in our study of *Cyclope neritea* was comparable with other studies of the *COI* sequences below the species level [e.g. *Scylla serrata*, Gopurenko *et al.* (1999); *Patelloida profunda*, Kirkendale & Meyer (2004); *Littorina* sp., Wilding *et al.* (2000)], we observed a marked difference when comparing native and recently established populations. We noticed a low within-population genetic diversity associated with a strong genetic structure in the native range of *C. neritea*, whereas the FAEC populations exhibited a large within-population genetic diversity associated with a moderate genetic structure. With the exception of Séné, our study revealed that FAEC populations did not originate from a unique introduction event from a highly diversified native

population but primarily from recurrent introductions from multiple sources. More specifically, in these FAEC populations, we observed large haplotypic and nucleotide diversities and the co-occurrence of two or three groups of highly divergent haplotypes. Multipeak mismatch curves and significant K -tests confirmed that these four FAEC populations are not in demographic equilibrium and that admixture events (i.e. mixing of individuals from populations that have evolved independently) have occurred. None of the 11 native populations exhibited a mixture of the haplotypes belonging to the three clades identified in most of the FAEC populations (Fig. 1). Moreover, we found all native populations were much less diversified than the recently established populations, with the most polymorphic native population (Venice Lagoon) being less diversified than any of the recently established populations. In conclusion, the native populations showed a phylogeographic pattern with a strong consistency between their geographic locations and their genetic identity, which was confirmed by AMOVA and the neighbour-joining tree. At the population level, the recently established range has an increased genetic diversity compared to the native range, and, on a more regional scale, the strong genetic structure of the natural range had been erased in the recently colonized areas.

