



HAL
open science

Evaluation des capacités adaptatives du bivalve *Macoma balthica* (L.) dans un contexte de changement global : analyse comparée des processus neutres et soumis à sélection

Vanessa Buren Becquet

► To cite this version:

Vanessa Buren Becquet. Evaluation des capacités adaptatives du bivalve *Macoma balthica* (L.) dans un contexte de changement global : analyse comparée des processus neutres et soumis à sélection. Sciences agricoles. Université de La Rochelle, 2011. Français. NNT : 2011LAROS323 . tel-00591636

HAL Id: tel-00591636

<https://theses.hal.science/tel-00591636>

Submitted on 27 Nov 2014

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

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

Evaluation des capacités adaptatives du bivalve *Macoma balthica* (L.) dans un contexte de changement global. Analyse comparée des processus neutres et soumis à sélection



THESE

présentée pour l'obtention du grade de

Docteur de l' Université de La Rochelle

par

Vanessa Becquet

Composition du jury

Mme Viviane Boulo
M Didier Aurelle
M Herman Hummel
M Paco Bustamante
Mme Pascale Garcia

Cadre Ifremer
Maître de Conférence
Professeur
Professeur
Professeur

Université de Montpellier II
Université de la Méditerranée
Netherlands Institute of Ecology, Yerseke
Université de La Rochelle
Université de La Rochelle

Sommaire

<i>Introduction Générale</i>	11
1- La littoralisation : conséquence de l'évolution des sociétés humaines ?	13
2- Quels impacts sur les écosystèmes naturels ?	14
2.1- La modification de l'habitat	15
2.2- La perturbation des cycles biogéochimiques	18
2.3- Le changement climatique	19
2.4- L'introduction d'espèces	23
3- Mécanismes adaptatifs et évaluation du risque sur les populations	25
3.1- La réponse individuelle	27
3.2- La réponse à l'échelle de la population	30
3.3- Cas particulier des populations en limite d'aire de répartition	30
4- Exemple du bivalve <i>Macoma balthica</i>	32
5- Objectifs de la thèse	35
<i>Partie I: Influence des contraintes anthropiques sur les processus démographiques des populations naturelles de <i>Macoma balthica</i></i>	39
Introduction	41
Chapitre I- Mise au point de nouveaux marqueurs moléculaires hypervariables	43
I.1- Choix des marqueurs moléculaires	45
I.1.1- Locus mitochondrial	45
I.1.2- Locus microsatellites	46
I.2- Description du marqueur mitochondrial	46
I.3- Mise au point des marqueurs microsatellites	47
I.3.1- Construction d'une banque enrichie en microsatellites	47
I.3.2- Séquençage et sélection des marqueurs microsatellites	48
I.3.3- Résultats et conclusion	50
Chapitre II- Etude d'un site anthropisé en limite d'aire de répartition: cas des populations de la baie de Marennes Oléron (France) soumises au réchauffement des eaux du golfe de Gascogne	51
II.1- Contexte actuel des populations de <i>Macoma balthica</i> en limite d'aire de répartition	53
II.2- Matériel et méthodes	55
II.2.1- Echantillonnage et extraction des ADN	55
II.2.2- Analyse des séquences obtenues avec le marqueur COI	56
II.2.3- Analyse des données nucléaires	57
II.2.3.1- Diversité génétique au sein des populations	57
II.2.3.2- Différentiation génétique et relations entre populations	57

II.3- Résultats	59
II.3.1- Structure mitochondriale	59
II.3.2- Diversité moléculaire	62
II.3.2- Structure génétique des populations	66
II.4-Discussion	68
II.4.1- Diversité génétique des populations situées en périphérie	69
II.4.2- Structure génétique en limite de répartition	70
II.4.2.1- <i>Dispersion à longue distance et barrière géographique</i>	70
II.4.3.1- <i>Rupture du flux de gènes dans le golfe de Gascogne</i>	71
II.5- Conclusion	71
Chapitre III- Etude d'un site anthropisé au coeur de l'aire de répartition. Cas des populations de la Baie de Gdansk (Pologne) soumises aux contaminants de la Vistule et aux contraintes physico-chimiques de la mer Baltique	73
III.1- La mer Baltique: un environnement particulier ?	75
III.2- Matériel et méthodes	76
III.2.1- Echantillonnage des populations	77
III.2.2- Conditions d'amplification et analyses statistiques	78
III.3- Résultats	78
III.3.1- Structure mitochondriale des populations de la mer Baltique	78
III.3.2- Diversité moléculaire en Baie de Gdansk	80
III.3.3- Structure génétique des populations de la Baie de Gdansk	82
III.3.4- Assignation des individus	82
III.4- Discussion et conclusion	85
III.4.1- Statut des populations de la mer Baltique	85
III.4.2- Structure génétique des populations de la Baie de Gdansk	85
III.4.2.1- <i>Importance de la thermo-halocline</i>	85
III.4.2.2- <i>Structuration des populations à l'embouchure de la Vistule</i>	86
III.5- Conclusion	88
<i>Partie II : Evaluation des capacités adaptatives de <i>Macoma balthica</i> et approche moléculaire de la sélection</i>	91
Importance de la sélection dans les processus adaptatifs	93
Chapitre I- Les nouvelles technologies de séquençage haut débit. Cas particulier du pyroséquençage 454	97
I.1-Description des principales technologies de séquençage massif ou NGS (<i>Next Generation Sequencing</i>) ; séquençage de nouvelle génération	99
I.2- La technique «454» ou pyroséquençage	101
I.3- Développement de la banque d'ADN génomique	101
I.3.1- Echantillonnage et extraction des ARN	101

I.3.2- Préparation de la banque d'ADNc et pyroséquençage	103
I.4- Analyses bioinformatiques des séquences obtenues	104
I.4.1- Tri et assemblage des séquences	104
I.4.2- Annotation des séquences	104
I.4.3- Comparaison des transcriptomes	107
Chapitre II- Résultats du pyroséquençage et mise en évidence de gènes candidats	109
II.1- Séquençage et profil des ADNc exprimés	111
II.2- Annotation fonctionnelle des transcriptomes	113
II.3- Activités peptidiques mises en évidence	113
II.2- Discussion et conclusion	113
II.2.1- Annotation du transcriptome de <i>Macoma balthica</i>	113
II.2.2- Mise en évidence de gènes impliqués dans le système de défense	114
Chapitre III- Adaptation différentielle selon les contraintes environnementales	117
III.1- Résultats	119
III.1.1-Comparaison des populations A et G	119
III.1.2-Comparaison des populations A et S	120
III.1.3-Comparaison des populations G et S	120
III.2- Discussion	122
III.3- Conclusion	123
<i>Conclusion Générale et Perspectives</i>	<i>125</i>
<i>Bibliographie</i>	<i>131</i>
<i>Annexes</i>	<i>147</i>

Préambule

Suite au programme européen BIOCOTBE pour lequel j'avais été recrutée en qualité de technicienne, j'ai souhaité reprendre mes études. Après la validation de mes acquis professionnels en octobre 2006, j'ai pu prétendre à un financement CIFRE auprès de l'ANRT qui m'a permis d'initier le présent projet. Néanmoins, fin 2007, j'ai obtenu un poste d'Ingénieur d'étude par concours externe au CNRS au sein du laboratoire LIENSS, et j'ai eu la chance de pouvoir continuer ma thèse en parallèle de mon activité professionnelle.

L'une des fonctions qui m'ont été confiées lors de ce recrutement étant le développement des approches de génomique dans le laboratoire, mon projet de thèse s'est tout naturellement intégré dans ces perspectives. Par ailleurs, il s'inscrit dans la thématique de l'équipe AMARE (Réponse des Animaux Marins à la variabilité Environnementale) et permet de faire le lien entre les différentes échelles d'intégration (molécule, individu et populations) dans la réponse adaptative.

En effet, les approches de génomique fonctionnelle, initiées dans ce travail constituent un outil fondamental pour la mise en place de futurs projets sur l'adaptation des populations dans un contexte de changement global et dans des perspectives de conservation de la biodiversité et d'évaluation du risque encouru par les populations naturelles dans les écosystèmes fortement anthropisés.

Depuis, j'ai intégré en 2009, l'ANR HIFLO (BLAN 08-1 320836, The genetic basis and history of adaptive differentiation in high gene flow marine species) dont l'objectif est l'étude du paradoxe de l'adaptation locale en milieu marin pour appréhender les mécanismes de l'adaptation contemporaine des espèces face aux changements environnementaux rapides engendrés par les pressions anthropiques (changements climatiques et la dégradation des milieux).

Introduction générale

	Population (Millions)	% de la population mondiale	% de la population côtière
Estuaires	1,599	27	71
Récifs coralliens	0,711	12	31
Mangroves	1,030	18	45
Herbiers	1,146	19	49
Total	5,596		

Tableau 1: Estimation de la répartition de la population dans les habitats côtiers (*Center for International Earth Science Information Network, 1995*). La somme des % est supérieure à 100 car certains habitats se chevauchent

1- La littoralisation, conséquence de l'évolution des sociétés humaines ?

Le littoral se définit comme l'espace de contact, l'interface entre la terre et la mer. Il est constitué d'un rivage, d'un avant-pays marin exploité et d'un arrière-pays continental dont les paysages, les aménagements et les activités, sont fortement marqués par la présence de la mer. Dans le monde entier, le littoral qui se répartit sur 1,6 million de kilomètres est un espace très convoité. Bien que très ancien, le processus de littoralisation (migration des populations vers les littoraux afin d'en exploiter les richesses halieutiques et de s'ouvrir sur le monde) s'est fortement développé depuis la deuxième guerre mondiale, avec actuellement, près de 40 % de la population mondiale vivant à moins de 100 km des côtes (MA¹ 2005).

Cent vingt trois pays dont les plus grandes mégapoles américaines, européennes et japonaises, sont ainsi concernés par un écosystème côtier (Burke *et al* 2001). La plupart des populations côtières vit à moins de 50 km d'un estuaire, d'une mangrove ou d'un récif corallien comme l'indiquent les résultats publiés par le CIESIN² en 1995 (Tableau 1) et la moitié des villes de plus de 500 000 habitants ont été construites à moins de 50 km de la côte (Kjerfve *et al* 2002) portant la densité mondiale sur les côtes à 99,6 personnes au kilomètre carré soit 3 fois plus qu'à l'intérieur des terres (données pour l'année 2000, Kay *et al* 2003). En France, un rapport de 2006 de l'IFEN³ indique que 10 % de la population vit dans les communes du littoral, celles-ci représentant seulement 4 % de l'ensemble du territoire français.

La littoralisation s'accompagne d'une véritable maritimisation de l'économie, favorisée par de nombreuses avancées technologiques et par le développement de voies maritimes. L'apparition de la machine à vapeur au 19^{ième} siècle et l'ouverture des canaux de Panama et de Suez, mais aussi le développement des conteneurs plus récemment, ont véritablement contribué à l'explosion des transports maritimes et au développement de nombreuses zones portuaires. Ces infrastructures permettent le transport de millions de tonnes de marchandises sur des milliers de kilomètres à travers le monde (5,8 milliards de tonnes de marchandises ont ainsi été transportées par voie maritime sur une distance moyenne de 7600 kilomètres en 2002, CNUCED⁴ 2007).

L'exploitation du milieu pour ses ressources a aussi largement contribué à la littoralisation, les écosystèmes côtiers figurant parmi les plus productifs au monde et rivalisant même avec les forêts tropicales en termes de productivité générale (matières

¹ MA: Millenium Assessment

² CIESIN: Center for International Earth Science Information Network

³ IFEN: Institut Français de l'ENvironnement

⁴ CNUCED: Conférence des Nations Unies sur le Commerce Et le Développement

premières et biens de consommation ; Primavera 1991, Spurgeon 1992, Barbier 1993). Outre les nombreux produits de la mer tels que poissons, crustacés et algues pêchés localement (Moberg et Ronnback 2003, Ronnback 1999), de nombreuses espèces d'intérêt commercial (crevettes, saumon, mérou, bar) utilisent ces habitats côtiers comme nurseries et y sont également pêchées. Le rendement commercial mondial de ces pêches est ainsi évalué à 34 milliards de dollars par an (MA 2005). La valeur commerciale d'un écosystème côtier peut d'ailleurs être calculée à partir de la valeur des produits de la mer qui y sont exploités, comme par exemple, celle des mangroves, comprise entre 750 et 16750 dollars par hectare (Ronnback 1999). Au niveau mondial, les écosystèmes côtiers les plus « rentables » sont sans conteste les récifs coralliens qui permettent notamment d'alimenter de nombreux restaurants asiatiques en produits de la mer (MA 2005).

Parallèlement à la pêche, l'aquaculture tient aussi une part importante dans l'exploitation des écosystèmes côtiers. En effet, un tiers de la consommation mondiale des produits de la mer provient de l'aquaculture (Lubchenco 2003). Ce secteur est florissant et rien qu'entre 1989 et 1998, sa production globale a doublé (Golburg 2001). Bien que le saumon et la crevette soient les espèces les plus « cultivées » au niveau mondial, de nombreux coquillages sont aussi exploités. Sur le littoral français, la conchyliculture représente une production moyenne de 200 000 tonnes par an (dont 130 000 tonnes d'huîtres) équivalent à environ 630 millions d'euros. Second producteur de coquillage en Europe, la France possède ainsi 53 000 concessions réparties sur environ 20 000 hectares (CNC⁵ 2006).

Le littoral n'est pas seulement exploité pour ses richesses alimentaires. La plaisance, la pêche récréative, la plongée sont des activités répandues sur l'ensemble des littoraux et leur impact économique est considérable (Farber 1988, King 1995, Morgan et Owens 2001). Le développement de ces activités touristiques grandissantes s'est accompagné par la mise en place de nombreuses infrastructures telles que des complexes hôteliers, des aéroports mais aussi des ports de plaisance, empiétant bien souvent sur le paysage côtier.

2- Quels impacts sur les écosystèmes naturels ?

La littoralisation n'est pas sans conséquences sur les écosystèmes et d'une manière générale, le développement des activités anthropiques se fait bien souvent au détriment de l'environnement. Ces dernières seraient ainsi responsables de l'augmentation considérable des extinctions d'espèces au cours des 300 dernières années, 1000 fois supérieures à celles enregistrées à partir des fossiles pendant toute l'humanité (UNEP⁶, 2009). Dans les 100 dernières années, 100 extinctions avérées d'oiseaux, de mammifères et

⁵ CNC: Comité National de la Conchyliculture

⁶ UNEP: United Nations Environment Programme

d'amphibiens ont été décrites et nombreuses sont les espèces encore menacées aujourd'hui (près de 12 % des espèces d'oiseaux et 23 % de mammifères selon l'IUCN⁷).

Si de nombreuses activités humaines sont susceptibles d'être à l'origine de ces extinctions, la transformation de l'habitat demeure toutefois le point de départ de ces dégradations, et ce, quelle que soit l'activité considérée (Vitousek *et al* 1997a). En effet une véritable spirale de dégradation de l'écosystème se met en place dès lors que l'on transforme un habitat (Figure 1), cette perturbation entraînant une cascade d'effets négatifs sur l'environnement (invasions biologiques, perturbation du climat...).

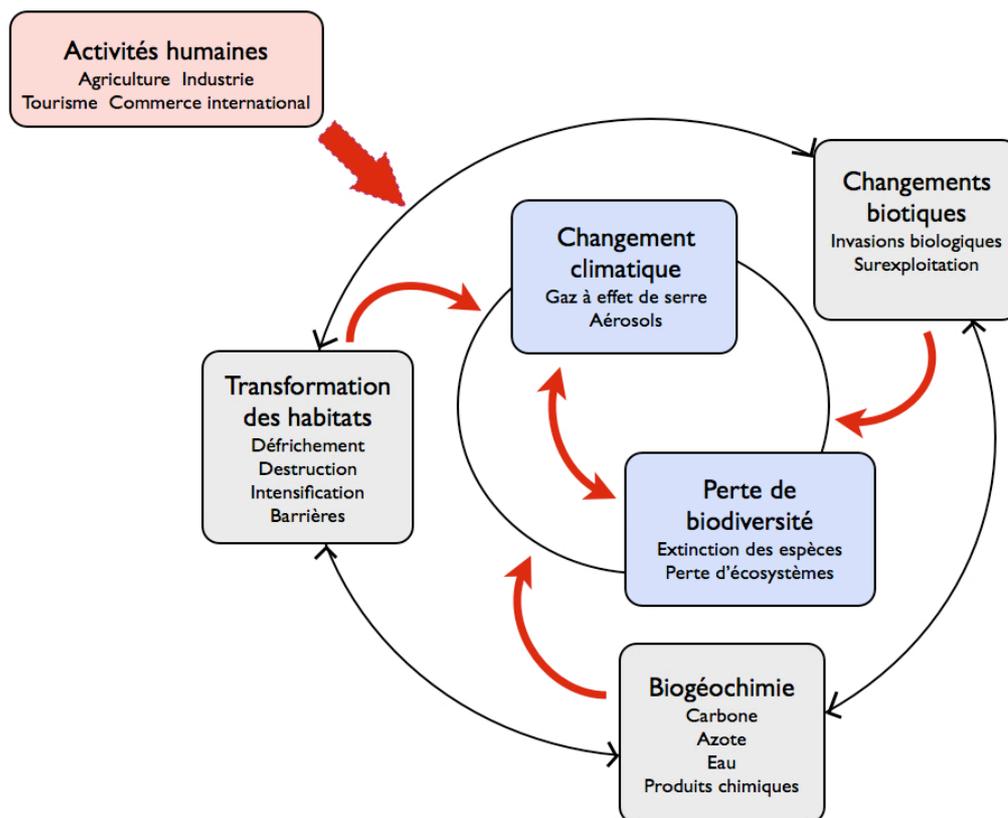


Figure 1: Conséquences des activités humaines sur les écosystèmes (inspiré de Lubchenco 1991)

2.1- La modification de l'habitat

La perte et la fragmentation de l'habitat représentent la cause majeure de perturbation de l'écosystème et d'atteinte à la biodiversité. Historiquement, l'habitat a

⁷ IUCN: International Union for Conservation of Nature

toujours été utilisé par l'homme que ce soit pour se nourrir (chasse, pêche) ou pour se loger. Cette exploitation du milieu a pris plusieurs formes, commençant avec la découverte du feu puis l'élevage des animaux, se développant rapidement avec la mise en place de l'agriculture et atteignant son point culminant avec l'ère industrielle.

La conversion du milieu en terres agricoles est ainsi une des premières causes de modification de l'habitat (Vitousek *et al* 1997a). Apparue au néolithique, l'agriculture, initialement un moyen de subsistance, s'est transformée peu à peu en une activité de production et de négoce. Son intensification après la 2nde guerre mondiale a alors permis aux agriculteurs d'accéder à des niveaux de production jamais atteints dans l'histoire humaine. La France qui produisait en 1950, 7 millions de tonnes de blé suffisant à peine à l'approvisionnement du pays, en produit aujourd'hui 40 millions de tonnes dont la moitié est exportée. Entre 1950 et 1980 (période de la révolution verte), davantage de surfaces ont été converties en terres cultivées qu'en 150 ans, entre 1700 et 1850. Un rapport de l'UNEP (2009) révèle qu'un quart de la surface terrestre est actuellement utilisée pour l'agriculture et que 10 à 20 % des prairies et forêts existantes aujourd'hui seront converties d'ici 2050.

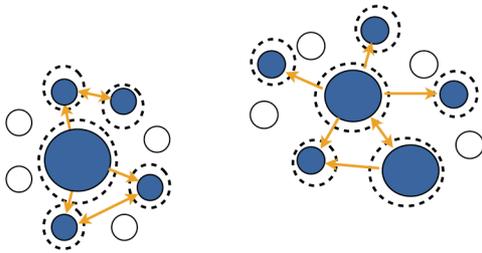
La surexploitation des habitats côtiers et le nombre grandissant d'activités récréatives ou professionnelles, nécessitent l'installation de réseaux routiers et le développement de nombreuses infrastructures (port, complexes hôteliers...) perturbant les écosystèmes naturels. Parmi eux, les mangroves et les récifs coralliens figurent probablement parmi les plus impactés (MA 2005). Le développement de l'aquaculture de la crevette (Ellison 2008) mais aussi de l'exploitation du bois (Valiela *et al* 2001) a ainsi provoqué la disparition de 38 % des mangroves mondiales au cours des vingt dernières années. De même, la pêche destructive (par les bombes ou le cyanure) a largement contribué à la destruction des récifs coralliens (notamment en Asie du sud-est) portant à 20 % la part de récifs détruits s'ajoutant aux 20 % supplémentaires actuellement fortement dégradés (MA 2005).

La destruction partielle ou totale d'un habitat, de même que la mise en place d'une barrière physique (e.g. une route, un barrage) vont favoriser le fractionnement de vastes écosystèmes en de nombreux fragments de taille réduite, isolés spatialement, et entraver ainsi le déplacement des individus. Or, la dispersion, processus fondamental dans l'expansion d'une espèce (Lomolino *et al* 2005), permet d'augmenter l'hétérozygotie au sein d'une population et de maintenir ainsi un niveau de diversité génétique favorable à l'adaptation. En outre, en s'éloignant de leur population natale, les individus optimisent leurs chances de trouver un partenaire sexuel et accèdent plus facilement à la ressource alimentaire, limitant ainsi les compétitions intra et interspécifique.

Plus généralement, la fragmentation des habitats tend à favoriser l'effet lisière et ainsi l'organisation en métapopulations (Levins 1969, Encadré 1). Les petites populations interconnectées entre elles mais éloignées du cœur de l'aire de répartition sont fragilisées. La réduction du flux de gènes ainsi générée par la présence de barrières et/ou l'isolement géographique peut conduire à la consanguinité. Ces populations de petite taille et de faible diversité génétique seront plus enclines à la dérive génétique et potentiellement moins bien adaptées.

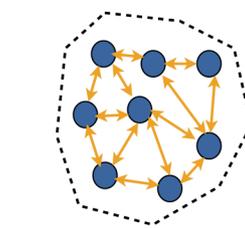
Les Métapopulations

Une métapopulation peut être définie comme une population de populations (Levins 1970), un groupe de populations locales interconnectées par des individus dispersants (Hanski 1991). Dans une métapopulation, les habitats sont fragmentés en portions plus ou moins favorables à la survie des individus. Il existe une variation spatiotemporelle de l'occupation des habitats, certains étant plus facilement occupés et d'autres plus souvent vides. Cet équilibre entre les extinctions locales et les recolonisations est maintenu par «l'effet rescue» défini par Brown (1977) qui consiste en l'immigration d'individus des populations environnantes vers les petites populations. Il existe différents modèles de population que Harrison (1991) a regroupé en quatre catégories décrites ci-dessous.



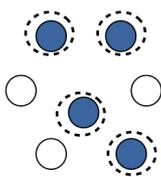
Modèle Continent-îles ou Source-puits

Les extinctions locales concernent uniquement certaines populations (populations îles et puits) et ont peu d'effet sur la persistance régionale (car les populations plus larges sont la source d'approvisionnement en migrants)



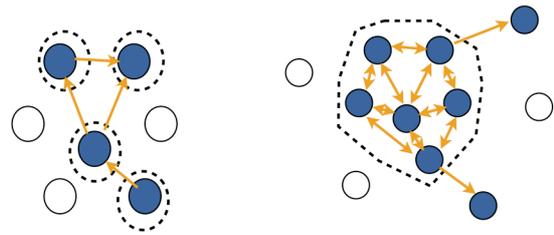
Modèle populations fragmentées

Variabilité spatiotemporelle de distribution des individus au sein des patches. Les habitats sont reliés entre eux par un fort taux de migration qui permet de les considérer comme une seule entité démographique. Il y a peu de risques d'extinction locale.



Modèle Populations fragmentées en déséquilibre

Les extinctions sont supérieures aux recolonisations, les extinctions locales sont liées au déclin régional de l'espèce. Conséquence d'une fragmentation, d'une réduction ou d'une détérioration de l'habitat à long terme, souvent liées à un impact anthropique.



Modèle Classique et intermédiaire

Dans le modèle classique de Levins (1969), la variation de la taille des habitats et le taux de migration entre eux sont modérés. Le modèle intermédiaire consiste en un modèle continent-îles fragmenté dans lequel la dispersion trop lente des individus du centre vers la périphérie provoque des extinctions locales.

Encadré 1: Les cercles pleins représentent des habitats occupés, les vides des habitats vacants. Les lignes pointillées représentent la limite des «populations». Les flèches indiquent une migration (colonisation)

2.2- La perturbation des cycles biogéochimiques

Le développement des activités liées à l'agriculture et à l'industrie a favorisé un relargage dans l'environnement de nombreux composés organiques de synthèse (Vitousek *et al* 1997a), dont la plupart sont toxiques, même à des concentrations faibles. L'exemple le plus probant est certainement l'utilisation des organochlorés, produits fortement utilisés dans les années 1950-60 dans l'agriculture (DDT⁸ des insecticides) ou dans l'industrie (PCB⁹). Bien qu'interdits à la vente depuis les années 1970, ces composés ont la particularité d'être persistants dans l'environnement et d'être détectés plusieurs décades après l'arrêt de leur utilisation (AFSSA¹⁰, Stow 1995). Ces molécules s'accumulent dans les sédiments marins et de rivière, et sont régulièrement bioaccumulées par de nombreux poissons et invertébrés (Coat 2005, Bocquené 2005), atteignant même les prédateurs supérieurs en s'amplifiant le long de la chaîne alimentaire. L'ajout de composés chimiques synthétiques dans l'environnement n'est pas la seule cause de la perturbation des cycles biogéochimiques. L'exploitation de l'habitat pour ses ressources a également entraîné une élévation de la charge en nutriments dans le milieu naturel. L'extraction et le brûlage du pétrole ont provoqué la libération massive de carbone fossile jusqu'alors prisonnier dans le sol, dans l'atmosphère. Ainsi, en 150 ans (du début de la révolution industrielle jusqu'aux années 1960), le taux moyen de carbone dans l'atmosphère a augmenté de 30 %, passant de 280 ppm à 362 ppm en 1957 (Schimel 1995). Parallèlement, la conversion des forêts et prairies en terres cultivées a considérablement réduit le captage du carbone atmosphérique par les végétaux (activité photosynthétique réduite, Encadré 2).

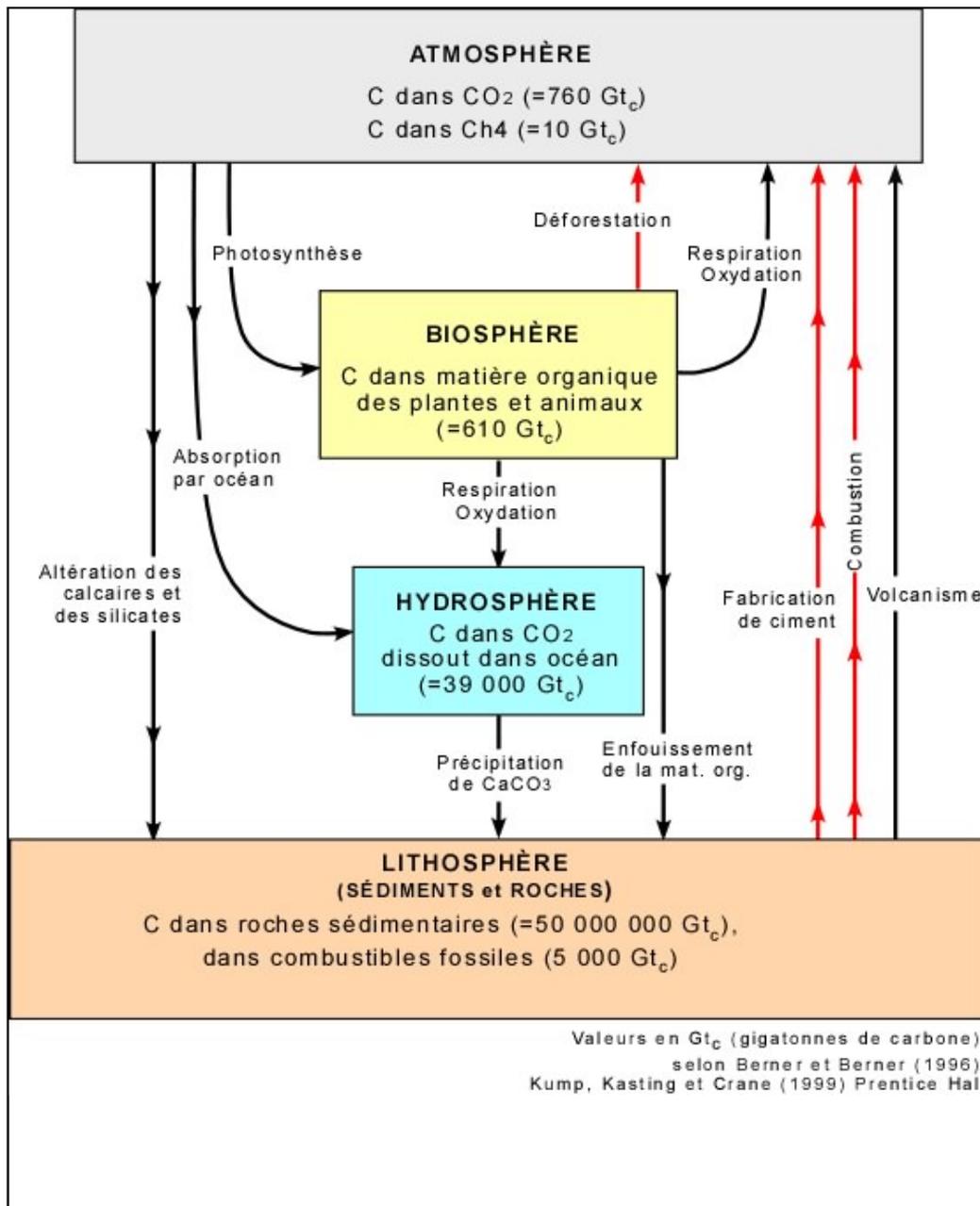
Le taux de carbone dans l'air est naturellement régulé par le taux d'azote dans le sol. Or depuis quelques décennies, ce taux d'azote tellurique a lui-même été modifié. La pollution routière et industrielle (dégageant essentiellement de l'oxyde d'azote) mais aussi l'agriculture (par l'apport massif d'engrais azotés) ont favorisé la saturation du sol en azote (IPPC¹¹ 1996), entraînant un déséquilibre nutritionnel chez les végétaux (notamment les arbres) avec une sensibilité accrue aux facteurs environnementaux (gel, sécheresse) et aux parasites (insectes, champignons). Les nombreux produits azotés ont contribué à l'acidification des sols via les pluies acides (Vitousek *et al* 1997b), conduisant à la réduction de leur fertilité et à la pollution des eaux et végétaux environnants. Dans le milieu aquatique, de nombreuses rivières ont ainsi été contaminées (Howarth *et al* 1996). Le déséquilibre en nutriments s'est transféré à la plupart des estuaires et des océans profitant à la prolifération de nombreuses algues (Hallegraeff 1993). Ces blooms alguaux de plus en plus nombreux sont responsables de l'eutrophisation du milieu marin, menaçant les ressources aquatiques (Nixon *et al* 1996).

⁸ DDT: DichloroDiphénylTrichloroéthane

⁹ PCB: PolyChloroBiphényle

¹⁰ AFSSA: Agence Française de Sécurité Sanitaire des Aliments

¹¹ IPPC: International Plant Protection Convention



Encadré 2: Cycle du carbone

2.3- Le changement climatique

La modification des niveaux d'azote et de carbone dans l'atmosphère au cours du siècle dernier, a favorisé l'augmentation des gaz à effet de serre, entraînant un dérèglement climatique sans précédent, caractérisé par une augmentation des précipitations

associées à une élévation importante de la température (GIEC¹² 2007). La température moyenne à la surface du globe a en effet augmenté de 0,74°C de 1906 à 2005, et une élévation globale de 2 à 6,4°C est prévue d'ici 2100. Parallèlement, la fonte massive des glaciers a entraîné une modification du niveau de la mer, qui est ainsi monté de 1,8 mm chaque année depuis 1961, ce phénomène s'accélégrant au cours des trente dernières années, passant à 3,1 millimètres par an.

Le réchauffement global s'accompagne de la perturbation des phénomènes météorologiques saisonniers tels que l'oscillation El Niño dans le Pacifique et la NAO¹³ dans l'Atlantique. L'oscillation El Niño est un phénomène qui favorise des hausses de la température atmosphérique associées à l'augmentation des précipitations, ceci en réchauffant la surface de l'eau habituellement froide des côtes du Pérou, du Chili et de l'Equateur. Ce phénomène de variation pluriannuel, appelé aussi ENSO (El Niño Southern Oscillation) a vu sa fréquence augmenter ces dernières années tout comme son homologue dans l'Atlantique nord (Hurrell et Van Loon 1997). La NAO est une variation naturelle du climat qui influence largement le climat de l'Europe de l'ouest, du nord de l'Afrique mais aussi de l'est des Etats-Unis (Figure 2). Ce phénomène qui affecte surtout la saison hivernale, dépend de la différence de pression atmosphérique entre l'anticyclone des Açores et la dépression islandaise. Selon que la différence de pression sera grande (NAO+) ou faible (NAO-), les hivers du nord ouest de l'Europe seront plutôt doux et humides ou plutôt froids et sec (Hurrell et Van Loon 1997). L'oscillation entre la NAO+ et la NAO- est régulière mais depuis les années 80, la fréquence des années NAO+ (comme celle des années El Niño) a augmenté (Visbeck *et al* 2001, Osborn 2006), provoquant des vagues de chaleur de plus en plus fréquentes (Figure 3) pouvant avoir des incidences néfastes sur les écosystèmes.

¹² GIEC: Groupe d'experts Intergouvernemental sur l'évolution du Climat

¹³ NAO: North Atlantic Oscillation

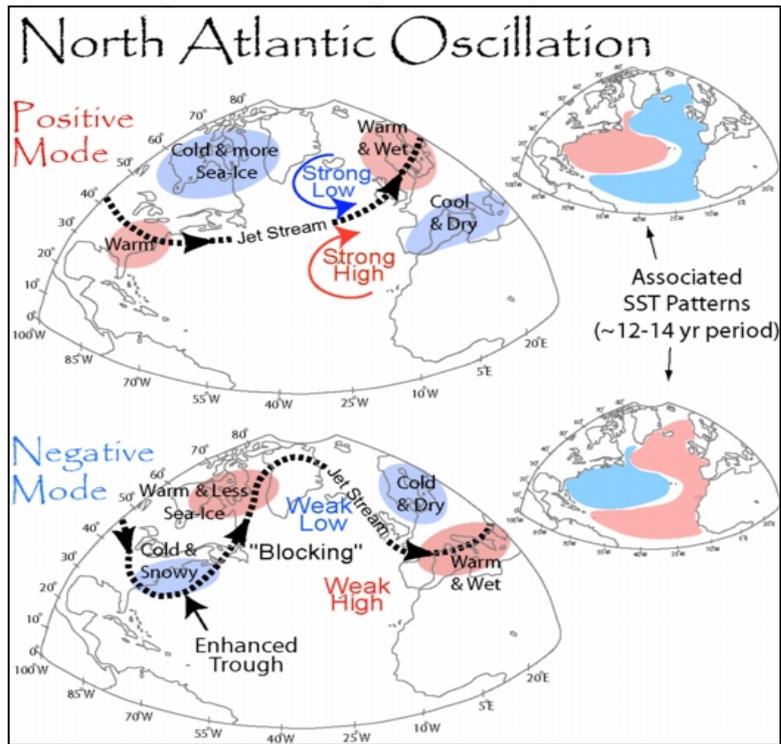


Figure 2: Evolution de l'indice NAO entre 1900 et 2000.
 Source: <http://www.cgd.ucar.edu/cas/jhurrell/indices.html> par Jim Hurrell

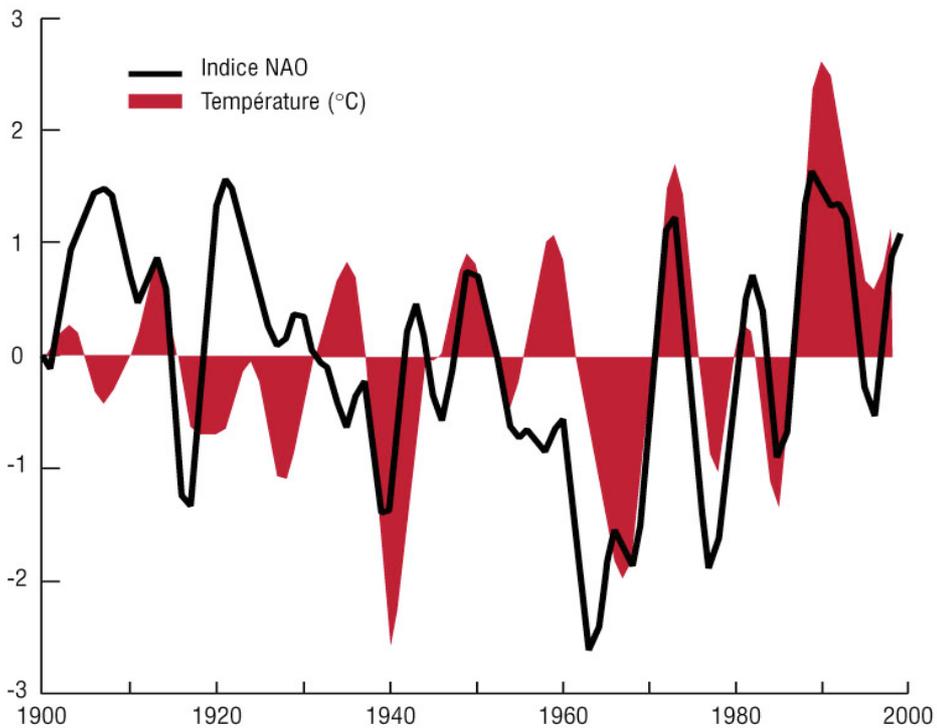


Figure 3 : Illustration des oscillations Nord atlantiques et températures associées des eaux de surfaces.
 Source : <http://airmap.unh.edu/background/nao.html> par James Bradburry et Cameron

Du fait de son importance dans la régulation du métabolisme des organismes, l'élévation rapide et, qui plus est, permanente de la température est susceptible d'avoir un impact au niveau des organismes ectothermes (capacité physiologique à compenser cette variation ou non), des populations et des communautés. De nombreuses études dont celle de Mac Carty (2001) ont été menées sur l'impact du réchauffement climatique montrant le rôle de ce paramètre dans la biologie et la dynamique des populations. Ainsi, les travaux de Janzen (1994) mettent en évidence une modification du sex-ratio des tortues peintes (*Chrysemys picta*) avec un fort déséquilibre en faveur des femelles, liée au réchauffement climatique.

La température peut également affecter l'abondance d'une espèce dans un milieu. Seather *et al* (2000) a ainsi montré que le nombre de cincles plongeurs (*Cinclus cinclus*) dans le sud de la Norvège avait augmenté en réponse à l'élévation de la température hivernale, la succession des hivers doux favorisant l'accès à la nourriture. A contrario, la population de puffin fuligineux dans l'ouest de l'Amérique du Nord a connu un déclin de 90 % entre 1987 et 1994 suite au réchauffement du courant Californien (Veit 1996 et 1997).

Le caractère hâtif des phénomènes printaniers liés à l'élévation de la température (Menzel *et al* 2006) peut aussi entraîner un décalage dans la phénologie des espèces. De nombreuses études sur la migration des oiseaux ont ainsi montré une arrivée précoce des oiseaux au printemps, notamment en Europe et en Amérique du Nord (Jian-bin 2005, Rubolini 2007b). Ce décalage phénologique serait favorisé par la présence précoce d'insectes et de végétaux dans leurs aires de passage et de reproduction (Schwartz 2006) mais également par de meilleures conditions de vol du fait des températures printanières élevées (Marra *et al* 2005, Sinelschikova 2007).

L'élévation de la température observée au cours de ces dernières années, impacte également l'aire de répartition des espèces. Une étude de Barry (1995) sur 45 espèces d'invertébrés marins des côtes californiennes a ainsi montré une migration vers le nord de la plupart de celles-ci, ainsi qu'un changement dans leur abondance (chez 35 espèces sur 45). Ce résultat issu d'une comparaison de la composition des communautés sur 60 ans (entre 1933 et 1993), semblait être corrélé avec le réchauffement des eaux de surface et l'augmentation de la fréquence des phénomènes ENSO (El Niño Southern Oscillation). Le changement dans la distribution mais aussi dans l'abondance des algues, du plancton et de nombreux poissons a été récemment confirmé par le GIEC dans son rapport sur l'impact des changements globaux sur les écosystèmes (2009). Par ailleurs, une étude récente de Parmesan et Yohe (2003) sur plus de 1500 espèces terrestres a aussi révélé une tendance croissante des animaux et des végétaux à migrer vers les pôles et en altitude.

De tels changements en termes d'abondance et de distribution sont à même de perturber la composition des communautés en modifiant l'équilibre proie/prédateur et les compétitions inter et intraspécifique. Par exemple, dans le sud de la Californie, de nombreux poissons récifaux à affinités tropicales ou subtropicales ont pu remonter vers le nord, entrant alors en compétition avec les espèces locales (Holbrook *et al* 1997).

Enfin, le réchauffement climatique ainsi que le renforcement du taux de CO₂ dans l'atmosphère ont également favorisé la modification de l'organisation de nombreuses communautés de végétaux, d'origine arctique (Chapin *et al* 1995), tempérée (Alward *et al* 1999) ou tropicale (Philips 1996). Les prairies arides du sud-ouest des USA ont ainsi été remplacées par un désert de broussailles suite à une perturbation des précipitations dans cette région (Brown 1997). Cette modification du paysage s'est accompagnée de l'extinction de nombreuses espèces d'animaux (coextinction).

2.4- L'introduction d'espèces

Le développement des transports et les échanges internationaux ont favorisé l'introduction de nombreuses espèces à travers le monde (Elton 1958, Di Castri *et al* 1990, Hobbs et Huenneke 1992, Ruiz *et al* 2000, *National Research Council* 2002). Les échanges commerciaux par voie maritime ont ainsi mis en place de véritables corridors d'invasion tels que la voie de l'Atlantique nord-ouest et des habitats côtiers de la Grande-Bretagne, celle entre les ports de l'Est asiatique et la baie de San Francisco ou bien celle des mers Ponto-Caspiennes responsable des récentes invasions de la région des Grands Lacs (Ricciardi et MacIsaac 2000). Alors que l'aire de répartition des espèces est naturellement contrainte par des barrières biogéographiques infranchissables (Lomolino *et al* 2005), des espèces de toute origine (végétale, animale ou microbienne) sont régulièrement introduites par l'homme, de manière volontaire ou non, dans des milieux qui leur était jusqu'alors inaccessibles.

Parmi ces nombreuses espèces introduites, seule une petite fraction (Williamson et Fitter 1996) est capable de s'établir dans un nouvel environnement (Mooney et Cleland 2001), de s'y reproduire efficacement puis de se disperser pour coloniser ce nouveau milieu (Boudouresque 1999). Malgré les multiples obstacles qui s'opposent au processus d'invasion, nombreuses sont les espèces introduites qui réussissent cette colonisation. Tous les phylums sont affectés par ces introductions, et à la fin des années 1990, les USA comptait une vingtaine d'espèces de mammifères introduites (Layne 1997), environ 10 % d'espèces d'oiseaux (Temple 1992), 53 espèces d'amphibiens et reptiles (Mc Coid et Kleberg 1995, Lafferty et Page 1997) mais aussi 4500 espèces d'arthropodes (OTA¹⁴ 1993).

Les invasions biologiques marines, restent encore sous-estimées (Reise *et al* 1999, Ruiz *et al* 2000). Toutefois le nombre d'espèces non indigènes (tous phylums confondus) sur les côtes des Etats-Unis est estimé à près de 400 (Ruiz *et al* 1997) et des évaluations font par ailleurs état de 70 espèces introduites en Australie (Ruiz *et al* 1997), 80 en Mer du Nord (Reise *et al* 1999), 52 en Grande-Bretagne (Eno 1996) et 35 en Mer Baltique (Leppäkoski 1994). L'implantation de ces espèces exogènes est doublement facilitée par (i) la multiplicité et la récurrence des introductions générées par les voies de transport maritimes

¹⁴ OTA: Office of Technology Assessment

(7000 espèces marines ont déjà été introduites au niveau mondial par le biais des eaux de ballast), (ii) les modes d'exploitation aquacole qui optimisent leur dispersion (les échanges d'huîtres sur le littoral Atlantique sont à l'origine de nombreuses introductions d'espèces, Gouletquer *et al* 2002, Wolff et Reise 2002).

Si seulement 1% des espèces introduites sont considérées comme néfastes pour leur environnement (Mooney et Cleland 2001), la rapidité à laquelle elles apparaissent, reste inquiétante. En 500 ans, 3 % de la surface terrestre (non recouverte de glace) ont ainsi été envahies (Mack 1985) et le réchauffement climatique des 40 dernières années a souvent favorisé les espèces introduites, plus plastiques et plus à même de coloniser un milieu que les espèces natives (Bradley *et al* 2009, Hellman *et al* 2008). Dans une étude de 2000, Pimentel a ainsi répertorié 120 000 espèces (tous taxons confondus) bien implantées sur leur nouveau territoire, et estimé à 314 milliards de dollars par an les dommages causés par celles-ci.