1706 B. SIMON-BOUHET, P. GARCIA-MEUNIER and F. VIARD

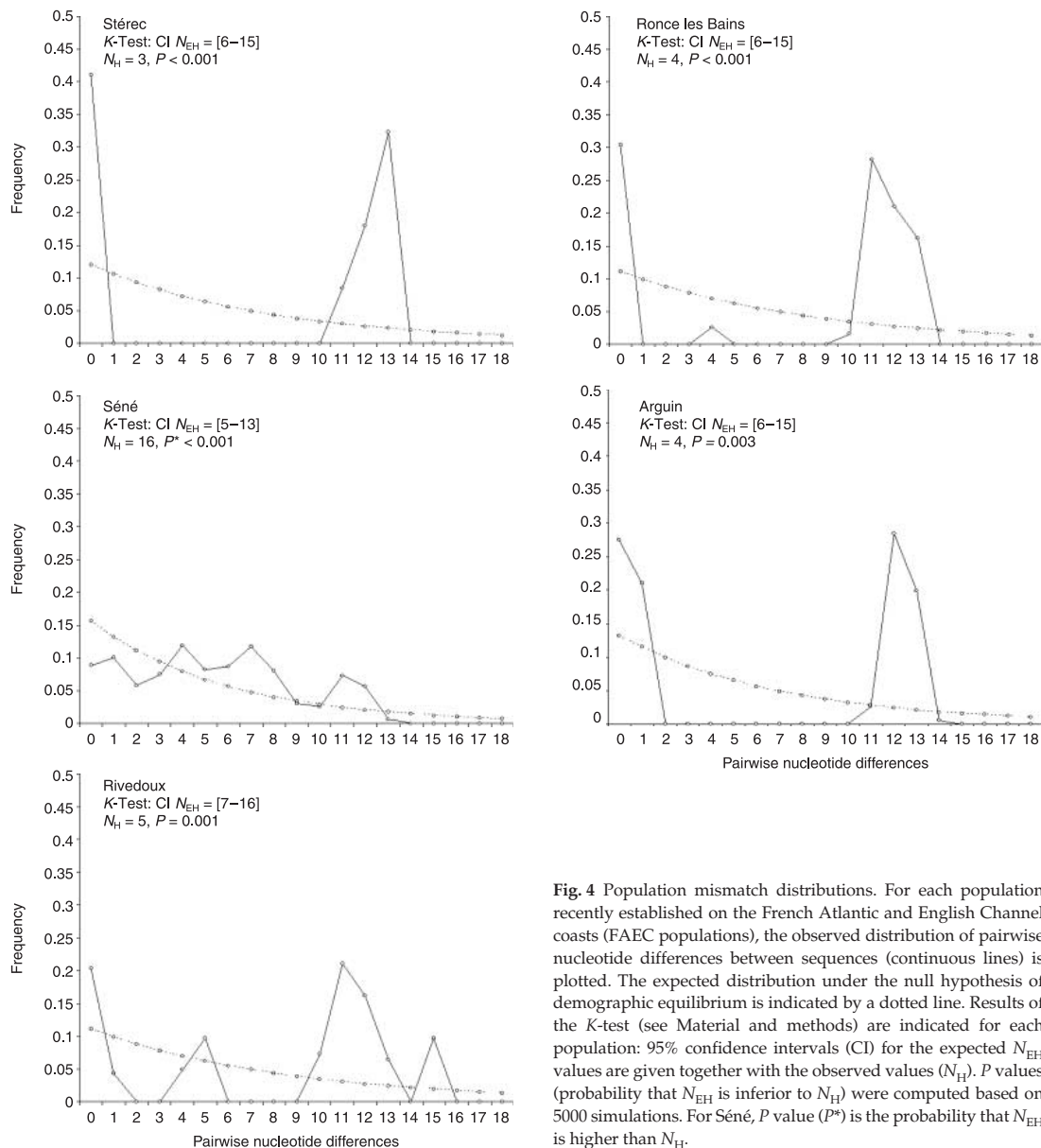


Fig. 4 Population mismatch distributions. For each population recently established on the French Atlantic and English Channel coasts (FAEC populations), the observed distribution of pairwise nucleotide differences between sequences (continuous lines) is plotted. The expected distribution under the null hypothesis of demographic equilibrium is indicated by a dotted line. Results of the K-test (see Material and methods) are indicated for each population: 95% confidence intervals (CI) for the expected N_{EH} values are given together with the observed values (N_H). P values (probability that N_{EH} is inferior to N_H) were computed based on 5000 simulations. For Séné, P value (P^*) is the probability that N_{EH} is higher than N_H .

As suggested earlier for Arguin (Bachelet *et al.* 2004), the results of our study shows that *C. neritea* populations recorded along the French coasts have been introduced since the 1970s and also strongly suggests that the introduction involved recurrent introduction events from several poorly diversified but highly divergent native populations. This is consistent with many genetic studies of terrestrial, fresh-

water and marine nonindigenous species (NIS) that failed to show the expected decrease of genetic polymorphism for numerous introduced species [Poaceae, Novak *et al.* (1993), brown mussel, Holland (2001), thiarid snail, Facon *et al.* (2003), slipper limpet, Dupont *et al.* (2003), Cuban lizard, Kolbe *et al.* (2004) or Japanese alga, Voisin *et al.* (2005)]. As observed by Kolbe *et al.* (2004), when native populations

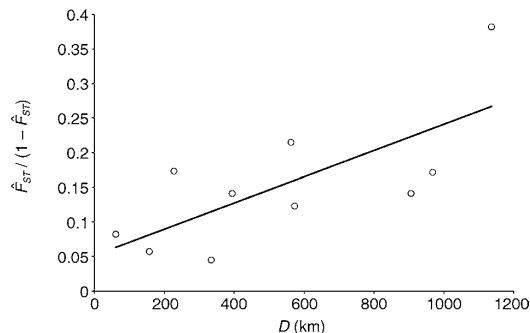


Fig. 5 Isolation by distance among the five French Atlantic and English Channel (FAEC) populations. The genetic distance $F_{ST}/(1 - F_{ST})$ for each population pair is plotted against geographic distance (D) in kilometres (see Materials and methods). The P value associated with a Mantel test for a positive correlation between genetic and geographic distances is given.

are monomorphic, the occurrence of multiple introductions can promote the settlement of NIS by transforming among-population variation in native ranges to within-population variation in introduced areas.

Alternative vector of introduction for Séné

Oyster exchanges, which are known to be responsible for numerous NIS introductions along the Atlantic coasts of Europe (Goulletquer *et al.* 2002; Wolff & Reise 2002), may have played a major role in the colonization of the FAEC coasts by *C. neritea* by introducing, near oyster beds, many individuals from various genetically differentiated populations. Such mechanisms and vectors have been suggested for other coastal marine introductions (see for instance *Ocenebrellus inornatus*, Martel *et al.* 2004). However, other shellfish cultures may also have contributed to the successful settlement of *C. neritea* along the FAEC coasts. This is exemplified by the particular genetic pattern of Séné, which showed: (i) a higher number of haplotypes than expected under the equilibrium hypothesis (K -test, Fig. 4); (ii) a position close to the populations from the Adriatic Sea in the neighbour-joining tree (Fig. 3); and (iii) the only nonsignificant genetic differentiation (55 pairwise Φ_{ST} comparisons) between the five recently established and the 11 native populations of our data set (i.e. between Séné and Venice Lagoon $\Phi_{ST} = 0.045$, $P = 0.072$; data not shown). A limited number of genetically diversified source populations, one of which having recently undergone a large demographic expansion, would seem a more probable explanation than the numerous introductions from less diversified but highly divergent populations that we observed in the other FAEC populations. This is consistent with the frequent exchanges of Japanese Manila clams, *Tapes philippinarum*, occurring

between the Venice Lagoon and shellfish culture sites of the Gulf of Morbihan, where Séné is located [J.-F. Auvray (SATMAR), personal communication]. This may be the principal or even the unique vector for the introduction of *C. neritea* in this area. This result illustrates the difficulty in generalizing the identification of a vector of introduction for a given NIS over its full range of introduction. As recently demonstrated for other marine NIS (e.g. the alga *Undaria pinnatifida*; Voisin *et al.* 2005), comprehensive studies of introduced populations are necessary for an inventory of all the factors/vectors that may promote the successful settlement of NIS. Nevertheless, all the French populations are characterized by at least one common feature: a large genetic diversity compared to populations from the native range. This has also been demonstrated in numerous successful NIS, suggesting the importance of propagule pressure for invasion success (Williamson 1996; Sax & Brown 2000).

Combination of natural spread and human-mediated introductions along the FAEC coasts

Human-mediated introductions associated with shellfish (i.e. oysters and/or Manila clams) transport are likely to have played a major role in the primary colonization process by *C. neritea* of the FAEC coasts. Recently established populations gradually appeared from south and moved northward and an isolation-by-distance pattern could be clearly seen between these populations (Fig. 5). This IBD may be explained either by secondary introductions between shellfish culture sites or by a natural spread from either the Iberian Peninsula or the primary site of introduction (i.e. Arguin). As we have no precise monitoring data about oyster exchanges between shellfish culture sites and contamination by *C. neritea*, we cannot definitively rule out secondary introductions. However, human-mediated exchanges between French shellfish culture sites occur randomly between very distant and noncontinuous locations. In particular, all French oysters farms (from the English Channel, Atlantic and even Mediterranean culture sites) rely on oyster spats only produced in the bays of Arcachon and Marennes-Oléron (Goulletquer & Heral 1997). If random human-mediated dispersal was the only process affecting *C. neritea* populations, the haplotypes should have been randomly mixed across the populations resulting in the absence of the apparent step-by-step process revealed by the IBD pattern. Given the low dispersal ability of the species, it is appealing to suggest a natural spread to explain the observed IBD although it is unlikely that such a pattern could have been produced over only 20–30 years. Nevertheless, natural spread could have occurred over a longer time period if the *C. neritea* populations had been present at low densities, and consequently unnoticed, along the FAEC coasts long before 1976. The patchy distribution of *C. neritea*

1708 B. SIMON-BOUHET, P. GARCIA-MEUNIER and F. VIARD

populations, even in its natural range, makes it difficult to detect even when population densities are not low.