Les introductions biologiques sont aujourd'hui considérées comme la deuxième cause d'extinction des espèces (UNEP 2009). Elles sont susceptibles de générer des changements de structure, de diversité et/ou de fonction dans les communautés envahies (Encadré 3). Les perturbations des paramètres physiques d'un environnement, sont susceptibles d'agir indirectement sur les communautés. Par exemple, des altérations du milieu dues à l'introduction de plantes exotiques ont été décrites (Vitousek 1990; Richardson *et al* 2000). Elles vont ainsi modifier la stabilité du sol, favoriser l'érosion, coloniser de nouveaux substrats ou affecter l'accumulation des ressources sur et dans le sol. En outre, en entrant en compétition pour l'espace et pour la ressource, les espèces exogènes peuvent avoir des incidences majeures sur la dynamique, la croissance, la reproduction voire la survie des espèces natives. En milieu marin, l'invasion massive et unilatérale de la Méditerranée (dans le sens mer Rouge – Méditerranée) par plus de 300 espèces d'algues, d'invertébrés et de poissons de la mer Rouge lors de l'ouverture du canal de Suez en 1869 (Por 1978), a provoqué des dommages importants au sein des communautés locales et facilité le déclin de nombreuses populations endémiques de Méditerranée (Galil 2000).

Dans certains cas, l'espèce native et l'espèce invasive vont pouvoir se reproduire menant à la formation d'individus hybrides. L'introgression des gènes exotiques dans le pool génétique natif sera d'autant plus rapide que le flux de gènes sera important. C'est par exemple le cas chez de nombreuses plantes où les gamètes femelles (graines) mais aussi mâles (pollen) vont facilement pouvoir se disperser. Des gènes favorables étant sélectionnés, les hybrides peuvent quelquefois se fixer et mener à de véritables nouvelles espèces. L'hybridation entre *Spartina alterniflora* (espèce invasive d'Amérique de l'Est introduite accidentellement en Angleterre à la fin du 19^{ème} siècle) et *Spartina maritima* (espèce locale) a ainsi donné lieu à une nouvelle espèce (*Spartina anglica*) particulièrement vigoureuse (Raybould *et al* 1991a) devenant elle-même par la suite une espèce invasive en Chine, Australie et Nouvelle-Zélande.

Les introductions multiples vont par ailleurs avoir tendance à maintenir une diversité élevée chez ces hybrides (Kolar et Lodge 2001) ce qui peut, à terme conduire à la diminution de la fitness de l'espèce native voire à son extinction (Rhymer et Simberloff 1996). Mc Millan et Wilcove (1994) ont ainsi montré que sur 24 espèces disparues aux Etats - Unis, 3 étaient le résultat d'une hybridation avec une espèce exotique.

Impacts des introductions biologiques

Conséquences écologiques

- Perte de la biodiversité native
- Modification de la structure des écosystèmes
- Modification du fonctionnement des écosystèmes
- Diminution de la qualité de l'eau

Conséquences évolutives

- Adaptation de l'espèce introduite
- Modifications génétiques non adaptatives (dérive) de l'espèce introduite
- Evolution de l'espèce native
- Hybridation entre espèces natives et introduites :
 - ⇒ pollution génétique des populations sauvages
 - ⇒ extinction de la population locale
 - ⇒ introgression
 - ⇒ spéciation

Conséquences économiques

- Coûts des nettoyages
- Coûts des traitements ou des mises en quarantaine
- Rendements de l'agriculture
- Interférence avec les ressources aquacoles
- Interférence avec les activités de pêche
- Industrie et énergie (dégâts des infrastructures)
- Perturbation du tourisme
- Altération des transports

Conséquences sanitaires et sociales

- Introduction d'espèces pathogènes
- Introduction d'espèces allergisantes
- Introduction d'espèces hôtes de pathogènes
- Introduction d'espèces créant un habitat favorable aux pathogènes

Encadré 3 : Impacts des introductions biologiques.

3- Mécanismes adaptatifs et évaluation du risque sur les populations

Le processus d'évolution se caractérise par sa nature double avec, d'une part le polymorphisme génétique comme facteur de diversification et d'autre part, la pression de sélection comme facteur d'adaptation. Ainsi, une sélection naturelle efficace va-t-elle s'exercer sur la variabilité génétique présente dans la population et amener le phénotype avec la meilleure valeur sélective à s'installer dans cet environnement. A l'échelle de l'espèce, les différentes adaptations locales des populations permettent de maintenir une diversité génétique susceptible d'être à son tour valorisée en cas de changement environnemental.

L'un des enjeux majeurs en biologie de la conservation est le maintien de la variabilité adaptative. Or nous avons montré plus haut les pressions que les actions anthropiques pouvaient exercer sur les écosystèmes naturels et leurs conséquences sur les populations, notamment en termes d'érosion de la variabilité. Dans un environnement instable, la majorité des espèces est généralement capable de s'adapter mais cet exercice peut s'avérer difficile lorsque le changement est trop rapide (e.g. changement climatique) ou trop complexe (e.g. changements globaux). La fragmentation des habitats, l'introduction d'espèces, la modification des paramètres physico-chimiques de l'environnement, sont autant d'agressions accumulées qui vont générer chez les espèces un stress qu'elles vont devoir gérer (Figure 4).

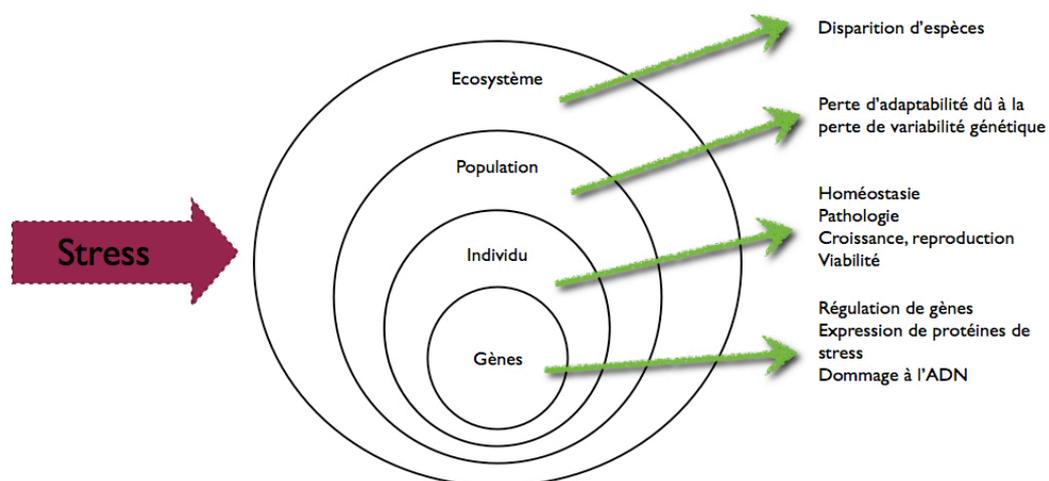


Figure 4 : Influence du stress à différents niveaux d'intégration

Le stress peut être défini comme une stimulation ponctuelle, susceptible de déclencher un ensemble de réactions non spécifiques de l'organisme, impliquant des réponses neuronales, neuroendocrines, physiologiques et comportementales. Face à des changements majeurs de l'environnement, différentes stratégies pourront être ainsi adoptées par les organismes, les populations et selon l'intensité et la durée du stress rencontré, une réponse sera mise en place à différents niveaux moléculaires, individuels, populationnels, voire à l'ensemble d'un écosystème.

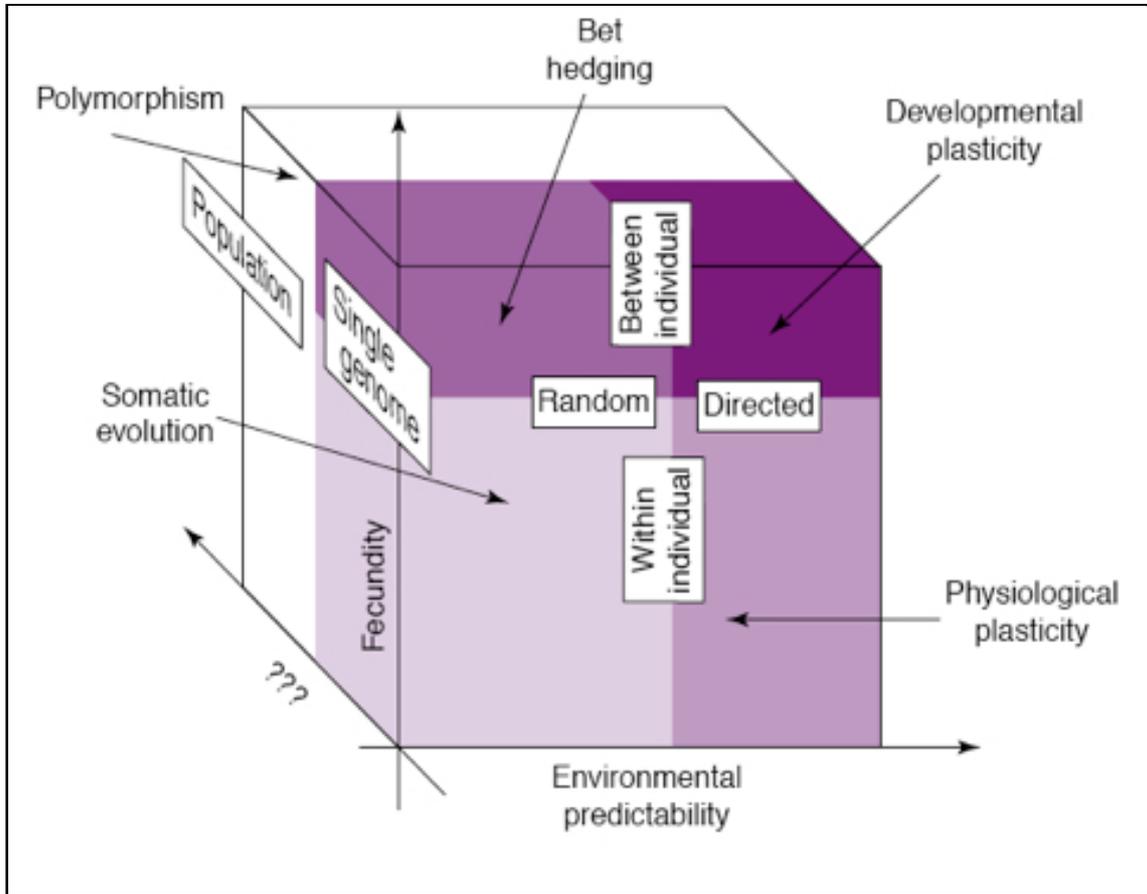
3.1- La réponse individuelle

Face à un changement environnemental, la mise en place d'une réponse adaptative sera étroitement liée à la rapidité mais aussi à la prédictibilité de ce changement (Encadré 4).

En réponse à des variations peu ou non prédictibles, l'organisme tentera dans un premier temps de maintenir les paramètres physico-chimiques nécessaires à sa survie (température, salinité...). Ce phénomène appelé homéostasie, est une réponse rapide de l'organisme mais peut toutefois ne pas suffire si la variation environnementale s'avère trop longue. Dans ce cas, les individus dont le milieu ne sera plus approprié à leur développement auront tendance à migrer vers des habitats plus adaptés (dans la mesure où ils sont capables de déplacement et/ou une niche écologique est disponible). Un environnement fluctuant aura tendance à affecter les juvéniles, moins robustes que les adultes. Dans ce cas, certaines espèces peuvent adopter une stratégie dite de bet-hedging qui consistera à «distribuer» des juvéniles dans différents milieux s'assurant ainsi qu'une partie au moins sera adaptée et pourra survivre.

Lors d'un changement hautement prédictible, comme l'alternance des saisons, les organismes pourront plus facilement anticiper et adopter différentes stratégies adaptatives. Le polyphénisme leur permettra ainsi d'obtenir plusieurs phénotypes à partir d'un seul génotype en fonction du contexte environnemental. Par exemple, chez les criquets pèlerin (*Schistocerca gregaria*), le juvénile pourra opter soit pour un phénotype solitaire à ailes courtes et monochrome soit pour un phénotype grégaire à ailes longues, très coloré, en fonction des signaux environnementaux qu'il aura perçu au cours de son développement. Ce processus est irréversible et en fonction des générations, il peut y avoir alternance des phénotypes (Sword *et al* 2000). La plasticité phénotypique est susceptible d'intervenir chez un même individu mais au fil de l'année. C'est le cas du phénomène de *life cycle staging* retrouvé chez certains vertébrés. La couleur du plumage du lagopède des saules (*Lapogus lapogus*) va ainsi changer en fonction de la couverture neigeuse de la toundra arctique (Pielou 1994). Ce phénomène réversible, permettra aux organismes de maximiser leur survie ainsi que leur succès reproducteur dans une vaste gamme environnementale.

Bases évolutives des stratégies adaptatives



Encadré 4: On s'attend à ce que les mécanismes d'adaptation évoluent selon les conditions. Les différents niveaux de prédictabilité et de fécondité (reflète la vitesse à laquelle la population croît dans un environnement changeant) peuvent avoir des impacts au sein et entre les individus et entraînent des réponses dirigées ou aléatoires. Certains facteurs (symbolisés par ???) peuvent favoriser une réponse individuelle plutôt qu'une réponse à l'échelle de la population. Source : Meyers 2002

Coûts potentiels et limites à la plasticité phénotypique

Coûts de la plasticité

Coûts de maintenance: coûts énergétiques des mécanismes sensoriels et de régulation liés à la plasticité

Coûts de production: coût de production de structures induites

Coûts d'acquisition de l'information: l'acquisition d'informations sur un environnement, peut être risqué et nécessite de l'énergie

Instabilité du développement: l'imprécision phénotypique peut réduire la fitness dans un contexte de sélection stabilisante

Coûts génétiques:

- linkage (des gènes promouvant la plasticité peuvent être liés à d'autres réduisant la fitness)
- pléiotropie (les gènes de la plasticité peuvent avoir un effet sur d'autres traits qualitatifs)
- épistasie (le locus régulant la production de la plasticité peut aussi modifier l'expression d'autres gènes)

Limites des bénéfices de la plasticité

Limite de la qualité de l'information: les organismes plastiques peuvent produire des phénotypes maladaptés si l'information qu'ils ont reçu de l'environnement était fausse ou si celui-ci a changé depuis.

Limite du décalage dans le temps: le décalage entre le changement environnemental et la réponse phénotypique peut induire une réduction de la fitness.

Limite du type de développement: des organismes fixés peuvent être plus à même que des organismes plastiques à produire des développements extrêmes.

Problème épiphénotypique: l'addition de phénotypes au cours du développement peut être moins efficace que le même phénotype intégré au cours du développement précoce.

Encadré 5 : Coûts et limites de la plasticité
Source : De Witts, 1998

La plasticité phénotypique a toutefois ses limites et peut avoir un coût supérieur au bénéfice apporté (DeWitt *et al* 1998, Encadré 5). Ainsi, le maintien des phénotypes favorables d'une génération à une autre (*via* la sélection) est dépendant de leur effet positif sur la fitness des individus. La plasticité peut être considérée comme un facteur adoucissant, permettant de réduire le coût de la sélection directionnelle.

3.2- La réponse à l'échelle des populations

Le génotype de chaque individu le prédispose plus ou moins bien à vivre dans un environnement donné. Au-delà d'une réponse individuelle et dans le cas d'une variation durable de l'environnement, la sélection naturelle va permettre de «trier» au fil des générations, les gènes les plus favorables, en agissant directement sur les fréquences alléliques. La sélection, effective seulement si elle permet d'augmenter la fitness des populations, sera d'autant plus efficace que le pool génétique disponible au sein de la population est variable. Cette diversité génétique est fortement liée aux flux de gènes existant entre populations.

Dans ce contexte, la fragmentation des habitats, et plus généralement le frein à la dispersion des individus, aura tendance à limiter le niveau de polymorphisme d'une population, et par voie de conséquence son potentiel évolutif. Par ailleurs, le morcellement en petites sous-populations va contribuer à la diminution de la taille efficace (nombre de reproducteurs limités) les rendant dès lors plus sensibles à la dérive génétique. Ce phénomène stochastique aura tendance à réduire les fréquences alléliques dans les populations et à augmenter la consanguinité (déficit en hétérozygotes) aboutissant à une perte de diversité qui aura elle aussi tendance à réduire la fitness des individus et ainsi limiter leur potentiel évolutif. Face à un changement environnemental supplémentaire, ces populations déjà réduites, seront moins bien adaptées et leur fitness sera de nouveau affectée (Figure 5).

3.3- Le cas particulier des populations en limite d'aire de répartition

L'aire de répartition d'une espèce est définie par de nombreux paramètres, d'origine biotique et abiotique. Le climat, la nature du sol, la présence d'eau mais aussi de prédateurs, de compétiteurs et de nourriture sont autant de facteurs qui vont favoriser ou

non l'implantation d'une espèce dans un milieu (Lomolino *et al* 2005). Ainsi, dès lors que tous les paramètres seront nécessaires à l'établissement d'une niche écologique, une espèce sera abondante. A contrario, l'absence d'un seul de ces facteurs suffira à réduire cette abondance (Brown 1984 et 1995, Hengeveld 1990, Lawton *et al* 1994).

En bordure d'aire de répartition, les individus seront souvent à la limite de leur tolérance et les populations potentiellement moins bien adaptées (Brown *et al* 1995, Thomas et Kunin 1999). Ces populations généralement petites et fragmentées (métapopulations) sont souvent plus sensibles à l'effet Allee et à la dérive génétique. En outre, leur taux de mutation souvent faible ne leur permet pas de maintenir un niveau de diversité génétique élevée (Alleaume 2006). Ces populations demeurent fréquemment connectées aux populations centrales mais le flux de gènes généré par les migrants venant du cœur de la répartition aura tendance à apporter des allèles délétères et à augmenter ainsi le fardeau génétique, limitant l'adaptation locale (Bolnick *et al* 2008).

Dans ces habitats marginaux, un changement environnemental rapide, (i.e. ingérable à l'échelle évolutive), et complexe (e.g. changements globaux) pourrait avoir un impact majeur sur la persistance des populations déjà instables (Figure 5).

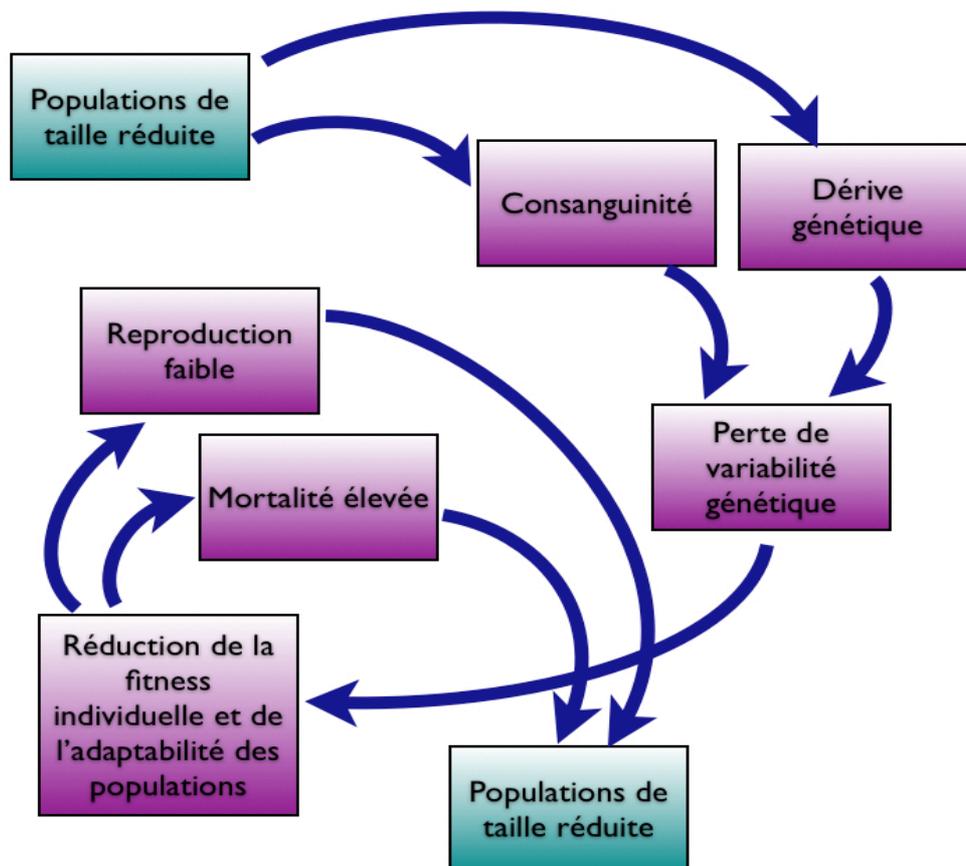


Figure 5: vortex d'extinction inspiré de Gilpin et Soulé (1986)

4- Exemple du bivalve *Macoma balthica*

Macoma balthica est un mollusque bivalve dont la taille varie de 5 mm à 2 cm et qui peut présenter différentes couleurs à l'âge adulte (Figure 6). Organisme benthique, il vit en milieu sablo-vaseux intertidal et subtidal peu profonds, généralement enfoui dans le sédiment à des profondeurs de 2 à 11 cm, mais également en milieu subtidal profond enfoui à plusieurs mètres, comme par exemple dans la mer Baltique. C'est un organisme dépositivore – suspensivore équipé de deux siphons qui lui permettent de filtrer l'eau à marée haute afin d'en extraire sa nourriture et d'excréter ses fèces.



Figure 6 : *Macoma balthica*

Son développement se caractérise par deux phases distinctes: une phase larvaire mobile et une phase adulte plutôt sessile. La phase larvaire est composée de 3 stades successifs qui se dérouleront en 2 ou 3 semaines (Caddy 1969). Dans un premier temps, la fécondation des œufs émis dans la colonne d'eau aboutit à une première larve appelée larve D (de part sa forme en D). Après deux jours, l'apparition d'un vélum (voile bordé de cils) va permettre à la larve de se déplacer et de se nourrir, ce stade dit « stade véligère » va durer à peu près 2 semaines. Après ce délai, un pied se forme lors du stade pédivéligère et progressivement le vélum disparaît. Finalement, 3 semaines après la fécondation, les larves se fixent dans le sédiment et vont se développer en juvéniles avec ses deux siphons (Figure 7).

Comme chez de nombreux bivalves, la durée de vie est inconnue, et semble varier en fonction de la zone géographique considérée. Ainsi dans l'Atlantique nord-est, les individus ont une durée de vie moyenne de 5 à 8 ans (Vogel 1959, Lammens 1967) mais des individus de 25 ans et plus, ont été découverts dans la région de Tvärminne en Finlande (Segerstråle 1960).

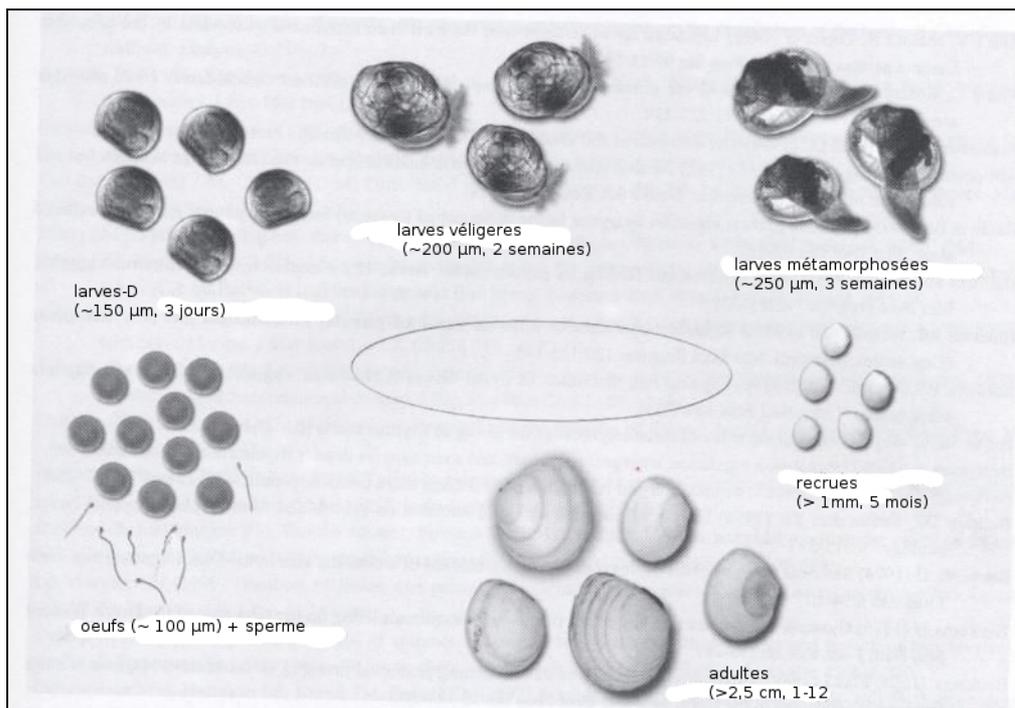


Figure 7 : Stades de développement de *Macoma balthica*
Source Oscar Bos

Organisme eurytherme et euryhalin, l'aire de répartition de *Macoma balthica* est vaste occupant l'ensemble de l'hémisphère nord de l'estuaire de la Gironde au sud jusqu'en Sibérie au nord (Meehan 1985) et dans des environnements saumâtres tels que la mer Baltique (Figure 8).

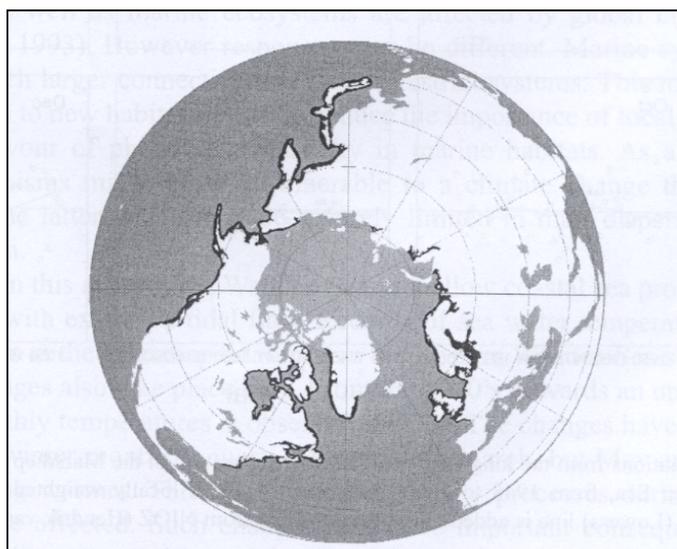


Figure 8 : Projection polaire de la distribution de *Macoma balthica*.
Source: Jan Drent

Initialement originaire du Pacifique nord (Meijer 1993), *Macoma balthica* a pu gagner l'océan Atlantique à l'ouverture du détroit de Béring, il y a environ 3,5 millions d'années. Présente sur les côtes de la mer du Nord au début du pléistocène (il y a environ 1,7 millions d'années), elle est rapidement devenue un taxon majoritaire de l'Europe du nord. La colonisation vers le sud s'est faite progressivement pour atteindre finalement les côtes de la péninsule ibérique (Otéro et Milan 1970). Au cours du dernier maximum glaciaire (il y a 18000 ans), elle a toutefois été confinée dans une zone dite refuge (Gomez et Lunt 2007) au sud de son aire de répartition (péninsule ibérique et golfe de Gascogne), les côtes de l'Europe du nord étant recouvertes de glace. La période interglaciaire qui a suivi (il y a environ 10000 ans) s'est alors accompagnée d'une recolonisation progressive des côtes nord est de l'Europe à partir des populations refuges du sud (Hewitt 1999).

Bien qu'encore largement répartie sur les côtes européennes, *M. balthica* voit son aire de répartition se déplacer vers le nord est. Sa limite sud se trouve désormais au niveau de l'estuaire de la Gironde (Bachelet 1980) alors que cette espèce était décrite il y a encore 40 ans sur les côtes nord ouest espagnoles et cette remontée vers le nord semble coïncider avec l'élévation concomitante de la température des eaux de surface dans le golfe de Gascogne (Goikoetxea 2009). De nombreuses études ont montré que *Macoma balthica* se révélait plus sensible aux changements de température que d'autres bivalves européens tels que *Cerastoderma edule*, *Crassostrea gigas*, *Mya arenaria* ou *Mytilus edulis* (Van der Meer 2006, Freitas 2007, Jansen 2007). Philippart *et al* (2003) a ainsi montré que l'élévation de la température en mer du nord de 1°C au cours des vingt dernières années (Van Aken 2008) s'est accompagnée d'un décalage de la période de ponte de *M. balthica* d'environ un jour par an depuis le milieu des années 80. Une étude de Beukema (2009) sur une période de 40 ans (des années 70 à nos jours) a par ailleurs corrélié l'effondrement des stocks de l'espèce dans la mer de Wadden à un recrutement plus faible qui serait directement lié à des températures saisonnières plus élevées que la moyenne. Globalement, des hivers plus doux et des étés plus chauds seraient à l'origine d'un déséquilibre énergétique réduit, se traduisant par la diminution de la biomasse des individus et réduisant ainsi les taux de reproduction et de survie. Cet impact sur les taux de reproduction et par conséquent de survie est d'autant accentué que la période propice au cours de laquelle les conditions de température et de disponibilité en nourriture sont optimales, est de plus en plus réduite (Beukema *et al* 2009).

La diminution des stocks de populations de *Macoma balthica* en mer de Wadden a impacté le réseau trophique de l'écosystème local. La succession d'hivers doux au cours des années 1988 à 1990 qui a réduit le recrutement des bivalves (dont *Macoma balthica*) a ainsi entraîné la diète, l'émigration de masse, mais aussi la mort de nombreux oiseaux limicoles prédateurs de ces bivalves (Beukema 1996, Camphuysen 2002). Par ailleurs, dans les Pertuis Charentais, lieu de séjour hivernal de nombreuses espèces d'oiseaux limicoles, des études menées au sein de notre laboratoire ont mis en évidence l'importance quantitative mais surtout qualitative de *Macoma balthica* dans le régime alimentaire des bécasseaux maubèche (*Calidris canutus*) et des barges à queue noire (*Limosa limosa*) (Quaintenne *et al* 2010 ; Robin, communication personnelle).

Macoma balthica apparaît comme un modèle d'étude particulièrement intéressant, puisque cette espèce qui occupe une place stratégique dans le réseau trophique des écosystèmes littoraux européens, semble être d'ores et déjà affectée par les changements globaux contemporains.

5- Objectifs de la thèse

De part leur productivité, la densité de population humaine hébergée et la multiplicité des activités, les zones côtières sont soumises à de forts enjeux de société générant des pressions variées, auxquelles les populations naturelles répondent par le biais de processus évolutifs complexes. Cependant depuis les années 1970, l'accélération des bouleversements environnementaux s'ajoute à l'exploitation intensive, et ces changements, à la fois rapides et additifs, impactent les potentiels adaptatif et évolutif des populations. Aussi, l'intensification des activités anthropiques sur les littoraux à forte densité démographique, met-elle aujourd'hui en péril les équilibres écosystémiques, ainsi qu'en témoignent le changement climatique, l'épuisement des ressources et l'érosion de la biodiversité, faisant clairement apparaître la notion de risque environnemental qu'il convient d'évaluer. Dans un contexte de conservation et de gestion des écosystèmes littoraux, la compréhension des mécanismes de l'adaptation contemporaine des espèces marines s'avère donc essentielle.

L'équipe AMARE, à laquelle j'appartiens au sein de l'UMR LIENSs 6250, se consacre à l'étude et la prédiction des effets des activités anthropiques sur l'adaptation et les performances des organismes marins à différentes échelles d'organisation. Si celle de l'organisme se prête davantage aux approches expérimentales, l'échelle des populations est principalement abordée par des approches de génétique des populations.

Les différentes forces évolutives (mutations, migrations, dérive génétique et sélection) modèlent le polymorphisme génétique et la divergence entre les populations sur lesquelles elles agissent, imprimant des signatures génétiques différentes selon leur nature. Des marqueurs moléculaires neutres issus du génome non codant et les outils statistiques s'appuyant sur les cadres conceptuels de la génétique des populations, permettent d'inférer à différentes échelles de temps et d'espace, le rôle de la dérive, des régimes de reproduction et des événements démographiques indépendamment des caractères adaptatifs. Or, il apparaît que la sélection pourrait jouer un rôle bien plus important dans la divergence entre deux populations que l'attendu de l'hypothèse neutraliste (Eyre-Walker 2006). Les avancées technologiques récentes dans le domaine de la génomique fonctionnelle, et notamment le développement du séquençage à haut débit, ouvrent aujourd'hui de nouveaux champs d'investigation en permettant de passer de l'étude d'un nombre restreint, à celle d'un grand

nombre de gènes, donnant ainsi accès au génome codant sur lequel s'exerce la sélection.

Dans une optique de conservation, et alors que la limite sud de répartition de *Macoma balthica* semble se décaler vers le nord est, nos objectifs sont (i) d'inférer la part des mécanismes démographiques et celle des mécanismes adaptatifs dans un tel déplacement d'aire, (ii) d'en identifier éventuellement l'origine (réchauffement ou qualité des eaux), selon une double approche de génétique des populations et de génomique fonctionnelle. Ce travail a été mené sur des populations naturelles échantillonnées dans trois baies européennes (Figure 9) aux caractéristiques spatiales et environnementales différentes :

- Une population en limite d'aire de répartition soumise à l'augmentation de la température des eaux de surface du golfe de Gascogne (baie de Marennes-Oléron, océan Atlantique, France).
- Une population au cœur de l'aire de répartition soumise à diverses contraintes physico chimiques et fortement impactée par la pollution industrielle (baie de Gdansk, mer Baltique, Pologne).
- Une population au cœur de l'aire de répartition prélevée en réserve naturelle faisant office de référent (baie de Somme, mer de la Manche, France).

La première partie de ce mémoire est ainsi consacrée à l'analyse des processus neutres (i.e. migration et dérive génétique), à l'aide de marqueurs moléculaires de type microsatellites et de l'ADN mitochondrial, afin de répondre aux questions suivantes :

- Quel est le statut des populations de *Macoma balthica* en limite d'aire ?
- Peut-on observer des signatures particulières en termes de diversité et de structure génétique des populations selon la nature de l'impact anthropique (degrés de contamination et/ou réchauffement) ? En particulier, des goulots d'étranglements (diminution importante et rapide du nombre d'individus) se sont-ils produits ?
- La présence de contaminants ou/et le réchauffement des eaux de surface ont-ils entraîné une rupture des échanges de gènes entre sous-populations créant des sous-groupes au sein d'une même baie ?
- Observe-t-on des signatures d'adaptation locale ?

La deuxième partie de ce mémoire se penche sur les processus sélectionnés grâce à l'accès au transcriptome de *Macoma balthica* permis par un séquençage massif, afin de :

- Mettre en évidence de nouveaux gènes candidats chez *Macoma balthica*, notamment des gènes impliqués dans la réponse générale au stress
- Comparer l'ensemble de transcrits (ARNm) issus de 3 populations naturelles

prélevées sur trois sites géographiquement éloignés, et soumis à des stress environnementaux différents.

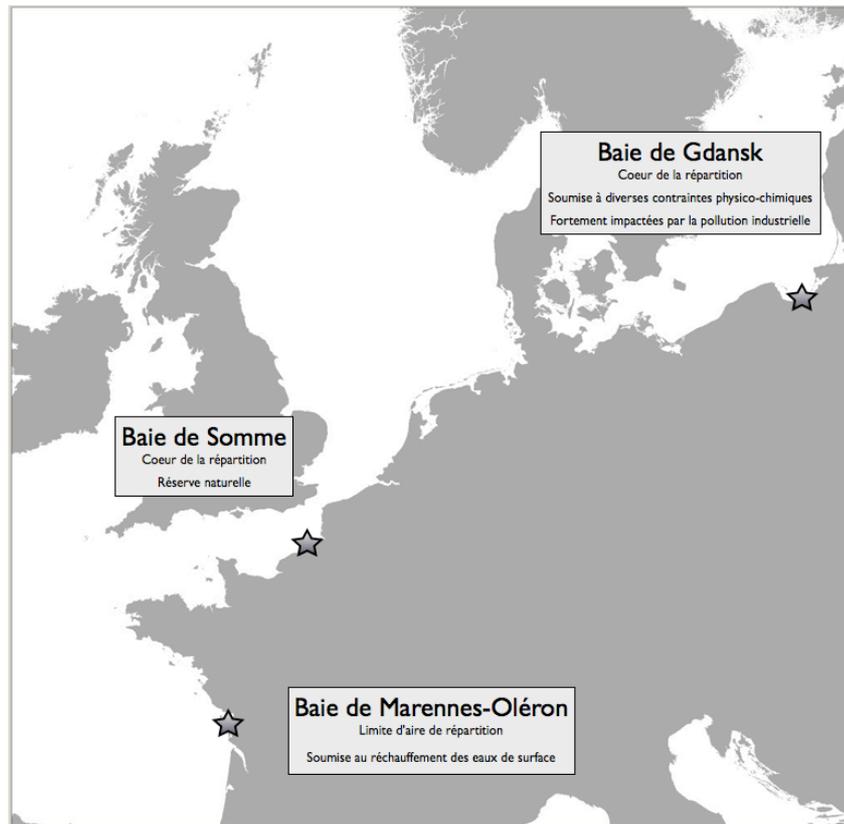


Figure 9 : Localisation des baies étudiées

Partie I

*Influence des contraintes anthropiques
sur les processus démographiques dans
les populations naturelles de *Macoma
balthica**

Introduction

De nombreuses études phylogéographiques ont été menées sur *Macoma balthica*, permettant de mettre en évidence une structuration des populations européennes. A l'aide d'allozymes, Väinola (2003) a ainsi mis en évidence deux groupes de populations, (i) une vaste population nord est Atlantique incluant les populations de la mer de Norvège, de la mer du Nord et des îles Britanniques, et (ii) une métapopulation structurée avec une forte diversité génétique et un nombre important d'allèles dans la partie nord de l'aire de répartition de l'espèce (océan pacifique, mer baltique et mer blanche). Ces deux ensembles pourraient correspondre à deux sous-espèces coexistant en Europe, la lignée Pacifique *Macoma balthica balthica* et la lignée Atlantique *Macoma balthica rubra*. Ces résultats ont été confirmés par Luttikhuisen *et al* (2003) et Nikula *et al* (2008) avec des marqueurs mitochondriaux. Luttikhuisen révèle également une signature mitochondriale spécifique pour les populations de *Macoma balthica* situées à l'extrême sud de l'aire de répartition (estuaire de la Gironde, France), avec un profil haplotypique divergent de ceux observés dans les autres populations européennes et pacifiques. A l'opposé, les travaux menés par Hummel *et al* (1995, 1997b) puis Strelkov *et al* (2007) avec des marqueurs allozymiques, ne montrent pas un tel isolement, ces populations étant similaires à celles de la mer de Norvège et de l'ouest de la mer de Barents. Toutefois, ces analyses ont été réalisées avec des marqueurs mitochondriaux et/ou des allozymes. Ces derniers sont potentiellement soumis à sélection et ne sont probablement pas les plus appropriés pour étudier l'impact des activités humaines sur les populations.

Dans la première partie de cette thèse, nous nous intéresserons aux processus évolutifs neutres qui agissent sur les schémas spatio-temporels de dispersion des populations de *Macoma balthica* soumises à des impacts anthropiques différents avec trois objectifs principaux :

- Développer des marqueurs moléculaires nucléaires neutres et hypervariables permettant l'étude de l'impact de processus contemporains sur la structure des populations ;
- Etudier précisément à différentes échelles spatio-temporelles, les populations en limite sud de répartition, comprendre leur structure génétique et définir leur statut (populations refuge ou en limite d'aire) ;
- Evaluer l'impact d'une forte activité anthropique (contaminations) sur la microstructure génétique de populations au coeur de l'aire de répartition.

Chapitre I

Mise au point de nouveaux marqueurs moléculaires hypervariables

I.1- Choix des marqueurs moléculaires

L'ensemble des résultats décrits dans ce chapitre a fait l'objet d'une publication dans la revue Conservation Genetics Ressources (présentée en fin de chapitre)

I.1.1- Locus mitochondrial

L'analyse de l'ADN mitochondrial (ADNmt) est couramment utilisée dans les études phylogéniques et phylogéographiques (Crozier 1990). Cet ADN présente en effet, de nombreux avantages. C'est une molécule circulaire, bicaténaire de 15000 à 20000 paires de bases (Boore 1999). Chaque cellule eucaryote contient plusieurs mitochondries et ainsi plusieurs copies identiques d'ADN mitochondrial (jusqu'à plus de 100000) facilitant ainsi son extraction et son amplification par PCR. Chaque molécule d'ADNmt animal est généralement constituée de 37 gènes codant pour 13 protéines, 2 ARN ribosomiaux (ARNr) et 22 ARN de transfert (ARNt), dépourvus d'introns, et d'une région non codante de 1 kilobase contenant les principales séquences de régulation et de transcription de la molécule d'ADNmt. Ce génome simple et de petite taille comparativement à celui du noyau, est un outil de choix pour l'étude de populations. Outre une grande conservation entre taxons qui facilite la comparaison de fragments homologues, le génome mitochondrial est haploïde et se transfère selon des modalités généralement uniparentale (SMI, Strict Maternal Inheritance, Birky 2001) à l'exception de certains bivalves comme la moule marine *Mytilus edulis* (Zouros 1992), la palourde *Venerupis philippinarum* ou la moule d'eau douce *Pyganodon grandis* qui possèdent 2 ADNmt distincts hérités de façon maternelle ou paternelle (DUI, Doubly Uniparental Inheritance, Breton 2007). Ces caractéristiques (hérédité maternelle et haploïdie) du génome mitochondrial génèrent une taille efficace plus faible que celle des gènes nucléaires. Ce génome est ainsi plus sensible à la dérive génétique et révèle d'autant mieux les événements démographiques subis par les populations, tels que les goulots d'étranglement et les effets fondateurs. Enfin, l'absence de recombinaison et une transmission clonale facilitent les reconstructions phylogénétiques. Par ailleurs, son taux d'évolution élevé (on s'accorde sur un taux d'évolution d'environ 2 % par million d'années) permet d'étudier les relations inter-spécifiques mais surtout intra-spécifiques.

I.1.2- Locus microsatellites

Les microsatellites sont des marqueurs nucléaires à hérédité biparentale très abondants chez les eucaryotes (Hancock 1999). Il y a ainsi près d'un microsatellite toutes les 30 à 100 kb chez les végétaux supérieurs (Morgante et Olivieri 1993). Ils sont généralement neutres et codominants (distinction des homozygotes et des hétérozygotes). Ils se caractérisent par des répétitions en tandem de motifs mono, di, tri ou tétra - nucléotides (Rassman 1991), un nombre de répétitions correspondant à un allèle, chaque allèle étant séparé d'un autre par une ou plusieurs mutations. Ces mutations sont généralement dues à un glissement de la polymérase lors de la réplication de l'ADN (Levinson et Gutman 1987, Eisen 1999). Ces marqueurs sont qualifiés d'hypervariables du fait de leur fort taux de polymorphisme (on peut ainsi trouver 10 à 30 allèles différents par locus microsatellites dans les populations naturelles, Dow et Ashley 1996) associé à un taux de mutation élevé (les taux de mutation des microsatellites sont de l'ordre de $5 \cdot 10^{-5}$ à $5 \cdot 10^{-3}$ alors qu'ils sont de 10^{-9} pour le reste du génome, Estoup et Angers 1998). Ce nombre élevé d'allèles va notamment permettre d'étudier les flux de gènes entre populations à des échelles spatio-temporelles fines.

I.2- Description du marqueur mitochondrial

L'amplification du fragment génique de la sous-unité 1 de la Cytochrome oxydase a été réalisée à partir d'amorces spécifiques de *Macoma balthica*, décrites par Luttikhuisen en 2003. Les réactions de PCR ont été réalisées dans un volume total de 50 µl, le mix étant composé de tampon 1X, de 1,85 mM de MgCl₂, 125 µM de dNTPs, 0,25 µM de chaque amorce, 1,6 U de Taq polymérase (Red Hot ® TAQ DNA Polymerase, Abgène) et de 10 ng d'ADN. Un thermocycleur MJResearch PTC 100 a été utilisé avec le programme PCR suivant : 3 min de dénaturation initiale à 95 °C, suivi par 35 cycles décomposés en une phase de dénaturation de 45 s à 95 °C, une hybridation de 60 s à 60 °C et une phase d'élongation de 90 s à 72 °C. Enfin une étape finale d'élongation de 5 min à 72 °C.

Le séquençage des produits PCR purifiés a ensuite été réalisé par la société Génoscreen (Campus Pasteur - 1 rue du Professeur Calmette - 59 000 Lille - France) avec un séquenceur automatique ABI PRISM® 3730 XL (Perkin-Elmer Applied Biosystems, Foster City, CA).