Cyclope neritea can be classified as a cryptogenic species as defined by Carlton (1996). Such 'not demonstrably native or introduced' species probably represent a large part of the newly recorded species (Carlton 1996). Ancestral *C. neritea* populations may have been present long before 1976 either because of natural migrations from the most northern native populations or because of ancient and unrecognized accidental introductions. For example, the oyster *Ostrea edulis* has been exploited in France since the early 20th century, with exchanges occurring between the Thau lagoon and the Bay of Marennes-Oléron (Hinard & Lambert 1928). Sax & Brown (2000) recently demonstrated that species that suddenly become abundant and widespread often do so only after having failed to establish from earlier multiple natural or accidental introductions. The recent sustainable settlement of large and detectable *C. neritea* populations (from 400 to > 1600 ind./m² in Rivedoux, Tardy *et al.* 1985) may have been promoted by two factors: (i) an accelerated rate of accidental introductions in the 1970s due to intensified oyster transfers between Mediterranean and Atlantic shellfish farms and (ii) the rapid increase of temperature in the Bay of Biscay. By weakening native communities, both factors (Dukes & Mooney 1999) may have favoured the settlement of introduced individuals and migrants coming from the south, thus creating a demographic reinforcement of cryptic *C. neritea* populations along the French Atlantic coasts.

Populations at the edge: introduction vs. expansion

The three populations located at the edge of the previously recognized native range of the species, namely San Vicente de la Barquera (SVB), Pedreña and Hossegor, were classified as populations of unknown status (Table 1). Although located at the northern edge of its distributional range, this area is classified as part of the natural range of *C. neritea*, as its presence has been regularly reported in the past (Sauriau 1991). Moreover, shellfish have been cultured in this area for a long time and the populations sampled were all located near oyster farming sites (*C. neritea* populations were not found elsewhere, B. Simon-Bouhet, personal observation). Therefore, these populations could be either natural or introduced, justifying the classification of *C. neritea* as a cryptogenic species. Some of our results support a native status for these study populations because, for example, the three populations were genetically close to other Iberian Peninsula populations (Fig. 3). However, the three study populations did not share exactly the same characteristics. Hossegor, which is located at the very edge of the natural range showed features that we observed in the set of native populations (set 3, Table 1): (i) the haplotypic and nucleotidic diversities were closer to the values observed

in set 3 than in recently established populations and (ii) the two haplotypes (H1 and H5) found in this population belonged to the same clade as native populations (clade 1; Figs 1 and 2). Conversely, the SVB population showed features that we observed in most of the FAEC populations: high genetic diversity and the occurrence of two highly divergent haplotypes, namely H1 and H18 (found in clade 1 and 2, respectively) that were never found together in native populations. In Pedreña, the low sampling size prevents us from carrying out fine analysis. We found many more empty shells than in all other populations and we caught only three living individuals, suggesting that in the past this population may have been large but is now declining. Our sampling along the Iberian Peninsula was based on a survey made by Sauriau (1991) whose records for the occurrence of *C. neritea* since 1920 were reviewed. However, *C. neritea* was not found in most of the sites cited by Sauriau (1991) showing both the instability of populations and the cryptic nature of *C. neritea* in this area. The appearance and disappearance of *C. neritea* populations is well known in this geographical area (Kisch 1950; Morton 1960) and may be partly due to the closeness of a biogeographic boundary between northern temperate species and southern subtropical species (Glémarec 1979). Such natural barriers are characterized by very peculiar patterns for species distribution, hybridization between species (e.g. Bierne *et al.* 2003) or genetic differentiation at the population level (Luttikhuisen *et al.* 2003; Jolly *et al.* 2005). A more detailed study of *C. neritea* populations in this area would be useful, especially because the warming of the Bay of Biscay may remove selective constraints, leading to reinforcement of introduced populations by natural migrations northwards.

In conclusion, our study showed that the introduced status of *C. neritea* along the FAEC coasts may be associated with several vectors (e.g. oysters or Manila clam) and four out of five populations exhibited a genetic pattern consistent with introductions from multiple sources. Recurrent introductions and admixtures of several evolutionary lineages could have been the starting point for a rapid demographic expansion in sites where this cryptogenic species may have been present at low densities in the past. Bachelet *et al.* (2004) suggested biological features that could make *C. neritea* a particularly competitive gastropod. The temperature increase in the Bay of Biscay may also facilitate its long-term settlement by allowing larger scale migrations northwards, reinforcing the potential for this gastropod to become a real invasive species. Genetic and/or ecological surveys are clearly needed to study those NIS located at the edge of their natural range.

Acknowledgements

The authors are grateful to Jacques Pigeot, Davide Tagliapietra, Francesco Patti and Jeroen Jansen for their help for collecting

samples. We acknowledge Morgan Perennou from the Ouest-Genopole®/Genomer platform at Station Biologique de Roscoff for technical assistance, Marie-Hélène Muller for providing a program she developed for simulating gene genealogies and Carolyn Engel and Claire Daguin for critical reading and improvements of the manuscript. This project is part of the INVABIO program of the Ministère de l'Écologie et du Développement Durable (project n°D4E/SRP/01115). F.V. acknowledges supports from the European Network of Excellence 'Marine Genomics' (contract n°505403), the CNRS (ATIP program) and the Région Bretagne (Program Accueil & Emergence). B. S.-B. benefited from a PhD grant from the Ministère de l'Éducation Nationale et de la Recherche.

References

- Bachelet G, Simon-Bouhet B, Desclaux C *et al.* (2004) Invasion of the eastern Bay of Biscay by the nassariid gastropod *Cyclope neritea*: origin and effects on the resident fauna. *Marine Ecology Progress Series*, **276**, 147–159.
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.
- Bierne N, Borsa P, Daguin C *et al.* (2003) Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M. galloprovincialis*. *Molecular Ecology*, **12**, 447–461.
- Bohonak AJ (2002) IBD (ISOLATION BY DISTANCE): a program for analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.
- Boulhic M, Tardy J (1986) *Cyclope neritea* (Linné, 1758): Gastéropode Prosobranchie nouvellement implanté sur les côtes de Charente-Maritime; éthologie, structure de la population. *Halictis*, **15**, 41–49.
- Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology*, **77**, 1653–1655.
- Carlton JT (2000) Global change and biological invasions in the oceans. In: *Invasive Species in A Changing World* (ed. Mooney HA, Hobbs RJ), pp. 31–53. Island Press, Washington, D.C.
- Carlton JT, Geller JB (1993) Ecological roulette: the global transport of non-indigenous marine organisms. *Science*, **261**, 78–82.
- Cassens I, Van Waerebeek K, Best PB *et al.* (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Molecular Ecology*, **12**, 1781–1792.
- Cohen AN, Carlton JT (1998) Accelerating invasion rate in a highly invaded estuary. *Science*, **279**, 555–558.
- Comes HP, Kadereit JW (1998) The effect of quaternary climatic changes on plant distribution and evolution. *Trends in Plant Sciences*, **3**, 432–438.
- Depaulis F, Veuille M (1998) Neutrality tests based on the distribution of haplotypes under an infinite-site model. *Molecular Biology and Evolution*, **15**, 1788–1790.
- Díaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapegue S (2004) Reduced female gene flow in the European flat oyster *Ostrea edulis*. *Journal of Heredity*, **95**, 510–516.
- Dukes JS, Mooney HA (1999) Does global change increase the success of biological invaders? *Trends in Ecology & Evolution*, **14**, 135–139.
- Dupont L, Jollivet D, Viard F (2003) High genetic diversity and ephemeral drift effects in a successful introduced mollusc (*Crepidula fornicata*: Gastropoda). *Marine Ecology Progress Series*, **235**, 183–195.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Facon B, Pointier J-P, Glaubrecht M *et al.* (2003) A molecular phylogeography approach to biological invasions of the New World by parthenogenetic thiarid snails. *Molecular Ecology*, **12**, 3027–3039.
- Galtier N, Depaulis F, Barton NH (2000) Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics*, **155**, 981–987.
- Gaylord B, Gaines SD (2000) Temperature or transport? Range limits in marine species mediated solely by flow. *American Naturalist*, **155**, 769–789.
- Glémarec M (1979) Les fluctuations temporelles des peuplements benthiques liées aux fluctuations climatiques. *Océanologica Acta*, **2**, 365–371.
- Gomez A, Lunt DH (in press) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity* (eds Weiss S, Ferrand N). Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Gomoiu M (1964) Biologisches Studium der Arten *Nassa reticulata* (L.) und *Cyclonassa neritea* (L.) im schwarzen Meer (rumänischer Küstenbereich). *Revue Roumaine de Biologie, Série de Zoologie*, **9**, 39–49.
- Gopurenko D, Hughes JM, Keenan CP (1999) Mitochondrial DNA evidence for rapid colonisation of the Indo-West Pacific by the mudcrab *Scylla serrata*. *Marine Biology*, **134**, 227–233.
- Gouletquer P, Heral M (1997) Marine molluscan production trends in France: from fisheries to aquaculture. In: *The History, Present Condition and Future of the Molluscan Fisheries of North and Central America and Europe Volume 3, Europe* (eds Mackenzie CL Jr, Burrell VG Jr, Rosenfield A, Hobart WL), pp. 137–164. NOAA technical report NMFS, Seattle, Washington.
- Gouletquer P, Bachelet G, Sauriau P-G, Noël P (2002) Open Atlantic coast of Europe — a century of introduced species into French waters. In: *Invasive Aquatic Species of Europe. Distribution, Impacts and Management* (eds Leppäkoski E, Gollasch S, Olenin S), pp. 276–290. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Groszolz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology & Evolution*, **17**, 22–27.
- Hamblin MT, Veuille M (1999) Population structure among African and derived populations of *Drosophila simulans*: evidence for ancient subdivision and recent admixture. *Genetics*, **153**, 305–317.
- Hinard G, Lambert L (1928) Tableau de l'ostréiculture française (deuxième partie). *Revue des travaux de l'office des pêches maritimes*, **1**, 61–127.
- Holland BS (2001) Invasion without a bottleneck: microsatellite variation in natural and invasive populations of the brown mussel *Perna perna* (L.). *Marine Biotechnology (New York)*, **3**, 407–415.
- Hummel H, Bogaards RH, Bachelet G *et al.* (2000) The respiratory performance and survival of the bivalve *Macoma balthica* (L.) at the southern limit of its distribution area: a translocation experiment. *Journal of Experimental Marine Biology and Ecology*, **251**, 85–102.
- IPCC (2001) *Climate change 2001: The scientific basis*. Technical summary of the working group I report (eds Albritton DL, Meira Filho LG). Cambridge University Press, Cambridge, England.