Les séquences obtenues ont été alignées avec le logiciel ClustalX (Thompson *et al* 1994), puis elles ont été manuellement vérifiées avec le logiciel BioEdit (Hall 1999).

I.3- Mise au point de marqueurs microsatellites

I.3.1- Construction d'une banque enrichie en microsatellites

L'ADN génomique de 4 échantillons prélevés en Mer de Wadden (Pays-Bas) a été extrait à partir d'environ 15 mg de muscle avec un kit Nucleospin tissue (Macherey Nagel). Une banque génomique de microsatellites, enrichie en motifs CA a ensuite été construite en suivant le protocole décrit par Billote *et al* (1999) (protocole détaillé en Annexe A). Brièvement, l'ADN total de chaque individu (environ 10 µg) a été digéré par l'enzyme RsaI (Promega) pendant 7h à 37°C. Les 4 ADN digérés ont ensuite été mélangés deux à deux et purifiés sur une colonne Nucleospin® (Macherey Nagel) avant d'être ligué à des adaptateurs RsaI permettant l'amplification par PCR des fragments obtenus. Le programme PCR consistait en une première phase de dénaturation de 4 min à 95°C suivi de 20 cycles de 30 s à 94°C, 1 min à 60°C et 1 min 30 à 72°C, terminé par une phase d'élongation finale de 5 min à 72°C.

L'enrichissement de la banque en motif AC a été réalisé grâce à un oligonucléotide biotynilé (AC)₁₀ qui a permis de sélectionner les fragments d'ADN porteurs de ce motif. Ces fragments ont ensuite été triés grâce à des billes magnétiques recouvertes de streptavidine (*Streptavidin Magnosphere® Paramagnetic Particles*; Promega) et isolés par lavages successifs de tampon SSC (Saline Sodium Citrate).

La fraction enrichie a ensuite été purifiée, amplifiée (Annexe A) et liguée dans un plasmide *pGEM®-T easy* (Promega). Ces plasmides ont ensuite été clonés dans des bactéries compétentes d'*Escherichia coli* DH5α. Les bactéries transformées (c'est à dire contenant le plasmide) ont ensuite été mises en culture sur du milieu LB-agar (*Luria Broth*) en présence d'ampicilline, d'IPTG et de Xgal, ces trois produits permettant de s'assurer d'une part que les bactéries contenaient bien le plasmide (résistant à l'ampicilline) mais aussi que le plasmide contenait bien l'insert (le fragment d'ADN enrichi en motif AC). Les colonies blanches (Annexe A) ont ensuite été repiquées en milieu LB liquide puis amplifiées par PCR avec les amorces spécifiques du plasmide *pGEM®-T easy* (Promega) afin de confirmer la présence de l'insert.

Chaque clone positif a ensuite été extrait avec un kit *Nucleospin Plasmid Quick Pure* de Macherey-Nagel selon le protocole du fournisseur. La concentration en ADN ainsi que la pureté ont ensuite été évaluées par mesure de la DO à 260 et 280 nm grâce à un spectromètre.

I.3.2- Séquençage et sélection des marqueurs microsatellites

Chaque clone extrait a été séquençé avec un kit *SequiTherm EXCEL™ II DNA Sequencing Kit* (Epicentre) sur un séquenceur à plaques Li-Cor NEN Global IR2 . Ce séquençage a permis de sélectionner les motifs microsatellites les plus pertinents (motif assez long et non entrecoupé, espace disponible de part et d'autre pour la définition d'amorces, Figure I.1).

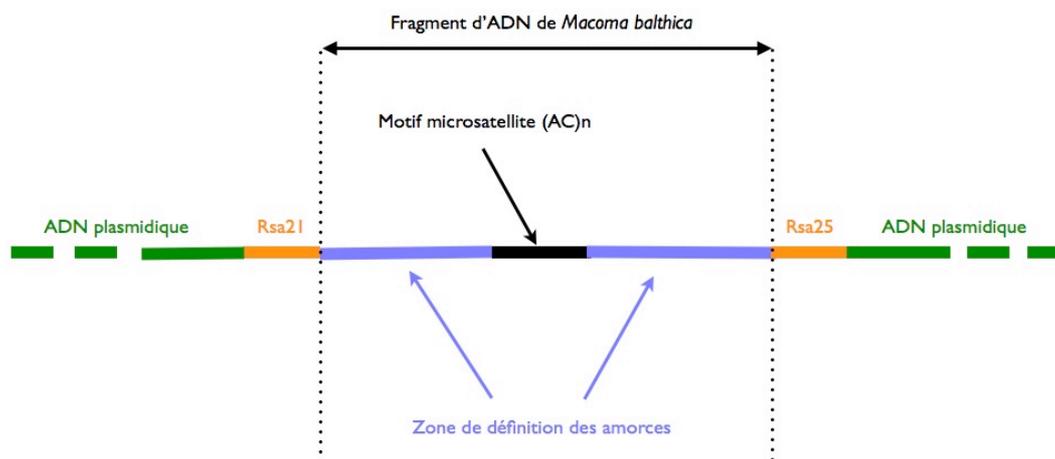


Figure I.1: Représentation schématique d'un fragment d'ADN de *Macoma balthica* obtenu par séquençage d'un clone positif : cas de figure idéal laissant un grand choix pour la définition d'amorces de part et d'autre du motif microsatellite situé au milieu du fragment d'ADN séquençé.

Les amorces ont été dessinées avec le logiciel OLIGO® v.6.0 (Rychlik & Rychlik 2000) et les amorces sens (F) ont été marquées avec une molécule fluorescente IRD700™ ou IRD800™. Chaque couple d'amorce a été testé sur 36 individus issus de deux populations éloignées géographiquement (15 individus de Fouras en France et 21 individus de la Mer de Wadden en Hollande) afin d'optimiser le choix des marqueurs microsatellites. La réaction d'amplification a été réalisée dans un volume final de 10 µl contenant environ 50 ng d'ADN, 0,25 µM d'amorce sens marquée en fluorescence, 0,25 µM d'amorce anti-sens non marquée, 1,5 mM de MgCl₂, 0,2 mM de chaque dNTP, du tampon PCR 1X et 0,5 U de *Thermoprime Plus DNA Polymerase* (ABgene). La réaction a été réalisée sur un thermocycleur MJ Research PTC-200 (MJ Research Inc.) avec un programme dit *touchdown* qui permet d'obtenir plus facilement la bonne température d'hybridation. Ce programme se décomposait ainsi: 3 min de dénaturation initiale à 95°C, suivi par 10 cycles consistant en une phase de dénaturation de 45 s à 95°C, une phase d'hybridation de 45 s à une température diminuant de 1°C à chaque cycle d'une température initiale T_i à une température finale T_f (Table I.1), puis une phase d'élongation de 45 s à 72°C. Ces 10 cycles étaient suivis de 30 cycles de 45 s à 95°C, une hybridation de 45 s à T_f, une élongation de 45 s à 72°C suivi d'une phase finale d'élongation de 7 min à 72°C

Genbank accession no.	Locus	Core Sequence	Primer sequences (5'-3')	Ti-Tf (C°)	Allele size range (bp)	N _{tall}	Fouras (N =15)				Waddensee (N =21)			
							N _{all}	H _e	H _o	F _{is}	N _{all}	H _e	H _o	F _{is}
EU 676331	Mb-mac 4	(TG)4CG(TG)3CA(TG)4	F: CTCATATCTTCACCCTAGA R: CCATTCCTGTGCATTAGCA	66-56	395-443	18	14	0,9	0,73	0,222	13	0,87	0,33	0,634
										(P = 0,0000)				(P = 0,0000)
EU 676330	Mb-mac 10	(TG)5GA(TG)2(TA)2(TG)7AC(TG)2	F: GGGTGTGATGGGATAATA R: TGGGGGCTACGAATAAGT	66-56	396-404	5	5	0,69	0,33	0,542	4	0,64	0,41	0,786
										(P = 0,0027)				(P = 0,0000)
EU676332	Mb-mac 17	(TG)14	F: TGTGGGATTTGAGATTTTTA R: TACATTTATTTTCATTACGAT	62-52	529-581	26	14	0,89	0,53	0,433	19	0,93	0,71	0,257
										(P = 0,0000)				(P = 0,0000)
EU 676333	Mb-mac 19	(TG)7CG(TG)2AGTGACC(TG)2	F: TCTTCTTATGTAGCGTGTT R: CCAGGGCGAGTTTTTCTT	64-54	331-373	17	14	0,91	0,8	0,154	12	0,88	0,52	0,427
										(P = 0,0409)				(P = 0,0005)
EU 676334	Mb-mac 40	(CA)5	F: GTGCGGACAAGAATACG R: TATTGGCTGTGCTCGTTTT	64-54	199-211	6	5	0,39	0,27	0,356	3	0,25	0,14	0,457
										(P = 0,3380)				(P = 0,0301)
EU 676335	Mb-mac 64	(TG)5	F: ATAATTTGTGGGTTGAGGT R: GTTTCGAGTTTCGCAGTCA	66-56	168-184	6	6	0,48	0,47	0,071	2	0,32	0,2	0,419
										(P = 0,1726)				(P = 0,3053)
EU 676337	Mb-mac 84	(CA)5	F: TATATCCCTTGATCGTTTT R: ACGTATGTTTTGTCCATGT	64-54	248-280	8	7	0,8	0,33	0,607	4	0,55	0,32	0,446
										(P = 0,0000)				(P = 0,0569)
GQ 406063	Mb-macB3-24	(CA)4CG(CA)4	F: AGATACGGACAGATTAACAC R: CGGACTCGGGTGTATGTAA	68-58	332-352	7	5	0,35	0,2	0,451	6	0,49	0,33	0,344
										(P = 0,0380)				(P = 0,1684)

Tableau I.x: Caractéristiques des huit marqueurs microsatellites isolés à partir de *Macoma balthica* (Numéro d'accension Genbank: EU676330-EU676337 and GQ 406053). T_i et T_f sont les températures d'hybridation initiales et finales de la PCR *touchdown*. N_{Tall} et H_T : nombre d' alleles et hétérozygotie totale pour l'ensemble de l'échantillon (N=36), N_{all} et H_E : nombre d' alleles et heterozygotie attendue à l'échelle de la population, F_{IS} : estimation de l'écart à l'équilibre de Hardy-Weinberg et P associées (test exact).

L'électrophorèse des produits PCR a ensuite été réalisée sur gel de polyacrylamide à 6,5% sur un séquenceur Li-Cor NEN Global IR2. Les marqueurs ont été retenus en fonction de leur polymorphisme intra et interpopulation.

L'indépendance des loci entre eux a été testé pour chaque paire de loci sur l'ensemble des populations avec le logiciel GENEPOP version 4.0.7 (Raymond et Rousset 1995), ceci en suivant un processus de MonteCarlo à Chaines de Markov avec les paramètres suivants: Dememorization number : 1000, Batch number : 100, Iteration per batch : 1000. Ce logiciel a également été utilisé pour estimer le nombre d'allèles sur l'ensemble des populations ($N_{t_{all}}$) et au sein de chaque population (N_{all}), le taux d'hétérozygotie attendue à l'équilibre de Weinberg (H_e), le taux d'hétérozygotie observée et le déficit en hétérozygote (F_{is}). La présence d'allèles nuls a également été testée avec le logiciel MICRO-CHECKER v.2.2.3 (Van Osterhout *et al* 2004)

I.3.3- Résultats et Conclusion

La construction de la banque a permis d'obtenir 172 clones positifs (contenant un insert). Sur ces 172 clones, 118 ont été séquencés permettant de définir 49 couples d'amorces. L'efficacité et la spécificité de ces couples d'amorces ont été testées par PCR et au final, 8 couples dont les caractéristiques sont décrites dans le Tableau I.1, ont été sélectionnés. Ces marqueurs sont tous polymorphes, présentant un nombre d'allèles total compris entre 5 et 26 et un nombre moyen de 11,62 allèles. Bien que la diversité nucléaire soit globalement élevée (H_e comprise entre 0,25 et 0,93) on peut noter une différence entre l'hétérozygotie observée (H_o) et celle attendue à l'équilibre de Hardy-Weinberg ce qui se traduit par des F_{is} assez élevés également (de 0,071 à 0,786). Le test effectué avec le logiciel microchecker a toutefois révélé la présence d'allèles nuls, notamment pour le marqueur Mb-macio qui présente un fort F_{is} dans les deux populations (0,542, $p = 0,0027$ et 0,786, $p = 0$ pour Fouras et Waddensee), qui pourrait expliquer ce fort déficit en hétérozygote (note de bas de page, explication). Par ailleurs, le test de Fisher réalisé avec Genepop a permis de montrer que les marqueurs étaient bien indépendants les uns des autres ($0,219 < p < 1,000$).

Microsatellite markers for the Baltic clam, *Macoma balthica* (Linné, 1758), a key species concerned by changing southern limit, in exploited littoral ecosystems

V. Becquet, I. Lanneluc, B. Simon-Bouhet and P. Garcia

Littoral, Environnement et Sociétés (LIENSS) UMR 6250 CNRS-Université de La Rochelle,
2 Rue Olympe de Gouges, 17042 La Rochelle cedex, France

Keywords: range shift, global warming, *Macoma balthica*, microsatellite

Corresponding author :

Vanessa Becquet

Institut du Littoral et de l'Environnement, 2 rue Olympe de Gouges, 17000 La Rochelle,
FRANCE

Fax : 0033546458264

vbecquet@univ-lr.fr

Abstract

During the last 30 years, the southern European range limit of the Baltic clam *Macoma balthica* strongly shifted northward. To understand the consequences of such a retreat on populations located near the edge of its current geographic range, we developed nine microsatellite markers and tested their usefulness on 36 individuals from two populations. All loci were found to be polymorphic, with five to twenty-six alleles per locus. These results appear promising to study the fine scale genetic structure and dynamics of *Macoma balthica* populations near the range limit.

Parmesan & Yohe (2003) recently showed that more than 1500 species had experienced switches in their biotic trends coherent with the twentieth century climate warming. In the northern hemisphere, a growing number of coastal marine mollusc species are concerned by northward shifts (Wetthey & Woodin 2008). However, consequences of such shifts on the function and structure of coastal ecosystems remain mostly unknown (Walther *et al.* 2002).

The Baltic clam *Macoma balthica* (Bivalvia: Tellinidae) is widely distributed in marine and estuarine soft-bottom habitats from temperate to arctic coastal waters in the North Atlantic and North Pacific oceans. Along the European coasts, its natural range is showing a strong shift toward the north-east and the species is rapidly vanishing from the Bay of Biscay (Jansen *et al.* 2007). As an important prey for migratory birds, macro-invertebrates and fish, *M. balthica* is one of the key-species along the European littoral (Philippart *et al.* 2003) and is thus of direct importance for fisheries and ecosystem conservation. Focusing on the current edge of its southern distribution, we aim to investigate the fine scale mechanisms of exchanges between unstable populations. Thus, we developed nine polymorphic

microsatellite markers and tested their usefulness by analysing two populations respectively sampled near the range limit and in the core of the species distribution.

Genomic DNA from four individuals from the Waddensee (Dutch coast) was isolated from < 15mg of muscle using a Nucleospin tissue kit (Macherey Nagel). A genomic library enriched for CA repeated microsatellites was constructed following the protocol described by Billote et al. (1999). Briefly, total DNA was digested with RsaI (Promega). After purification on a Nucleospin Column (Macherey Nagel), 3 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 was carried out using biotinylated (AC)₁₀ probes attached to streptavidin-coated magnetic beads (Streptavidin Magsphere® Paramagnetic Particles; Promega). The purified enriched fraction was ligated into pGEM®-T easy (Promega) and transformed into *Escherichia coli* DH5α competent cells. Recombinants with appropriate insert sizes were determined by polymerase chain reaction (99 clones). From a total of 67 positive clones sequenced using a SequiTherm EXCEL™ II DNA Sequencing Kit (Epicentre) on a Li-Cor GEN Global IR2 DNA sequencer, 37 primer pairs were designed using OLIGO® v.6.0 (Rychlik & Rychlik 2000). Forward primers were fluorescently-labelled with infra-red fluorescent dye IRD700™ or IRD800™ for screening. The genomic DNA for genotyping was obtained using Nucleospin tissue kit (Macherey Nagel). Amplification conditions were optimized for 9 loci (Table 1) for which preliminary tests showed unambiguous patterns. Each reaction was performed in 10 µl and contained approximately 50 ng of DNA, 0.25 µM of fluorescently labelled forward primer, 0.25 µM of unlabelled reverse primer, 1.5 mM of MgCl₂, 0.2 mM of each dNTP, 1X PCR buffer [75 mM Tris-HCl, pH 8.8 at 25°C, 20 mM

(NH₄)₂SO₄ and 0.01% Tween 20] and 0.5 U of Thermoprime Plus DNA Polymerase (ABgene), using a MJ Research PTC-200 thermocycler (MJ research Inc.). Thermal cycling conditions followed a touch-down PCR procedure: 3 min of an initial denaturation step at 95°C followed by 10 cycles of denaturation at 95°C for 45s, annealing at temperature decreasing by 1°C per cycle from T_i to T_a (Table 1) for 45s and extension at 72°C for 45s, then followed by 30 cycles at 95°C for 45s, annealing at final T_a for 45s and extension at 72°C for 45s. 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 nine loci was tested in two populations of *M. balthica* [Fouras, N= 15 and Waddensee, N=21]. No linkage disequilibria across loci were detected using GENEPOP version 4.0.7 (Raymond & Rousset 1995). This software was also used to estimate the expected heterozygosities and adequacy of genotypic proportions to Hardy-Weinberg expectations (Table 1). Over the 36 individuals analysed, the number of alleles ranged from 5 to 26 and expected heterozygosity from 0.13 to 0.91. At the population level, deviations from Hardy-Weinberg equilibrium were variable across loci and populations. The loci Mb-mac 10 showed a significant heterozygote deficiency for both populations, which might be due to null alleles. This hypothesis was tested using MICRO-CHECKER v.2.2.3 (Van Oosterhout 2004). It was confirmed for these loci that the heterozygote deficiency were not related to deviation from panmixia or genotyping errors but possibly to the presence of null alleles. Furthermore, all populations were found to be polymorphic [e.g. mean number of alleles: 8.00 and 7.22 in Fouras and Waddensee respectively].

Although care should be taken to accommodate null alleles, these microsatellites should prove useful for analyzing fine scale genetic structures, population dynamics, and relationships within and between local populations of *Macoma balthica*.

References

- 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.
- Jansen JM, Pronker AE, Bonga SW, Hummel H (2007) *Macoma balthica* in Spain, a few decades back in climate history. *Journal of Experimental Marine Biology and Ecology*, **344**, 161-169
- Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37-42.
- Philippart CJM, Van Aken HM, Beukema JJ, Bos OG, Cadée GC, Dekker R (2003) Climate-related changes in recruitment of the bivalve *Macoma balthica*. *Limnology and Oceanography*, **48**, 2171-2185.
- Raymond M, Rousset F (1995) GENEPOP (version1.2): a population genetics software for exact tests and ecumenicism. *The Journal of Heredity*, **86**, 248-249.
- Rychlik W, Rychlik P (2000) Oligo primer analysis software. Version 6.0.1. Cascade, Colorado. Molecular Biology Insights, Inc.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538.

Walther GR, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin JM, Hoegh-Guldberg O, Bairlein F (2002) Ecological responses to recent climate change. *Nature*, **416**, 389-395

Wethey DS, Woodin SA (2008) Ecological hindcasting of biogeographic responses to climate change in the European intertidal zone. *Hydrobiologia*, **606**, 139-151.

Acknowledgements

We would like to thank Claire Daguin (Evolution et Génétique des Populations Marines, Roscoff) for improvement of the protocol. V. Becquet was supported by a grant from the Association Nationale de la Recherche Technique.

Chapitre 2

*Etude d'un site anthropisé en limite
d'aire de répartition.*

*Cas des populations de la Baie de
Marennes Oléron (France) soumises au
réchauffement des eaux du Golfe de
Gascogne.*

II.1- Contexte actuel des populations de *Macoma balthica* en limite d'aire de répartition

L'ensemble des résultats décrits dans ce chapitre a fait l'objet d'une publication soumise à la revue Conservation Genetics (Annexe C)

Au cours des dernières décennies, les pressions anthropiques exercées sur les populations naturelles ont considérablement augmenté (Vitousek *et al* 1997a, Walker et Kendrick 1998) générant fragmentation de l'habitat, déplacement d'aire de répartition des espèces et difficultés d'adaptation. Ces effets ont été accélérés au cours des 40 dernières années par les changements globaux (Bell et Collins 2008), trop récents et trop rapides pour que l'adaptation se mette en place. Le stress généré par les activités humaines (pollution, perte et fragmentation de l'habitat, espèces invasives...) susceptible de perturber voire d'interrompre la connectivité entre les populations, a été renforcé par l'augmentation de la température (Gordon 1998, Schroter *et al* 2005). En effet entre 1906 et 2005, la température moyenne à la surface de la terre a augmenté de 0,74 °C, en partie du fait de l'augmentation de l'émission de gaz à effets de serre (GIEC 2007). Dans l'environnement marin, la température moyenne des eaux de surface a augmenté de 0,23 °C par décennie entre 1977 et 2007 dans l'Atlantique nord est (Goikoetxea *et al* 2009). En réponse à cette élévation de température, l'aire de répartition de nombreuses espèces s'est modifiée, et une étude de Parmesan et Yohe a ainsi montré que près de 1500 espèces terrestres s'étaient ainsi déplacées vers les pôles ou en altitude. *Macoma balthica* a ainsi connu un déplacement de son aire de répartition vers le nord est au cours des dernières décennies, les populations de la péninsule ibérique ayant probablement disparu suite à l'élévation de la température dans le golfe de Gascogne (Jansen 2007).

La réponse des populations situées en limite d'aire à de telles pressions environnementales demeure actuellement inconnue (Walther *et al* 2002). Il est toutefois probable que ces populations présentent des caractéristiques structurales et génétiques susceptibles de limiter leur capacité d'adaptation (Eckert *et al* 2008). Ainsi, les populations en limite de répartition présentent la plupart du temps un taux de croissance réduit (Brown *et al* 1995, Thomas et Kunin 1999). Elles sont souvent fragmentées, avec un taux de mutation réduit, et plus facilement soumises à la dérive génétique et à l'effet Allee (Courchamp *et al* 1999, Alleaume-Benharira *et al* 2006). Parallèlement, le flux de gènes asymétrique des populations centrales vers les populations périphériques limite la plupart du temps la mise en place d'une adaptation locale (Ronce et Kirkpatrick 2001, Bolnick *et al* 2008). Dans ce contexte, les changements environnementaux associés à une diversité génétique supposée réduite chez les populations en limite d'aire pourraient augmenter leurs risques d'extinction (Lawton 1994, Vucetich et Wayne 2003), l'équilibre entre extinctions et recolonisations dans

ces régions étant plus fragile, et le renouvellement des populations plus difficile (Lande *et al* 2003). Dans de telles conditions, la vitesse à laquelle se produit le changement environnemental, représente la limitation principale à la mise en place d'un processus adaptatif.

Macoma balthica a fait l'objet de plusieurs études phylogéographiques (Hummel *et al* 1997b, Väinöläe 2003, Luttikhuisen *et al* 2003, Strelkov 2007 *et al*, Nikula *et al* 2008). Toutefois, l'attention portée sur les populations en limite d'aire de répartition est restée marginale avec l'examen de seulement deux populations dans le golfe de Gascogne (Gironde et/ou Loire, Hummel, Luttikhuisen et Strelkov). Ce site est identifié comme l'une des quatre zones refuges en Europe durant le dernier maximum glaciaire il y a 18000 ans (Gomez et Lunt 2007). Quand la glace s'est retirée, les espèces ont recolonisé les régions plus au nord par effets fondateurs successifs (Hewitt 1999), avec pour conséquence attendue une baisse de la diversité génétique de la zone refuge vers les populations nouvellement établies (Petit *et al* 2003). Cependant, l'étude menée par Hummel *et al* (1997b) à l'aide de marqueurs allozymiques a montré au contraire une augmentation de la diversité (mesurée par l'hétérozygotie observée et attendue, ainsi que par le nombre d'allèles) du sud vers le nord, la plus faible diversité ayant été trouvée dans la population de Gironde (la plus méridionale). Ainsi, deux attendus théoriques s'opposent pour les populations de *Macoma balthica* situées en limite d'aire de répartition : d'une part, leur situation en limite d'aire de répartition les prédispose à une diversité faible, d'autre part leur situation dans une zone refuge glaciaire récente les prédisposent à une diversité élevée.

Dans ce contexte, il était important d'identifier le statut de ces populations en limite d'aire. Nous avons donc décrit précisément la distribution, la structure et la diversité génétique de ces populations afin d'évaluer leur potentiel évolutif dans un environnement sélectif variable (changement climatique, pression anthropique, Moritz 1994, Crandall *et al* 2000). Dans une première étape, les processus historiques ont été appréhendés par l'analyse mitochondriale de populations échantillonnées le long des côtes françaises, de la mer du Nord, de la Manche et du Nord de l'Océan Atlantique. Dans un deuxième temps, les événements plus contemporains tels que l'histoire démographique récente des populations mais aussi les flux de gènes entre populations ont été étudiés à l'aide du set de marqueurs microsatellites décrit au Chapitre 1, en portant une attention particulière aux populations situées en limite sud actuelle de répartition.

II.2 Matériel et méthodes

II.2.1- Echantillonnage et extraction des ADN

Pour cette étude, nous avons bénéficié d'un échantillonnage à échelle européenne réalisé entre 2003 et 2006 dans le cadre du programme européen BIOCOMBE (contrat EVK3-2001-00146) qui nous a permis d'obtenir de nombreuses populations, de la limite de répartition dans le sud-ouest de la France jusqu'à la mer de Barents en Russie (Figure I.2). Par la suite, dans le but d'une étude plus fine de la limite d'aire de distribution, plusieurs sites de la façade atlantique ont également été échantillonnés au cours des années 2007 et 2008. Au final, 19 sites européens dont 12 en France ont été étudiés, avec une moyenne de 30 individus par site (Tableau I.2)

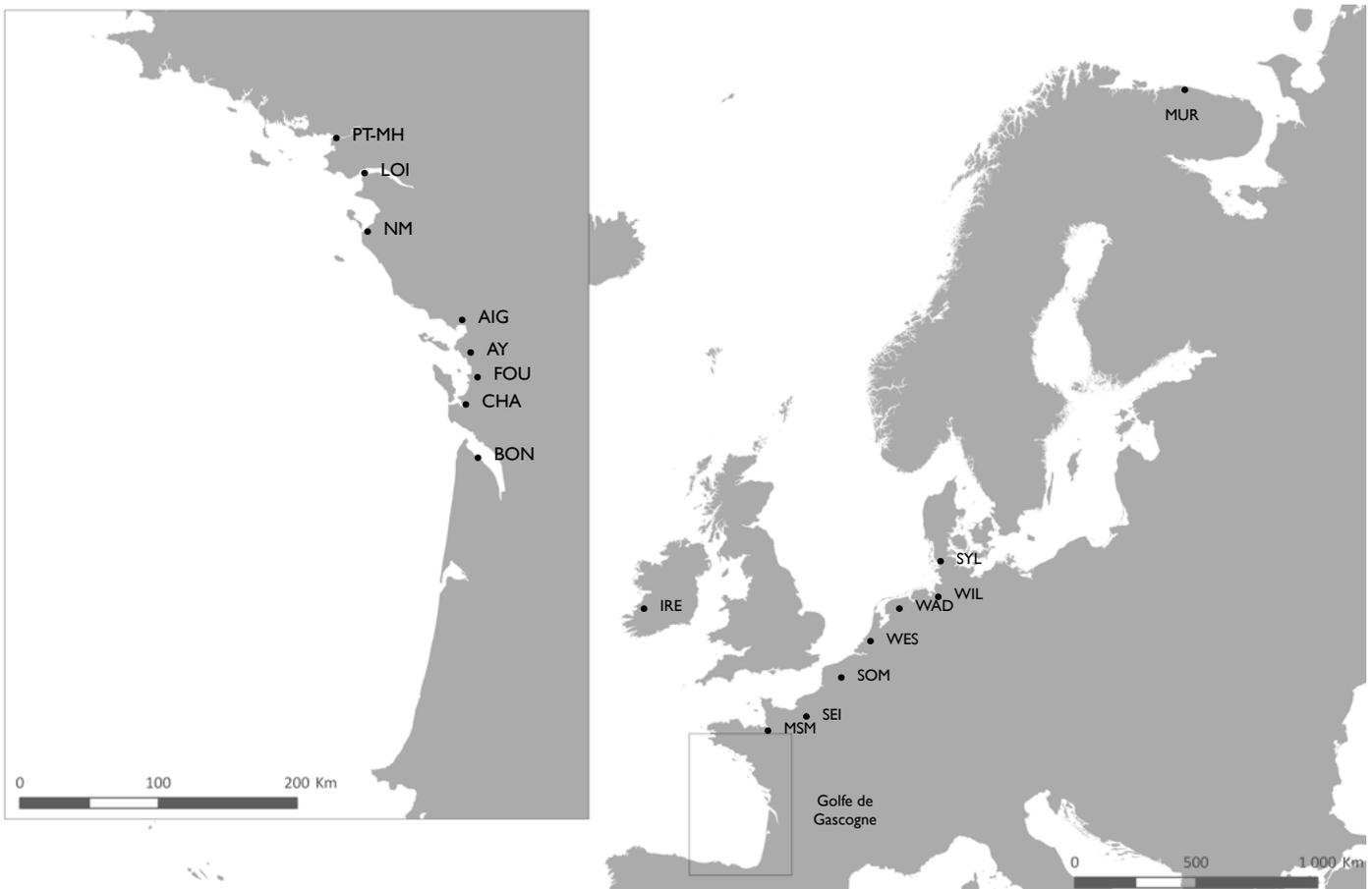


Figure I. 2 : Localisation des sites échantillonnés

Site	Code	GPS coord.N	GPS coord. W	N _{mito}	N _{nuc}
(1) Bonne Anse	BON	45.5523	-0.9294	16	13
(2) Charente	CHA	45.9583	-1.0501	15	25
(3) Fouras	F	45.984	-1.0925	8	36
(4) Yves	Y	46.0116	-1.0561	0	39
(5) Aytre	AY	46.126	-1.1306	13	50
(6) Aiguillon	AIG	46.1617	-1.1298	31	16
(7) Noirmoutiers	NM	46.9035	-2.1671	20	40
(8) Loire	LOI	47.2675	-2.1715	24	47
(9) Pont-Mahe	PT-MH	47.4442	-2.4608	16	23
(10) Mont Saint Michel	MSM	48.438	-1.5153	30	30
(11) Seine	SEI	49.4029	0.1205	32	29
(12) Somme	SOM	50.2146	1.6227	29	28
(13) Westerschelde	WES	51.3792	3.6272	31	27
(14) Balgzand	WAD	52.9301	4.7953	31	28
(15) Wilhemshaven	WIL	53.4802	8.0641	32	28
(16) Sylt	SYL	54.7789	8.2954	40	40
(17) Ireland	IRE	53.257	-9.1201	36	31
(18) Murman	MUR	69.3073	33.556	20	15
Total				424	573

Tableau I.2: nombre d'individus analysés avec le marqueur mitochondrial (N_{mito}) et les marqueurs nucléaires (N_{nuc})

Ces échantillons ont été conservés dans de l'éthanol absolu ou à -20°C . L'extraction de l'ADN a ensuite été réalisée avec un kit d'extraction *DneasyTM Tissue Kit* (Qiagen, Allemagne) à partir d'environ 15 mg de tissu musculaire prélevés au niveau du pied de l'animal. Chaque ADN a été amplifié par PCR avec le marqueur mitochondrial COI et les huit marqueurs microsatellites retenus, selon les protocoles détaillés au Chapitre 1. Au final 573 individus ont été génotypés avec les huit marqueurs microsatellites et 424 avec le marqueur mitochondrial (Tableau I.2).

II.2.2- Analyse des séquences obtenues avec le marqueur COI

Le nombre d'haplotypes H , la diversité haplotypique H_d (Nei 1987), le nombre de sites polymorphes S , et la diversité nucléotidique π (i.e. nombre moyen de différences nucléotidiques entre paires de séquences) ont été calculés pour chaque population avec le logiciel DnaSP v.5 (Librado et Rozas, 2009). Nous avons examiné la structure génétique totale et régionale des populations avec une analyse de la variance moléculaire (AMOVA, Excoffier *et al* 1992) avec le logiciel ARLEQUIN v.3.11. (Excoffier *et al* 2005). Les indices de fixation (Wright 1951) analogues aux paramètres F_{st} , F_{ct} et F_{sc} , ont été calculés pour analyser la différenciation génétique entre populations sur l'ensemble de l'aire étudiée (Φ_{st}), entre les groupes de populations dans l'aire de répartition de l'espèce (Φ_{ct}) et entre les populations au sein des groupes (Φ_{sc}) comme défini par Excoffier *et al* (1992). Les Φ -statistiques tiennent compte à la fois des fréquences haplotypiques et des distances moléculaires entre haplotypes. Les indices Φ ont été testés statistiquement par une permutation non-paramétrique disponible dans ARLEQUIN v.3.11. Les analyses phylogénétiques au niveau intraspécifique ont été réalisées avec le logiciel Network (Bandelt *et al* 1999) qui construit des réseaux d'haplotypes basés sur un algorithme appelé *median-joining* (Cassens *et al* 2003). La sélection des phylogénies intraspécifiques les plus probables a été effectuée grâce à un algorithme de maximum de parcimonie. De tels réseaux dessinent les relations évolutives (basées sur les événements mutationnels) entre les différents haplotypes observés sur l'ensemble du jeu de données. Un unique réseau synthétisant cette information a finalement été représenté et a été interprété en parallèle d'une carte de distribution des haplotypes dans les différentes populations étudiées.

Des sets de populations ont été définis à partir des similarités mitochondriales observées au niveau populationnel afin de tester des différences génétiques entre groupes de populations. Pour comparer les moyennes, des tests unilatéraux ont été réalisés dans R v.2.11.1 (R Development Core Team, 2010) utilisant soit le test t de Student soit le test de Wilcoxon quand la loi normale n'était pas respectée.

II.2.3- Analyse des données nucléaires

L'amplification de l'ensemble des échantillons a été réalisée selon le même protocole que celui décrit au Chapitre 1 pour la sélection des marqueurs microsatellites, la température d'hybridation T_i variant selon les amorces de 68°C à 62°C, les 30 cycles se faisant alors à une température d'hybridation fixe comprise entre 58°C et 52°C. L'électrophorèse sur gel de polyacrylamide a permis de génotyper tous les individus avec tous les marqueurs. La taille des allèles pour chaque marqueur a été déterminée à l'aide du logiciel SAGA-GT (v3.1: Automated microsatellite Analysis Software, LI-COR Biosciences) et d'un marqueur de poids moléculaire de référence.

II.2.3.1- Diversité génétique au sein des populations

Le nombre d'allèles par population et par locus (N_{all}), l'hétérozygotie observée (H_o) et attendue à l'équilibre de Hardy-Weinberg (H_e) ont été calculés avec le logiciel Genetix 4.05.2. (Belkhir et al 1996-2004). La richesse allélique (R_{all}), permettant de s'affranchir des différences de la taille d'échantillonnage entre les populations a été calculée avec le logiciel FSTAT 2.9.3.2 (Goudet 2002). N_{all} , R_{all} , H_o et H_e moyen et total ont aussi été calculés pour chaque set de populations (Golfe de Gascogne et Populations-centrales). Comme pour les données mitochondriales, des tests unilatéraux ont été réalisés pour comparer les valeurs moyennes de chaque set de populations (même procédure que celle décrite précédemment).

II.2.3.2- Différentiation génétique et relations entre populations

La différenciation génétique par paire de populations (F_{st}) et l'écart à l'équilibre de Hardy-Weinberg (F_{is}) ont été calculés avec le logiciel Genetix 4.05.2 (Belkhir et al 1996-2004). Dans le but d'estimer les relations entre les populations à l'échelle Européenne, les distances génétiques de Cavalli-Sforza et Edwards (1967) ont été calculées à partir des fréquences alléliques avec le logiciel PHYLIP (Felsenstein 2004).

Un arbre consensus de populations a été construit selon la méthode du neighbor-joining (NJ) (Saitou et Nei, 1987) avec le logiciel CONSENSE (Margush et McMorris, 1981), 1000 bootstraps ayant été compilés par le logiciel Seqboot (Felsenstein 1985), disponibles dans PHYLIP 3.69. Parce que l'hypothèse de départ qui considère que les populations sont à l'équilibre mutation-dérive, peut être biaisée avec l'utilisation des statistiques F_{st} si certaines populations ont subi des événements démographiques majeurs (goulots d'étranglement...), nous avons ensuite testé la structure des populations par une analyse Bayésienne disponible dans le «package» R Geneland (Guillot *et al* 2005). Ce package permet de s'affranchir du regroupement à priori des populations et prend en compte les coordonnées géographiques des échantillons ainsi que leur génotypes à plusieurs loci pour identifier des groupes à l'équilibre de Hardy-Weinberg. Pour chaque série, le nombre de groupes ou de classes (K) à l'équilibre de Hardy-Weinberg est estimé avec le maximum de vraisemblance. Ensuite, la valeur de K la plus vraisemblable est fixée et la probabilité postérieure d'appartenir à chaque classe est analysée pour chaque individu de l'ensemble du jeu de données. Finalement, une carte de ces probabilités postérieures est générée. Au total, 30 séries indépendantes ont été réalisées pour s'assurer de la cohérence des résultats avec les paramètres suivants: 150 000 itérations, taux maximum du processus de Poisson = 550, nombre maximum de noyaux de la tessellation de Poisson-Voronoi = 1650.

II.3- Résultats

II.3.1- Structure mitochondriale

Un fragment de 313 paires de base du gène COI a été séquencé pour 424 individus à 17 sites (Tableau I.2). Nous avons identifié 20 sites polymorphes définissant 19 haplotypes (N° d'accension Genbank: HM756170 à HM756189). Les relations entre ces 19 haplotypes sont détaillées dans le réseau d'haplotype présenté à la Figure I.3. Ce réseau est structuré en 4 clades composés des haplotypes les plus fréquents (94,7 % de la somme des fréquences des haplotypes H1 à H4). Parmi ces quatre haplotypes majoritaires, trois sont séparés par une seule mutation (H1 à H3) alors que l'haplotype le plus distant (H4) est séparé par trois mutations. Les 15 autres haplotypes sont équitablement répartis autour des 4 haplotypes majoritaires à des fréquences moindres. La structure globale du réseau est cohérent avec la localisation géographique des sites d'échantillonnage: les haplotypes H1 et H2 sont uniquement retrouvés dans les populations du Golfe de Gascogne, l'haplotype H3 correspond essentiellement à des individus de la Manche et de la mer du Nord alors que l'haplotype le plus distant (H4) est caractéristique des populations d'Irlande et de Murman (Figures I.4 et I.5).

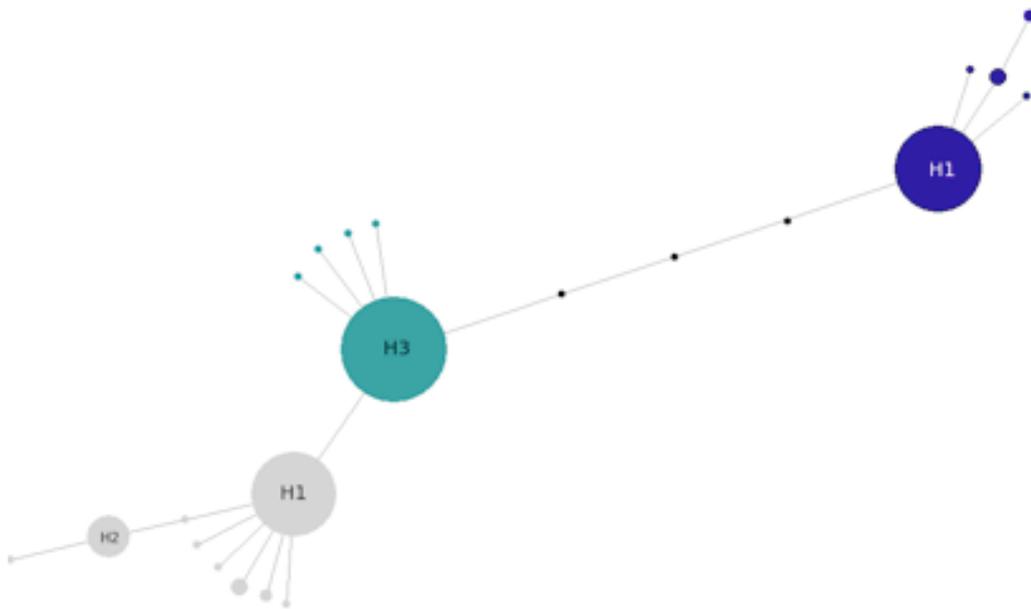


Figure I.3: Réseau présentant les relations phylogénétiques entre les haplotypes de *Macoma balthica*. H1 et H2 sont spécifiques aux populations du golfe de Gascogne. H3 et H4 sont caractéristiques des Populations Centrales. Chaque haplotype est séparé des autres par une mutation (non représentée ici) à l'exception de H1 et H3 qui sont séparés par trois mutations (points noirs).

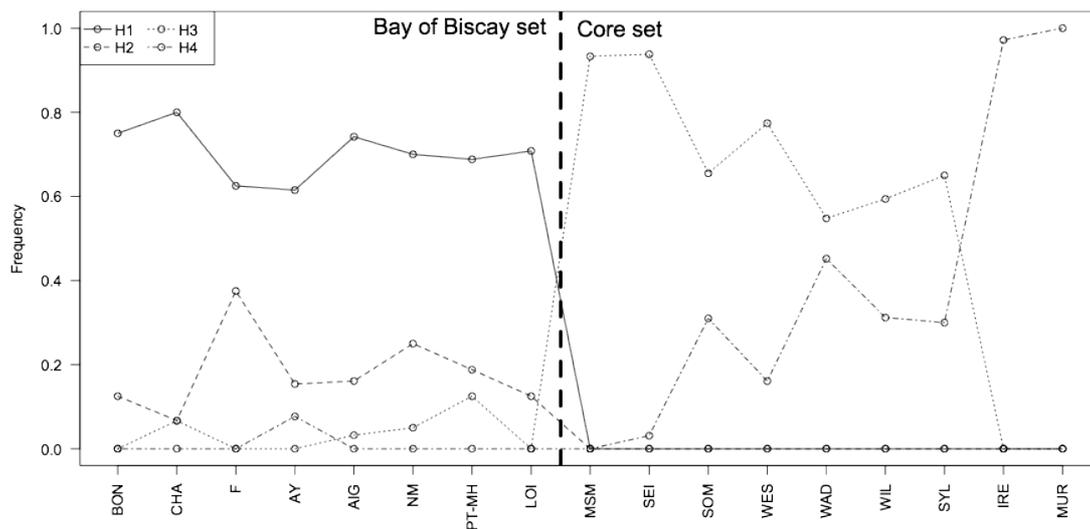


Figure I.4: Distribution des fréquences des haplotypes majoritaires le long des côtes européennes. A l'exception de l'Irlande, les sites sont classés par latitude croissante le long de l'axe x.