1710 B. SIMON-BOUHET, P. GARCIA-MEUNIER and F. VIARD

- Jolly MT, Jollivet D, Gentil F, Thiebaut E, Viard F (2005) Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the north coast of France. *Heredity*, **94**, 23–32.
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.
- Kirkendale LA, Meyer CP (2004) Phylogeography of the *Patelloida profunda* group (Gastropoda: Lottidae): diversification in a dispersal-driven marine system. *Molecular Ecology*, **13**, 2749–2762.
- Kisch BS (1950) Growth stages of *Cyclope neritea* (L.). *Journal of Conchology*, **23**, 108.
- Kolbe JJ, Glor RE, Rodriguez Schettino L *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.
- Le Roux A, Maheo R, Dagusan J (1988) Présence et particularités écologiques de *Cyclope neritea* (Linné 1758) (Gastéropode prosobranchie) dans le Golfe du Morbihan. *Haliotis*, **18**, 310.
- Luttkhuizen PC, Drent J, Baker AJ (2003) Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. *Molecular Ecology*, **12**, 2215–2229.
- Martel C, Viard F, Bourguet D, Garcia-Meunier P (2004) Invasion by the marine gastropod *Ocenebrellus inornatus* in France I. Scenario for the source of introduction. *Journal of Experimental Biology and Ecology*, **305**, 155–170.
- Morton J (1960) The habits of *Cyclope neritea*, a style-bearing stenoglossan gastropod. *Proceedings of the Malacological Society of London*, **34**, 96–105.
- Muller M-H (2002) Flux de gènes et domestication: l'histoire évolutive d'un complexe sauvage-cultivé — Approches expérimentales et théorique chez la luzerne pérenne (*Medicago sativa*), Biologie, physiologie des organismes et des populations, Thesis, Université des Sciences et Techniques du Languedoc.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Novak SJ, Mack RN, Soltis PS (1993) Genetic variation in *Bromus tectorum* (Poaceae): introduction dynamics in North America. *Canadian Journal of Botany*, **71**, 1441–1448.
- Pigeot J (1988) Présence du Gastéropode Nassaridé *Cyclope neritea* (Linné, 1788) dans le bassin de Marennes-Oléron. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**, 717–719.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- Rozas J, Rozas R (1999) DNASP (version 3): an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.
- Saitou N, Nei M (1987) The neighbor-joining method: a new method of reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.
- Sauriau P-G (1989) *Cyclope neritea* (Linné, 1758) dans le bassin de Marennes-Oléron, Bilan des observations de 1984 à 1986. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**, 995–1007.
- Sauriau P-G (1991) Spread of *Cyclope neritea* (Mollusca: Gastropoda) along the north-eastern Atlantic coasts in relation to oyster culture and to climatic fluctuations. *Marine Biology*, **109**, 299–309.
- Sax DF, Brown JH (2000) The paradox of invasion. *Global Ecology and Biogeography*, **9**, 363–371.
- Schneider S, Roessli D, Excoffier L (2001) ARLEQUIN: A software for population data analysis Version 2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.
- Stachowicz JJ, Terwin JR, Whitlatch RB, Osman RW (2002) Linking climate change and biological invasions: ocean warming facilitates nonindigenous species invasions. *Proceedings of the National Academy of Sciences, USA*, **99**, 15497–15500.
- Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.
- Tardy J, Gaillard J, Portères G (1985) *Cyclope neritea* (Linné, 1788), une espèce de gastéropode nassaridé nouvelle pour les côtes du centre-ouest Atlantique de France. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**, 391–396.
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.
- Voisin M, Engel C, Viard F (2005) Differential shuffling of native genetic diversity across introduced region in a brown alga: aquaculture vs. maritime traffic effects. *Proceedings of the National Academy of Sciences, USA*, **102**, 5432–5437.
- Wilding CS, Grahame J, Mill PJ (2000) Mitochondrial DNA *Col* haplotype variation in sibling species of rough periwinkles. *Heredity*, **85**, 62–74.
- Williamson M (1996) *Biological Invasions*. Chapman & Hall, London.
- Wolff WJ, Reise K (2002) Oyster imports as a vector for introduction of alien species into northern and western coastal waters. In: *Invasive Aquatic Species of Europe. Distribution, Impact and Management* (eds Leppäkoski E, Gollasch S, Olenin S). Kluwer Academic Publishers, Dordrecht/Boston/London.
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

This work is part of B. Simon-Bouhet's doctoral research. His thesis aims at studying the basis of recent changes in the geographical distribution of the gastropod *C. neritea*. P. Garcia-Meunier uses molecular genetics to study the effects of environmental processes on genetic variability in coastal species. F. Viard study historical and contemporary dispersal of marine algae and invertebrates, including introduced species, through population genetics approaches.

RANGE EXPANSION OF *C. NERITEA*: AN mtDNA SURVEY 1711

Appendix

Haplotype definition (H1 to H29: GenBank accession numbers: AY789970–AY789991 and DQ087210–DQ087216) and number of occurrence (N). Only polymorphic sites are indicated and positions are given in number of base pairs.

Haplotypes	N	11122222222333333333334444445 22445506912356689912234778890224770 159140344284044622970387472981250693
H1	263	AACACAATCGAAGCGCGGCTATATACGGTTGAGAT
H2	3	-----A-----
H3	3	-----A-----
H4	9	-----T-----
H5	5	-----C-----
H6	1	---G-----G---C---C---
H7	2	---G-----G---C---C---
H8	1	---TG-----G---C---C---
H9	1	---TG-----A---G---C---C---
H10	4	G---TG-----A---G---C---C---
H11	1	G---TG-----A---G---C---C---
H12	1	G---TG-----G---C---C---
H13	1	G---TG---G-----G---C---C---
H14	1	---GG-----G---C---C---
H15	1	---TGG-----T-T-----G---C---C---
H16	13	---TGG-----T-T-----CG-G---C---C---
H17	35	---TGG-----T-T-A---CG-G---C---C---
H18	119	---TGG-----T-T---T---G-G---C---C---GC
H19	6	---TGG-----T-T---T---G-GA-C---C---GC
H20	1	G---TG-C-----A---G---C---C---
H21	1	G---TG-C---GG-----A---G---C---C---
H22	97	G---TG-C---GG-----A---G---C-CA---C
H23	1	G---TG-C---GG-----A---G---C-CA---C
H24	7	G---GTG-C---GG-----A---G---C-CA---C
H25	1	G---TG-C---GG---AA-----G---C-CA---C
H26	3	G---TG-C---GG-----G---C-CA---C
H27	4	G---TG-C-AGGA-----G---C-CA---C
H28	1	---TG-----G-G-AC-C-GA---
H29	8	-G---TG---T-----T---TCG-GC---CCC---C