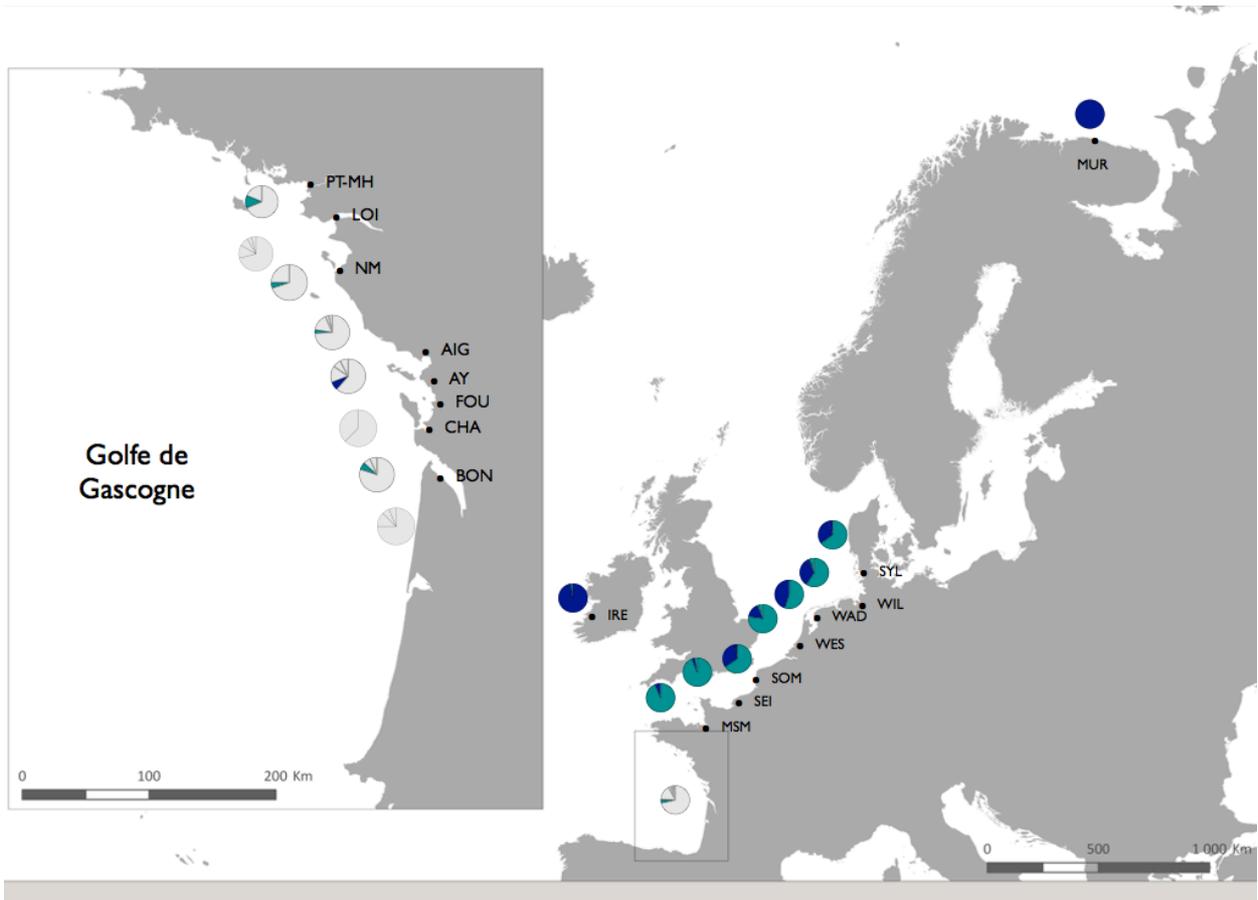


Figure I. 5 : Distribution européenne des haplotypes

II.3.2-Diversité moléculaire

Deux sets de population ont été définis à partir de la structure mitochondriale afin de tester les différences de diversité moléculaire potentielles entre les populations présentant un profil haplotypique contrasté. Le set Golfe de Gascogne est composé des 9 sites les plus au sud (populations 1 à 9, Tableau I.3) alors que le set Populations Centrales comprend les 9 populations restantes (populations 10 à 18, Tableau I.3).

Sur l'ensemble des données (à l'exception de la population de Murman qui est monomorphe), le nombre d'haplotypes H et de sites ségréants S varient respectivement de 2 à 5 (moyenne = 3,35, écart-type = 1,32) et de 2 à 7 (moyenne = 3,71, écart-type = 1,65, Tableau I.3). Si on compare la diversité mitochondriale entre les deux sets de populations, le nombre moyen d'haplotypes (Tableau I.3) est significativement supérieur dans le set Golfe de Gascogne (test Wilcoxon unilatéral, $W = 55$, $p = 0,033$). Le même résultat a été obtenu pour la diversité nucléotidique (test Wilcoxon unilatéral, $W = 57$, $p = 0,023$) avec une différence encore plus flagrante entre les sets de populations: la diversité nucléotidique moyenne dans le set Golfe de Gascogne ($\pi = 0,03153 \pm 0,02646$) étant plus de 10 fois supérieure à celle du set Populations Centrales ($\pi = 0,00275 \pm 0,00192$, Tableau I.3). Au niveau nucléaire, un génotype multilocus a été obtenu pour 542 individus collectés sur 18 sites d'échantillonnage (Tableau I.2). Aucun déséquilibre de liaison entre les marqueurs n'a été trouvé quand on s'intéresse à l'ensemble des populations (p comprise entre 0,2308 et 1,0000). Le nombre total d'allèles (N_{all}) varie de 6,62 (Bonne Anse) à 11,25 (Aytré) avec un total de 20,87 sur l'ensemble des populations (Tableau I.3).

L'analyse de la richesse allélique (R_{all}) qui varie de 4,0295 (11) à 4,7716 (6) nous permet d'affirmer que ces différences sont uniquement dues à la taille de l'échantillon (moyenne sur l'ensemble des populations et écart-type = 4,39 \pm 0,23). Une différence très significative (test t de Student, $t = 5,933$, $df = 16$, $p = 1,051e-5$) apparaît quand on compare les valeurs moyennes de R_{all} pour chaque set de populations, la valeur de R_{all} dans le set Golfe de Gascogne ($R_{all} = 4,58 \pm 0,15$) étant supérieure à la valeur de R_{all} dans le set Populations Centrales ($R_{all} = 4,21 \pm 0,12$, Figure I.6a). Des résultats similaires sont obtenus sur l'hétérozygotie observée et attendue (Figure I.6b) : peu de variation pour les valeurs de H_o et de H_e à l'échelle de l'ensemble des populations (moyenne sur les 18 populations et écart-type : 0,366 \pm 0,0425 pour H_o et 0,6325 \pm 0,0347 pour H_e) mais les deux indices sont significativement différents quand on regarde les valeurs obtenues pour chaque set, les indices étant supérieurs dans le set Golfe de Gascogne ($t = 2,513$, $df = 16$, $p = 0,012$ pour H_o , et $t = 2,246$, $df = 16$, $p = 0,020$ pour H_e). Le déficit en hétérozygote est relativement élevé, comme attendu chez les mollusques marins (valeur moyenne de F_{is} et écart-type : 0,424 \pm 0,054). Le même niveau de déficit en hétérozygote est observé dans les 2 sets de populations (valeur moyenne de F_{is} et écart-type : 0,4187 \pm 0,0491 pour le set Golfe de Gascogne et 0,4305 \pm 0,0603 pour le set Populations Centrales) et le test t de Student n'a révélé aucune différence entre les moyennes obtenues pour les deux sets ($t = 0,457$, $df = 16$, $p = 0,654$).

Bay of Biscay	Mitochondrial data					Nuclear data					
	N	S	H	Hd	π	N	Nall	Rall	Ho	Hs	Fis
(1)	16	4	4	0.442	0.00269	13	6.625	4.3806	0.304	0.577	0.5027
(2)	15	3	5	0.476	0.00188	25	8.526	4.5149	0.368	0.630	0.4325
(3)	-	-	-	-	-	39	11	4.7357	0.375	0.661	0.4437
(4)	8	2	2	0.536	0.05357	36	1.75	4.6560	0.368	0.640	0.4342
(5)	13	7	5	0.628	0.07179	50	11.25	4.7440	0.434	0.659	0.3339
(6)	31	4	5	0.434	0.00219	16	9.625	4.7716	0.408	0.670	0.4207
(7)	20	2	3	0.468	0.04184	37	9.125	4.4706	0.430	0.676	0.3764
(8)	24	5	5	0.493	0.03913	47	10.25	4.5025	0.381	0.674	0.4444
(9)	16	2	3	0.508	0.03917	23	7.375	4.4446	0.423	0.654	0.3794
Total	143	11	11	0.464	0.0411	286	18.25	4.5800	0.395	0.668	0.4305
Mean	17.9	3.63	4	0.498	0.0315	31.78	8.392	4.5801	0.388	0.649	0.4186
(S.D.)	7.08	1.768	1.195	0.062	0.0265	13.16	2.932	0.1475	0.041	0.031	0.0491
Core population	N	S	H	Hd	π	N	Nall	Rall	Ho	Hs	Fis
(10)	30	4	2	0.129	0.00154	30	8.375	4.2451	0.376	0.618	0.4087
(11)	32	4	3	0.123	0.00127	29	8.125	4.0295	0.333	0.599	0.4598
(12)	29	4	3	0.490	0.00441	28	8.375	4.2860	0.329	0.602	0.4681
(13)	31	5	4	0.385	0.00290	27	8.125	4.3306	0.303	0.626	0.5298
(14)	31	3	2	0.512	0.00460	28	8.750	4.3304	0.357	0.639	0.4570
(15)	32	6	5	0.565	0.00491	28	8.125	4.2112	0.382	0.607	0.3864
(16)	40	4	3	0.497	0.00448	40	9.500	4.2814	0.343	0.622	0.4613
(17)	36	4	2	0.056	0.00067	31	7.625	4.1177	0.379	0.558	0.3354
(18)	20	0	1	0	0	15	7.750	4.0522	0.293	0.673	0.3681
Total	281	11	10	0.523	0.07684	256	15.50	4.2110	0.358	0.628	0.4434
Mean	31.2	3.78	2.78	0.306	0.00275	28.44	8.306	4.2093	0.344	0.616	0.4305
(S.D.)	5.40	1.64	1.20	0.225	0.00192	6.386	0.559	0.1157	0.033	0.031	0.0604
All populations	N	S	H	Hd	π	N	Nall	Rall	Ho	Hs	Fis
Total	424	18	19	0.720	0.08984	542	20.875		0.3740	0.6500	0.4426
Mean	24.9	3.71	3.35	0.396	0.01630	30.11	8.305	4.395	0.3660	0.6325	0.4246
(S.D.)	9.15	1.65	1.32	0.192	0.02300	10.18	0.165	0.230	0.0425	0.0347	0.0537

Tableau I. 3 : N: nombre d'échantillons; S: sites ségrégeants; H: nombre d'haploype; Hd: diversité haplotypique; π : diversité nucléotidique; Nall: nombre d'allèles; Rall: richesse allélique; Ho: hétérozygotie observée; Hs: diversité génétique et Fis: déficience en hétérozygote

L'analyse de la richesse allélique (R_{all}) qui varie de 4,0295 (11) à 4,7716 (6) nous permet d'affirmer que ces différences sont uniquement dues à la taille de l'échantillon (moyenne sur l'ensemble des populations et écart-type = 4,39 +/- 0,23). On observe toutefois une différence hautement significative (test t de Student, $t = 5,933$, $df = 16$, $p = 1,051e-5$) quand on compare les valeurs moyennes de R_{all} pour chaque set de populations, la valeur de R_{all} dans le set Golfe de Gascogne ($R_{all} = 4,58 +/- 0,15$) étant supérieure à la valeur de R_{all} dans le set Populations Centrales ($R_{all} = 4,21 +/- 0,12$, Figure I.6a).

On obtient des résultats similaires quand on regarde l'hétérozygotie observée et attendue (Figure I.6b) : peu de variation pour les valeurs de H_o et de H_s à l'échelle de l'ensemble des populations (moyenne sur les 18 populations et écart-type : 0,366 +/- 0,0425 pour H_o et 0,6325 +/- 0,0347 pour H_s) mais les deux indices sont significativement différents quand on regarde les valeurs obtenues pour chaque set, les indices étant supérieurs dans le set Golfe de Gascogne ($t = 2,513$, $df = 16$, $p = 0,012$ pour H_o , et $t = 2,246$, $df = 16$, $p = 0,020$ pour H_s). Le déficit en hétérozygote est relativement élevé, comme attendu chez les mollusques marins (valeur moyenne de F_{is} et écart-type : 0,424 +/- 0,054). Le même niveau de déficit en hétérozygote a été retrouvé dans les deux sets de populations (valeur moyenne de F_{is} et écart-type : 0,4187 +/- 0,0491 pour le set Golfe de Gascogne et 0,4305 +/- 0,0603 pour le set Populations Centrales) et le test t de Student n'a révélé aucune différence entre les moyennes obtenues pour les deux sets ($t = 0,457$, $df = 16$, $p = 0,654$)

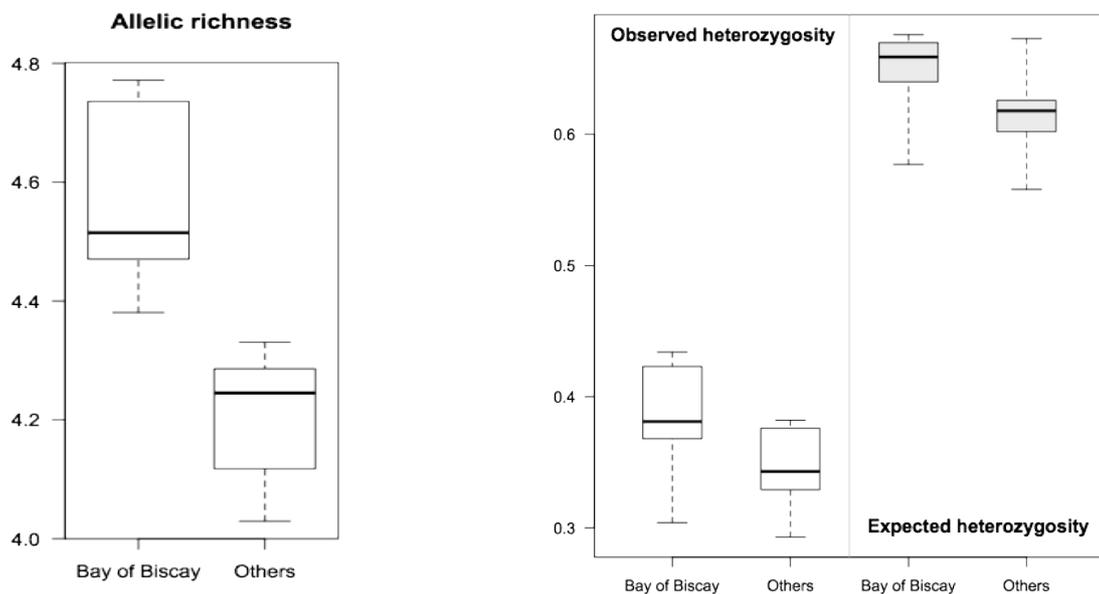


Figure I.6a: Comparaison des richesses alléliques entre les deux sets de population.
Figure I.6b: Comparaisons des hétérozygoties entre les deux sets de population.

<i>Site</i>	(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1)	0.000																
(2)	-0.046	0.000															
(3)	-0.017	0.041	0.000														
(5)	-0.047	-0.035	-0.018	0.000													
(6)	-0.037	-0.037	0.002	-0.032	0.000												
(7)	-0.029	0.006	-0.075	-0.029	-0.015	0.000											
(8)	-0.036	-0.009	-0.060	-0.040	-0.026	-0.054	0.000										
(9)	-0.035	-0.036	0.053	-0.030	-0.016	0.023	0.011	0.000									
(10)	0.492*	0.601*	0.374*	0.368*	0.520*	0.444*	0.431*	0.547*	0.000								
(11)	0.507*	0.622*	0.386*	0.379*	0.538*	0.460*	0.447*	0.564*	-0.021	0.000							
(12)	0.509*	0.627*	0.387*	0.382*	0.541*	0.462*	0.450*	0.567*	-0.024	-0.032	0.000						
(13)	0.442*	0.528*	0.342*	0.325*	0.466*	0.398*	0.385*	0.490*	0.012*	0.014	0.023	0.000					
(14)	0.509*	0.569*	0.446*	0.395*	0.525*	0.485*	0.480*	0.540*	0.288*	0.302*	0.318*	0.146*	0.000				
(15)	0.456*	0.517*	0.383*	0.344*	0.471*	0.425*	0.418*	0.491*	0.193*	0.204*	0.218*	0.072	-0.019	0.000			
(16)	0.461*	0.526*	0.388*	0.347*	0.477*	0.431*	0.426*	0.497*	0.180*	0.190*	0.203*	0.059	-0.007	-0.026	0.000		
(17)	0.862*	0.909*	0.836*	0.782*	0.880*	0.864*	0.868*	0.878*	0.869*	0.884*	0.891*	0.767*	0.472*	0.529*	0.571*	0.000	
(18)	0.891*	0.938*	0.866*	0.813*	0.909*	0.894*	0.899*	0.906*	0.910*	0.925*	0.931*	0.812*	0.533*	0.584*	0.626*	0.000	0.000

Tableau I.4: Fst par paires de populations obtenues. Les valeurs avec * sont significatives.

II.3.3-Structure génétique des populations

L'estimation des F_{st} à partir du marqueur mitochondrial (Tableau I.4) n'a pas montré de différences significatives entre les populations du golfe de Gascogne (populations 1 à 9) mais ces populations sont clairement différenciées des populations Irlande et Murman d'une part et des populations du set Populations Centrales d'autre part. Les résultats de l'AMOVA confirment ce regroupement : en effet la forte structure génétique observée sur l'ensemble des populations ($\Phi_{st} = 0,613$, $p = 0$) est essentiellement due à une forte différenciation génétique entre ces trois groupes ($\Phi_{ct} = 0,587$, $p = 0$) plutôt qu'à la différenciation entre populations au sein des groupes ($\Phi_{sc} = 0,0613$, $p = 0$). La structure obtenue à partir des marqueurs nucléaires est similaire à celle obtenue avec les données mitochondriales avec toutefois une légère différence. La structure générale ($F_{it} = 0,4425$) est essentiellement due à un fort déficit en hétérozygotes au sein des populations ($F_{is} = 0,4189$) alors que les populations entre elles ne sont pas fortement différenciées ($F_{st} = 0,0407$). Malgré tout, l'arbre consensus de populations, construit avec le logiciel PHYLIP (Figure I.7) montre que le set golfe de Gascogne est la plupart du temps isolé des autres populations (dans 57,7 % des cas sur les 1000 arbres simulés). A l'intérieur de ce groupe, une microstructure a été mise en évidence avec deux sous-groupes distincts : un sous-groupe Marennes-Oléron (MO, populations 1 à 6) et un sous-groupe Loire (LOI, populations 7 à 9). Ces deux sous-groupes ont été observés dans respectivement 73,6 % et 94 % des 1000 arbres simulés. Cette microstructure n'avait pas été mise en évidence avec les données mitochondriales. Afin de s'affranchir des groupes réalisés de manière manuelle (et donc potentiellement arbitraire) avec les F_{st} classiques, nous avons utilisé l'approche probabiliste du logiciel Genland (Guillot et al 2005). Les 30 simulations effectuées ont toutes abouti au même résultat (Figure I.8) : apparition d'une forte barrière au flux de gènes au sein du Golfe de Gascogne entre les populations 6 et mise en évidence de deux populations à l'équilibre de Hardy-Weinberg: le groupe Marennes-Oléron (populations 1 à 6) clairement isolé des populations situées plus au nord (populations 7 à 18) toutes regroupées dans un second groupe.

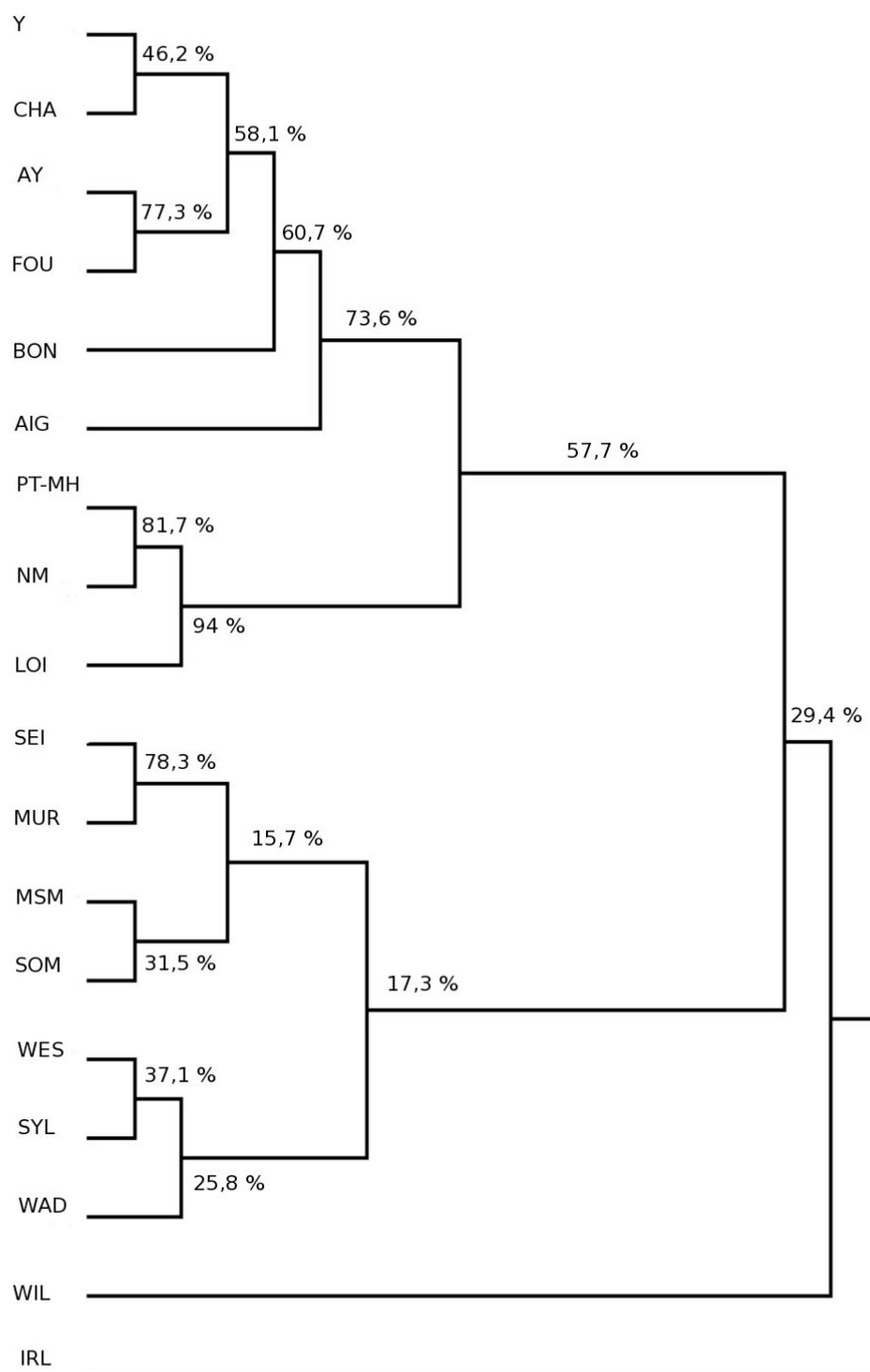


Figure I.7: Arbre consensus *neighbour joining*. Les distances génétiques ont été calculées à partir des fréquences alléliques des microsatellites (1000 bootstrap)

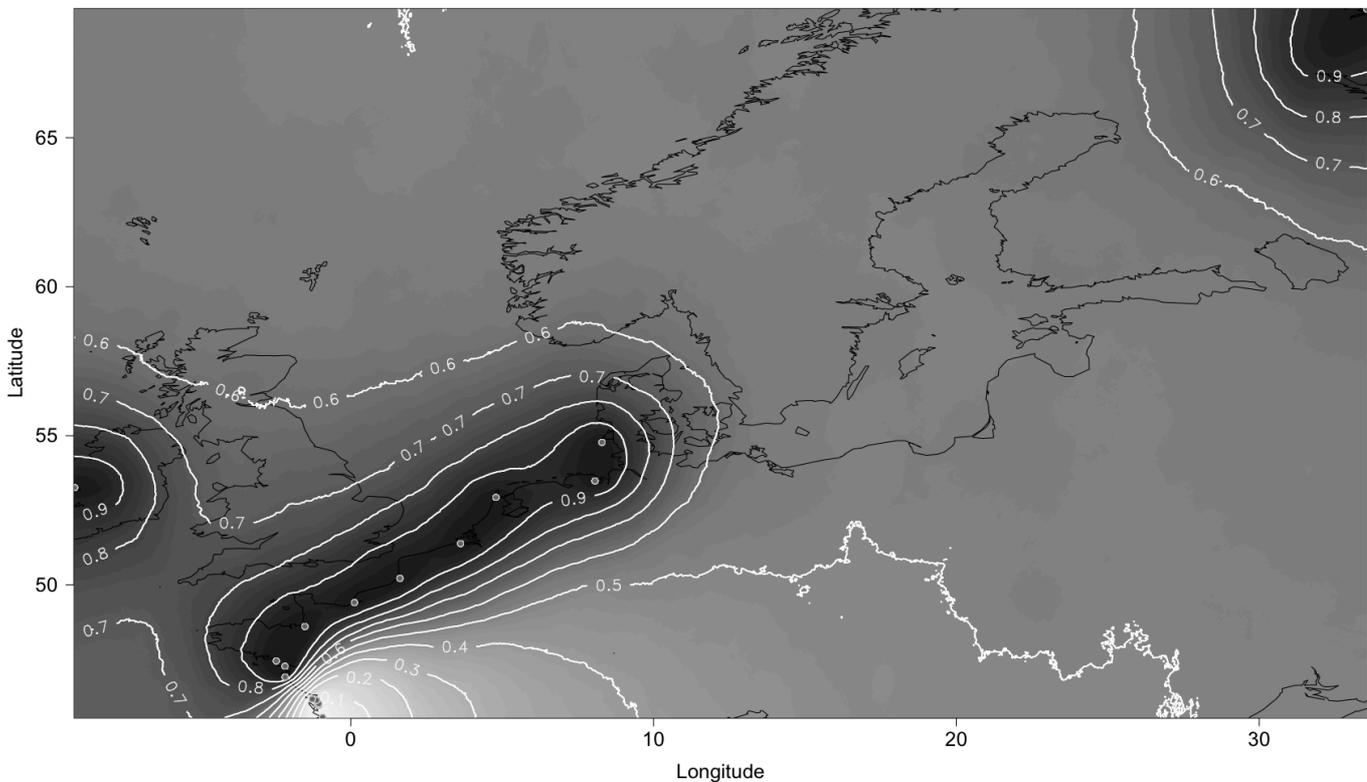


Figure I.8: Carte des probabilités postérieures d'appartenance à la première population à l'équilibre de Hardy Weinberg identifiée par Geneland. Les zones claires correspondent aux probabilités les plus faibles et les zones sombres aux probabilités les plus fortes. Les lignes blanches correspondent aux isolignes de probabilités, les lignes noires aux côtes et les points gris représentent les sites d'échantillonnage.

II.4- Discussion

La distribution biogéographique des populations naturelles de *Macoma balthica* a déjà été abordée avec divers marqueurs. Cependant, les populations situées au sud de l'aire de répartition étant représentées la plupart du temps par un site unique. Dans cette étude, un échantillonnage plus conséquent de cette région a été étudié, ce qui nous a permis de porter toute notre attention sur les populations théoriquement instables et fragiles des bordures d'aires de répartition.

II.4.1-Diversité génétique des populations périphériques

Dans le modèle classique *abundant-center*, l'attendu théorique dit que les espèces sont moins abondantes en limite d'aire qu'au coeur de la répartition, avec des populations de plus en plus petites et fragmentées au fur et à mesure que l'on s'approche de la bordure de l'aire (Lawton 1993, Vucetich et Waite 2003). L'organisation génétique en est généralement affectée, en terme de diversité et de structure. Par exemple, de nombreuses études ont montré que les populations périphériques exhibaient une faible diversité génétique, accompagnée d'une forte différenciation génétique (métapopulations structurées), d'une taille efficace réduite et d'un isolement géographique par rapport aux populations plus centrales (El Mousadik et Petit 1996, Durka 1999, Bouzat et Johnson 2004).

Près de l'actuelle limite sud de répartition, les populations de *Macoma balthica* semblent différer de ce modèle biogéographique (au regard des marqueurs mitochondriaux et nucléaires). Les populations du Golfe de Gascogne présentent une diversité génétique significativement élevée par rapport au set Populations Centrales ($p = 0,0122$ pour l'hétérozygotie et $p = 1,051e-5$ pour la richesse allélique). Par ailleurs, en dépit d'un déficit global en hétérozygotes ($F_{is} = 0,4305$) cohérent avec les valeurs observées pour de nombreux mollusques marins, les populations du Golfe de Gascogne ne semblent pas présenter d'autofécondation au vu de la forte diversité génétique observée dans ces populations. Ce résultat surprenant pourrait être lié aux processus historiques de dispersion de ce bivalve. En effet, *Macoma balthica* a pu coloniser l'océan Atlantique jusqu'aux côtes espagnoles à l'ouverture du détroit de Béring au Pléistocène. Lors du dernier maximum glaciaire (il y a 18000 ans), toutes les côtes de l'Europe du nord étaient recouvertes de glace à l'exception du golfe de Gascogne, les régions espagnoles devenant alors une des quatre zones refuge tempérées (Gomez et Lunt 2007). Les populations de ces zones refuge sont probablement à l'origine d'une grande partie de la diversité intraspécifique des biotas tempérés actuels (Petit et al 2003) et les sites les plus septentrionaux ont pu être reconquis après cette période glaciaire (Austin et al 2004, Howes *et al* 2006). Hewitt (1999) a par ailleurs clairement démontré que les espèces tempérées du sud de l'Europe présentaient de nombreux génomes géographiquement distincts et possédant qui plus est, une grande variété d'allèles. Cette observation est également valable dans les populations les plus méridionales de *Macoma balthica*, caractérisées par deux haplotypes majoritaires (H1 et H2) spécifiques du golfe de Gascogne. De plus, la forte diversité nucléaire observée dans les populations du Golfe de Gascogne est essentiellement due à la présence d'un grand nombre d'allèles privés. Ainsi 44 %, 47 % et 69 % des allèles obtenus respectivement aux loci mac40, mac10 et mac84 sont des allèles retrouvés uniquement dans les populations les plus au sud. Celles-ci ont en effet au total 42 allèles rares et spécifiques sur l'ensemble des 8 loci alors que les populations centrales n'en ont que 19.

Ces résultats sont cohérents avec une expansion rapide des populations de la zone refuge en Espagne et dans le golfe de Gascogne vers le nord après la dernière glaciation.

II.4.2-Structure génétique en limite de répartition

Les marqueurs mitochondriaux et nucléaires montrent une structure génétique significative de nos populations. Toutefois, cette structure diffère légèrement en fonction des marqueurs considérés: la structure de l'ADN mitochondrial révèle une barrière géographique claire entre les populations du Golfe de Gascogne (populations 1 à 9) et celles situées plus au nord (Figure I.2), alors que les résultats obtenus avec les marqueurs nucléaires mettent en évidence une limite supplémentaire aux flux de gènes à l'intérieur du golfe de Gascogne (entre les populations 6 et 7, Figure I.7).

Au sein des groupes par contre, les populations partagent toutes les mêmes caractéristiques et ce, quelle que soit l'analyse considérée (mitochondriale ou nucléaire). Ainsi les populations du Golfe de Gascogne ont un profil mitochondrial similaire: chacune d'entre elles possède deux des haplotypes les plus fréquents (H1 et H2), haplotypes non retrouvés dans le set «Population Centrales». En outre, les haplotypes H3 et H4 sont rarement observés dans les populations du Golfe de Gascogne alors qu'ils sont les plus fréquents dans les populations centrales (à l'exception de la population Murman qui est monomorphe, Figure I.2). Ce modèle mitochondrial est cohérent avec une recolonisation post-glaciaire des côtes du nord de l'Europe selon deux voies distinctes : (i) du sud au nord à partir de la zone refuge ibérique et (ii) du nord au sud à partir de l'océan Pacifique.

Concernant les résultats nucléaires, l'arbre neighbor-joining mais aussi les simulations bayésiennes (Figures I.7 et I.8) montrent également une solide barrière génétique au sein du Golfe de Gascogne (valeurs de bootstrap égales à 73% et 94 % pour respectivement le sous-groupe Marennes-Oléron et le sous-groupe Loire). Cette barrière n'avait pas été mise en évidence lors précédentes études avec des marqueurs allozymiques (Hummel 1997b), ceux-ci n'étant probablement pas assez résolutifs.

Cette structure observée pose alors deux questions (i) comment expliquer une structure si forte pour un organisme capable de se disperser sur de très longues distances pendant sa phase larvaire pélagique ? (ii) comment expliquer la différence de localisation de cette barrière en fonction des marqueurs utilisés ?

II.4.2.1-Dispersion à longue distance et barrière géographique

L'utilisation des deux types de marqueurs nous a permis d'identifier une forte structure génétique au niveau de la Bretagne (côtes françaises les plus à l'ouest délimitant le sud de la Manche). Une telle structure observée entre des populations de *Macoma balthica* si proches géographiquement, est inattendue au regard des traits d'histoire de vie de l'espèce.

En effet, son premier stade de développement est composé d'une phase larvaire de 2 à 5 semaines au cours desquelles les larves, portées par les courants océaniques sont susceptibles de parcourir de longues distances (Beukema et De Vlas 1989) avant de se fixer à l'âge adulte (Caddy 1969, Drent 2002). En l'absence d'une barrière physique ou génétique, on s'attend à ce que ce stade larvaire pélagique maintienne un certain flux de gènes entre les populations du sud et du nord de la Bretagne. Deux sortes de barrières géographiques pourraient expliquer cette structure génétique globale (à savoir, l'isolement génétique des populations les plus méridionales par rapport aux populations centrales). Premièrement, la Bretagne est connue pour être une barrière biogéographique marine. Leppäkoski *et al* (2002) ont ainsi montré que les effets du Gulf Stream le long des côtes européennes pouvaient résulter en une forte barrière biogéographique autour de la Bretagne pour une variété de phylum marins, limitant l'échange interspécifique entre les zones du nord et du sud. Deuxièmement, l'hydrodynamique locale pourrait influencer la phase larvaire et maintenir ainsi les populations du sud isolées des populations du centre de la répartition de *Macoma balthica*. Un modèle de dispersion larvaire testé sur le polychète marin *Pectinaria koreni* a ainsi confirmé l'existence d'une forte rétention larvaire au sein d'une baie entraînant alors de l'auto-recrutement (Ellien *et al* 2000). Ces mécanismes d'isolement post-zygotique pourraient exister dans la baie semi-fermée de Marennes-Oléron, et à plus grande échelle dans les Pertuis Charentais (populations 1 à 6), zone semi-fermée de 1300 km² délimité par deux îles d'environ 30 kilomètres de long éloignées l'une de l'autre par quelques kilomètres seulement.

II.4.2.2- Rupture au flux de gènes dans le golfe de Gascogne.

En plus de la barrière biogéographique présentée ci-dessus, une barrière génétique au flux de gènes pourrait aussi être une explication à la forte structure observée et aux résultats contrastés obtenus avec les deux types de marqueurs (Lemaire *et al* 2005).

La position de la barrière au flux de gènes mise en évidence avec les deux types de marqueurs, diffère de quelques centaines de kilomètres. Comme souligné plus haut, la Bretagne est une barrière biogéographique marine majeure. Dans ce contexte, les 3 populations du groupe Loire (populations 7 à 9) qui sont assignées à différentes régions selon le marqueur moléculaire utilisé (dans le set Golfe de Gascogne par le marqueur mitochondrial ou dans le set Populations Centrales par les marqueurs nucléaires) se trouvent peut-être situées dans une zone de limite au flux de gènes. La recolonisation postglaciaire à partir des deux stocks isolés de *Macoma balthica* a pu être ralentie aux alentours de cette barrière biogéographique, expliquant les différences entre modèles mitochondriaux et nucléaires de part et d'autre de la Bretagne. De plus, la migration peut être une cause de déséquilibre cyto-nucléaire, particulièrement quand le déséquilibre existe dans le groupe de

migrants (Asmusen et Arnold 1991, Asmussen et Orive 2000), ce qui est probable lorsque les migrants sont issus de deux (ou plus) groupes génétiquement distincts (Latta *et al* 2001). Ainsi, considérant que les populations de *Macoma balthica* du centre et de la limite sud de l'aire de répartition sont des populations divergentes, il n'est pas surprenant d'observer un désaccord entre les signatures mitochondriales et nucléaires dans une zone, où l'admixture de deux types de populations est censée se produire (zone de transition décrite par Lemaire *et al* 2005).

En outre, un flux de gènes sexe-biaisé peut également induire un isolement prézygotique (Crochet *et al* 2003). Mais les flux de gènes mitochondriaux et nucléaires entre ces deux régions, estimés à la fois par la méthode des allèles privés (Slatkin 1985) et la méthode F_{st} (Wright 1951) se sont avérées similaires, laissant supposer que la migration est équitable entre mâles et femelles dans cette zone.

II.5- Conclusion

Deux résultats majeurs résultent de l'étude des populations de *Macoma balthica* en limite sud d'aire de répartition:

- Ces populations présentent une diversité génétique élevée probablement due à leur localisation près d'une zone refuge glaciaire. Cette diversité peut laisser présager d'un potentiel évolutif également élevé face à un stress supplémentaire.
- Les marqueurs nucléaires ont permis de mettre en évidence une forte barrière au flux de gènes au sein même du Golfe de Gascogne. Cette barrière n'avait jusqu'alors jamais été mise en évidence probablement parce que les marqueurs utilisés n'étaient pas adaptés et l'échantillonnage des populations en limite d'aire trop restreint (une seule population analysée dans les travaux de Luttikhuisen *et al* (2003) et Nikula *et al* (2008)).

Chapitre 3

*Etude d'un site anthropisé au cœur de
l'aire de répartition.*

*Cas des populations de la baie de
Gdansk (Pologne) soumises aux
contaminants de la Vistule et aux
contraintes physico-chimiques de la
mer Baltique*

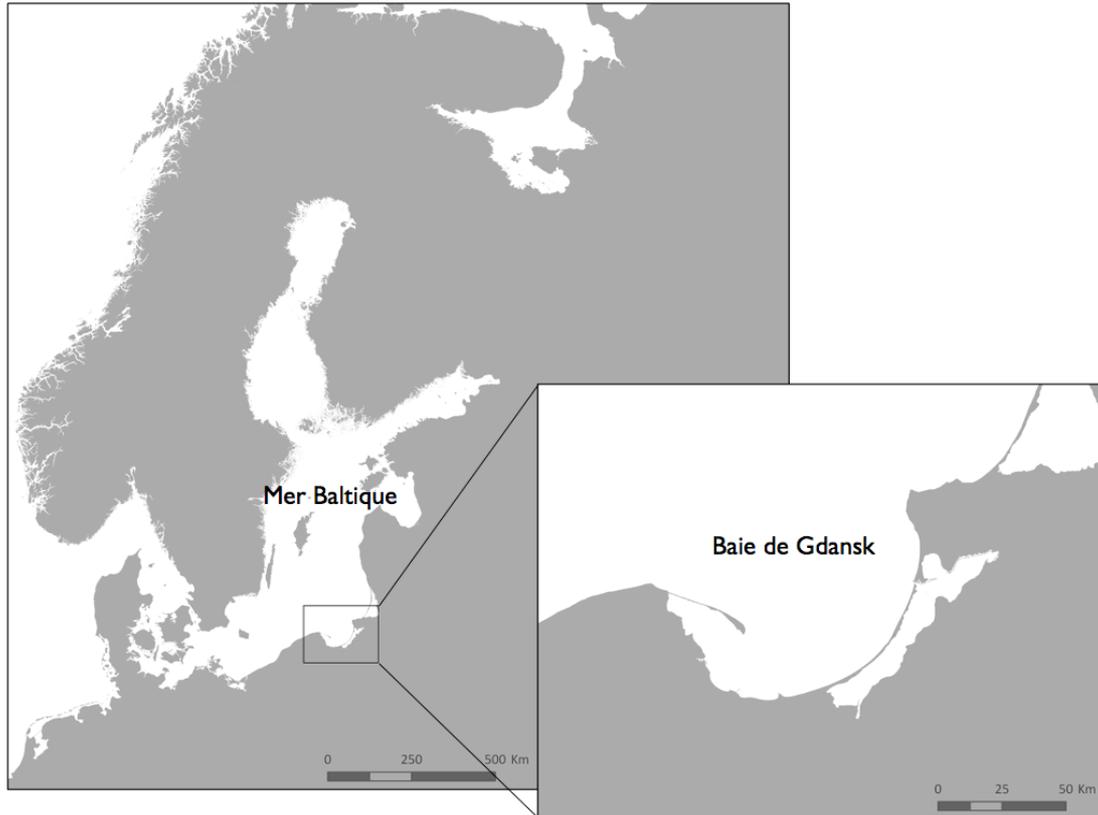


Figure I.9 : Localisation de la baie de Gdansk dans la mer Baltique (Pologne).

Site	Code	Nmito	Nnuc
Panache de la Vistule	SW 5	19	18
Panache de la Vistule	SW 10	20	18
Panache de la Vistule	SW 20	20	23
Panache de la Vistule	SW 30	26	10
Panache de la Vistule	SW 40	23	5
Panache de la Vistule	SW 60	18	8
Baie de Puck	MW 30	19	21
Baie de Gdansk	GN 40	24	12
Presqu'île de Hel	HELL 45	29	17
Presqu'île de Hel	HELL 60	28	30
Presqu'île de Hel	HELL 85	5	8
Mecklenburg Bight	MEC	27	—
Sylt	SYL	41	—
Murman	MUR	20	—

Tableau I.5 : nombre d'individus analysés avec le marqueur mitochondrial (Nmito) et les marqueurs nucléaires (Nnuc).

III.1- La mer Baltique: un environnement marin particulier ?

La mer Baltique est une mer semi-fermée ayant peu d'échanges avec le système océanique voisin et se distinguant des autres bassins océaniques mondiaux par plusieurs aspects. Peu profonde, elle se caractérise (i) par des eaux saumâtres avec une salinité moyenne 5 fois inférieure à celle des océans et mers mondiales (environ 7 PSU pour la mer Baltique contre 35 pour les autres systèmes marins), et (ii) une stratification de l'eau en fonction de la température et de la salinité (Rheinheimer 1995). Ainsi, au-delà d'une profondeur variant de 20 à 50 mètres, l'eau est plus froide et plus salée. Ces paramètres s'accompagnent également d'une diminution de l'apport en oxygène et il est probable que les ressources alimentaires varient également au-delà de cette zone.

En dehors de ces caractéristiques physico-chimiques, la mer Baltique est également soumise à une forte pollution liée à une activité humaine intense dans son bassin versant. De nombreuses études ont en effet montré l'impact des apports constants d'eau douce en provenance de la Vistule. Ce fleuve, située au sud de la mer Baltique dans la baie de Gdansk (Figure I.9), draine en effet 54 % de la Pologne et longe des zones industrielles et urbaines comme Katowice, Cracovie ou Varsovie (Rainbow *et al* 2004, Szefer 2002). Elle se charge de nombreux éléments qui sont déversés de manière constante dans la mer Baltique et des métaux comme le cuivre, le zinc ou l'argent, sont retrouvés en concentration importante à son embouchure (Sokolowski *et al* 1998, 2002 et 2007).

Malgré ces conditions environnementales extrêmes, *Macoma balthica* demeure une espèce dominante de la macrofaune benthique de la mer Baltique (Rumohr *et al* 1996). Présente aussi bien en milieu peu profond sub-littoral qu'en zone profonde aphotique (jusqu'à 91 m de profondeur, Janas 1998), elle est également décrite au nord de la mer Baltique, région où la salinité avoisine seulement les 3 PSU et dans laquelle peu d'espèces locales sont capables de vivre (Väinölä 1989). Dans cette région, *Macoma balthica* demeure par ailleurs constamment submergée, la mer Baltique n'étant pas soumise au phénomène des marées. En outre, les populations de *Macoma balthica* de la mer Baltique présentent des différences génétiques significatives avec celles de l'Océan Atlantique. Väinölä en 2003 puis Strelkov en 2007 ont en effet montré que les populations de *Macoma balthica* de la mer Baltique résultaient de l'hybridation du génome de la sous-espèce Pacifique *Macoma balthica balthica* (60%) et de celui de la sous-espèce Atlantique *Macoma balthica rubra* (40%). Une étude plus récente de Nikula *et al* (2008) a par ailleurs mis en évidence une zone hybride au niveau du détroit de Kattegat (entre la mer du nord et la mer Baltique). Cette zone de transition s'étend dans la mer Baltique jusqu'à la baie de Gdansk et maintient la connexion entre un taxa isolé (e.g. mer Baltique) et une des lignées parentales (e.g. Atlantique).

Après l'étude de populations de *Macoma balthica* subissant un réchauffement climatique en limite d'aire de répartition, nous avons choisi de nous pencher sur les

populations de la baie de Gdansk localisées au cœur de l'aire de répartition et contraintes à l'apport permanent de métaux lourds par la Vistule et à une thermo-halocline. Ces différents stress sont probablement à l'origine du taux de métabolisme limité observé chez certaines de ces populations (autour de 6mg de poids sec par cm³), taux voisin de celui des populations en limite d'aire de répartition (Hummel *et al* 2000). Les travaux de Hummel suggèrent que les populations de la mer Baltique et plus précisément celles de la baie de Gdansk, sont en limite de tolérance physiologique (comme les populations du sud de la répartition).

Dans ce troisième chapitre, notre objectif est d'analyser la part des processus démographiques (notamment la rupture des flux de gènes) dans la réponse aux pressions environnementales. Dans un premier temps, nous avons mené une analyse phylogéographique avec un marqueur mitochondrial à l'échelle européenne intégrant des populations de l'océan Atlantique et de la mer du nord mais aussi de la zone hybride. Dans un second temps, une étude génétique à micro-échelle a été menée avec des marqueurs microsatellites sur des populations de la baie de Gdansk.

III.2- Matériel et méthodes

III.2.1- Echantillonnage des populations

Pour l'analyse phylogéographique à l'échelle européenne, nous avons utilisé les mêmes populations de Murman et de Sylt que celles étudiées au Chapitre 2 (Tableau I.5). L'échantillonnage de la population de Mecklenburg Bight dans le détroit de Kattegat été réalisé en même temps que ces populations, soit au cours de l'année 2006. Les populations de la baie de Gdansk ont été échantillonnées entre 1997 et 1999 à 9 sites décrits à la Figure I.10.

Ces sites ont été choisis selon plusieurs critères environnementaux: le site de la baie de Puck (Mw30) est caractérisée par une forte eutrophisation et une forte hypoxie. Les populations du site de Hel (Hell45, Hell60 et Hell85) sont quant à elles situées à différentes profondeurs à l'entrée de la baie de Gdansk. Le site de la baie de Gdansk proprement dit (Gn40), se trouve sur une des principales voies maritimes reliant les ports de Gdynia et de Gdansk à la mer Baltique.

Enfin, des populations (Sw) ont été prélevées à l'embouchure de la Vistule selon un gradient de pollution et de profondeur (de 5 à 60 mètres) (Figure I.11).

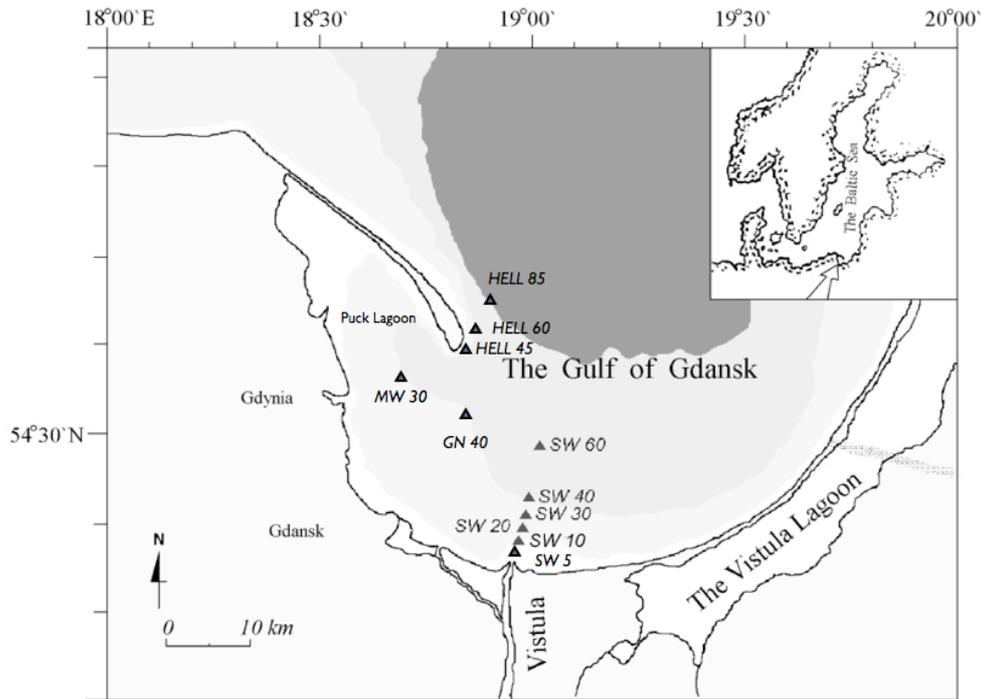


Figure I.10 : Sites d'échantillonnage dans la baie de Gdansk.

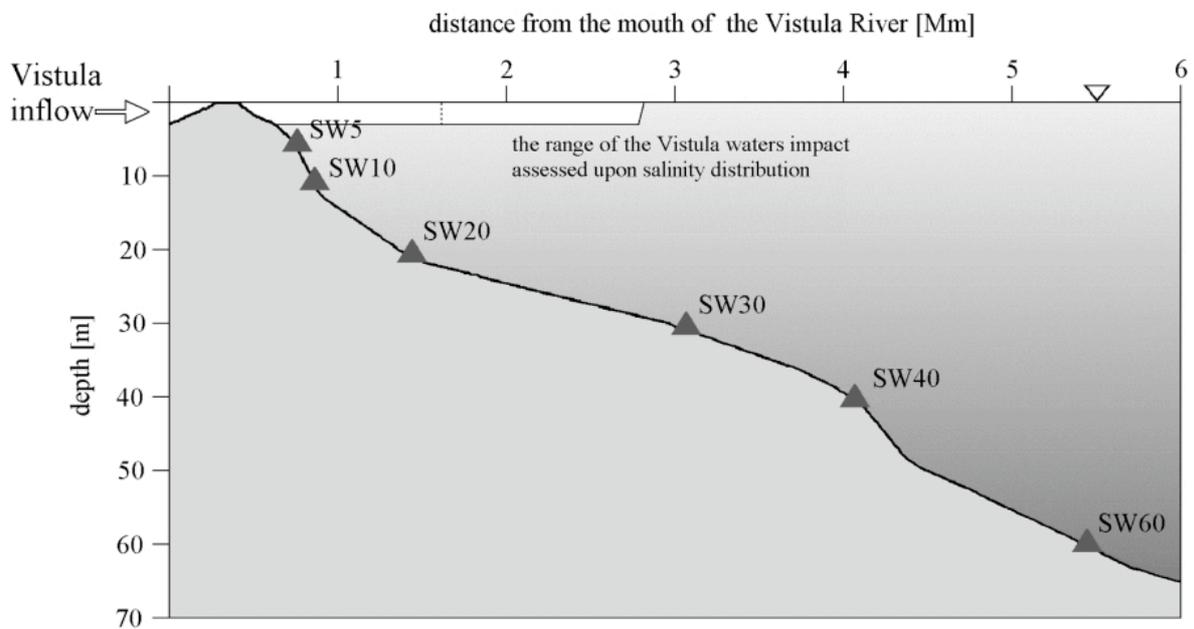


Figure I.11 : Sites d'échantillonnage dans le panache de la Vistule

III.2.2- Conditions d'amplification et analyses statistiques mises en oeuvre

Les individus ont été conservés dans de l'éthanol à 95 % jusqu'à l'extraction de l'ADN. Le protocole d'extraction et les conditions d'amplification avec les marqueurs mitochondrial COI et microsattellites, sont les mêmes que celles décrites au Chapitre 2. Seuls 6 marqueurs microsattellites (mac4, mac19, mac10, mac64, mac84, mac40) ont été utilisés dans cette étude, les marqueurs mac17 et macb324 n'étant pas indépendants des autres ($0,143 < p < 1,000$). La plupart des analyses statistiques réalisées sur les données obtenues avec le marqueur mitochondrial et les marqueurs microsattellites, sont elles aussi décrites au Chapitre 2.

La structure génétique globale (F_{it} , F_{is} , F_{st}) ainsi que la différenciation génétique par paires de populations (F_{st}) ont été calculées avec le logiciel Genetix (Belkhir 1996-2004) par la méthode de Weir et Cokerham (1984). La significativité des valeurs de F_{st} a été testée par permutations globales des individus (1000 permutations).

L'assignation des individus au sein de la baie de Gdansk a été réalisée avec le logiciel GenClass v2.0.g (Piry *et al* 2004) selon la méthode bayésienne de Rannala et Mountain (1997) sur la base des données microsattellites. Ce logiciel permet de calculer, pour chaque individu, la population source la plus vraisemblable ainsi que les probabilités d'appartenance à chacune des populations sources.

Des groupes de populations ont été réalisés en fonction de la profondeur d'échantillonnage pour l'analyse de la variance moléculaire (AMOVA). Trois groupes ont ainsi été déterminés: le groupe 1 regroupe les populations Sw5, Sw10 et Sw20, le groupe 2 regroupe les populations Mw30, Sw30, Sw40 et Hell45 et le groupe 3 est constitué des populations prélevées en dessous de 50 mètres (Hell60, Hell85, Sw60).

III.3- Résultats

III.3.1- Structure mitochondriale des populations de la mer Baltique

Un fragment de 334 paires de base du gène COI a été séquencé sur 268 individus échantillonnés sur 14 sites le long des côtes du nord de l'Europe. Nous avons identifié 32 sites polymorphes définissant 36 haplotypes. Les relations entre ces 32 haplotypes sont présentées dans le réseau d'haplotype (Figure I.12).

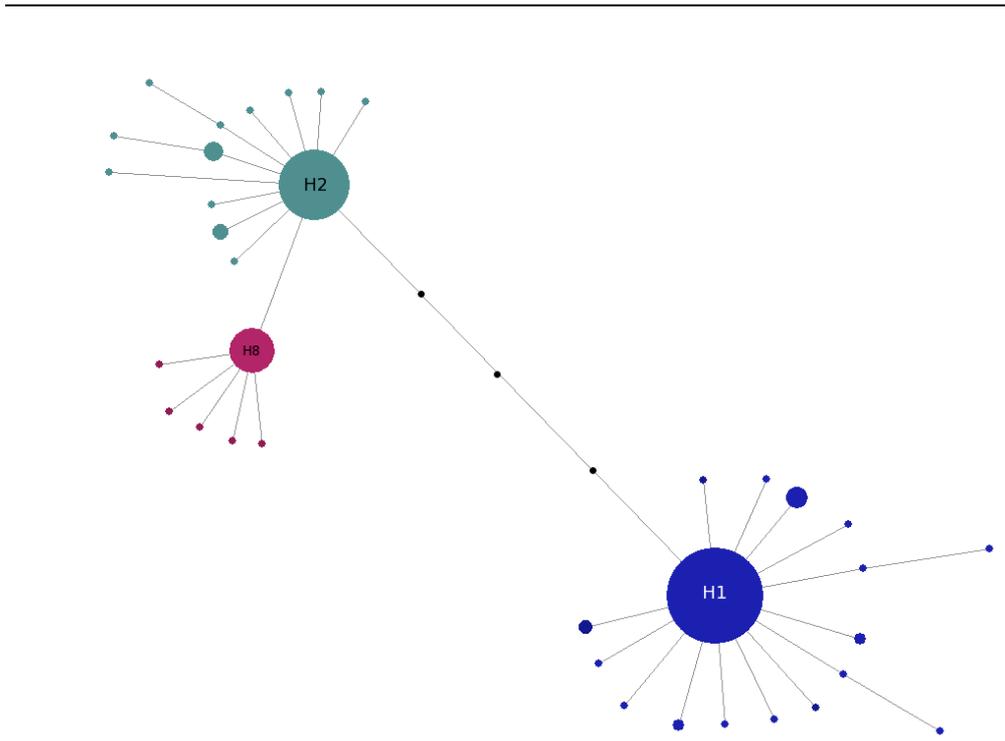


Figure I.12 : Réseau présentant les relations phylogénétiques entre les haplotypes de *Macoma balthica*. H1 et H2 sont caractéristiques de la lignée parentale Atlantique (*Macoma macoma rubra*). H8 est spécifique des populations de la baie de Gdansk.

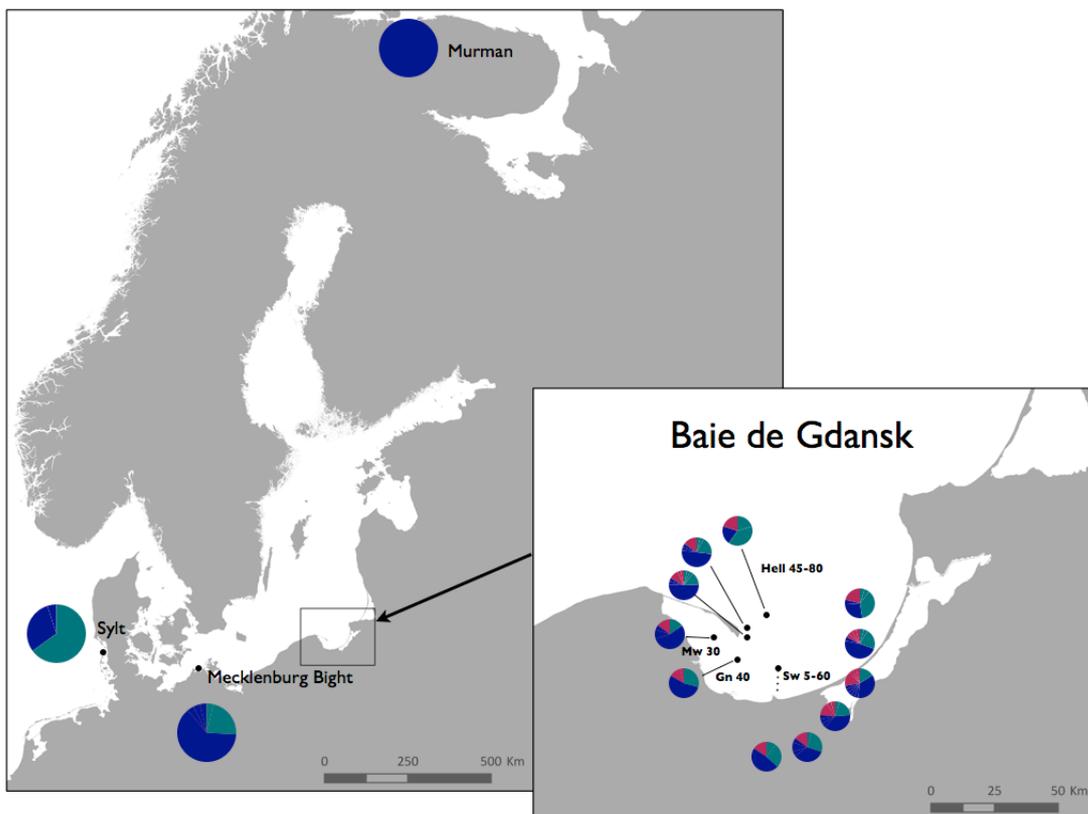


Figure I.13 : Distribution des haplotypes majoritaires dans le nord de l'Europe et au sein de la baie de Gdansk

Ce réseau est structuré en 3 clades composés des haplotypes les plus fréquents (H1, H2 et H8). Parmi ces trois haplotypes majoritaires, deux sont séparés par une seule mutation (H2 et H8) alors que l'haplotype le plus distant (H1) est séparé par trois mutations. Les 33 autres haplotypes sont équitablement répartis autour des 3 haplotypes majoritaires à des fréquences moindres. Les haplotypes H1 et H2 se retrouvent dans l'ensemble des populations (à l'exception de Murman qui est monomorphe) mais généralement, la fréquence de l'haplotype H1 est supérieure à celle de l'haplotype H2 (exception faite pour Hell85 et Sw60 où l'haplotype H2 est dominant). L'haplotype H8 est spécifique des populations de la baie de Gdansk (Figures I.12 et I.13). Notons toutefois que la population de Mecklenburg Bight, située à l'entrée de la mer Baltique, ne possède pas cet haplotype.

III.3.2- Diversité moléculaire en Baie de Gdansk

Quand on considère uniquement les populations de la Baie de Gdansk (Tableau I.6), le nombre d'haplotypes H et de sites ségrégeants S varient respectivement de 4 à 9 (moyenne = 6,18 écart-type = 2,08) et de 4 à 11 (moyenne = 7,36 écart-type = 2,37, table 6). La diversité haplotypique (H_d) quant à elle, varie de 0,583 (Gn40) à 0,9 (Hell85) avec une valeur moyenne de 0,7381 (écart-type = 0,0871) et une valeur totale pour la baie de 0,724. La diversité nucléotidique varie de 0,0625 (Hell85) à 0,08542 (Sw10) pour une valeur totale de 0,0681 (valeur moyenne = 0,06726, écart type= 0,00954).

Un génotype multilocus a été obtenu pour 170 individus collectés (Tableau I.6) avec 6 marqueurs microsatellites. Aucun déséquilibre de liaison entre les marqueurs n'a été trouvé sur l'ensemble des populations (p comprise entre 0,21480 et 1,0000). Le nombre total d'allèles (N_{all}) varie de 5,5 (Sw40) à 6,833 (Hell60 et Mw30) avec un total de 11,16 sur l'ensemble des populations (Tableau I.6). Le nombre d'allèles varie peu d'une population à une autre (moyenne = 5,072, écart-type = 1,365). Ce résultat est confirmé par l'analyse de la richesse allélique (R_{all}) qui varie de 1,528 (Sw40) à 1,620 (Hell45), ceci même pour les populations à faible effectif (Hell85, Sw40 et Sw60, moyenne sur l'ensemble des populations et écart-type = 1,5729 +/- 0,033).

La variation observée sur l'ensemble des populations pour l'hétérozygotie observée (H_o), (moyenne sur les 11 populations et écart-type: 0,395 +/- 0,0731 pour H_o) est essentiellement due à deux valeurs extrêmes obtenues pour les populations Sw5 (H_o = 0,2818) et Sw30 (H_o = 0,5354). Cette variation est moindre pour l'hétérozygotie attendue qui varie de 0,4396 à 0,5906 (pour respectivement Sw40 et Hell45) et dont la valeur moyenne s'élève à 0,5386 (écart-type = 0,048). Le déficit en hétérozygote est relativement élevé, comme attendu chez les mollusques marins (valeur moyenne de F_{is} et écart-type: 0,334 +/- 0,135).

Population	Mitochondrial					Nucléaire					
	N	S	H	Hd	π	N	Nall	Rall	Ho	Hs	Fis
HELL 85	5	5	4	0,900	0,0625	8	4,33	1,580	0,4048	0,5114	0,3276
HELL 60	28	10	9	0,751	0,0733	30	6,83	1,592	0,4085	0,578	0,3155
HELL 45	29	9	9	0,709	0,0645	17	6,16	1,620	0,3945	0,5906	0,3735
GN 40	24	4	3	0,583	0,0546	12	4,33	1,513	0,3131	0,4926	0,4173
MW 30	20	6	5	0,674	0,0623	21	6,83	1,609	0,4097	0,5883	0,3826
SW 5	19	5	4	0,708	0,0643	18	5,14	1,557	0,2818	0,5344	0,5389
SW 10	21	11	7	0,800	0,0854	18	5,83	1,598	0,4613	0,5862	0,2547
SW 20	19	10	8	0,825	0,0818	23	5,83	1,576	0,3702	0,5652	0,3731
SW 30	23	7	7	0,759	0,0699	10	3,83	1,578	0,5354	0,5186	0,0491
SW 40	23	8	7	0,652	0,0627	5	2,5	1,528	0,4472	0,4396	0,1827
SW 60	18	6	5	0,758	0,0582	8	4,16	1,551	0,3175	0,5202	0,4643
Total	233	24	30	0,724	0,0681	170	11,16	1,611	0,2810	0,6000	0,5580
Moyenne	20,81	7,36	6,18	0,738	0,067	15,45	5,07	1,572	0,3940	0,5380	0,3340
Ecart-type	6,35	2,37	2,08	0,087	0,009	7,59	1,36	0,033	0,0730	0,0480	0,1350

Tableau I.6 : Diversité mitochondriale et nucléaire pour chaque population

III.3.3- Structure génétique des populations de la baie de Gdansk

Le Tableau I.7 montre les valeurs de F_{st} par paires de populations estimées à partir du marqueur mitochondrial. La quasi-totalité des tests exacts se sont révélés non significatifs ($p \geq 0,2$). Toutefois, il faut noter une légère différenciation ($0,01802 < p < 0,03604$) entre la population de Sw60 et celles de Hell45, Gn40, Mw30 et Sw40 qui ont toutes été prélevées à une profondeur équivalente (30-40 mètres). Les résultats de l'AMOVA, réalisée entre les trois groupes correspondant aux différentes profondeurs, révèlent l'absence de structure génétique globale sur l'ensemble des populations ($\Phi_{st} = 0,0519$, $p = 0,520$) mais également entre populations au sein des groupes ($\Phi_{sc} = -0,0210$, $p = 0,893$). Toutefois, une légère structure génétique significative se révèle entre les groupes sur l'ensemble de la baie ($\Phi_{ct} = 0,025$, $p = 0,013$).

La structure générale ($F_{it} = 0,352$) observée avec les marqueurs nucléaires est essentiellement due à un fort déficit en hétérozygotes au sein des populations ($F_{is} = 0,345$), la différenciation entre populations étant faible ($F_{st} = 0,010$). Les résultats obtenus pour la différenciation génétique par paire de populations (Tableau I.8) sont globalement peu significatifs. Toutefois, la population Sw60 présente une divergence significative avec les populations Hell 60, Sw5, Sw10, Sw20, Sw30, GN40 et Mw30. L'arbre réalisé à partir des distances génétiques entre populations (Figure I.14) confirme également cette nette différenciation entre Sw60 et les autres populations. On peut noter que les fortes valeurs de F_{st} obtenues entre la population Sw60 et les populations Gn40, Sw30 et Mw30 (respectivement 0,200, 0,232 et 0,150) confirment également le résultat obtenu avec le marqueur mitochondrial.

III.3.4-Assignation des individus

Les résultats d'assignation des individus aux différentes populations de la baie montrent un fort taux d'auto-recrutement (Figure I.15). Selon les populations, l'auto-assignation varie de 30 % (Mw30) à 87 % (Sw60) avec une valeur moyenne de 57 %. A l'exception de Hell85 (50 %), l'auto-recrutement croît avec le niveau de profondeur ($r_2 = 0,65$, $p = 0,039$). Cette tendance est également confirmée avec les seules populations de la Vistule (Sw5 à Sw60, Figure I.16) : de 40 % pour Sw5 à 87 % pour Sw60. A partir de 40 mètres, le pourcentage d'assignation (compris entre 40 et 50 % pour les individus prélevés entre 5 et 30 mètres, passe à 80 % et 87 % pour respectivement Sw40 et Sw60.

	HELL60	HELL45	HEL85	GN40	MW30	SW5	SW10	SW20	SW30	SW40	SW60
HELL60	0.00000										
HELL45	-0.01739	0.00000									
HEL85	0.05796	0.15176	0.00000								
GN40	-0.02997	-0.03163	0.13006	0.00000							
MW30	-0.01728	-0.03199	0.18467	-0.03331	0.00000						
SW5	-0.03117	-0.00370	0.01170	-0.02231	0.00718	0.00000					
SW10	-0.03661	-0.01053	0.01056	-0.02645	-0.00986	-0.04463	0.00000				
SW20	-0.03658	-0.00852	0.01496	-0.02399	-0.00427	-0.04374	-0.04311	0.00000			
SW30	-0.03402	-0.02120	0.06716	-0.03187	-0.02544	-0.02605	-0.03479	-0.03402	0.00000		
SW40	-0.02277	-0.03314	0.14030	-0.03778	-0.03870	-0.01374	-0.01918	-0.01370	-0.02766	0.00000	
SW60	0.05574	0.13104*	-0.10736	0.10549*	0.15653*	0.01697	0.02271	0.02496	0.06734	0.12271*	0.00000

Tableau I.7 : Fst par paires de population obtenues à partir des données mitochondriales. Les valeurs en gras avec * sont significatives.

En outre, une part moindre des individus issus des populations Sw30, Sw40 et Sw60 sont recrutés dans les populations Sw5, Sw10 et Sw20 mais aucun individu de ces trois dernières populations ne sont recrutés dans les populations plus profondes (Sw30 à Sw60).

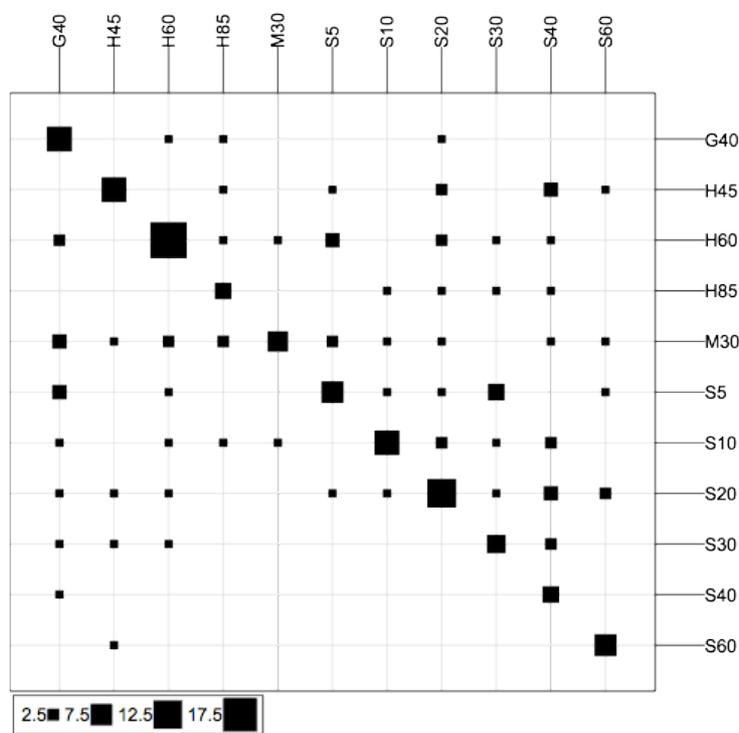


Figure I.15 : Résultats des tests d'auto-assignation des individus des populations de la baie de Gdansk. Pour chaque population, le nombre d'individus assignés à chacune des sources potentielles est représenté par des carrés de taille différente (échelle au bas de la figure).

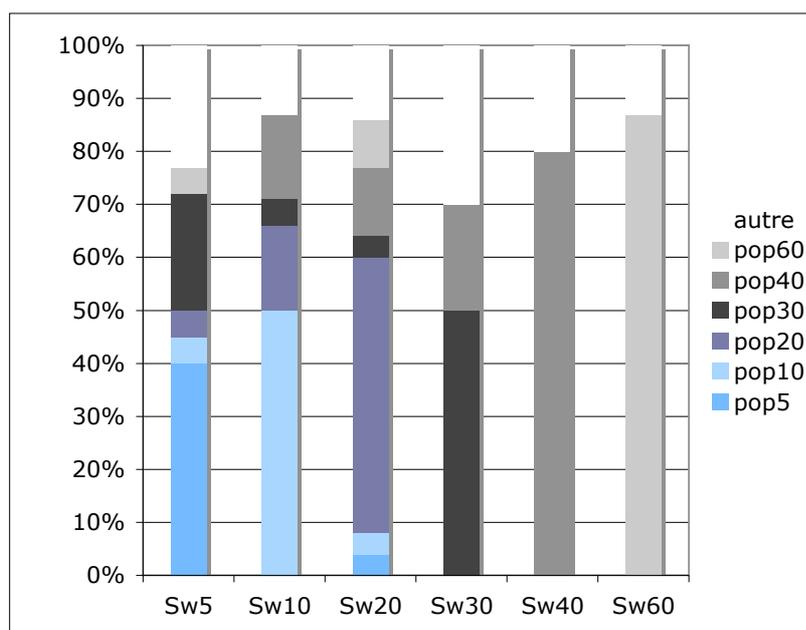


Figure I.16 : Assignment des individus des populations de la Vistule. La catégorie "autre" correspond aux autres populations de la baie de Gdansk

III.4-Discussion

III.4.1-Statut des populations de la mer Baltique

L'analyse de la structure génétique des populations à l'aide du marqueur mitochondrial a permis de mettre en évidence une faible divergence entre les populations de la mer Baltique, d'une part, et celles de l'océan Atlantique et de la mer du nord, d'autre part (seulement 3 mutations et 1 mutation pour respectivement, Murman et Sylt, Figure I.12). Parallèlement, elle a révélé un haplotype spécifique des populations du sud de la mer Baltique (H8). Ces résultats sont en accord avec ceux obtenus précédemment par Strelkov en 2007 et Nikula en 2008, et cet haplotype H8 correspond sans doute à un haplotype issu de l'hybridation entre les génomes Pacifique et Atlantique. Concernant la population de Mecklenburg Bight, celle-ci se trouvant plus au coeur de la zone de transition entre la mer du Nord et la mer Baltique (détroit de Kattegat), il est probable que l'apport régulier de migrants issus de la mer du Nord et de l'océan Atlantique d'une part, et la plus forte salinité de cette zone, d'autre part, aient empêché la sélection de l'haplotype H8.

III.4.2- Structure génétique des populations de la Baie de Gdansk

III.4.2.1- Importance de la thermo-halocline

L'étude de la structure des populations de la baie de Gdansk à l'aide de marqueurs mitochondriaux et nucléaires, a permis de mettre en évidence une structuration des populations en fonction de la profondeur d'échantillonnage.

Bien que les données obtenues à partir du marqueur mitochondrial n'aient pas révélé de structure génétique forte (seulement quatre tests exacts de différenciation de populations deux à deux s'avèrent significatifs, $0,01802 < p < 0,03604$, Tableau I.7), elles ont tout de même permis de différencier la population Sw60 prélevée à 60 mètres de profondeur, des populations prélevées entre 30 et 40 mètres (Hell45, Gn40, Mw30 et Sw40).

Les valeurs de F_{st} obtenues avec les marqueurs microsatellites confirment cette

tendance révélant une divergence significative entre la population Sw60 et les populations Sw30, Gn40 et Mw30 (F_{st} respectifs 0,232, 0,200 et 0,150, Tableau I.8). Par ailleurs, une faible mais significative structuration apparaît dans l'AMOVA, entre les trois groupes de populations établis en fonction de la profondeur d'échantillonnage ($\Phi_{ct} = 0,025$, $p = 0,013$).

En outre, les tests d'auto-assignation révèlent un taux d'auto-recrutement de 30 à 87 % selon les populations avec des niveaux différents selon la profondeur d'échantillonnage. Ainsi, les populations prélevées entre 5 m et 30 m de fond présentent un taux d'auto-recrutement moyen de 44% alors que le taux moyen observé au-delà de 40 mètres, est de 72 % (Figure I.15 et I.16). En théorie, les larves sont susceptibles de se disperser sur de grandes distances. Le brassage résultant des échanges entre groupes de larves d'origines différentes, conditionne les niveaux de différenciation et de structure génétiques entre populations (Roughgarden *et al* 1988, Eckman 1996, Gaylord et Gaines 2000). Or le brassage n'est pas toujours effectif et les larves sont susceptibles d'être recrutées localement en raison (i) des mouvements spatio-temporels des masses d'eau qui peuvent influencer leur trajectoire (Gaines *et al* 2003, Siegel *et al* 2003), (ii) des conditions de survie différentes liées à des contraintes environnementales variables dans l'espace et le temps (Gaines et Bertness 1992, Morgan *et al* 2000, Ellien *et al* 2004) et (iii) un succès reproducteur aléatoire (Hedgecock 1994a, 1994b). La présence d'une thermo-halocline dans la Baie de Gdansk à une profondeur comprise entre 40 et 50 m (Cyberska 1998, Rheinheimer 1995, Rumohr *et al* 1996, Hummel *et al* 2000) pourrait représenter une contrainte environnementale suffisante pour limiter les échanges larvaires et ainsi expliquer les différents niveaux d'auto-recrutement entre les populations situées en deçà et au delà de la thermo-halocline.

III.4.2.2- Structuration des populations à l'embouchure de la Vistule

Les populations échantillonnées dans le panache de la Vistule sont soumises à un gradient environnemental: les populations les plus proches de l'embouchure de la Vistule (Sw5 et Sw10) sont directement confrontées aux effluents de ce fleuve (Figure I.11), se trouvant ainsi dans des eaux peu salées et enrichies en métaux lourds. Au fur et à mesure que l'on s'éloigne de cette zone (au-delà de Sw20), les populations sont de moins en moins soumises aux apports anthropiques mais sont alors confrontées à une salinité plus élevée. Au-delà de la thermo-halocline (40 mètres), les populations qui sont plus éloignées de l'embouchure de la Vistule sont dans un milieu plus salin et surtout plus froid mais sont, à priori peu affectées par ses rejets pollués. Dans cette zone toutefois, les organismes restent confrontés à une pollution importante due à la remise en suspension de nombreux éléments (notamment métalliques) à partir du sédiment par le biais des flux benthiques (Janas 2004, Sokolowski *et al* 2007). En outre, les espèces vivant dans les parties les plus profondes de la

baie de Gdansk (et plus généralement de la mer Baltique) sont également soumises à un déficit en oxygène dissous, cette hypoxie (voire anoxie sur certains sites), résultat d'une forte eutrophisation dû à l'anthropisation massive (Laine 2003), s'accompagnant généralement d'une production en sulfure d'hydrogène (Janas 2004).

Globalement, les populations de notre étude ont toutes une diversité génétique élevée, révélée aussi bien par le marqueur mitochondrial ($0,546 < \mu < 0,854$) que les marqueurs microsatellites ($0,4396 < Hs < 0,5862$). La structure globale de ces populations ($F_{it} = 0,333$) est majoritairement due à un déficit en hétérozygotes au sein de ces populations ($F_{is} = 0,322$) et à une faible différenciation entre populations ($F_{st} = 0,017$).

On remarque cependant que la population Sw60 semble se distinguer des autres populations moins profondes. En effet, exception faite de la population Sw40 pour laquelle l'effectif est très faible ($N = 5$), il existe une différenciation génétique entre la population Sw60 et l'ensemble des autres populations de la Vistule ($0,084 < F_{st} < 0,232$). Cette distinction de la population Sw60 est renforcée à la fois par l'arbre établi à partir des distances génétiques mais aussi par la forte auto-assignation (87 %) mesurée au sein de cette population. La population Sw60 présente en outre une forte diversité génétique, ceci au niveau mitochondrial et nucléaire ($\mu = 0,0582$ et $Hs = 0,5202$). Ainsi la population Sw60 est polymorphe, génétiquement structurée, et les flux de gènes la reliant aux autres populations de la Vistule et plus largement à l'ensemble des populations de la baie de Gdansk, sont relativement limités. Toutes les conditions - un flux de gènes réduit et la fidélité à un habitat favorisé par une hétérogénéité spatiale du milieu (Hasting 1983) - sont ainsi réunies pour favoriser l'installation d'un processus d'adaptation locale permettant de maintenir des génotypes adaptés localement (et ainsi des phénotypes) via une sélection naturelle différentielle.

Il est toutefois surprenant de constater que les populations qui se trouvent juste à l'embouchure de la Vistule et donc à priori fortement impactées par la pollution ne présentent aucune structure génétique marquée. Le flux de gènes existant entre ces populations et plus généralement avec les autres populations de la baie de Gdansk peuvent expliquer cette absence de signature moléculaire.

Le flux de gènes n'est pas forcément un obstacle à l'adaptation locale même si la sélection naturelle divergente reste la force dominante dans le phénomène d'adaptation (Kawecki et Ebert 2004). L'adaptation locale peut en effet être conservée, malgré l'existence d'un flux de gènes dès lors que la reproduction des résidents avec les immigrants reste limitée. En effet, l'apport de gènes extérieurs dans une population adaptée localement, peut contrer cette sélection et ainsi limiter l'adaptation locale. Dans le cas où les immigrants sont plus rares que les locaux, la reproduction préférentielle des immigrants avec les résidents pourra affecter la fitness des hybrides issus de ce croisement selon deux cas de figure: (i) les performances de l'hybride seront inférieures à la moyenne des performances des parents ou (ii) les hybrides seront avantagés (cas d'hétérosis, Ebert *et al* 2002).

La diminution de la fitness peut résulter de (i) la dilution des gènes associés à l'adaptation locale, d'autant plus probable si ils sont récessifs ou que l'hétérozygote est désavantagé, ou de (ii) la rupture des complexes de gènes co-adaptés (phénomènes d'épistasie), susceptible également de réduire la fitness des génotypes recombinants, à la deuxième et troisième génération d'hybrides (*outbreeding depression*) générant alors une sélection contre les gènes immigrants pour des raisons indépendantes de l'adaptation locale (Burton 1990, Armbruster *et al* 1997, Fenster et Galloway 2000).

Les échanges de migrants entre les populations de la Vistule sont asymétriques (Figure 1.16). Ainsi, le flux de gène existe des populations situées au delà de 30m vers les populations situées à 5, 10 ou 20 mètres (en moyenne 24 % des migrants) mais n'est pas observé dans le sens inverse. Les populations situées de part et d'autre des 30 mètres de profondeur pourraient être organisées en structure source-puits (Pulliam 1988), les populations Sw30, Sw4 et Sw60 envoyant l'excès de migrants dans des populations périphériques (Sw5, 10 et 20). Toutefois, la forte diversité génétique observée dans ces populations ne semble pas corroborer cette hypothèse.

La différenciation génétique observée entre la population Sw60 et le reste des populations de la Vistule, mais aussi le flux de gènes asymétrique existant entre les populations Sw30, Sw40, Sw60 et les populations moins profondes, montrent l'existence d'une hétérogénéité génétique au sein de ces populations. Bien que proches les unes des autres, elles sont toutefois localisées le long d'un gradient environnemental complexe qui pourrait favoriser une structure génétique en mosaïque (*chaotic genetic patchiness*). Ce concept, décrit par Johnson et Black en 1982, désigne en effet une hétérogénéité génétique pouvant être plus forte au sein de populations proches (c'est le cas des populations de la Vistule) qu'entre populations plus éloignées (e.g. entre populations de la baie de Gdansk). Fréquemment observé chez les invertébrés marins (Watts *et al* 1990, Hedgecock 1994a, Edmands *et al* 1996, David *et al* 1997), ce phénomène, qui peut être la conséquence d'une sélection naturelle susceptible d'agir avant ou après le recrutement des larves, est étroitement lié aux conditions environnementales dans lesquelles elles vont évoluer (Johnson et Black 1982, 1984, Watts *et al* 1990, Hedgecock 1994a, b).

III.5- Conclusion

Trois résultats importants ressortent de l'étude des populations du sud de la mer Baltique.

- Ces populations, localisées dans une zone hybride et en contact permanent avec une des lignées parentales, sont majoritairement composées de l'haplotype

parental (haplotype H_I de l'Atlantique) avec une plus faible proportion de l'haplotype Baltique (correspondant en fait à une hybridation précédente entre la lignée Pacifique et la lignée Atlantique).

- Dans la baie de Gdansk, la différenciation génétique observée entre les populations situées à 60m de profondeur et les populations situées à 40m ainsi que les résultats d'assignation des individus nous permettent de préciser la profondeur de la thermo-halocline aux alentours de 40 mètres.

- La population située au-delà de la thermocline dans le panache de la Vistule semble présenter un profil d'adaptation locale. Toutefois cette population étant soumise à de multiples stress d'origines naturelle (gradient de salinité, de température) et anthropique (eutrophisation, contamination aux métaux lourds) dont les effets sont potentiellement additifs, il nous est impossible de préciser davantage les processus de sélection naturelle et d'adaptation en cours, par les approches de génétique des populations.

Partie II

*Evaluation des capacités adaptatives
de *Macoma balthica* et approche
moléculaire de la sélection.*

Importance de la sélection dans les processus adaptatifs

Face aux contraintes de l'environnement, les caractères phénotypiques les mieux adaptés dans une population sont favorisés par la sélection naturelle. Toutefois, les pressions environnementales étant variées sur l'ensemble de l'aire de répartition d'une espèce, les différentes actions de la sélection naturelle sont susceptibles de générer des différenciations entre populations, reflet d'une adaptation locale avec émergence des phénotypes présentant les meilleurs fitness (Kawecki et Ebert 2004). Des facteurs anthropiques comme les variations climatiques notamment, sont considérés comme des pressions de sélection susceptibles d'induire des adaptations locales (Thomas 2005, Joshi *et al* 2001, Jump et Penuelas 2005, Huey *et al* 2000). Dans le domaine marin, des phénomènes d'adaptation local du poisson extrémophile (*Poecilia mexicana*) à un milieu enrichi en sulfure d'hydrogène ont par exemple été décrits (Plath *et al* 2007). La compréhension de l'impact de tels changements sur les populations, et plus particulièrement de leur capacité à répondre à ces contraintes, est donc fondamentale dans une démarche de conservation.

Ce potentiel adaptatif des populations est directement lié au niveau de diversité génétique additive disponible pour des caractères adaptatifs, c'est à dire à la variabilité transmissible des gènes contrôlant l'expression des caractères adaptatifs (Soulé 1987, Howe et Brunner 2005). C'est donc sur la part héritable des caractères corrélés avec le succès reproducteur que la sélection agit pour produire des changements adaptatifs, ceci selon trois modalités: (i) la sélection directionnelle qui va favoriser le développement d'un caractère (situé dans la partie extrême de la distribution) au détriment d'un autre, (ii) la sélection stabilisante qui aura tendance à éliminer les formes extrêmes d'un caractère au bénéfice des individus moyens, (iii) la sélection disruptive qui va favoriser les formes les plus extrêmes d'un caractère. La sélection directionnelle fait généralement suite à un changement environnemental majeur (réchauffement global, augmentation des polluants dans l'air, l'eau, maladie affectant une population) et une étude de Frankham et Weber (2000) a par ailleurs permis de mettre en évidence les conséquences que ce processus avait eu sur de nombreuses espèces (modification des taux de reproduction et de croissance, modification du comportement, tolérance aux métaux lourds...). Le processus de sélection disruptive est plus fréquemment observé dans le cas des populations fragmentées. Les contraintes sélectives agissant dans des directions opposées sont susceptibles de générer d'importants niveaux de divergence entre de telles populations (Rueffler *et al* 2006), favorisant l'installation d'adaptation locale, voire à terme d'un isolement reproducteur.

Traditionnellement la recherche de gènes impliqués dans la variation des caractères adaptatifs a été abordée par la cartographie de loci à effet quantitatif (QTL), basée sur des associations statistiques significatives entre marqueurs moléculaires et variation de caractères quantitatifs au sein de pedigrees (Borevitz et Chory 2004), mais plus rarement au sein de populations naturelles (Slate 2005). Néanmoins, de nombreuses études

ont été menées afin d'évaluer le potentiel adaptatif des individus dans un environnement variable (O'Donnel *et al* 2009, Usenko *et al* 2008, Kong *et al* 2005) et le développement de la génomique a aujourd'hui ouvert la voie vers l'étude des bases moléculaires de l'adaptation.

Avec l'arrivée de la PCR quantitative dans les années 1980, de nombreuses fonctions géniques ont pu être corrélées à des variations des paramètres environnementaux, le principe étant de mesurer l'expression différentielle de gènes candidats¹⁵ chez des organismes maintenus dans des conditions expérimentales variables et contrôlées. La PCR quantitative s'avère être la méthode la plus appropriée pour quantifier l'expression de gènes chez les espèces dans lesquelles des gènes candidats sont d'ores et déjà identifiés. Dans les années 1990, une autre démarche, appelée banque soustractive (SSH¹⁶) a permis d'élargir le nombre de gènes analysés pour un paramètre environnemental donné, grâce à l'analyse ciblée des régions différentiellement exprimées. Bien que largement répandues, ces études sont souvent le reflet d'une réponse individuelle dans des conditions d'expérimentation parfois très éloignées de celles que rencontre l'organisme dans son milieu naturel. Elles consistent la plupart du temps à évaluer la génotoxicité de certaines molécules sur les fonctions physiologiques des organismes ou bien à différents stades de développement, de sorte que le potentiel adaptatif et évolutif de la population dans l'écosystème impacté reste inconnu. De ce fait, les approches menées sur des populations naturelles dans leur écosystème et non pas en milieu contrôlé, apparaissent comme particulièrement pertinentes.

Le modèle biologique *Macoma balthica* a fait l'objet de nombreuses analyses éco-physiologiques et éco-toxicologiques. Ces travaux ont mis en évidence chez cette espèce, une forte sensibilité à l'augmentation de température (Jansen *et al* 2007) et une grande tolérance aux variations de salinité (Väinölä et Varvio 1989) ainsi qu'aux environnements pollués notamment par les métaux lourds (Sokolowski *et al* 1998, 2002, 2007, Hummel *et al* 2000).

En l'absence de données génomiques disponibles chez cette espèce et désireux de mener nos investigations sur des populations en conditions naturelles, nous avons été confronté à un choix méthodologique pour inférer les bases moléculaires de l'adaptation. La réalisation d'une banque soustractive SSH sur des individus issus des trois sites a été envisagée puis rejetée du fait (i) de la difficulté de prélever des individus dans les mêmes conditions physiologiques (stade de développement, maturité) sur des sites aussi éloignés géographiquement, (ii) du risque de perdre des transcrits rares du fait de la normalisation de la banque d'ADNc.

Notre choix s'est finalement porté sur les nouvelles technologies de séquençage à haut débit permettant l'analyse en parallèle de la quasi totalité des gènes exprimés dans un

¹⁵ Gènes candidats: tout gène susceptible de déterminer des différences dans les caractères entre organismes, d'être impliqué dans des fonctions physiologiques.

¹⁶ SSH: *Subtractive Suppressive Hybridization* (Banque soustractive); permet de révéler uniquement les gènes différentiellement exprimés par comparaison de transcriptomes.

tissu, voire un organisme entier. Un pyroséquençage 454 a ainsi été mené sur *Macoma balthica*, avec un double objectif:

- Identifier des gènes candidats chez *Macoma balthica*, notamment des gènes impliqués dans la réponse générale au stress ;
- Comparer l'ensemble de transcrits (ARNm) issus de 3 populations naturelles prélevées sur trois sites soumis à des stress environnementaux différents.

Chapitre I

*Les nouvelles méthodes de séquençage
haut débit.*



Cas particulier du pyroséquençage 454

I.1- Description des principales technologies de séquençage massif ou NGS (*Next Generation Sequencing* ; séquençage de nouvelle génération)

De nouvelles méthodes d'analyses du génome sont apparues sur le marché depuis environ 4 ans (Metzker 2010) après 20 années de suprématie de la méthode de Sanger pour le séquençage automatique. Ces technologies dites NGS¹⁷, apparues en nombre sur le marché, permettent de séquencer massivement et de manière consécutive des milliers de séquences courtes (300 à 800 pb) issues de banques d'ADN (génomique, ADNc...). Nous ne décrivons ici que les méthodes les plus couramment utilisées (454/Roche, Solexa/Illumina, APG/Life (SOLiD)).