Résumé

L'aire de distribution des espèces évolue en fonction de facteurs biotiques et abiotiques, y compris sur de courtes échelles de temps. Récemment, deux facteurs ont vu leur rôle s'accroître : (i) la température (en relation avec le changement climatique global) et (ii) les activités humaines responsables notamment d'altérations des habitats, de pollutions et d'introductions biologiques.

Au cours de cette thèse, nous nous sommes intéressés à l'évolution récente de l'aire de distribution du gastéropode nassariidae *Cyclope neritea*, apparu il y a trente ans sur les côtes françaises de l'Atlantique et de la Manche, dans le prolongement direct de son aire de distribution naturelle. Autour de la Péninsule Ibérique, une étude de terrain a confirmé l'instabilité des populations de cette espèce dans cette région. En outre, la discontinuité d'habitat, et les propriétés biologiques de l'espèce semblent indiquer que l'expansion récente de l'aire de distribution de *Cyclope neritea* serait le fait d'introductions biologiques. L'analyse du polymorphisme d'un locus mitochondrial (fragment de 533 paires de bases du gène codant pour la Cytochrome Oxydase I) et de quatre locus microsatellites nous a révélé une forte structure phylogéographique, attendue chez une espèce à développement direct. Cette structure nous a permis de démontrer que l'apparition de populations en bordure de l'aire naturelle de l'espèce met principalement en jeu des processus d'introductions récurrents trouvant leur source dans plusieurs régions génétiquement divergentes de l'aire naturelle. Cependant, l'hypothèse d'une présence de populations peu denses et cryptiques issues d'une dispersion naturelle Sud-Nord ne peut pas être totalement exclue. L'évolution récente de l'aire de distribution de *Cyclope neritea* témoigne ainsi de la synergie entre processus naturels et anthropiques sur des échelles de temps historiques et contemporaines.

Mots-clés: Aire de distribution, expansion naturelle, introductions biologiques, milieu côtier, phylogéographie, flux de gènes, mollusque, *Cyclope neritea*.

Abstract

Species' range is influenced by both biotic and abiotic factors. Among them, two are of a growing influence : (i) temperature (in relation with global change) and (ii) human activities responsible for habitat alterations, pollutions and biological introductions.

We studied the recent evolution of the species range of *Cyclope neritea*, a nassariid gastropod that appeared along the French Atlantic and English Chanel coasts thirty years ago, at the edge of its previously recognised native range. A field survey around the Iberian Peninsula confirmed that populations located in this area are unstable over time. Moreover, habitat discontinuities and biological properties of the model species makes likely the hypothesis of a recent introduction to explain the sudden range expansion of *Cyclope neritea*. By studying mitochondrial polymorphism (533 base pair fragment of the COI gene) and four polymorphic microsatellite loci, we showed a strong phylogeographic structure in its native range, as expected for this direct- developing species. Thanks to this structure, we demonstrated that the appearance of new populations near the range limit of the species was mainly caused by massive and recurrent human-mediated introductions of individuals coming from several highly divergent sources in the native range. Nevertheless, we could not exclude the possibility of a historical presence of cryptic populations (low density) originating from natural spread that could be now reinforced by the temperature increase in the Bay of Biscay. The recent evolution of the range of *Cyclope neritea* is thus an example of the synergistic action of both natural and human-mediated processes, acting on long and short time scales.

Keywords: Species range, natural expansion, biological introductions, coastal areas, phylogeography, gene flow, mollusc.