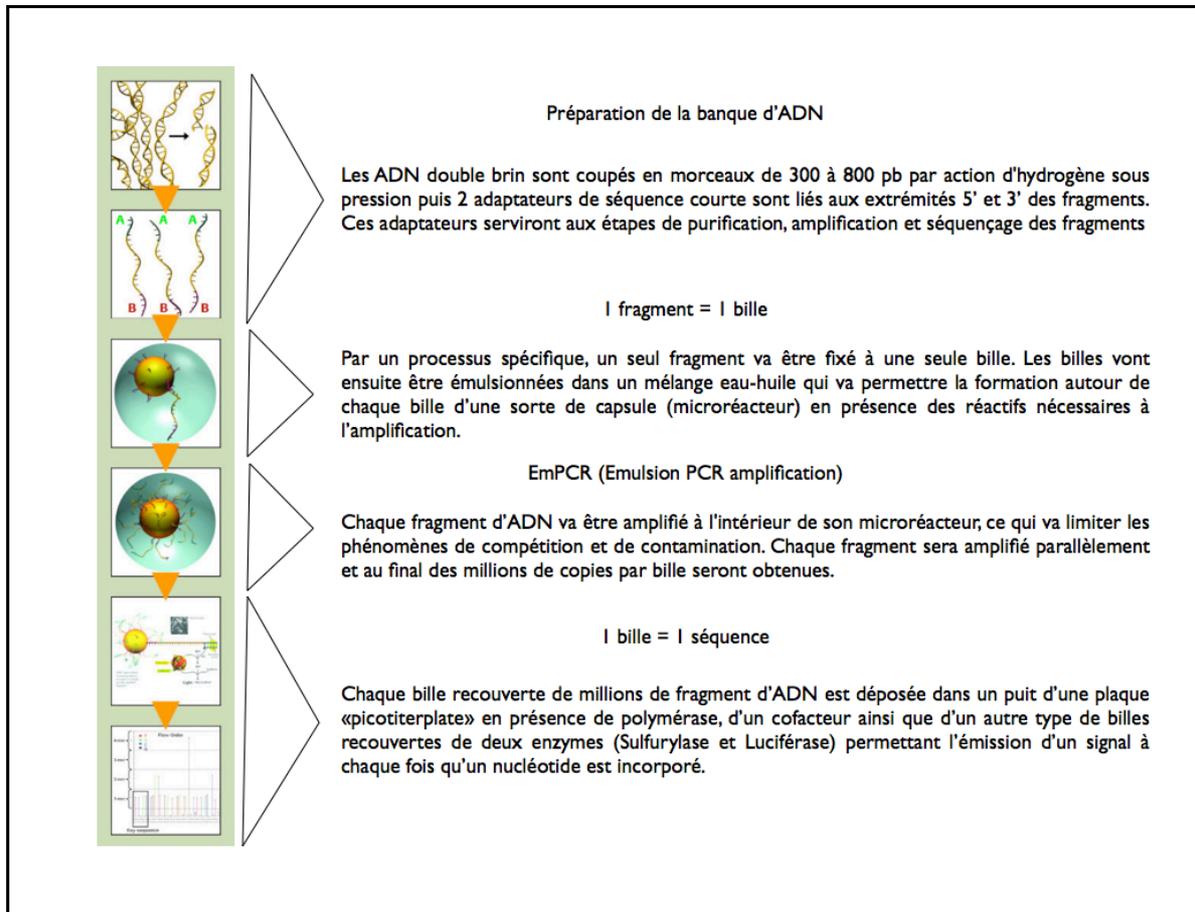
Leur point fort est l'absence de clonage, les échantillons étant directement et individuellement séquencés après ligation ou immobilisation sur une surface solide ou un support. Les techniques disponibles se distinguent essentiellement par les méthodes de préparation des échantillons et de séquençage, cette dernière étape, elle-même associée à la méthode d'acquisition du signal par imagerie (protocoles détaillés pour chaque technologie en Annexe A).

La préparation des échantillons passe généralement par une étape d'amplification, soit par *emulsion* PCR (EmPCR) soit par amplification sur phase solide (Annexe A).

La technique de séquençage et d'imagerie reste en revanche spécifique à l'appareil et la société qui le commercialise. Trois types de procédés sont ainsi utilisés selon les appareils considérés: (i) la méthode dite CRT (*Cyclic Reversible Termination*) utilisée par le Solexa (Illumina), consiste en une incorporation simultanée des 4 nucléotides fluorescents dont l'émission de signal se fait par déprotection (Annexe A), (ii) la méthode SBL (*Sequencing By Ligation*) spécifique du système SOLiD (*Applied Biotechnologies*) dans lequel la fluorescence des nucléotides (incorporés simultanément) est révélée par clivage (Annexe A) et (iii) la méthode du pyroséquençage caractéristique du 454 (Roche) qui permet l'ajout d'un nucléotide unique non marqué puis l'émission d'un signal par une cascade de réactions enzymatiques.

Dans cette thèse, nous avons choisi la méthode dite «454», qui présente l'avantage d'obtenir des séquences de taille 3 à 6 fois supérieure à celle des fragments obtenus avec les deux autres technologies et qui, par ailleurs, semblait être la technique la plus appropriée pour l'étude *de novo* des génomes (Metzker 2010). Cette technologie est plus précisément décrite dans le matériel et méthode ci-après.

¹⁷ NGS: méthodes de séquençage de nouvelle génération en comparaison avec la méthode de Sanger jusque là utilisée et décrite comme la méthode de première génération



Encadré 6 : Technologie du pyroséquençage 454

I.2- La technique «454» ou pyroséquençage

La technologie 454 *Life Sciences* (*Four Five Four* pour *Facile Functional Findings*, Margulies *et al* 2005) a été conçue par J. Rothberg au sein de la compagnie 454 *Life Sciences* (Branford, USA) spécialisée dans la fabrication de séquenceur de gènes à haut débit (cette société intervient notamment dans le projet de séquençage du génome de l'homme de Néanderthal avec le Max Plant Institut). Cette approche innovante repose sur l'intégration de plusieurs technologies de pointe: le pyroséquençage, les plaques en fibre optique picotitrées qui contiennent 1,6 millions de puits, la PCR en émulsion dans des micro-réacteurs permettant 300 000 réactions de PCR en parallèle couplée à des procédés informatiques pour l'acquisition, le traitement et l'analyse des images. Ainsi, le pyroséquençage permet de traiter avec un seul instrument plus de 20 millions de bases nucléotidiques par cycle de 4 heures et de générer ainsi, avec plus de 99 % d'exactitude, autant de données que 100 séquenceurs capillaires haut-débit.

La technique «454» consiste à séquencer en parallèle des fragments d'ADN liés à des billes, chaque bille étant associée à un fragment d'ADN unique (1 fragment = 1 bille = 1 séquence). Ce processus se décompose en 4 étapes majeures (Encadré 6):

- (i) la préparation de la banque d'ADN consistant à liguer des adaptateurs aux fragments d'ADN.
- (ii) la ligation de chaque fragment à une bille (*DNA capture beads*) afin de les isoler (1 fragment = 1 bille).
- (iii) l'amplification de chaque fragment d'ADN.
- (iv) le séquençage de chaque fragment (1 bille = 1 séquence).

Dans le cadre de cette étude, nous avons fait appel à la société *Beckman Coulter Genomics* qui a réalisé le pyroséquençage sur un appareil 454 *Life sciences GS FLX Titanium* de chez Roche.

I.3- Développement de la banque d'ADN génomique

I.3.1- Echantillonnage et extraction des ARN

Trois populations de *Macoma balthica* (10 individus par population) ont été prélevées en Janvier 2008 sur les sites des baies de Marennes-Oléron, baie de Somme et baie de Gdansk (Figure II.1). Afin de garantir la conservation de leur ARN, les individus ont été

conservés entiers, dès leur collecte dans du *RNAlater* (Sigma) afin de perméabiliser les tissus et de stabiliser les ARN cellulaires *in situ*.

L'ARN total de 10 individus par population a été extrait selon la méthode du Trizol (Annexe A). Tout le matériel destiné à l'extraction a été décontaminé (destruction des RNAses avec une solution d'H₂O₂ à 3%). Pour la préparation des solutions, de l'eau DEPC (DiEthyl PyroCarbonate) également dépourvue de RNases a été préparée et utilisée à la place de l'eau milliQ classique.

Une fois extraits, les ARN totaux ont été dosés au nanodrop à deux longueurs d'ondes, 260 nm (acides nucléiques) et 280 nm (longueur d'onde d'absorption des protéines). Les rapports DO₂₆₀/DO₂₈₀ ont été calculés, et seuls les ARN pour lequel le ratio était compris entre 1,8 et 2 ont été conservés (ceci permettant à priori d'éliminer les ARN les moins purs). La qualité des ARN obtenus a été vérifiée par électrophorèse sur gel d'agarose 1% (par observation des bandes des ARN ribosomiaux 18S et 28S).

I.3.2- Préparation de la banque d'ADNc et pyroséquencage

454

Les 10 ARN extraits pour chaque population ont été groupés avant l'étape de reverse-transcription (préparation des ADNc). Cette étape a été réalisée avec le kit de synthèse d'ADNc, *MINT cDNA synthesis kit* (Evrogen). L'utilisation de ce kit permet de travailler directement sur les ARN totaux grâce à l'utilisation d'amorces terminées par une queue polyA capables de sélectionner uniquement les ARNm et de s'affranchir des ARN ribosomiaux et de transfert. Par ailleurs, la dernière étape du protocole permet d'obtenir des ADNc double brin, plus stables. Une fois synthétisés, la qualité des ADNc a été vérifiée sur un gel d'agarose à 1,2% dans du tampon TAE 1X. Ils ont ensuite été purifiés avec le kit *NucleoSpin® Extract II* et dosés au Nanodrop.

Les ADNc de chaque population (environ 20 µg) ont ensuite été envoyés à la société *Beckman Coulter Genomics* qui a procédé au pyroséquencage selon la méthode décrite précédemment. Lors de la préparation de la banque d'ADNc, chaque population a été identifiée et marquée (étiquette A pour les échantillons de Marennes-Oléron, étiquette S pour les échantillons de Somme et étiquette G pour ceux de Gdansk). Les ADNc ont ensuite été groupés pour le séquençage. L'étiquette unique portée par les ADNc de chaque population a permis par la suite d'identifier les séquences correspondant à chacune d'entre elle.

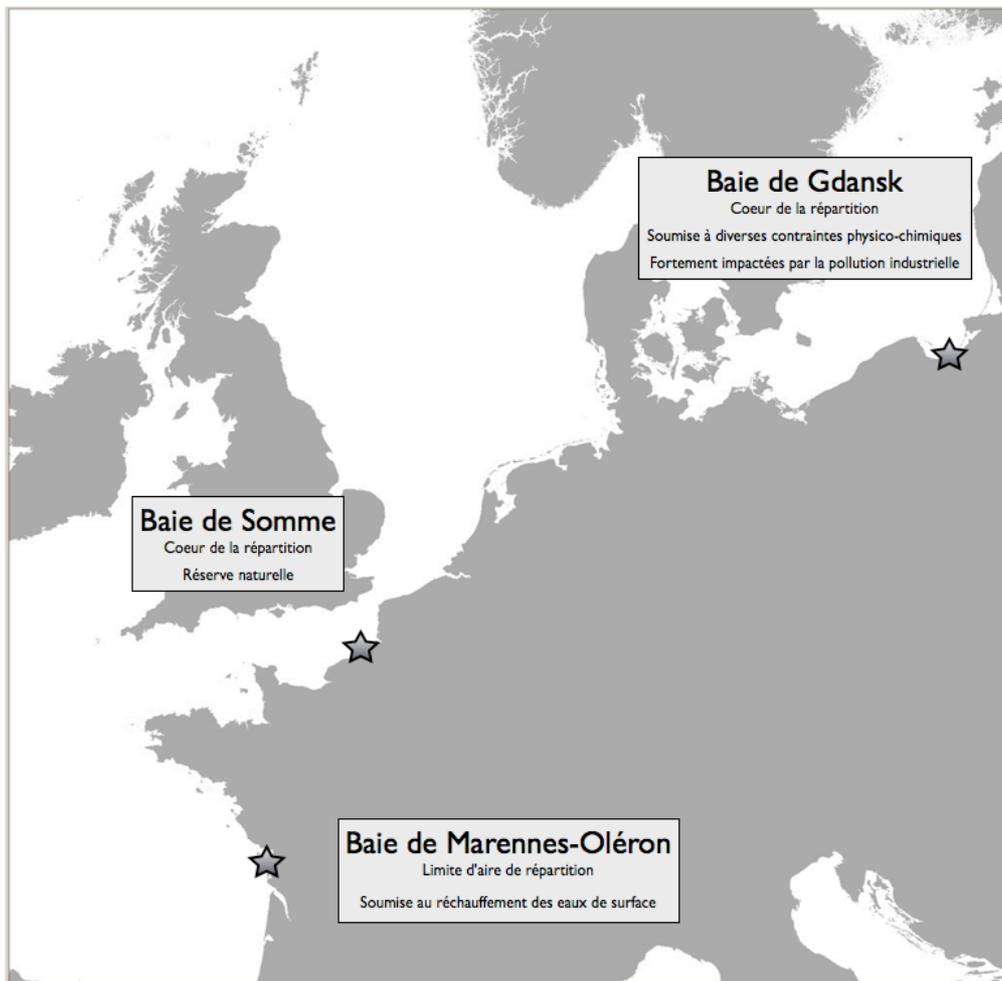


Figure II.1 : Sites d'échantillonnage pour l'analyse génomique

I.4-Analyses bioinformatiques des séquences obtenues

I.4.1-Tri et assemblage des séquences

Les séquences obtenues par pyroséquençage ont d'abord été purgées (i) des séquences amorces utilisées lors de la reverse transcription des ARNm en ADNc, (ii) des séquences inférieures à 100 paires de base à l'aide du script *seqclean* disponible à l'adresse suivante <http://compbio.dfci.harvard.edu/tgi/software/>, et (iii) des séquences correspondant à de l'ADN ribosomal (équivalent à 76000 séquences) à l'aide d'un script développé par Khalid Belkir (UMR 5554, Montpellier II). Pour chaque population, les séquences ont été assemblées, avec le logiciel Mira (<http://sorceforge.net/projects/mira-assembler/>) afin d'obtenir des *contigs*¹⁸ (Encadré 7). Il est en effet impossible, malgré les technologies actuelles, de séquencer en une fois des molécules d'ADN aussi longues que des chromosomes. La stratégie consiste donc à effectuer un grand nombre de séquençages de courts fragments afin de reconstituer ces séquences. La quantité de séquençage réalisé est généralement plusieurs fois supérieure à la taille du génome considéré, ces séquençages redondants (jusqu'à 8 à 10 fois pour les très grands génomes) permettant ainsi de raccorder les séquences les unes aux autres au cours de l'étape dite d'assemblage et de s'assurer de la qualité du résultat de chaque lecture.

Les grands génomes sont caractérisés par des séquences répétées en plusieurs endroits du génome, qui sont susceptibles de provoquer des assemblages erronés entre deux séquences issues de régions distantes du génome. Ces fragments répétés sont donc «masqués» au cours de l'assemblage (notés lrc par Mira). Une fois cet assemblage réalisé, l'ensemble des *contig* obtenu a été débarrassé des *contigs* notés «lrc» considérés comme peu fiables par Mira. Cette étape a été réalisée à l'aide d'un script écrit par Audrey Rohfrischt (Post-doctorante LIENSS, ANR Hi-Flo).

I.4.2-Annotation des séquences

Le projet *Gene Ontology* (*The Gene Ontology Consortium* 2000) a été mis en place pour développer un vocabulaire (termes GO) destiné à décrire les produits géniques en termes de processus biologique, de composition cellulaire et de fonction moléculaire indépendamment de l'espèce considérée.

¹⁸ Contig: Séquence continue d'ADN générée par l'assemblage de fragments de séquences chevauchantes d'un chromosome

Du séquençage à l'assemblage construction des contigs

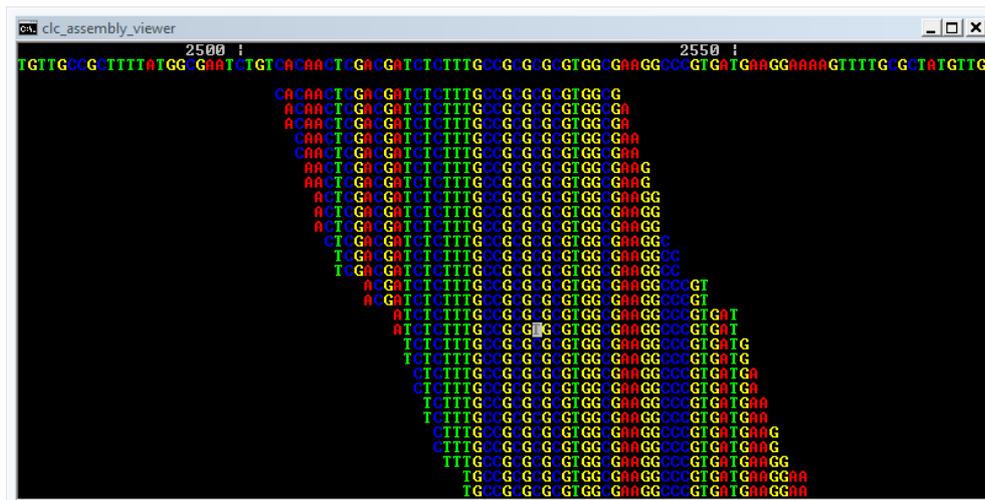
Etape 1 : fractionner le fragment à séquencer en sous-fragments

Etape 2 : effectuer un nombre de séquençage assez important de manière à ce que ces séquençages mis bout à bout, représentent 10 fois la longueur de la séquence du fragment initial ; chaque base du fragment initial doit apparaître dans 10 lectures en moyenne

Etape 3 : comparaison des séquences c'est à dire alignement des parties qui se recouvrent partiellement (parties chevauchantes)

Etape 4 : assemblage des séquences chevauchantes en enchaînements plus grands appelés *contig*

Le *coverage* représente le nombre de séquences qui contribuent à l'information à une position donnée du *contig*; le niveau de *coverage* est relatif au nombre total de séquences dans le *contig*.



Encadré 7 : Construction des *contigs*

C'est au cours de l'étape d'annotation que les *contigs* obtenus par assemblage des séquences sont ainsi assignés à des comportements cellulaires. Trois catégories structurelles sont définies :

- la composition cellulaire qui peut être une structure anatomique ou un groupe de produits de gènes (ribosome, protéasome ...)
- la fonction moléculaire qui décrit les activités qui interviennent au niveau moléculaire mais sans préciser ni où ni quand l'action prend place ;
- le processus biologique qui se définit par une série d'événements accomplis par un ou plusieurs assemblages de fonction moléculaire. En comparaison avec un processus moléculaire, le processus biologique est composé de plusieurs étapes.

L'attribution d'une fonction génique à chaque séquence a été réalisée avec le logiciel Blast2go (Conesa *et al* 2005) selon 3 étapes successives (Figure II.2) :

La première étape consiste à comparer les séquences obtenues à une base de données accessible en ligne sur le site de NCBI à l'adresse suivante <http://www.ncbi.nih.gov/BLAST>. Cet outil (*Basic Local Alignment Search Tool*) retrouve des séquences similaires et calcule une valeur statistiquement significative d'homologie (*e-value*). L'algorithme blastx permet de comparer chaque séquence considérée à une base de données de séquences protéiques, ceci en utilisant tous les cadres de lecture. De nombreuses séquences homologues sont ainsi assignées aux séquences obtenues qui peuvent être triées par le logiciel Blast2go en déterminant une valeur significative seuil ainsi qu'un nombre maximum de séquences homologues à attribuer. Dans notre étude, les valeurs par défaut ont été appliquées, à savoir *Blast Expect value* = 1.10^{-3} et *Number of Blast Hits* = 20.

La seconde étape est un processus de mapping qui permet de retrouver tous les termes GO (*Gene Ontology*) associés à ces séquences homologues. Quatre *mapping* sont réalisés en parallèle à partir des numéros d'accèsion et des identifiants des bases de données (GI, *GenBank Identifier*) issus du résultat du *Blast*, afin d'attribuer aux séquences homologues un nom de gène et une fonction protéique. Ces étapes sont réalisées à l'aide de plusieurs bases de données disponibles dans NCBI ou directement dans la base de données GO.

Enfin, une troisième étape d'annotation proprement dite permet de sélectionner, parmi cet ensemble de termes GO, ceux qui sont les plus appropriés à nos séquences, par application de plusieurs filtres sélectionnant uniquement les séquences homologues dont le niveau de similarité est suffisant. Dans notre étude, les valeurs par défaut ont été appliquées, à savoir *Pre-e-value-Hit-Filter* = 6, *Pre-Similarity-Hit-filter* = 30, *Annotation Cut-off* = 55 et *GO-weight* = 5.

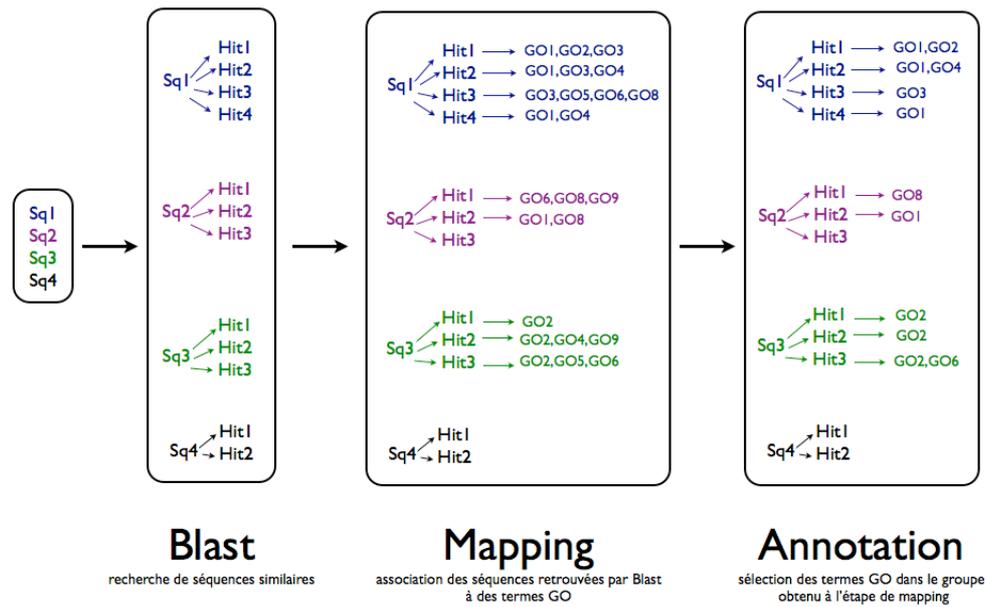


Figure II.2 : Etapes successives réalisées dans Blast2go pour l'annotation des séquences

I.4.3-Comparaison des transcriptomes

L'ensemble des séquences obtenues pour chaque population a été comparé par un test exact de Fisher avec le logiciel GOSSIP (Blüthgen *et al* 2005) disponible dans Blast2go (Conesa *et al* 2005). Ce logiciel permet de tester l'enrichissement d'un set de séquences en termes GO vis à vis d'un set de référence. Nous avons ainsi comparé les sets de séquence de chaque population deux à deux (A *versus* G, A *versus* S, G *versus* S) afin d'évaluer un différentiel d'expression entre population. Ces différents tests ont été réalisés uniquement sur les séquences annotées avec le logiciel Blast2go (Conesa *et al* 2005). Une correction multiple a été réalisée sur les tests selon la méthode décrite par Blüthgen *et al* (2005).

Chapitre 2

*Resultats du pyroséquençage et mise en
évidence de gènes candidats*

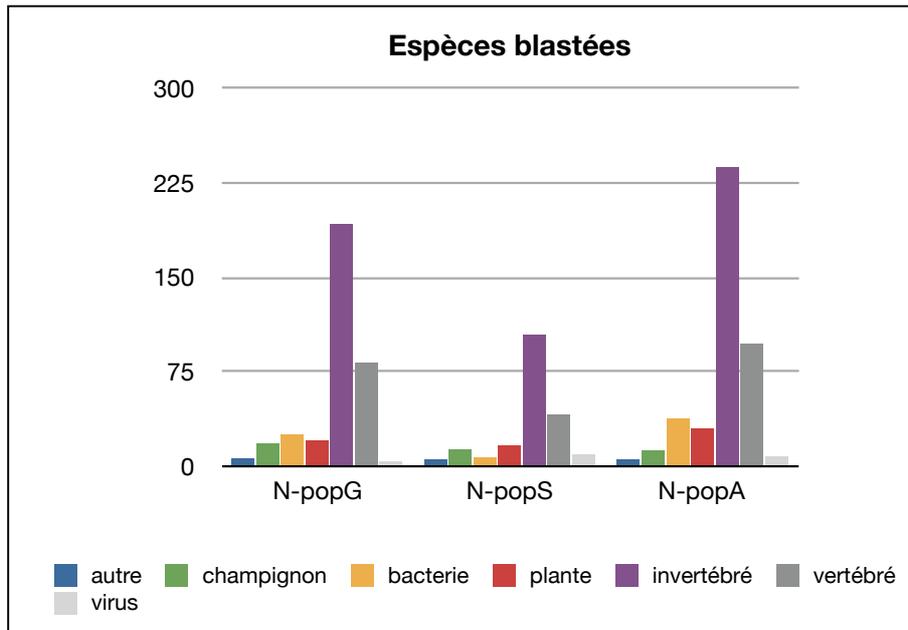


Figure II.3 : Taxons *blastés* dans NCBI pour les populations G, A et S.

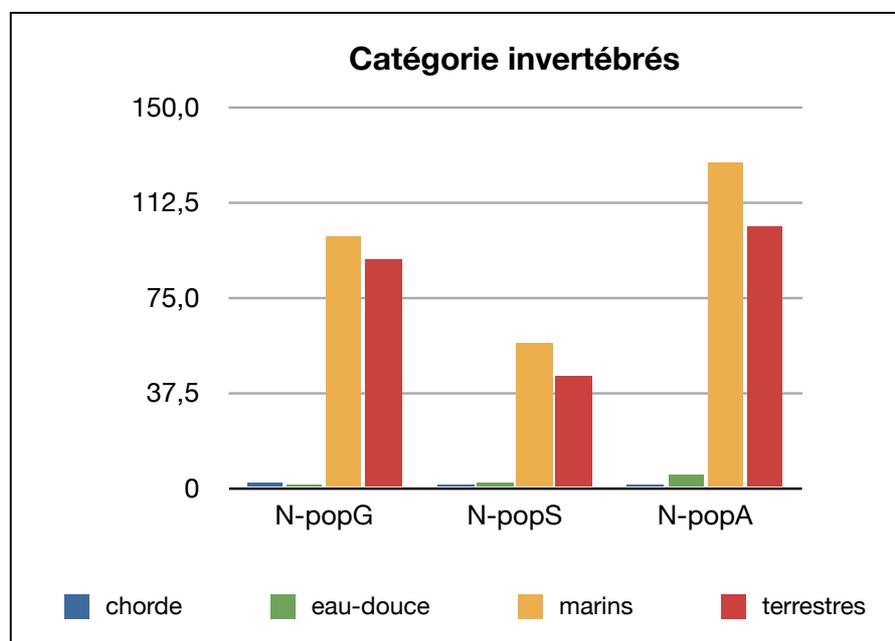


Figure II.4 : Catégories d'invertébrés *blastés* pour chaque population.

La première étape de notre analyse moléculaire de l'adaptation de *Macoma balthica* a consisté à rechercher des gènes candidats, notamment impliqués dans la réponse générale au stress, grâce au pyroséquençage du transcriptome.

II.1- Séquençage et profil des ADNc exprimés

Le pyroséquençage du transcriptome de *Macoma balthica* a permis d'obtenir 871 962 séquences dont 44 % appartiennent à la population de la baie de Marennes Oléron (population A), 31 % à la population de la baie de Gdansk (population G) et 25 % à la population de la baie de Somme (population S) (Tableau II.1). 30221 contigs et singletons d'une longueur variant de 318 à 4541 paires de bases ont été obtenus, les populations A, G et S ayant respectivement 10568, 8939 et 10714 séquences. La recherche d'homologie à partir de ces séquences, sur l'ensemble des données disponibles dans Genbank, a permis d'associer environ 30 % de notre set de séquences à des gènes orthologues¹⁹.

	Population A	Population G	Population S	Total
Séquences totales	383271	270880	217811	871962
Après nettoyage	209180	181766	176521	567467
Contigs	21124	17682	19498	58304
Contigs (sans «lrc») et singletons	10568	8939	10714	30221
Blast	3631	3394	2834	9859
Mapping	3584	2932	2811	9327
Annotation	2351	1831	1373	5555

Tableau II.1 : Efficacité du pyroséquençage

Sur la totalité des séquences homologues, plusieurs phylums ont été reconnus avec toutefois une majorité pour les invertébrés (environ 50 %) et les vertébrés (environ 20 %), les séquences présentant une homologie avec des séquences issues de végétaux, bactéries et virus étant nettement moins nombreuses (respectivement 8 %, 7 % et 3 %) (Figure II.3).

¹⁹ Orthologues : gènes qui occupent la même région du génome chez deux espèces différentes

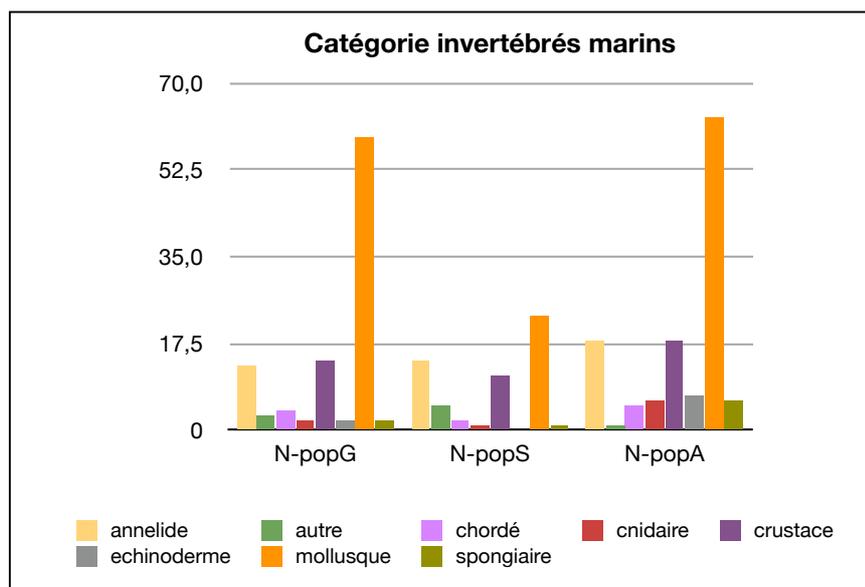


Figure II.5 : Catégories d'invertébrés marins *blastées* pour chaque population.

En outre, les séquences homologues à des séquences d'invertébrés, concernent majoritairement des invertébrés marins (environ 50 % quelle que soit la population considérée) et plus particulièrement les séquences de mollusques (de 40 % à 60 % selon les populations, Figures II.4 et II.5).

II.2- Annotation fonctionnelle des transcriptomes: *Gene Ontology*

L'annotation des transcriptomes a révélé 18 catégories de processus biologiques pour les populations A et S, et 19 pour la population G. Les processus majoritaires dans les trois populations sont les processus métaboliques, cellulaires et la régulation biologique (Annexe B).

Au niveau de la fonction moléculaire, le nombre de catégories diffère selon les populations. Ainsi, les populations A et G présentent respectivement 60 et 52 catégories alors que la population S n'en a que 27 (Annexe B). Parmi ces catégories, de nombreuses unités géniques sont liées à l'activité oxydoréductrice dans les populations A et G.

II.3- Activités peptidiques mises en évidence

Une analyse plus précise de la fonction moléculaire attribuée à chaque séquence nous a permis d'isoler de nombreuses protéines impliquées dans la réponse au stress. Ainsi sur les 5555 séquences annotées sur l'ensemble des populations, 940 correspondent à des protéines susceptibles d'être impliquées dans des mécanismes de réponse au stress, les plus fréquentes étant des protéases (30 %) intervenant dans la régulation du stress oxydatif. En plus des séquences codant pour des protéines ayant un rôle dans la réponse générale au stress (7 %), dans la réponse à une substance chimique (2 %), dans les dommages à l'ADN (3 %) ou dans l'apoptose (2 %), neuf *Heat Shock Protein* sont observées ainsi que des protéines intervenant dans le catabolisme telles que des séquences codant pour le protéasome, les facteurs de régulation associés et de nombreuses ubiquitines (29 différentes) (Annexe B)

La réponse inflammatoire est représentée par 8 % des séquences, avec notamment de nombreuses cathepsines (9 différentes) ainsi que leurs précurseurs. Plus de 11 catégories différentes de lectines sont retrouvées dans les séquences liées à la réponse immunitaire (13 %). Parmi les protéines intervenant dans les phénomènes d'oxydoréduction, on peut compter 13 ferritines, 12 glutathion peroxydases mais seulement 7 séquences codant pour des *Super Oxyde Dismutase* (SOD), 1 seule pour les catalases (CAT).

II.4- Discussion et conclusion

II.4.1- Annotation du transcriptome de *Macoma balthica*

Le pyroséquencage 454 nous a permis d'obtenir 30221 *contig* et singletons générant plus de 5000 séquences annotées. Toutefois, la majorité des fragments géniques séquencés dans notre étude n'ont pas montré d'homologie avec des gènes connus, et seul 30 % des *contig* et singletons ont pu être associés à un gène répertorié. Ce résultat est toutefois cohérent avec ceux généralement obtenus chez les mollusques (Boutet *et al* 2004, Huvet *et al* 2004, Bultelle *et al* 2002), principalement du fait de la divergence élevée existant entre les acides aminés des mollusques et ceux des taxa classiquement étudiés en génomique. En outre, l'absence d'orthologues peut s'expliquer par le fait que (i) ce sont souvent de nouvelles familles de gènes qui sont découvertes et que (ii) peu d'espèces de mollusques sont actuellement étudiées en génomique.

Il n'est pas question dans le contexte de cette thèse de prétendre à l'étude de la fonctionnalité du transcriptome de *Macoma balthica* dans son ensemble. Toutefois, les 30 % du génome identifiés représentent près de 10000 gènes dont on connaît tout ou partie du rôle dans le fonctionnement cellulaire. La plupart d'entre eux interviennent dans les processus cellulaires, les voies métaboliques et les régulations biologiques (ceci quelle que soit la population considérée, Annexe B).

II.4.2- Mise en évidence de gènes impliqués dans le système de défense

Ce travail a mis en évidence plus de 300 gènes présentant une homologie significative ($4,78.10^{-149} < e\text{-value} < 6,31.10^{-7}$) avec des gènes de défense identifiés chez des organismes modèles (Annexe B). Parmi eux, les protéases et leurs inhibiteurs ont un rôle majeur dans le métabolisme des protéines, le remodelage cellulaire et la défense contre les pathogènes (Hiemstra 2002, Dumas *et al* 2005, Hamilton *et al* 2008). Certaines de ces protéases ont un rôle capital dans la réaction immunitaire, notamment les cathepsines retrouvées en nombre dans notre étude (CTSA, CTSB, CTSC, CTSD, CTSF, CTSF, CTSK, CTSL, CTSS, CTSZ).

De nombreuses séquences codant pour le protéasome ainsi que pour plusieurs ubiquitines ont été également observées. Le protéasome est un complexe enzymatique qui dégrade les protéines défectueuses (mal repliées, dénaturées ou obsolètes). Le protéasome le plus courant est le protéasome 26S (retrouvé dans notre étude), qui intervient dans le cycle cellulaire, l'expression génique et la réponse au stress oxydant. Avant d'être détruites par le protéasome, les protéines sont au préalable «marquées» par des ubiquitines. Nous avons ainsi isolé 5 ubiquitines différentes (USP2, USP3, USP14, USP16 et USP10-Like).

De nombreuses c-type lectines ont également été mises en évidence par notre étude. Les lectines sont des protéines qui interviennent dans l'adhésion inter-cellulaire. Leur rôle dans la réponse immunitaire innée chez les animaux est largement reconnu, qui plus est chez les invertébrés (Arason 1996, Yang *et al* 1996, Ewart et Johnson 2001). Elles se lient spécifiquement aux carbohydrates spécifiques de certains agents infectieux pathogènes. De nombreuses autres protéines intervenant dans la réponse immunitaire innée chez les invertébrés (Cousteau 2009) ont été également mises en évidence dans cette étude: les FREP (*Fibrinogen RElated Protein*), les peptidoglycanes ainsi que les protéines de réponse aux bactéries Gram⁻ sont connues pour intervenir dès la phase de reconnaissance du non-soi. Outre les lectines qui participent à l'activation et à l'amplification de la réponse, nous pouvons noter la présence de deux protéines apparentées aux cytokines présentes chez les vertébrés, l'*allograft inflammatory factor 1* et le *macrophage migration inhibitory factor*, déjà décrits chez le gastéropode *Biomphalaria glabrata* (Cousteau, 2009).

Les ferritines qui sont des protéines de stockage du fer, sont considérées comme la 1^{ère} ligne de défense immunitaire de l'organisme (Liu et Theil 2005). Elles ont une activité antibactérienne, antivirale, antifongique mais aussi anti-inflammatoire et anti-oxydante. Il a par ailleurs été montré que leur taux augmentait de manière significative lors de la prolifération des tumeurs. Une étude de 2004 (Larade et Storey) sur la Littorine (*Littorina littorea*) a également montré une forte implication des ferritines en réponse à l'anoxie.

De nombreux gènes identifiés sont susceptibles d'être impliqués dans la réponse générale au stress. Ils comprennent par exemple 8 protéines de chocs thermiques ou *Heat Shock Protein* (HSP4, HSP10, HSP9, HSP22, HSP60, HSP70 ET HSP90). Les HSP sont des protéines extrêmement conservées (Lindquist 1986), dites chaperonnes dont le rôle est d'empêcher l'accumulation de protéines anormales en rétablissant une conformation normale ou en les guidant vers le protéasome pour destruction (c'est le cas de HSP90). Si ces protéines sont synthétisées durant tout le développement d'un organisme (Feige *et al* 1996), leur production est fortement induite par des conditions toxiques telles qu'une exposition virale, aux métaux lourds, mais aussi une élévation de la température (Santoro 2000). Parmi elles, HSP70 est considérée comme un biomarqueur de stress et de toxicité dans des organismes très variés (Ryan et Hightower 1996, Radjev et Sharp 2000). Nous avons également pu révéler des protéines de réponse au stress oxydatif, parmi lesquelles de nombreuses peroxydases telles que les glutathion peroxydase (GPX2, 3 et 6) ou les thiorédoxine peroxydase mais aussi des glutathion S-transférases (GST) ainsi que des superoxyde dismutase (SOD) et une catalase.

Les données de séquençage du transcriptome de *Macoma balthica* ont donc permis de mettre en évidence des gènes candidats impliqués dans la réponse générale au stress, et au terme de cette étude, plus de 300 gènes candidats ont ainsi été décrits.

Chapitre 3

Adaptation différentielle selon les contraintes environnementales

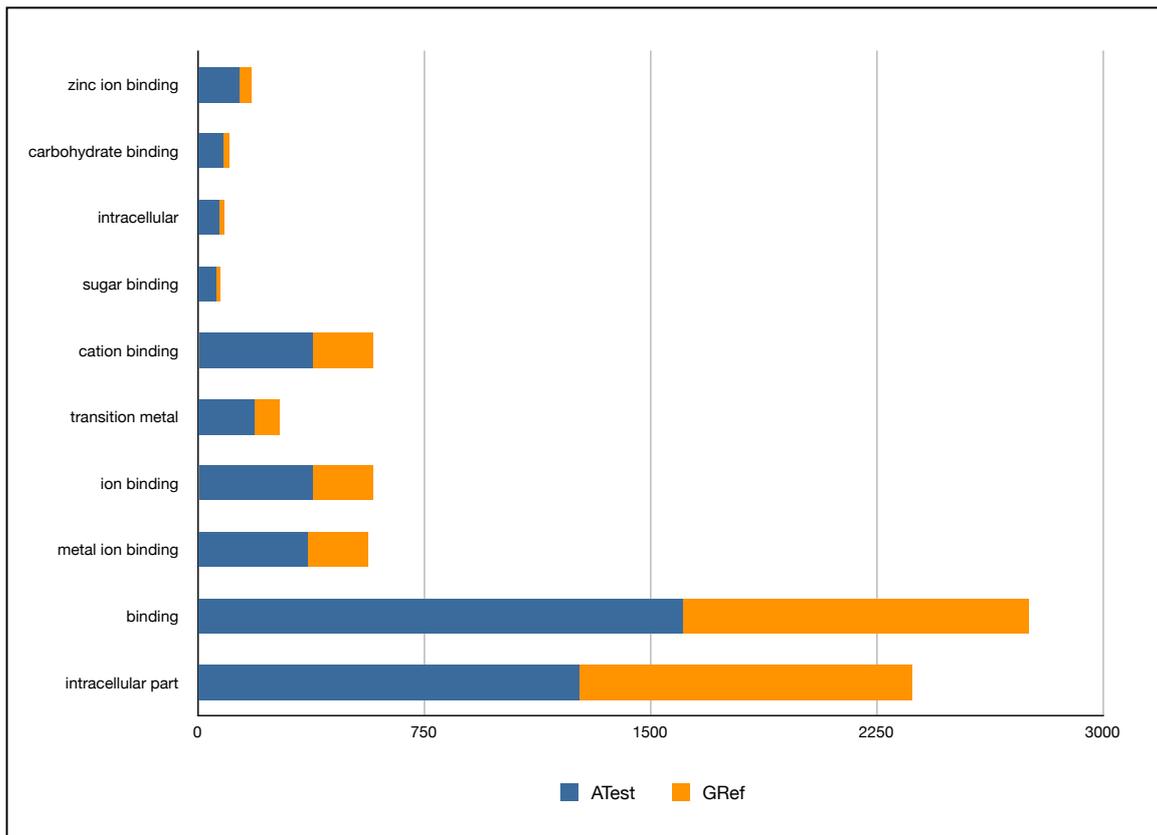


Figure II.6 : Comparaison des populations A et G. Test de Fisher, corrections multiples (p, FDR et FWER<0,5)

La deuxième étape de l'analyse moléculaire de l'adaptation de *Macoma balthica*, se veut comparative. Grâce au marquage initial des populations, nous sommes en mesure d'identifier la population d'origine des transcrits (ARNm) et ainsi susceptibles de relier des conditions environnementales à un niveau d'expression génique. Notre objectif est de mettre en évidence des réponses moléculaires différentes selon les populations et leurs contraintes anthropiques.

III.1- Résultats

La comparaison des populations deux à deux a permis de mettre en évidence des gènes différentiellement exprimés. Cette comparaison s'est faite selon trois niveaux de lecture en fonction de la significativité du test de Fisher. Ainsi une population est considérée comme significativement enrichie en termes GO lorsque tous les indices de correction (FRD, FWER et p) sont supérieurs à 0,05. A l'inverse, l'enrichissement est faiblement significatif si seulement deux de ces indices, voire un seul (FWER et p ou seulement p) sont supérieurs à 0,05.

III.1.1- Comparaison des populations A et G

Les 2351 séquences issues de la population A ont été comparées aux 1831 séquences de la population G. A partir de ce set de séquences, 89 catégories ont été comparées par un test de Fisher (Figure II.6). Sur ces 89 catégories de termes GO, seulement 10 étaient différentiellement exprimées ($2,39.10^{-7} < \text{FDR} < 0,007$, $5,97.10^{-8} < \text{FWER} < 0,002$, $7,86.10^{-10} < p < 1,23.10^{-5}$), 9 termes GO étant surexprimés dans la population A (soit 10 % de la totalité des catégories comparées) et seulement 1 (environ 1 %) dans la population G.

Toutefois, en considérant les termes GO pour lesquels seule p est inférieure à 0,05, 20 termes GO supplémentaires sont surexprimés dans la population A (soit 23 %) et 59 dans la population G (soit 66 %) (Annexe B). Parmi ces 79 termes GO (20 de la population A et 59 de la population G), il faut noter des fonctions moléculaires peptidiques intéressantes. Ainsi, dans la population A, diverses activités de lyse des protéines (activité inhibitrice des peptidases (p = 0,012), activités métalloendopeptidase et métallopeptidase (p = 0,021 et 0,028), activité phosphatase-tyrosine (p = 0,017) et protéolyse (p = 0,037) s'expriment. Dans la population G, on remarque parmi ces termes GO surexprimés, des protéines intervenant

dans la réponse au stress ($p = 0,008$), ou aux dommages à l'ADN ($p = 0,012$), aux organismes ou aux stimuli chimiques ($p = 0,013$ et $0,019$).

III.1.2- Comparaison des populations A et S

La comparaison des 2351 séquences de la population A aux 1373 séquences de la population S a permis de mettre en évidence 28 termes GO significativement différentiellement exprimés sur un total de 189 fonctions moléculaires comparées (soit environ 15%). Sur ces 28 fonctions moléculaires, 18 sont surexprimées dans la population A (soit environ 10%) et 10 dans la population S (environ 5 %) (Figure II.7).

Vingt fonctions moléculaires peuvent être considérées en restreignant la significativité du test de Fisher à $FDR < 0,05$ et $p < 0,05$ (mais $FWER > 0,05$) (Annexe B). Parmi celles-ci, 4 sont surexprimées dans la population A (processus métaboliques et biosynthétiques).

En considérant uniquement les valeurs pour lesquelles p est significative ($FWER$ et $FDR < 0,05$), 141 fonctions moléculaires supplémentaires (soit 74 % des fonctions comparées) sont alors différentiellement exprimées. Sur ces 141 fonctions, la régulation de la réponse au stress et du système immunitaire, sont sous-exprimées dans la population A (respectivement 6 séquences pour la population A contre 26 et 25 pour la population S, $p = 0,021$ et $0,028$).

III.1.3- Comparaison des populations G et S

Sur les 3204 séquences testées (1831 séquences de la population G et 1373 séquences de la population S), 281 fonctions moléculaires peuvent être comparées (Figure II.8). Sur ces 281 fonctions moléculaires, 47 sont différentiellement exprimées (soit environ 17%, FDR , $FWER$ et $p < 0,05$), 28 étant surexprimées dans la population G et 19 dans la population S. En retenant uniquement les valeurs pour lesquelles p et la FDR sont $< 0,05$ ($FWER > 0,05$), 42 fonctions moléculaires supplémentaires peuvent être considérées (environ 15%) quasi toutes surexprimées dans la population G (33 contre 9 pour la population S) (Annexe B). Parmi elles, notamment, la réponse aux dommages à l'ADN ($p = 4,62.10^{-4}$ et $FDR = 0,007$) ainsi que la réponse cellulaire au stress ($p = 0,0015$ et $FDR = 0,02$), surexprimées dans la population G. Enfin, si on s'intéresse uniquement aux fonctions moléculaires pour lesquelles, seule la valeur p est significative ($> 0,05$), 192 fonctions moléculaires sont exprimées différentiellement dans les 2 populations.

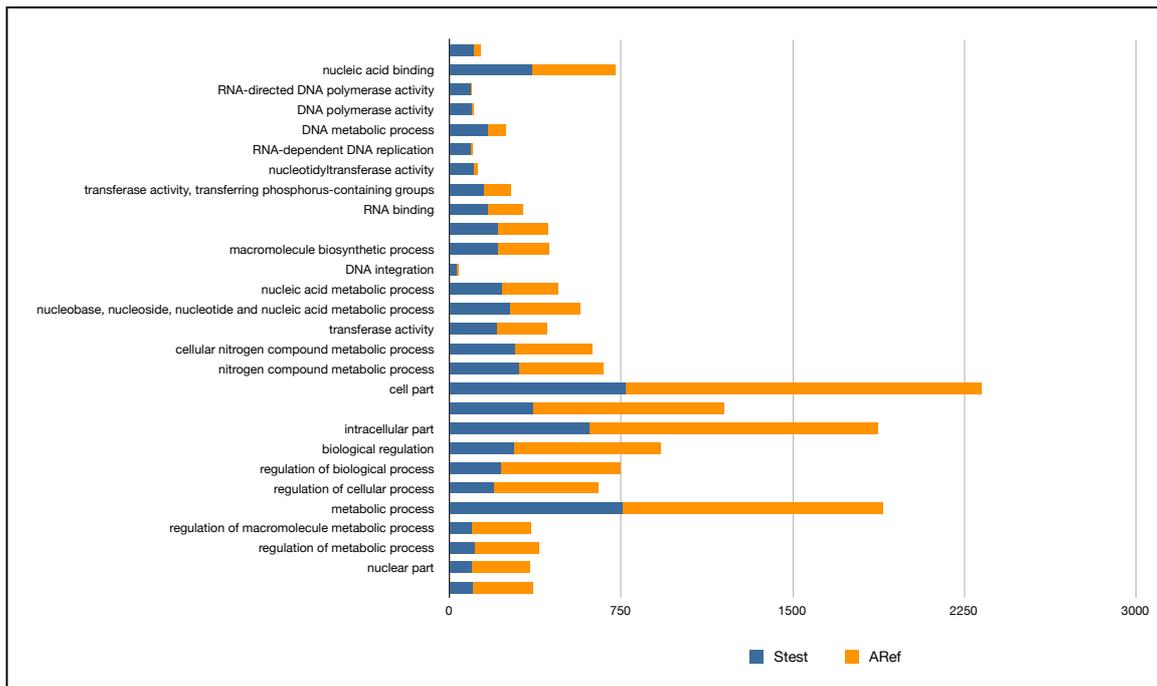


Figure II.7 : Comparaison des populations A et G. Test de Fisher, corrections multiples (p , FDR et FWER $< 0,5$)

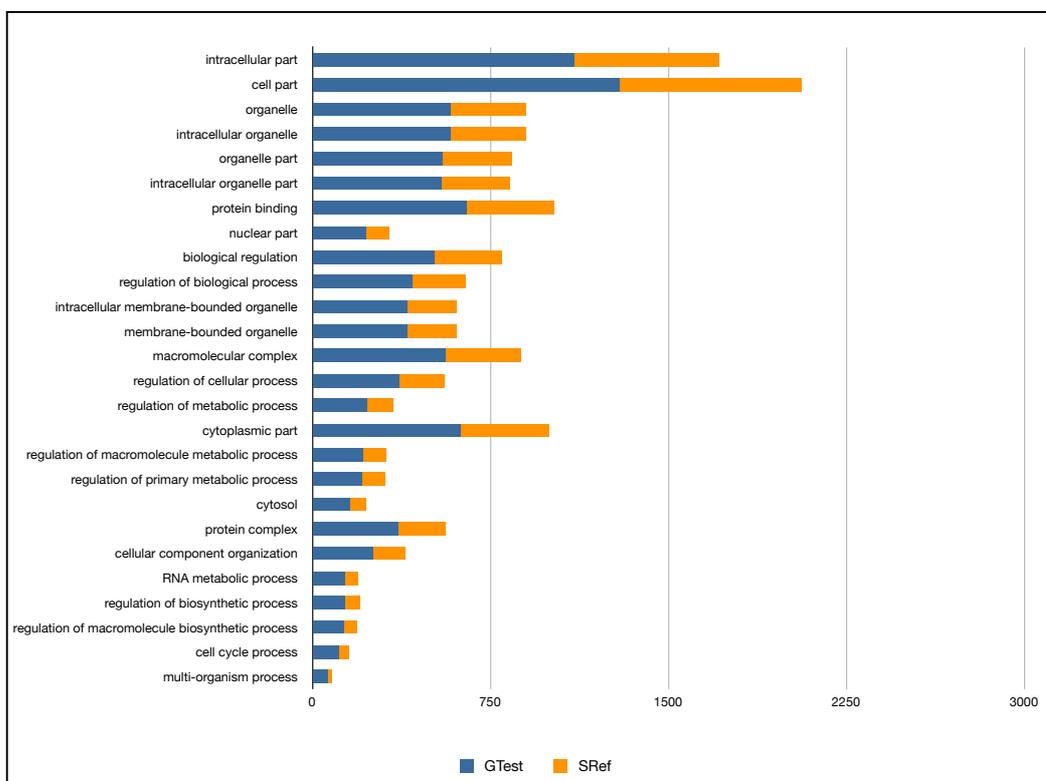


Figure II.8 : Comparaison des populations G et S. Test de Fisher, corrections multiples (p , FDR et FWER $< 0,5$)

Parmi celles-ci, les peptidases sont plutôt sous-exprimées dans la population G (métallopeptidase, serine-type endopeptidase) mais l'activité réponse cellulaire aux substances organiques est surexprimée ($p = 0,037$).

III.2- Discussion

Le marquage des ADNc en fonction des populations, avant le pyroséquençage nous a permis de comparer les transcriptomes d'individus provenant des différentes régions européennes (populations A, G et S). L'étude de l'enrichissement en termes GO d'une population par rapport à une autre, n'a pas mis en évidence de différentiel d'expression des gènes concernés par le stress franc et significatif. En effet, les gènes qui sont différentiellement exprimés de manière significative selon les populations (valeurs de FDR, FWER et $p < 0,05$) concernent essentiellement ceux liés au métabolisme général des cellules.

Néanmoins, cette étude nous a permis de dégager des tendances liées à chaque population. Ainsi, dans le transcriptome de la population de Gdansk (G), les gènes de réponse générale au stress, de réparation de l'ADN et d'homéostasie par rapport au métal, semblent davantage exprimés que dans les population de Marennes Oléron (A) et de la baie de Somme (S). Or les populations de *Macoma balthica* de la baie de Gdansk sont soumises à des paramètres physico-chimiques extrêmes (faible salinité, hypoxie) ainsi qu'à une forte pollution par les métaux lourds liés à l'activité industrielle de cette région (Sokolowski *et al* 2002, Szefer 2002). Il n'est donc pas surprenant de constater une surexpression des gènes intervenant dans la réponse cellulaire au stress, aux substances organiques ou encore aux dommages à l'ADN.

En revanche, une surexpression faiblement significative des protéines liées à la réponse à la température (HSP) peut être observée dans la population de Marennes Oléron. Or cette baie, qui représente la limite sud de l'aire de répartition de *Macoma balthica*, est soumise à l'élévation de la température des eaux de surface du Golfe de Gascogne depuis une cinquantaine d'années (Goikoetxea *et al* 2009). Par ailleurs, les populations de cette Baie sont à la limite de leur tolérance à la température (Jansen *et al* 2007). Les tests réalisés dans notre étude, malgré leur faible significativité, semblent confirmer la sensibilité de *Macoma balthica* à l'augmentation de la température.

Enfin, l'absence de surexpression de termes GO liés au stress dans la population de la baie de Somme semble confirmer la relative préservation de ce site en réserve naturelle, choisi comme référence pour cette étude. Toutefois, si ces populations sont moins soumises aux perturbations anthropiques, elles subissent en revanche une forte pression de prédation du fait de la grande fréquentation de cette baie par les prédateurs supérieurs et notamment les oiseaux limicoles. Une étude de Edelaar (2000) sur la profondeur d'enfouissement des

bivalves en réponse à la prédation des oiseaux limicoles mettait ainsi en évidence une différence de comportement et d'indice de conditions en fonction de la prédation. Nos résultats (peu d'expression de gènes de stress) suggèrent que la réponse à la prédation n'a pas de bases génétiques. Cette réponse comportementale (enfouissement plus profond face à un stress environnemental) a d'ailleurs été observée chez *Macoma balthica* en fonction du niveau d'hypoxie de l'eau dans la baie de Gdansk (Sokolowski *et al* 2008).

Toutefois, nous avons mis en évidence dans cette population un résultat assez surprenant dans une réserve naturelle: un enrichissement faiblement significatif en terme GO associé aux radiations ionisantes, ce terme n'ayant été détecté dans aucune des deux autres populations. Or la réserve naturelle de la baie de Somme est située à moins de 100 km au nord de la centrale nucléaire EDF de Penly (76).

III.3- Conclusion

Si l'adaptation aux contraintes environnementales ne s'exprime pas toujours de manière fortement significative, nos résultats semblent toutefois cohérents avec les caractéristiques environnementales de chacun des trois sites.

Le fait que les gènes impliqués dans la réponse générale au stress et les dommages à l'ADN soient surexprimés dans la population de la baie de Gdansk, laisse supposer de plus fortes contraintes environnementales que dans les autres populations. Cependant, notre approche méthodologique ne nous permet pas d'évaluer la part de chacune des contraintes environnementales (réchauffement, salinité, eutrophisation et pollution) dans ces réponses adaptatives. En effet notre choix de travailler *in situ* sur des populations naturelles ne nous permet pas de nous affranchir des interactions entre différentes pressions, de leur caractère additif et de leur ancienneté. Seules des approches en milieu contrôlé, nous donneraient ce niveau d'information. La comparaison des transcriptomes entre populations, telle qu'elle a été réalisée dans cette étude, constitue ainsi une première étape d'analyse. Elle a permis de sélectionner un certain nombre de gènes candidats dans les processus d'adaptation locale et devra être poursuivie par des approches plus larges telles que la recherche de polymorphisme et la détection d'*outliers*²⁰.

²⁰ outlier : marqueur neutre hors norme, potentiellement sous sélection

*Conclusion générale
et Perspectives*

Les bivalves marins présentent des caractéristiques démographiques (importante fécondité, grandes tailles de populations et fort potentiel de dispersion par le biais d'une phase larvaire planctonique passive, Thornson 1950, Kinlan et Gaines 2003) qui les distinguent de la plupart des organismes terrestres (Strathmann 1985, Palumbi 1992). De ce fait, les concepts et théories de la génétique des populations prévoient pour ce type d'organisme une diversité génétique forte, une faible différenciation génétique entre populations, une efficacité accrue de la sélection et une adaptation locale coûteuse, puisqu'à priori les échelles de dispersion surpassent l'hétérogénéité spatiale de l'environnement. Or ces attendus théoriques ne sont pas toujours observés. L'adaptation est en principe limitée par les flux de gènes qui s'opposent aux effets de la sélection locale. Néanmoins, nombreux sont les organismes marins qui maintiennent des polymorphismes d'adaptation aux variations locales de leur environnement selon trois hypothèses (i) une dispersion plus limitée qu'attendu (ii) un polymorphisme adaptatif moins important en milieu marin (iii) une longue histoire de conditions démographiques hors équilibre (colonisation, expansion, fragmentation) qui favorise l'émergence et le maintien d'un polymorphisme adaptatif. Cette observation paradoxale représente donc un véritable défi aux théories classiques de l'adaptation locale, et sa compréhension est fondamentale pour appréhender les mécanismes de l'adaptation contemporaine des espèces marines, notamment face aux changements environnementaux rapides engendrés par les pressions anthropiques (changements climatiques et dégradation des milieux).

Notre travail sur le bivalve marin *Macoma balthica* a apporté quelques éléments de compréhension à ces processus d'adaptation en milieu marin. Les approches de la génétique des populations associées à celles de la génomique ont permis d'appréhender les bases génétiques de la différenciation adaptative. Cette espèce clef des littoraux européens, se caractérise par une aire de répartition très étendue et une grande tolérance aux variations environnementales. Toutefois, la récente remontée vers le nord de sa distribution semble indiquer que *Macoma balthica* ait atteint ses limites adaptatives dans certaines régions. Ses caractéristiques biologiques et son rôle important dans le réseau trophique, en font un modèle de choix pour étudier les processus évolutifs d'adaptation dans un contexte d'écosystèmes fortement anthropisés.

Dans la première partie de cette thèse, nous avons inféré l'incidence de l'histoire démographique de cette espèce (colonisation, fragmentation, flux géniques), sur l'existence et le maintien de différenciations génétiques adaptatives. Les concepts de la génétique des populations et l'utilisation de marqueurs moléculaires neutres (ADN mitochondrial et microsatellites) sur des populations soumises à des contraintes environnementales différentes (baie de Marennes-Oléron et baie de Gdansk), ont permis de mettre en évidence un certain nombre de résultats :

- Les populations de *Macoma balthica* situées en limite sud d'aire de répartition et soumises au réchauffement des eaux du golfe de Gascogne, présentent une diversité génétique neutre élevée, supérieure à celle des populations situées au cœur de la répartition. Ces observations contraires à l'attendu théorique de populations marginales en périphérie de distribution (Eckert *et al* 2008), permettent de clarifier le statut de ces populations qui sont probablement issues d'une zone refuge lors du dernier maximum glaciaire (Gomez et Lunt 2007) ;

- Les populations prélevées au sein de la baie de Marennes Oléron sont panmictiques ;

- Il existe une forte barrière au flux de gènes entre les populations du golfe de Gascogne et les populations plus septentrionales, susceptible de favoriser l'isolement des populations en limite d'aire, les rendant à priori plus vulnérables aux changements environnementaux ;

- Les populations de la baie de Gdansk, soumises à des conditions environnementales particulièrement contraignantes (salinité, eutrophisation et contaminants), sont également caractérisées par une forte diversité génétique neutre ;

- Dans la baie de Gdansk, il existe une limite au flux de gènes au niveau de la thermo-halocline dont nous avons pu préciser la profondeur aux alentours des 40 mètres ;

- Dans l'embouchure de la Vistule, fortement contaminée par les métaux lourds notamment, une faible mais significative structuration génétique est observée et certaines des populations présentent des signes d'adaptation locale.

Dans la deuxième partie du travail, nous avons bénéficié des progrès récents dans le séquençage haut-débit des génomes pour identifier chez *Macoma balthica*, des régions génomiques influencées par la sélection. Le séquençage massif du transcriptome, réalisé par pyroséquencage 454 sur trois populations soumises à des contraintes environnementales différentes (baie de Marennes-Oléron, baie de Somme et baie de Gdansk), a apporté les premiers éléments de compréhension des bases moléculaires de l'adaptation en conditions naturelles:

- des gènes candidats impliqués dans la réponse au stress, ont été mis en évidence, notamment des *Heat Shock Protein* (HSP), des cathepsines et des protéines impliquées dans les mécanismes de réparation de l'ADN ;

- des différentiels d'expression certes faibles, de ces gènes ont été observés *in situ* selon les conditions environnementales. Ainsi les populations situées en limite d'aire de répartition, soumises à un réchauffement des eaux de surface dans le golfe de Gascogne, semblent présenter des taux de HSP plus élevés. Parallèlement, les gènes surexprimés dans

les populations de la baie de Gdansk seraient plutôt impliqués dans la réparation à l'ADN et la réponse aux stimuli chimiques.

Cette thèse a complété et affiné les connaissances sur l'histoire démographique d'un mollusque marin à fort flux génique. Nos travaux constituent la première étape d'analyse des bases moléculaires de la sélection chez une espèce clef des écosystèmes littoraux fortement anthropisés. Cette approche originale, une des rares à s'intéresser à des populations naturelles analysées *in situ*, a permis de dégager des preuves de sélection naturelle divergente en milieu environnemental contrasté.

La méthode des *scan* F_{st} (Beaumont et Nichols 1996) permet de détecter des *outliers* (loci affectés par une association environnement-génotype) par le génotypage d'effectifs importants d'individus à l'aide de différents types de marqueurs moléculaires. Cette approche nous permettrait d'associer les niveaux de différenciation génétique entre populations à une ou plusieurs variables écologiques (modèle GEA, *Genetic Environment Association*, Hedrick 1976). A cet effet, dans le cadre de l'ANR Hi-Flo, des polymorphismes nucléotidiques ont d'ores et déjà été identifiés à partir de l'assemblage des données du pyroséquençage (14464 SNPs).

Par ailleurs, nous ne pouvons nous affranchir d'une approche plus précise qui nous permettra d'évaluer l'impact des différentes pressions à l'échelle de l'organisme. C'est pourquoi, l'analyse de l'expression des gènes candidats en milieu contrôlé est également prévue. Ainsi, nous envisageons de développer une puce à ADN composée de plusieurs gènes candidats susceptibles d'être différenciellement exprimés dans un contexte de réchauffement d'une part, et de forte contamination d'autre part. Deux échelles d'expérimentation seront abordées, en mésocosme et par des transplantations de populations en milieu naturel. Cette approche nous permettrait également d'évaluer les effets cumulatifs des pressions anthropiques.

Bibliographie

- Alleaume-Benharira** M. et al. (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *J. Evol. Biol.* 19, 203-215.
- Alward** R.D., Detling J.K., Milchunas D.G. (1999) Grassland vegetation changes and nocturnal global warming. *Science*, 283: 229-231.
- Arason** G. (1996) Lectins as defence molecules in vertebrates and invertebrates. *Fish and Shellfish Immunology* 6: 277-289.
- Armbruster** P., Bradshaw W.E., Holzapfel C.M. (1997) Evolution of the genetic architecture underlying fitness in the pitcher-plant mosquito, *Wyeomyia smithii*. *Evolution* 51: 451-458.
- Bachelet** G. (1980) Growth and recruitment of the tellinid bivalve *Macoma balthica* at the southern limit of its geographical distribution, the Gironde Estuary (SW France). *Mar Biol* 59:105-117
- Bandelt** H.J., Forster P. and Rohl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, Vol 16, 37-48
- Barbier** E.B. (1993) Sustainable Use of Wetlands—Valuing Tropical Wetland Benefits—Economic Methodologies and Applications. *Geographical Journal*, 159 : 22-32.
- Barry** J.P., Baxter C.H., Sagarin R.H., Gilman S.E. (1995). Climate-related longterm faunal changes in a California rocky intertidal community. *Sciences* 267: 672-675.
- Beaumont** M.A., Nichols R.A. (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proceedings Biological Sciences* 263:1619-1626.
- Belkhir** K, Borsa P, Chikhi P, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous windows pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Universités de Montpellier II, Montpellier, France.
- Becquet** V., Lanneluc I., Simon-Bouhet B. and Garcia P. (2009) Microsatellite markers for the Baltic clam, *Macoma balthica* (Linné, 1758), a key species concerned by changing southernlimit, in exploited littoral ecosystems. *Conservation Genet Ressources*
- Bell** & Collins 2008 Adaptation, extinction and global change. *Evolutionary applications*, 1: 3-16
- Berner** E.K., Berner R.A. (1996) *Global Environment*. Prentice-Hall, Upper Saddle River, 2ème édition, p365.
- Beukema** J.J. (1993) Successive changes in distribution patterns as an adaptive strategy in the bivalve *Macoma balthica* (L.) in the Wadden Sea. *Helgoländer Meeresuntersuchungen*, 47: 287-304.
- Beukema** J.J, Dekker R, Jansen J.M, (2009) Some like it cold: populations of the tellinid bivalve *Macoma balthica* (L.) suffer in various ways from a warming climate. *Mar Ecol Prog Ser* 384: 135-145.
- Billote** N., Lagoda P., Risterucci A.M., Baurens F.C. (1999) Microsatellite-enriched libraries: applied methodology for the tech development of SSR markers in tropical crops. *Fruits*, 54, 277-288.
- Birky** C.W. (2001) The Inheritance of Genes in Mitochondria and Chloroplasts: Laws, Mechanisms, and Models. *Annual review of Genetics* 35: 125-148.
- Blüthgen** N., Brand K., Cajavec B., Swat M., Herzel H., Beule D. (2005) Biological Profiling of Gene Groups utilizing Gene Ontology. *Genome Informatics* 16: 106-115.
- Bolnick** (2008) Evidence for asymmetric migration load in a pair of ecologically divergent stickleback populations. *Biological Journal of the Linnean Society*, 2008, 94, 273-287.
- Boore** J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Research* 27: 1767-1780.
- Borevitz** J.O. and Chory J. (2004) Genomics tools for QTL analysis and gene discovery. *Current Opinion in Plant Biology* 7: 132-136.
- Boudouresque** C.F., Ribera M.A. (1994) Les introductions d'espèces végétales et animales en milieu marin - conséquences écologiques, économiques et problèmes législatifs. In: First International Workshop on *Caulerpa taxifolia*. Ed. CF Boudouresque C.F., Meinesz A. and Gravez V. GIS Posidonie publishers, 29-102.
- Boudouresque** C.F. (1999) The Red Sea-Mediterranean link: unwanted effects of canals. *Proceedings Norway/UN Conference on Alien Species*, 1-5 July 1996, pp 107-115.

- Bouzat J.L.**, Johnson K. (2004). Genetic structure among closely spaced leks in a peripheral population of lesser prairie-chickens. *Molecular ecology*, 13, 499-505.
- Bradley B.A.**, Oppenheimer M., Wilcove D.S. (2009) Climate change and plant invasions: restoration opportunities ahead? *Global Change Biology*, 15: 1511-1521.
- Brown J.H. & Kodric-Brown A** (1977) Turnover Rates in Insular Biogeography: Effect of Immigration on Extinction. *Ecology*, 58:445-449.
- Brown J.H.** (1984) On the relationship between abundance and distribution of species. *American naturalist*, 124: 255-279.
- Brown J.H.**, Mehlman D.W., Stevens G.C. (1995) Spatial variation in abundance. *Ecology*, 76: 2028-2043.
- Brown J.H.**, Valone T.J., Curtin C.G. (1997) Reorganization of an arid ecosystem in response to recent climate change. *PNAS*, 94: 9729-9733.
- Burke L.**, Kura Y., Kassem K., Ravenga C., Spalding M., McAllister D. (2001) Pilot Assessment of Global Ecosystems: Coastal Ecosystems. World Resources Institute (WRI), Washington, D.C. (USA), 94 pp
- Boutet, I.**, Tanguy A., Moraga D., 2004. Response of the Pacific oyster *Crassostrea gigas* to hydrocarbon contamination under experimental conditions. *Gene* 329, 147-157
- Breton S.**, Doucet Beaupré H., Stewart D.T., Hoeh W.R., Blier P.U. (2007) The unusual system of doubly uniparental inheritance of mtDNA: isn't one enough?. *Trends in Genetics* 23: 465-474.
- Bultelle F.**, Panchout M., Leboulenger F., Danger J.M., 2002. Identification of differentially expressed genes in *Dreissena polymorpha* exposed to contaminants. *Mar. Environ. Res.* 54, 385-389
- Burton R.S.** (1990) Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. *Evolution* 44: 1814-1822.
- Caddy J.F.** (1969). Development of mantle organs, feeding and locomotion in postlarval *Macoma balthica* (L.) (Lamellibranchiata) *Canadian Journal of Zoology* 47: 609-617.
- Cassens I.**, Van Waerebeek K., Best P.B., Crespo E.A., Reyes J. and Milinkovitch M.C. (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Molecular Ecology*, 12 (7) : 1781-1792.
- Cavalli-Sforza L.L.**, Edwards A.W.F (1967) Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet* 19 : 233-257.
- Chapin F.S.**, Shaver G.R., Goblin A.E., Nadelhoffer K.J., Laundre J.A. (1995) Responses of arctic tundra to experimental and observed changes in climate. *Ecology*, 76: 694-711.
- Coles S.L.**, DeFelice R.C., Eldredge L.G., Carlton J.T. (1999) Historical and recent introductions of non-indigenous marine species into Pearl Harbor, Oahu, Hawaiian Islands. *Marine Biology*, 135: 147-158.
- Conesa A.**, Götz S., García-Gómez J.M., Terol J., Talón M., Robles M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics Applications Note* 21: 3674-3676.
- Courchamp F.** et al. (1999) Inverse density dependence and the Allee effect. *Trends Ecol. Evol.* 14, 405-410.
- Cousteau C.** (2009) Immunity in parasite-vector snails. *Médecine Sciences* 25: 399-403.
- Crozier R.** (1990) From population genetics to phylogeny : uses and limits of mitochondrial DNA. *Australian Systematic Botany*, 3: 111-124.
- Cyberska B.**, Lauer Z., Trzosinska A. (1998) Environmental conditions in the Polish zone of southern Baltic sea during 1997. Institute of Meteorology and Water Management, Marine Branch Materials, Gdynia, pp. 211-240.
- CIESIN** (Center for International Earth Science Information Network) (2003) Gridded population of the world (GPW) Version 3 beta. [online]. Available at <http://quin.unep-wcmc.org/MA/index.cfm> (with username and password).
- CNC** (Comité National de la Conchyliculture) (2006)
- CNUCED** (2007) World investment report. New York, USA et Geneve, Suisse.
- D'Antonio C.M.** (2000) Fire, plant invasions, and global changes. Pages 65-93 in H. A. Mooney and R. J. Hobbs, editors. *Invasive species in a changing world*. Island Press, Covelo,

California, USA.

David P., Perdieu M.A., Pernot A.F., Jarne P. (1997) Fine-grained spatial and temporal population genetic structure in the marine bivalve *Spisula ovalis*. *Evolution* 51: 1318-1322.

DeWitt T.J., Sih A., Sloan Wilson D. (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution* ,13: 77-81.

Di Castri F., Hansen A.J., Debussche M. (1990) On invading species and invaded ecosystems: the history of historical chance and biological necessity. In *Biological Invasions in Europe and the Mediterranean Basin* (eds DI Castri, F., Hansen A.J., Debussche M.) pp 3:16, Kluwer Academic Publishers, Dordrecht, Netherlands.

Doumas S., Kolokotronis A., Stefanopoulos P. (2005) Anti-Inflammatory and Antimicrobial Roles of Secretory Leukocyte Protease Inhibitor. *Infection and Immunity* 73: 1271-1274.

Dow B.D., Ashley M.V. (1996) Microsatellite analysis of seed dispersal and parentage of slappings in bur oak, *Quercus macrocarpa*. *Molecular ecology* 5: 615-627.

Drent J. (2002) Temperature responses in larvae of *Macoma balthica* from a northerly and southerly population of the European distribution range. *Journal of Experimental Marine Biology and Ecology*, 275, 117-129.

Dulvy N.K., Rogers S.I., Jennings S., Stelzenmüller V., Dye S.R., Skoldal H.R. (2008). Climate change and deepening of the North Sea fish assemblage: a biotic indicator of warming seas. *Journal of Applied Ecology* 45 : 1029-1039.

Durka W. (1999). Genetic diversity in peripheral and subcentral populations of *Corrigiola litoralis* L. (Illecebracea). *Heredity*, 83, 476-484

Ebert D., Haag C., Kirkpatrick M., Riek M., Hottinger J.W., Pajunen V.I. (2002) A selective advantage to immigrant genes in a *Daphnia* metapopulation. *Science* 295: 485-488.

Eckert C.G., Samis K.E., Loughheed S.C. (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular ecology*, 17, 1170-1188

Eckman J.E. (1996) Closing the larval loop: linking larval ecology to the population dynamics of marine benthic invertebrates. *Journal of Experimental Marine Biology and Ecology* 200: 207-237.

Edelaar P. (2000) Phenotypic plasticity of burrowing depth in the bivalve *Macoma balthica*: experimental evidence and general implications. In: *The Evolutionary Biology of the Bivalvia*, Vol. 177 (eds Harper EM, Taylor JD, Crame JA) pp. 451-458. Geological Society: Special Publications, London, UK.

Edjung G. and Bonsdorff E. (1992) Predation on the bivalve *Macoma balthica* by the isopod *Saduria entomon*: laboratory and field experiments. *Marine Ecology Progress Series*, 88, 207-214

Edmands S.P., Moberg E., Burton R.S. (1996) Allozyme and mitochondrial DNA evidence of population subdivision in the purple sea urchin *Strongylocentrotus purpuratus*. *Marine Biology* 126: 443-450.

Eisen J.A. (1999) Mechanistic basis for microsatellite instability. *Microsatellites, Evolution and Applications* (ed Goldstein D.B, Schlötterer C.) Oxford University Press, New York. pp 34:48.

Ellien C., Thiebaut E., Dumas F., Salomon J.C., Nival P. (2004) A modeling study of the respective role of hydrodynamic processes and larval mortality on larval dispersal and recruitment of benthic invertebrates: example of *Pectinaria koreni* (Annelida: Polychaeta) in the Bay of Seine (English Channel). *Journal of Plankton Research* 26: 117-132.

Ellison A.M. (2008) Managing mangroves with benthic biodiversity in mind: Moving beyond roving banditry. *Journal of Sea Research*, 59 : 2-15.

Elton C. (1958) *The ecology of invasions by animals and plants*, Chapman and Hall. London, England.

El Mousadik & Petit (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, 92, 832-839

Eno N.C. (1996) Non-native marine species in British waters: effects and controls. *Aquatic Conservation*, 6: 215-228.

Estoup A. & Angers B. (1998) Microsatellites and minisatellites for molecular ecology : theoretical and empirical considerations. In : *Advances in molecular ecology*, pages 55-86. NATO

press, Amsterdam.

Excoffier L., Smouse P.E. and Quattro J.M. (1992) Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics*, Vol 131, 479-491

Excoffier L., Laval G. and Schneider S. (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online*. 2005; 1: 47-50.

Eyre-Walker A. (2006) The genomic rate of adaptive evolution. *Trends in Ecology and Evolution*, 21: 569-575.

Ewart K.V., Johnson S.C., Ross N.W. (2001) Lectins of the innate immune system and their relevance to fish health. *Journal of Marine Science* 58: 380-385.

Farber S. (1988) The Value of Coastal Wetlands for Recreation—an Application of Travel Cost and Contingent Valuation Methodologies. *Journal of Environmental Management* 26: 299-312.

Fenster C.B., Galloway L.F. (2000) Population differentiation in an annual legume: genetic architecture. *Evolution* 54: 1157-1172.

Feige U., Morimoto R. I., Yahara I., Polla B. S. (1996) Stress-inducible cellular responses. Feige U., Morimoto R. I., Yahara I., Polla, B. S. (Eds). Birkhauser-verlag : Basel, 1996, 492p

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6.

Finerty J.P. (1980) The population ecology of cycles in small mammals: mathematical theory and biological fact. New Haven, CT: Yale University press.

Frankham R., Weber K. (2000) Nature of quantitative genetic variation. In: *Evolutionary Genetics: From molecules to Man* (eds. Singh R, Krimbas CB), pp 351-368; Cambridge, UK: Cambridge University Press.

Freitas V., Campos J., Fonds M., Van der Meer H.W. (2007) Potential impact of temperature change on epibenthic predator-bivalve prey interaction in temperate estuary. *Journal of Thermal Biology*, 32: 328-340.

Gaines S.D., Gaylord B., Largier J.L. (2003) Avoiding current oversights in marine reserve design. *Ecological Applications* 13: S32-S46.

Gaines S.D., Bertness M.D. (1992) Dispersal of juveniles and variable recruitment in sessile marine species. *Nature* 360: 579-580.

Galil B.S. (2000) A sea under siege – alien species in the Mediterranean. *Biological Invasions*, 2 : 177-186.

Gaylord B., Gaines S.D. (2000) Temperature or transport ? TRange limits in marine species mediated solely by flow. *American Naturalist* 155: 769-789.

GIEC (2007). Bilan des changements climatiques: Contribution des groupes de travail I, II et III au quatrième rapport d'évaluation du Groupe d'Experts Intergouvernemental sur l'évolution du climat (Equipe de rédaction principale : Pachauri R.K., Resinger A.) GIEC. Genève, Suisse. 103pp.

Gilpin M.E., Soulé M.E. (1986) Minimum viable populations: processes of species extinctions. Pages 19-34 in M.E Soulé, editor. *Conservation Biology. The science of scarcity and diversity*. Sinauer Associates, Sunderland, Massachusetts.

Goikoetxea N., Borja A., Fontán A., González M., Valencia V. (2009) Trends and anomalies in sea-surface temperature, observed over the last 60 years, within the southeastern Bay of Biscay. *Continental Shelf Research* 29 (2009) 1060-1069

Gómez A, Lunt A.H, (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula, in : Wein S., Ferrand N., (eds), *Phylogeography of Southern European Refugia*, 155-188.

Goldburg R.J., Elliot M.S., Naylor R.L. (2001) *Marine Aquaculture in the United States: Environmental Impacts and Policy Options*. Prepared for the Pew Oceans Commission, Arlington, VA (USA).

Gordon D.R. (1998) Effect of invasive, non-indigenous plant species on ecosystem processes: lessons from Florida. *Ecological Applications*, 8(4), pp. 975-989

Hall T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 41:95-98

Hallegraeff G.M. (1993) A review of harmful algal blooms and their apparent global increase. *Phycologia*, 32: 79-99.

- Hamilton** G., Colbert J.D., Schuettelkopf A.W., Watts C. (2008) Cystatin F is a cathepsin C-directed protease inhibitor regulated by proteolysis. *European Molecular Biology Organization* 27: 499-508.
- Hancock** J.M. (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. *Microsatellites, Evolution and Applications* (ed Goldstein D.B, Schlötterer C.) Oxford University Press, New York. pp 1-9.
- Hanski** I. (1999) *Metapopulation Ecology*. Oxford University Press, Oxford.
- Harrison** S (1991) Local extinction in a metapopulation context: an empirical evaluation. *Biological Journal of Linnean Society*, 42: 73-88.
- Hartl** D.L. (1994) *Génétique des populations*. Médecine-Sciences, éditions Flammarion. 305p.
- Hasting** A. (1983) Can spatial variation alone lead to selection for dispersal ? *Theoretical Population Biology* 24: 244-251.
- Hedgecock** D. (1994a) Does variance in reproductive success limit effective population sizes of marine organisms ? Pages 122-134 in A.R. Beaumont eds. *Genetics and evolution of aquatic organisms*. Chapman and Hall, London, UK.
- Hedgecock** D. (1994b) Temporal and spatial genetic structure of marine animal populations in the California Current. *California Cooperative Oceanic Fisheries Investigations Reports* 35: 73-81.
- Hedrick** P.W., Ginevan M.E., Ewing E.P. (1976) Genetic Polymorphism in Heterogeneous Environments *Annual Review of Ecology and Systematics* 7:1-32
- Hellmann** J.J., Byers J.E, Bierwagen B.G, Dukes J.S (2008) Five Potential Consequences of Climate Change for Invasive Species. *Conservation Biology*, 22:534-543.
- Hengeveld** H.G. (1990) Global climate change: implications for air-temperature and water supply in Canada. *Transactions of the American Fisheries Society*, 119: 176-182.
- Hewitt** G.M (1999) Post-glacial re-colonization of European biota. *Biol J Linn Soc* 68 : 87-112.
- Hiemstra** P.S., (2002) Novel roles of protease inhibitors in infection and inflammation *Biochemical Society Transactions* 30: 116-12.
- Hobbs** R.J. & **Huenneke** L.F (1992) Disturbance, Diversity, and Invasion: Implications for Conservation. *Conservation Biology*, 6 : 324-337.
- Holbrook** S.J., Schmitt R.J., Stephens J.S. (1997) Changes in a assemblage of temperate reef fishes associated with a climate shift. *Ecological Applications*, 7: 1299-1310.
- Howarth** R.W., Billen G., Swaney D., Townsend A., Jaworski N., Lajtha K., Downing J.A., Elmgren R., Caraco N., Jordan T. et al. (1996) Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. *Biogeochemistry*, 35: 75-139.
- Howe** G.T., Brunner A.M. (2005) An evolving approach to understanding plant adaptation. *New Phytologist* 167: 1-5.
- Hulscher** J.B. (1973) Burying depth and trematode infection in *Macoma balthica*. *Netherlands Journal of Sea Research*, 6: 141-156.
- Hummel** H. (1985) Food intake of *Macoma balthica* (Mollusca) in relation to seasonal changes in its potential food on a tidal flat in the Dutch Wadden Sea. *Neth. J. Sea Res.* 19, 52-76.
- Hummel** H., Bogaards R.H., Amiard-Triquet C., Bachelet G., Desprez M., Marchand J., Rybarczyk H., Sylvand B., de Wit Y., de Wolf L. (1995) Uniform variation in genetic traits of a marine bivalve related to starvation, pollution and geographic clines. *J exp mar Biol Ecol* 191 : 133-150
- Hummel** H., Amiard-Triquet C., Bachelet G., Desprez M., Marchand J., Sylvand B., Amiard J.C., Rybarczyk H., Bogaards R.H., de Wolf L. (1997a) Comparison of ecophysiological, biochemical and genetic traits in the estuarine bivalve *Macoma balthica* from areas between the Netherlands and its southern limits (Gironde): geographic clines parallel effects of starvation and copper exposure. L.E. Hawkins & S. Hutchinson (eds.). *Proceeding 30th European Marine Biological Symposium*, University of Southampton, pp 15-20
- Hummel**, H., Bogaards, R., Bek, T., Polishchuk, L., Amiard- Triquet, C., Bachelet, G., Desprez, M., Strelkov, P., Sukhotin, A., Naumov, A., Dahle, S., Denisenko, S., Gantsevich, M., Sokolov K, de Wolf, L., (1997b). Sensitivity to stress in the bivalve *Macoma balthica* from the most northern (Arctic) to the most southern (French) populations: low sensitivity in Arctic populations because of genetic adaptations? *Hydrobiologia* 355, 127-138.

- Hummel H.**, Bogaards R., Bek T., Polishchuk L., Sokolov K., Amiard-Triquet C., Bachelet G., Desprez M., Naumov A., Strelkov P. (1998) Growth in the bivalve *Macoma balthica* from its northern to its southern distribution limit: a discontinuity in North Europe because of genetic adaptations in Arctic populations? *Comp Biochem Physiol A* 120 : 133-141
- Hummel H.**, Sokolowski A., Bogaards R., Wolowicz M. (2000a) Ecophysiological and genetic traits of the Baltic clam *Macoma balthica* in the Baltic: differences between populations in the Gdańsk Bay due to acclimatization or genetic adaptation? *Hydrobiologia* 85: 621-637.
- Hummel H.**, Bogaards R.H., Bachelet G., Caron F., Sola J.C., Amiard-Triquet C. (2000b) 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.
- Huvet A.**, Herpin, A., Degremont, L., Labreuche, Y., Samain, J.F., Cunningham, C., 2004. The identification of genes from the oyster *Crassostrea gigas* that are differentially expressed in progeny exhibiting opposed susceptibility to summer mortality. *Gene* 343, 211-220
- Hurrell J.W.** & **Van Loon H.** (1997) Decadal Variations in Climate Associated with the North Atlantic Oscillation. *Climatic Change*, 36: 301-326.
- Huey R.B.**, Gilchrist G.W., Carlson M.L., Berrigan D., Serra, L. 2000. Rapid evolution of a geographic cline in size in an introduced fly. *Science*. 287: 308-309.
- IFEN** (Institut Français de l'ENvironnement) (2006) L'environnement en France. France
- IPPC** (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,.
- Janas U.** (1998) The influence of oxygen deficiency and the presence of hydrosulphur on macrozoobenthos of the Gulf of Gdansk. Ph.D. Dissertation. University of Gdańsk, Gdańsk.
- Janas U.**, Wocial, J. and Szaniawska, A. (2004). Seasonal and annual changes in the macrozoobenthic populations of the Gulf of Gdańsk with respect to hypoxia and hydrogen sulfide. *Oceanologia* 46, 85-102.
- Jansen J.M.**, Pronker A.E., Bonga S.W., Hummel H. (2007) *Macoma balthica* in Spain, a few decades back in climate history. *Journal of Experimental Marine Biology and Ecology*, 344, 161-169
- Janzen F.J.** (1994) Climate change and temperature-dependant sex determination in reptiles. *Proceedings of the National Academy of Science*, 91: 7487-7490.
- Johnson M.S.**, Black R. (1982) Chaotic patchiness in an intertidal limpet, *Siphonaria* sp. *Marine Biology* 70: 157-164.
- Joshi J.**, Schmid B., Caldeira M.C., et al. (18 co-authors). 2001. Local adaptation enhances performance of common plant species. *Ecology Letters*. 4: 536-544.
- Jules E.S.**, Shahani P. (2003) A broader ecological context to habitat fragmentation: Why matrix habitat is more important than we thought? *Journal of Vegetation Science* 14: 459-464.
- Jump A.S.**, Penuelas J. 2005. Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters* 8: 1010-1020.
- Kawecki T.J.**, Ebert D. (2004) Conceptual issues in local adaptation. *Ecology letters* 7: 1225-1241.
- Kay R.**, Alder J., Brown D., Houghton P. (2003) Management Cybernetics: A New Institutional Framework for Coastal Management. *Management Coastal* 31: 213-227.
- Kimura M.** (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular Evolution*. Vol 16 (2)
- Kinlan B.P.**, Gaines S.D. (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology* 84:2007-2020.
- King O.H.** (1995) Estimating the value of marine resources: A marine recreation case. *Ocean & Coastal Management* 27: 129-141.
- Kjerfve B.**, Wiebe W.J., Kremer H.H., Salomons W., Marshall Crossland J.I.C., Morcom N., Harvey N., Marshall Crossland J.I.O. (2002) Caribbean Basins: LOICZ Global Change Assessment and Synthesis of River Catchment/Island-Coastal Sea Interactions and Human Dimensions; with a desktop study of Oceania Basins. LOICZ-IPO, Texel (Netherlands), 174 pp.

- Knick** S.T., Dobkin D.S., Rotenberry J.T., Schroeder M.A., Vander Haegen W.M., Van Riper C. III (2003) Teetering on the edge or too late? Conservation and research issues for avifauna of sagebrush habitats. *The Condor*, 105: 611-634.
- Kolar** C.S., Lodge D.M. (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution*. 16: 199-204.
- Kong** L., Anderson J.M., Ohm H.W. (2005) Induction of wheat defense and stress-related genes in response to *Fusarium graminearum*. *Genome*, 48: 29-40.
- Koutsikopoulos** C., Beillois P., Leroy C. & Taillefer F. (1998) Temporal trends and spatial structures of the sea surface temperature in the Bay of Biscay. *Oceanologica Acta* 21:335-344.
- Kump** L.R., Kasting J.F. et Crane R.G. 1999. *The Earth System*. Prentice Hall, Upper Saddle River, New Jersey, 351p
- Lafferty** K.D. & **Page** C.J. (1997) Predation of the endangered tidewater goby, *Eucyclogobius newberryi*, by the introduced African clawed frog, *Xenopus laevis*, with notes on the frog's parasites. *Copeia*, 3: 589-592.
- Laine** A.O., Sandler H., Andersin A.B., Stigzelius J. (1997) Long-term changes of macrozoobenthos in the Eastern Gotland Basin and the Gulf of Finland (Baltic Sea) in relation to the hydrographical regime. *Journal of Sea Research* 38: 135-159.
- Laine** A.O. (2003) Distribution of soft-bottom macrofauna in the deep open Baltic Sea in relation to environmental variability. *Estuarine Coastal and Shelf Science*, 57: 87-97.
- Lammens** J.J. (1967) Growth and reproduction of a tidal flat population of *Macoma balthica* (L.). *Netherlands Journal of Sea Research*, 3: 315-382.
- Lande** R., Engen S, Sæther B.E (2003) *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, New York.
- Larade** K., Storey K.B. (2004) Accumulation and translation of ferritin heavy chain transcripts following anoxia exposure in a marine invertebrate. *Journal of Experimental Biology*. 207: 1353-1360 .
- Lawton** J.H., Nee S., Letcher A.J., Harvey P.H. (1994) Animal distributions: Patterns and processes. In *Large-scale ecology and conservation biology*. eds Edwards P.J, May R.M, Webbs N.R, London: Blackwell scientific Publications, 41:58
- Layne** J.N. (1997) Nonindigenous mammals. Pages 157-186 in Simberloff D, Schmitz DC, Brown TC, eds. *Strangers in Paradise*. Washington (DC): Island Press.
- Leppäkoski** E. (1994) Non-indigenous species in the Baltic Sea. In C. F. Boudouresque, F. Briand & C. Nolan (eds), *Introduced species in European coastal waters*, pp. 60-66. European Commission, Bruxelles.
- Levins** R. (1969) Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. *Bulletin of the ESA*, 15, 237:240.
- Levinson** G., Gutman G.A. (1987) High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acides Research* 15: 5323-5338.
- Librado** P. and **Rozas** J. (2009) DNAsp v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452
- Lidicker** W.Z. Jr (1988) Solving the enigma of microtine «cycles». *Journal of mammalogy*, 69: 225-235.
- Lindquist** S. (1986) The heat-shock response. *Annu. Rev. Biochem.*, 55, 1151-1191
- Liu** X. & **Theil** E.C. (2005) Ferritins: Dynamic Management of Biological Iron and Oxygen Chemistry. *Accounts of Chemical Research* 38: 167-175
- Lomolino** M., Riddle B., Brown J. (2005) *Biogeography*. Sinauer Associates, Inc, Sunderland, Massachusetts, third edition.
- Lubchenco** J. (2003) The Blue Revolution: A Global Ecological Perspective. *World Aquaculture Magazine*, 34(4). Guest editorial.
- Luttikhuisen** P.C., Drent J., van Delden W., Piersma T. (2003a) Spatially structured genetic variation in a broadcast-spawning bivalve: quantitative versus molecular traits. *Journal of Evolutionary Biology*, 16, 260-272.
- Luttikhuisen** P.C., Drent J. and Baker A.J. (2003b). Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. *Molecular Ecology* 12, 2215-2229

- Mack** R. N. (1985) Invading plants: their potential contribution to population biology. In: Studies on plant demography: a festschrift for John L. Haper. Ed. J While. Academic Press, London, 127-142.
- Malecki** R.A., Blossey B., Hight S.D., Schroeder D., Kok L.T., Coulson J. (1993) Biological control of purple loosestrife. *BioScience*, 43:680-686.
- Margulies** M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380.
- Margush** T, McMorris F.R (1981) Consensus-trees. *Bulletin of Mathematical Biology* 43 : 234-244.
- Marra** P.P., Francis C.M., Mulvihill R.S., Moore F.R. (2005) The influence of climate on the timing and rate of spring bird migration. *Oecologia*, 142: 307-305.
- McCarty** J.P (2001) Ecological consequences of recent climate change. *Conservation Biology*, 15: 320-331.
- McCoid** M.J., Kleberg C. (1995) Non-native reptiles and amphibians. Pages 433-437 in LaRoe E.T., Farris G.S., Puckett C.E., Doran P.D., Mac M.J., eds. *Our Living Resources: A Report to the Nation on the Distribution, Abundance, and Health of US Plants, Animals, and Ecosystems*. Washington (DC): US Department of the Interior, National Biological Service.
- McMillan** M., Wilcove D. (1994) Gone but not forgotten: why have species protected by the endangered species act become extinct? *Endangered Species Update*, 11, 5.
- Millenium Ecosystem Assesment** (2005) *Ecosystems and Human Well-Being: synthesis*. Island Press, Washington, DC.
- Mattila** J. & **Bonsdorff** E. (1998) Predation by juvenile flounder (*Platichthys flesus* L.): a test of prey vulnerability, predator preference, switching behaviour and functional response. *Journal of Experimental Marine Biology and Ecology*. Vol 227, Issue 2, 30 September 1998, Pages 221-236.
- Meehan** B.W. (1985) Genetic comparison of *Macoma balthica* (Bivalvia, Telinidae) from the eastern and western North Atlantic Ocean. *Marine Ecology Progress Series*, 22: 69-76.
- Menzel** A. *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, 12: 1969-1976.
- Metzker** M.L. (2010) Sequencing technologies the next generation. *Nature reviews Genetics* 11: 31-46.
- Meijer** T (1993) Stratigraphical notes on *Macoma* (Bivalvia) in the southern part of the North Sea Basin and some remarks on the arrival of Pacific species. *Scripta Geologica, Special Issue*, 2, 297-312.
- Meyers** L. *et al.* (2002) Fighting change with change: adaptive variation in an uncertain world. *Trends in Ecology and Evolution*, 17: 551-557
- Moberg** F., Ronnback P. (2003) Ecosystem services of the tropical seascape: interactions, substitutions and restoration. *Ocean & Coastal Management* 46: 27-46.
- Mooney** H.A., Cleland E.E. (2001) The evolutionary impact of invasive species. *PNAS* 98: 5446-5451.
- Morgan** L.E., Wing S.R., Botsford L.W., Lundquist C.J., Diehl J.M. (2000) Spatial variability in red sea urchin (*Strongylocentrotus franciscanus*) recruitment in northern California. *Fisheries Oceanography* 9: 83-98.
- Morgan** C., Owens N. (2001) Benefits of water quality policies: the Chesapeake Bay. *Ecological Economics*, 39: 271-284.
- Morgante** M. & **Olivieri** A.M. (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* 3: 175-182.
- Morin** N. (1995) Vascular plants of the United States. Pages 200-205 in LaRoe E.T., Farris G.S., Puckett C.E., Doran P.D., Mac M.J., eds. *Our Living Resources: A Report to the Nation on the Distribution, Abundance, and Health of US Plants, Animals, and Ecosystems*. Washington (DC): US Department of the Interior National Biological Service.
- Moritz** C. (1994) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution*
- Morse** L.E., Kartesz J.T., Kutner L.S. (1995) Native vascular plants. Pages 205-209 in LaRoe E.T., Farris S., Puckett C.E., Doran P.D., Mac M.J., eds. *Our Living Resources: A Report to the Nation on the Distribution, Abundance, and Health of US Plants, Animals, and Ecosystems*. Washington (DC): US Department of the Interior National Biological Service.

- Nei M.**, Maruyama T., Chakraborty R. (1975) The bottleneck effect and genetic variability in populations. *Evolutions* 29: 1-10.
- Nei M.** (1987) *Molecular evolutionary genetics*. Columbia University Press, New-York.
- Nikula R.**, Strelkov P. and Väinölä R. (2008) A broad transition zone between an inner Baltic hybrid swarm and a pure North Sea subspecies of *Macoma balthica* (Mollusca, Bivalvia). *Molecular Ecology* (2008) 17, 1505-1522
- Nixon, S.W.**, Granger, S.L., Nowicki, B.L. (1996) An assessment of the annual mass balance of carbon, nitrogen, and phosphorus in Narragansett Bay. *Biogeochemistry*, 31: 15-61.
- National Council Research** 2002
- O'Donnell M.**, Hammond L., Hofmann G. (2009) Predicted impact of ocean acidification on a marine invertebrate: elevated CO₂ alters response to thermal stress in sea urchin larvae. *Marine Biology*, 156: 439-446
- Osborn T.J.** (2006) Recent variations in the winter North Atlantic Oscillation. *Weather*, 61: 353-355
- OTA** (1993) *Harmful Non-Indigenous Species in the United States*. Washington (DC): Office of Technology Assessment, US Congress.
- Otero J.H.** & Milan F.J. (1970) Distribucion de los moluscos: gasteropodos y pelecipodos, marinos, de las costas de Galicia. *Cuad. Biol.* 1, 79-93.
- Palumbi S.R.** (1992) Marine speciation on a small planet. *Trends in Ecology and Evolution*. 7:114-118.
- Parmesan C.**, Yohe G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421, 37-42
- Petit et al.** 2003 Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science* 6 June 2003: Vol. 300. no. 5625, pp. 1563 - 1565
- Philippart C.J.M.**, Van Aken H.M., Beukema J.J., Bos O.G., Cadée G.C., Dekker R. (2003) Climate-related changes in recruitment of the bivalve *Macoma balthica*. *Limnology and Oceanography*, 48, 2171-2185.
- Phillips O.L.** (1996) Long-term environmental change in tropical forests: increasing tree turnover. *Environmental Conservation*, 23: 235-248.
- Pielou E. C.** (1994) *A naturalist's guide to the arctic*. University of Chicago Press, Chicago, IL.
- Piersma T.**, Beukema J.J. (eds) (1993) Tropic interactions between shorebirds and their invertebrate prey. *Neth J Sea Res* 31:299-512
- Piry S.**, Alapetite A., Cornuet J.M., Paetkau D., Baudouin L., Estoup A. (2004) GeneClass2 : a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, 95: 536-539.
- Plath M.**, Hauswaldt J.S., Moll K., Tobler M., Garcia De Leon F.J., Schlupp I., Tiedemann R. (2007) Local adaptation and pronounced genetic differentiation in an extremophile fish, *Poecilia mexicana*, inhabiting a Mexican cave with toxic hydrogen sulphide. *Molecular Ecology* 16: 967-976
- Por F.D.** (1978) Lessepsian migration: the influx of Red Sea biota into the Mediterranean. Springer, Berlin, Germany.
- Primavera J.H.** (1991) Intensive Prawn Farming in the Philippines Ecological, Social, and Economic-Implications. *Ambio*, 20: 28-33.
- Quaintenne G.**, Van Gils J.A., Bocher P., Dekinga A., Piersma T. (2010) Diet selection in a molluscivore shorebird across Western Europe: does it show short- or long-term intake rate-maximization? *Journal of Animal Ecology* 79: 53-62
- R Development Core Team** (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>
- Rajdev S.**, Sharp F.R. (2000) Stress proteins as molecular markers of neurotoxicity. *Toxicologic Pathology*. 28: 105-112.
- Rainbow P.S.**, Fialkowski W., Sokolowski A., Smith B.D., Wolowicz M. (2004). Geographical and seasonal variations of trace metal bioavailabilities in the Gulf of Gdańsk, Poland using mussels (*Mytilus trossulus*) and barnacles (*Balanus improvisus*) as biomonitors. *Marine Biology*

Rassmann K. & Schlottere C. (1991) Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, 12: 113-118.

Rannala B. & Mountain J. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, 94: 9197-9201.

Raybould A.F., Gray A.J., Lawrence M.J., Marshall D.F. (1991a) The evolution of *Spartina anglica* C.E. Hubbard (Gramineae): origin and genetic variability. *Biology Journal of the Linnean Society*. 43: 111-126.

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.

Reise K., Gollasch S., Wolff W.J. (1999) Introduced marine species of the North Sea coasts. *Helgoland Marine Research*, 52: 219-234.

Rheinheimer G. (1995) *Meereskunde der Ostsee*. 2nd ed. Springer verlag, Berlin.

Rhymer J.M., Simberloff D. (1996) Extinction by Hybridization and Introgression. *Annual Review of Ecology and Systematics*, 27: 83-109.

Ricciardi A., MacIsaac H.J. (2000) Recent mass invasion of the north American Great Lakes by Ponto-Caspian species. *Trends in Ecology and Evolution*, 15: 62-65.

Richardson D.M., Pysek P., Rejmanek M., Barbour M.G., Dane Panetta F., West C.J. (2000) Naturalization and Invasion of Alien Plants: Concepts and Definitions. *Diversity and Distributions*, 6: 93-107.

Ronce & Kirckpatrick (2001) When sources become sinks: migrational meltdown in heterogeneous habitats *Evolution*, Vol. 55, No. 8, pp. 1520-1531

Ronnback P. (1999) The ecological basis for economic value of seafood production supported by mangrove ecosystems. *Ecological Economics*, 29 : 235-252.

Roughgarden J., Gaines S., Possingham H. (1988) Recruitment dynamics in complex life cycles. *Science* 241: 1460-1466.

Rueffler C., Van Dooren T.J.M., Leimar O. (2006) Disruptive selection and then what ? *Trends in Ecology and Evolution*. 21: 238-245.

Rubolini D., Møller A.P., Rainio K., Lehikoinen E. (2007b) Intraspecific consistency and geographic variability in temporal trends of spring migration phenology among European bird species. *Climate research*, 35: 135-146.

Ruiz G.M., Carlton J.T., Grosholz E.D., Hines A.H. (1997) Global invasions of marine and estuarine habitats by non-indigenous species: mechanisms, extent and consequences. *American Zoology*, 37: 621-632.

Ruiz G.M., Rawlings T.K., Dobbs F.C., Drake L.A., Mullady T., Huq A., Colwell R.R. (2000) Global spread of microorganisms by ships: ballast water discharged from vessels harbours a cocktail of potential pathogens. *Nature*, 408: 49-50.

Rumohr H., Bonsdorff E., Pearson T.H. (1996) Zoobenthic succession in Baltic sedimentary habitats. *Archive of Fishery and Marine Research* 44: 179-214.

Ryan J.A., Hightower L.E. (1996) Stress proteins as molecular biomarkers for environmental toxicology. *EXS* 77, 411-424.

Rychlik W., Rychlik P. (2000) Oligo primer analysis software. Version 6.0.1. Cascade, Colorado. Molecular Biology Insights, Inc.

Saitou N. & Nei M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, Vol 4, 406-425

Santoro M.G. *Heat shock factors* and the control of the stress response. *Biochem. Pharmacol.*, 2000, 59, 55-63.

Scarlatto O.A. (1981) *The Bivalve Molluscs of the Boreal Western Pacific*. Nauka, Leningrad.

Schimel D.S. (1995) Terrestrial ecosystems and the carbon cycle. *Global Change Biology*, 1: 77-91.

Schroter (2005) Ecosystem Service Supply and Vulnerability to Global Change in Europe. *Science*, 310: 1333-1337.

Seather B.E., Tufto J., Jerstad K., Rostad O.W., Skatan J.E. (2000) Population

- dynamical consequences of climate change for a small temperate songbird. *Science*, 287: 854-856.
- Segerstrale S.G. (1960) Investigations on Baltic populations of the bivalve *Macoma balthica* (L.). Part I. *Soc. Sci. Penn. Comm. Biol.*, 32 : 1-72
- Siegel D.A., Kinlan B.P., Gaylord B., Gaines S.D. (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series* 260: 83-96.
- Sinelschikova A., Kosarev V., Panov I., Baushev A.N. (2007) The influence of wind conditions in Europe on the advance in timing of the spring migration of the song thrush (*Turdus philomelos*) in the south-east Baltic region. *International Journal of Biometeorology*, 51: 431-440.
- Sokolowski A., Wolowicz M., Hummel H., Bogaards R. (1998) Physiological responses of *Macoma balthica* to copper pollution in the Baltic. *Oceanologica Acta* 22: 431-439.
- Sokolowski A., Fichet D., Garcia-Meunier P., Radenac G., Wolowicz M., Blanchard G. (2002) The relationship between metal concentrations and phenotypes in the Baltic clam *Macoma balthica* (L.) from the Gulf of Gdańsk, southern Baltic. *Chemosphere* 47: 475-484.
- Sokolowski A., Wolowicz M., Hummel H. (2007) Metal sources to the Baltic clam *Macoma balthica* (mollusca: Bivalvia) in the southern Baltic Sea (the Gulf of Gdańsk). *Marine Environmental Research* 63: 236-256.
- Sokołowski** A., Pawlikowski K., Wołowicz M., Garcia P., Namieśnik J. (2008) Shell deformations in the Baltic clam *Macoma balthica* from southern Baltic sea (the Gulf of Gdansk): hypotheses on environmental effects. *AMBIO: A Journal of the Human Environment* 37: 93-100
- Soule** M.E., editor. **1987**. *Viable populations for conservation*. Cambridge University Press, Cambridge, England
- Spurgeon** J.P.G (1992) The Economic Valuation of Coral Reefs. *Marine Pollution Bulletin*, 24: 529-536.
- Stenseth** N.C., Ims R.A (1993) *Biology of Lemmings*. London: Academic Press.
- Stow** C., Carpenter R., Madenjian C.P., Eby L.A., Jackson L.J (1995) *Bioscience*, 45: 752.
- Strathmann** R.R. (1985) Feeding and Nonfeeding Larval Development and Life-History Evolution in Marine Invertebrates. *Annual Review of Ecology and Systematics*. 16:339-361
- Strelkov** P., Nikula R. and Väinölä R. 2007 *Macoma balthica* in the White and Barents Seas: properties of a widespread marine hybrid swarm (Mollusca: Bivalvia). *Molecular Ecology* (2007)16, 4110-4127
- Sword** G.A., Simpson S.J., El Hadi O.T., Wilps H. (2000) Density-dependent aposematism in the desert locust. *Proceedings Biological Sciences*, 267: 63-68.
- Szefer** P., Glasby G.P., Pempkowiak J., Kaliszan R. (1995) Extraction studies of heavy-metal pollutants in surficial sediments from the southern Baltic Sea of Poland. *Chemical Geology* 120: 111-126.
- Szefer P.** (2002) *Metals, metalloids and radionuclides in the Baltic ecosystem*. Elsevier, Amsterdam.
- Temple** S.A. (1992) Exotic birds: A growing problem with no easy solution. *The Auk*, 109: 395-397.
- The Gene Ontology Consortium.** (2000) Gene ontology: tool for the unification of biology. *Nature Genetics*.25: 25-9.
- Thomas**, C.D. and **Kunin**, W.E. (1999) The spatial structure of populations. *J. Anim. Ecol.* 68, 647-657.
- Thomas** C.D. (2005) Recent evolutionary effects of climate change. In: Lovejoy TE, Hannah L, editors. *Climate Change and Biodiversity*. New Haven, Connecticut: Yale University Press. p. 75-88
- Thompson** D.G., Stuckey R.L., Thompson E.B. (1987) Spread, Impact, and Control of Purple Loosestrife (*Lythrum salicaria*) in North American Wetlands. Washington (DC): US Fish and Wildlife Service. Fish and Wildlife Research Report no.2.
- Thompson** J.D., Higgins D.G. and Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, Vol. 22, No. 22 4673-4680
- Thorson** G. (1950) Reproductive and Larval Ecology Of Marine Bottom Invertebrates. *Biological Reviews* 25:1-45.
- Usenko** C.Y., Harper S.L., Tanguay R.L. (2008) Fullerene C₆₀ exposure elicits an

oxidative stress response in embryonic zebrafish. *Toxicology and Applied Pharmacology*, 15: 44-55.

Väänölä R., Varvio S.L. (1989) Biosystematics of *Macoma balthica* in northwestern Europe. in: Ryland JS, Tyler PA (eds). *Reproduction, Genetics and Distribution of Marine Organisms*. Proceedings of the 23rd European Marine Biology Symposium. Olsen and Olsen, Fredensborg, pp 309-317.

Väänölä R. (2003) Repeated trans-Arctic invasions in littoral bivalves: molecular zoogeography of the *Macoma balthica* complex. *Marine Biology* (2003) 143: 935-946

Valiela I., Bowen J.L., York J.K. (2001) Mangrove Forests: One of the World's Threatened Major Tropical Environments. *BioScience*, 51: 807-815.

Van Aken H.M. (2008) Variability of the water temperature in the western Wadden Sea on tidal to centennial time scales. *Journal of Sea Research*, 60: 227-234.

Van der Meer J., (2006) An introduction to Dynamic Energy Budgets (DEB) models with special emphasis on parameter estimation. *Journal of Sea Research*, 56, 85-102.

Van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535-538.

Van Wilgen B.W., Richardson D.M. (1985) The Effects of Alien Shrub Invasions on Vegetation Structure and Fire Behaviour in South African Fynbos Shrublands: A Simulation Study. *Journal of Applied Ecology*, 22: 955-966.

Veit R.R., Pyle P., McGowan J.A. (1996) Ocean warming and long-term change in pelagic bird abundance within the California current system. *Marine Ecology Progress Series*, 139: 11-18.

Veit R.R., McGowan J.A., Ainley D.G., Wahls T.T., Pyle P. (1997) Apex marine predator declines ninety percent in association with changing oceanic climate. *Global Change Biology*, 3: 23-28.

Villemant C., Haxaire J., Streito J.C. (2006) Premier bilan de l'invasion de *Vespa velutina* Lepeletier en France (Hymenoptera, Vespidae). *Bulletin de la Société entomologique de France*. 111: 535-538.

Visbeck M.H., Hurrell J.W., Polvani L., Cullen H.M. (2001). The North Atlantic Oscillation: Past, present, and future. *PNAS*, 98 : 12876-12877.

Vitousek P.M. (1990) Biological invasions and ecosystem processes: towards an integration of population biology and ecosystem studies. *Oikos*, 57: 7-13.

Vitousek P.M., Mooney H.A., Lubchenco J., Melillo J.M. (1997a) Human Domination of Earth's Ecosystems. *Science*, 277: 494-499.

Vitousek P.M., Aber J.D., Howarth R.W., Likens G.E., Matson P.A., Schindler D.W., Schlesinger W.H., Tilman D.G. (1997b) Human Alteration Of The Global Nitrogen Cycle: Sources And Consequences. *Ecological Applications*, 7:737-750.

Vucetich J.A., Waite T.A (2003) Spatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. *Conservation Genetics*, 4 : 639-645.

Walker D. and Kendrick G. (1998). Threats to macroalgal diversity: marine habitat destruction and fragmentation, pollution and introduced species. *Botanica Marina*, 41 (1) : 105-112.

Walther G.R., Post E., Convey P., Menzel A., Parmesan C., Beebee T.J.C., Fromentin J.M., Hoegh-Guldberg O., Bairlein F. (2002) Ecological responses to recent climate change. *Nature*, 416, 389-395

Watts R.J., Johnson M.S., Black R. (1990) Effects of recruitment on genetic patchiness in the urchin *Echinometra matchaei* in Western Australia. *Marine Biology* 105: 141-151.

Weir B. & Cockerham C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370.

Wetthey D.S., Woodin S.A. (2008) Ecological hindcasting of biogeographic responses to climate change in the European intertidal zone. *Hydrobiologia*, 606, 139-151.

Whisenant S.G. (1990). *Changing Fire Frequencies on Idaho's Snake River Plains: Ecological and Management Implications*. Logan (UT): US Department of Agriculture, Forest Service, Intermountain Research Center. General Technical Report INT-276.

Vogel K. (1959) Wachstumsunterbrechungen bei Lamellibranchiaten und Brachiopoden. *Neues Jahrbuch für Geologie und Paläontologie, Abhandlung B*, 109: 109-129.

World Conservation Union for threats of extinction, 2006

Wright S. (1951) The genetical structure of populations *Annals of Eugenics*, Vol. 15 , pp. 323-354.

Yang R.Y., Hsu D.K., Liu F.T. (1996) Expression of galectin-3 modulates T-cell growth and apoptosis. *Proceedings of the National Academy of Sciences* 93:6737-674

Zouros E., Freeman K.R., Ball A.O., Pogson G.H. (1992) Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature* 359: 412-414.

Annexes

*A- Compléments
méthodologiques*

Réalisation d'une banque de microsatellites enrichie en motif CA

Digérer 5 à 10 µg d'ADN avec RsaI (coupe CA/GT)

40 µl ADN (2 à 4 µg)

30 U de RsaI (3 µl)

10 µl de tampon (fourni avec l'enzyme)

1 µl de BSA (fournie avec l'enzyme)

H₂O qsp 100 µl

Digestion 7H00 à 37°C

Purification sur colonne Macherey nagel (Nucléospin)

Ligation des adaptateurs Rsa 21 et Rsa 25

Rsa 21 5' CTCTTGCTTACGCGTGGACTA 3'

Rsa 25 5' TAGTCCACGCGTAAGCAAGAGCACA 3'

Rsa 21 (1µM) 10 µl

Rsa 25 (1µM) 10 µl

ADN digéré 20 µl

T4 ligase (3U/µl) 3 µl

Tampon 10X 10 µl

H₂O qsp 100 µl

Ligation une nuit à 18°C

Purification sur colonne Macherey nagel (Nucléospin)

PCR Post Ligation

ADN purifié 7 µl

Tampon 10X 5 µl

MgCl₂ (25 mM) 6 µl

Rsa 21 (10 µM) 4,7 µl

DNTP (10 mM) 1 µl

Taq (5U/µl) 1 µl

H₂O qsp 50 µl

Programme PCR	95°C	4 min	} 20 cycles	sec
	94°C	30		
	60°C	1 min		
	72°C			
1min30				
	72°C	5 min		

Purification sur colonne Macherey nagel (Nucléospin)

Enrichissement en motif AC grâce à 1 oligo biotinylé en 5' (AC)10

dilution de l'ADN purifié (50 µl) dans 400 µl H₂O

dénaturation 10 min à 95°C

ajout 150 pmole (1,5 µl d'1 solution 100 µM) de l'oligo + 13 µl tampon SSC 20X (0.5X final)

incuber 20 min à T° ambiante avec agitation légère

préparation des billes magnétiques streptavidine

3 lavages successifs avec 300µl de tampon SSC 0,5X

reprise des billes dans 100 µl de tampon SSC 0,5X

mélanger 100 µl de billes + ADN + oligo biotinylé

incuber 10 min à T° ambiante

4 lavages avec 300 µl de tampon SSC 0,1X)

resuspendre dans 100 puis 150 µl d'H₂O

Purification sur colonne Macherey nagel (Nucléospin)

PCR post-enrichissement

ADN purifié	10 µl
Tampon 10X	10 µl
MgCl ₂ (25 mM)	12 µl
Rsa 21 (10 µM)	8,7 µl
DNTP (10 mM)	2 µl
Taq (5U/µl)	2 µl
H ₂ O qsp	50 µl

Programme PCR	95°C	1 min	} 20 cycles
	94°C	40 sec	
	60°C	1 min	
	72°C	2min	
	72°C	5 min	

Purification sur colonne Macherey nagel (Nucléospin)

Ligation au plasmide pGEM-T easy

VT = 10 µl

Plasmide (50 ng/µl)	1µl
---------------------	-----

Tampon 2X	5µl
-----------	-----

ADN	3µl
-----	-----

Ligase (3U/µl)	1µl
----------------	-----

Ligation une nuit à 4°C

Transformation de bactéries compétentes DH5α avec le plasmide

décongeler 1 tube de bactéries compétentes dans la glace

ajouter 2 µl du plasmide, agiter en douceur en tapotant le tube

incuber 30 min minimum à 0°C

choc thermique au bain marie 1min30 (maxi) à 42°C et replacer à 0°C

ajouter 600 µl de LB, incuber 30 à 45 min à 37°C

étaier 200 µl sur des boites LB-agar contenant ampicilline (résistance conférée par le plasmide), IPTG(inducteur permettant l'utilisation de Xgal si le plasmide est vide), Xgal

incubation overnight à 37°C (étuve)

NB : si après incubation le nombre de colonie est trop important (et qu'on ne peut pas les isoler les unes des autres) réétaier les bactéries transformées en diminuant le volume (augmenter si pas assez de colonies)

Extraction ARN au Trizol

Homogénéiser avec un pilon en téflon dans 1 µtube RNase free 1 à 10 mg tissu dans 800 µl de Trizol : le volume tissulaire ne doit pas être supérieur à 10 % du volume total

S'il reste du tissu non lysé, centrifuger 10 min à 12000 g à 4°C et récupérer le surnageant (la pelote contient l'ADN de haut poids moléculaire, les membranes et les polysaccharides)

Incuber le tissu homogénéisé à T° ambiante pour permettre la dissolution complète des complexes nucleoprotidiques.

Ajouter 0,2 ml de chloroforme par ml de trizol et fermer les tubes. Agiter vigoureusement les tubes manuellement pendant 15 s et incuber 2 à 3 min à T° ambiante.

Centrifuger 10 minutes à 12000 g à 4°C ; le mélange est séparé en 1 phase inférieure rouge (phenol-chloroforme), une interphase et une phase supérieure incolore aqueuse. Cette phase contient l'ARN et représente environ 60 % du volume de trizol utilisé au départ.

Transférer la phase aqueuse dans 1 tube propre, ajouter 5 à 10 µg de RNase free glycogen (celui-ci va coprécipiter avec l'ARN)

Précipiter l'ARN en mélangeant avec de l'isopropanol (0,5 ml pour 1 ml de trizol). Incuber 10 min à T° ambiante et centrifuger 10 minutes à 12000 g à 4°C. Après la centrifugation, le précipité d'ARN forme une pelote le long du tube.

Ôter le surnageant, laver la pelote avec de l'éthanol 75% (1 ml pour 1 ml de trizol).

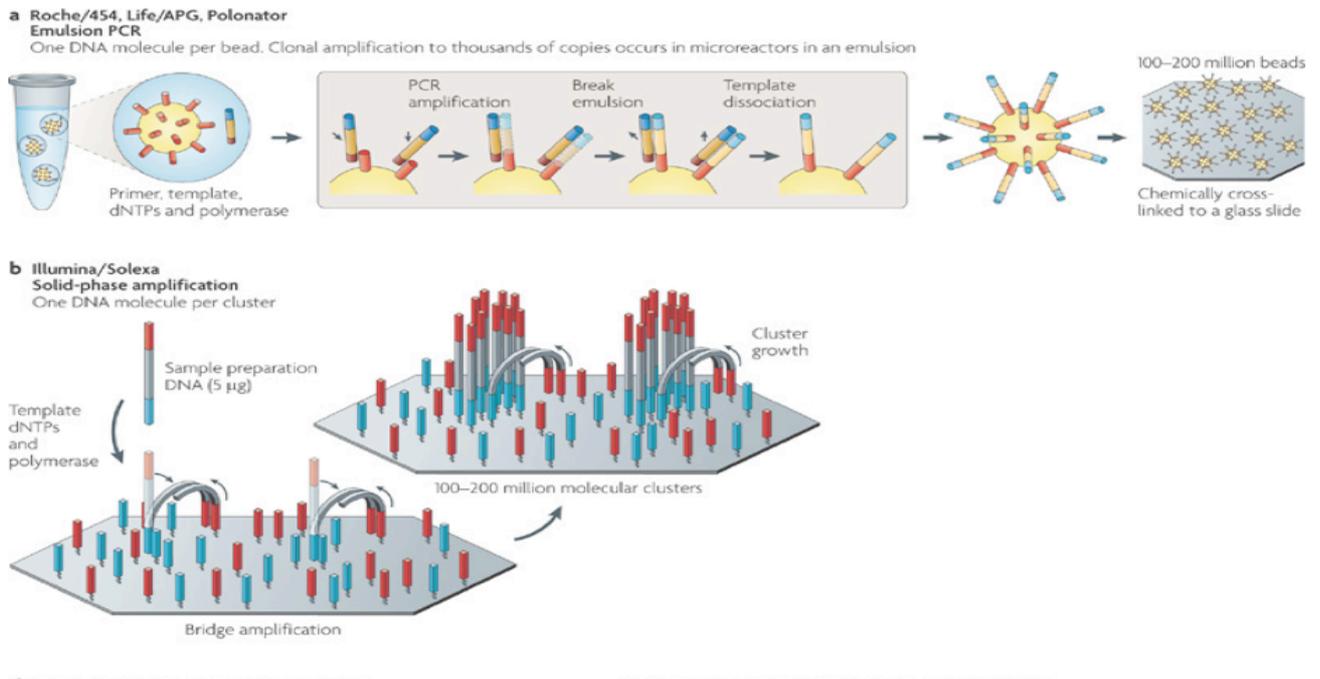
Vortexer et centrifuger 5 min à 7500 g à 4°C. Sécher l'ARN à l'air libre (maxi 10 minutes). La pelote ne doit pas être complètement sèche afin d'éviter une solubilité partielle (ce qui entraîne 1 ratio A260/280 < 1,6).

Dissoudre l'ARN dans de l'eau RNase free ou du SDS 0,5% (pas de SDS en cas d'utilisation de l'ARN dans des réactions enzymatiques) en pipetant plusieurs fois. Incuber 10 min à 55°C.

Conserver à -80°C.

Technologies NGS (*Next Generation Sequencing*)

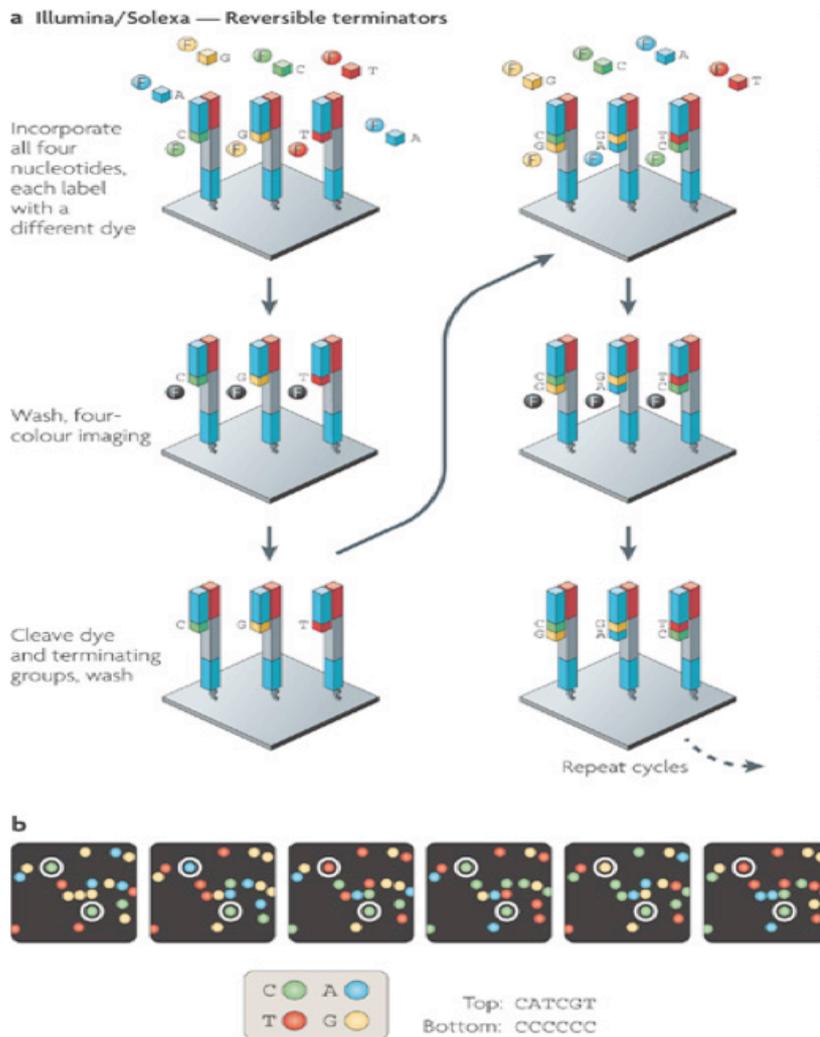
Immobilisation des échantillons



Stratégies d'immobilisation des échantillons et amplification: Dans l'*emulsion* PCR (emPCR, a), l'ADN est encapsulé dans des billes qui sont attachées chimiquement sur un support en verre (Life/APG) ou déposé dans des plaques picotitrées (Roche/454). L'amplification en phase solide (Illumina/Solexa, b) se décompose en deux étapes: une hybridation initiale suivi d'une élongation simple brin, puis une amplification "en pont" de l'échantillon immobilisé grâce aux amorces adjacentes

Source: M. Metzker (2010) Sequencing technologies: the next generation. Nature Review Genetics

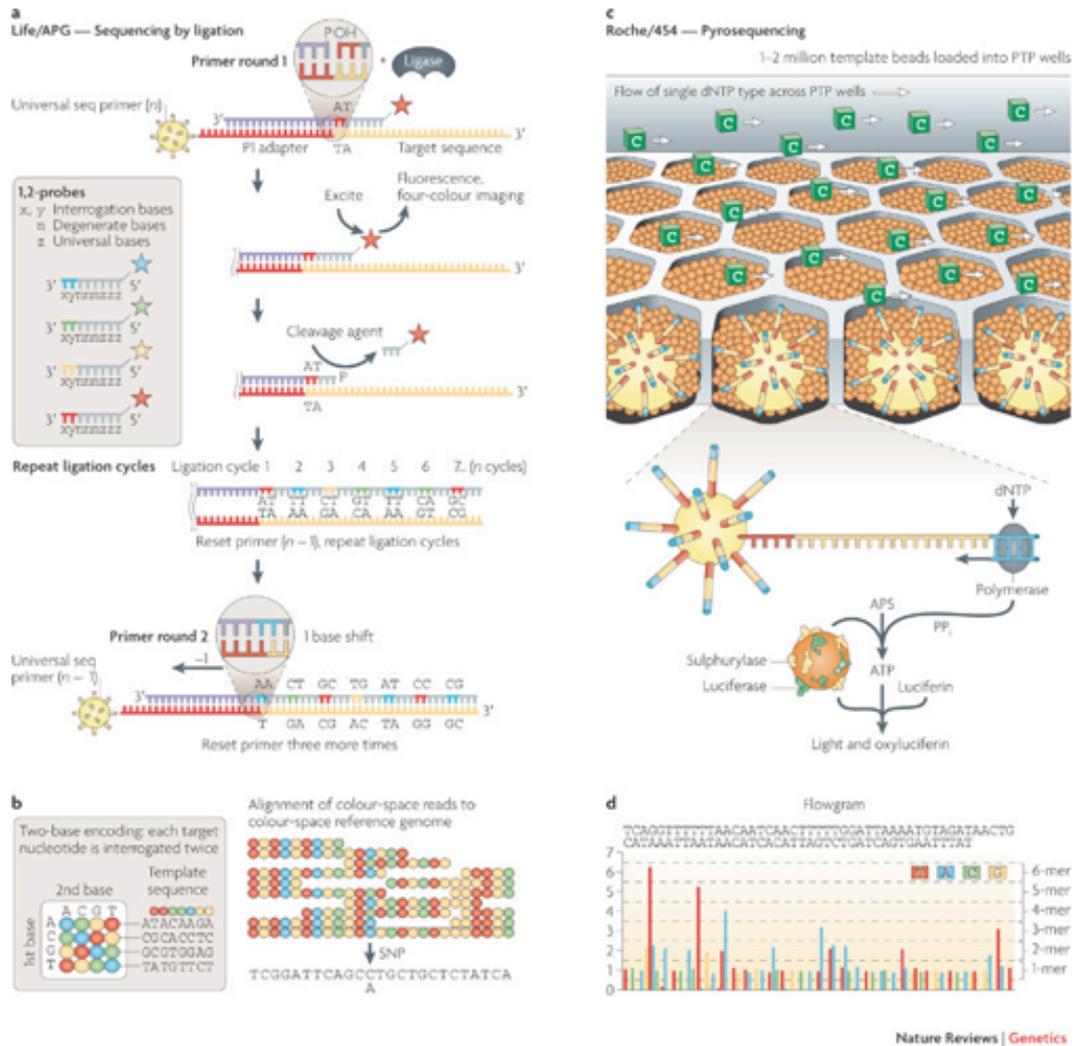
Technologies NGS (Next Generation Sequencing) Méthode CRT (Cyclic Reversible Termination)



La méthode CRT (Cycling Reversible Termination) à quatre couleurs est utilisée par la technologie Illumina / Solexa. Chaque nucléotide a une terminaison -O-azidométhyl réversible en 3'. Une fois la lecture réalisée, une étape de clivage permet d'éliminer les marqueurs fluorescents et de régénérer le groupe 3'-OH en utilisant l'agent réducteur tris (2-carboxy) phosphine (TCEP).

Source: M. Metzker (2010) Sequencing technologies: the next generation.
Nature Review Genetics

Technologies NGS (Next Generation Sequencing) Pyroséquence et méthode SBL (*Sequencing By Ligation*)



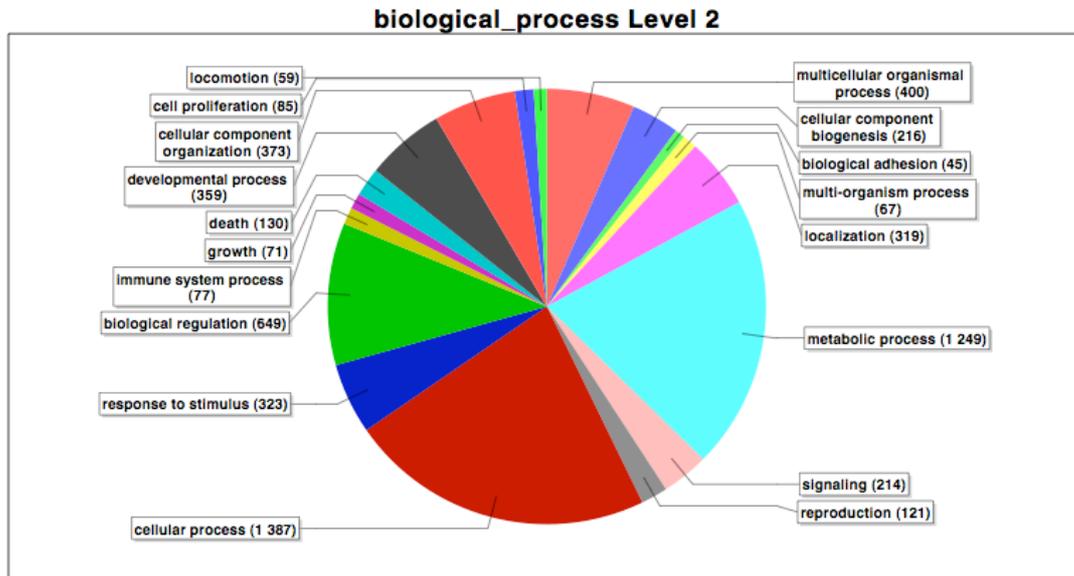
Source: M. Metzker (2010) Sequencing technologies: the next generation.
Nature Review Genetics

*B- Résultats de la Partie II,
Chapitre 2*

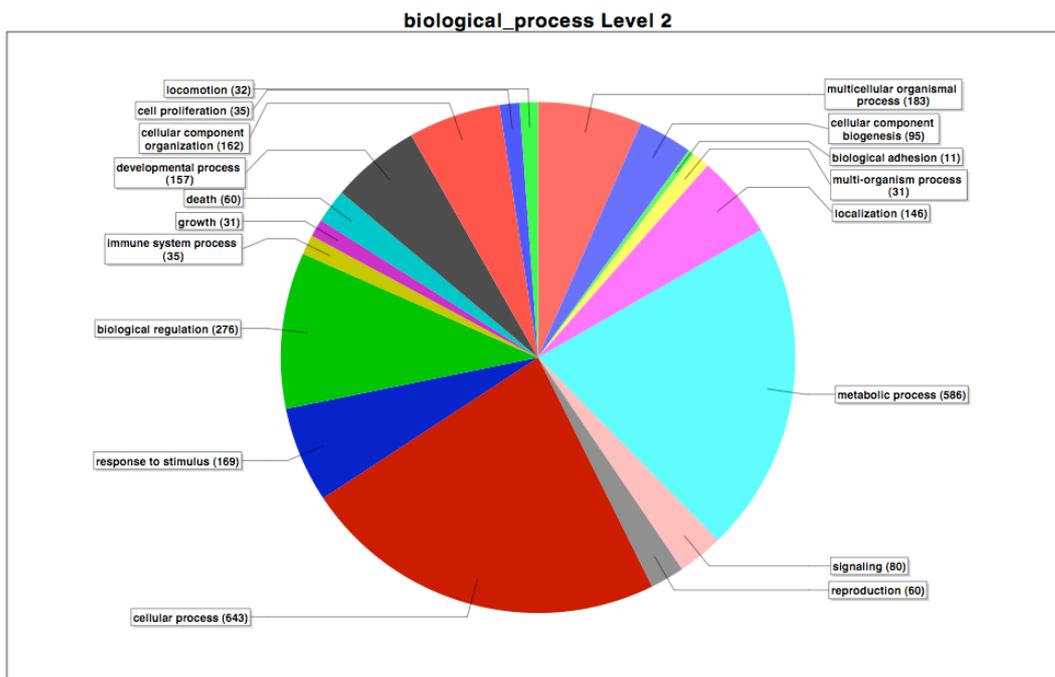
*Annotation fonctionnelle des
transcriptome : Gene Ontology*

*Activités peptidiques mises en
évidence*

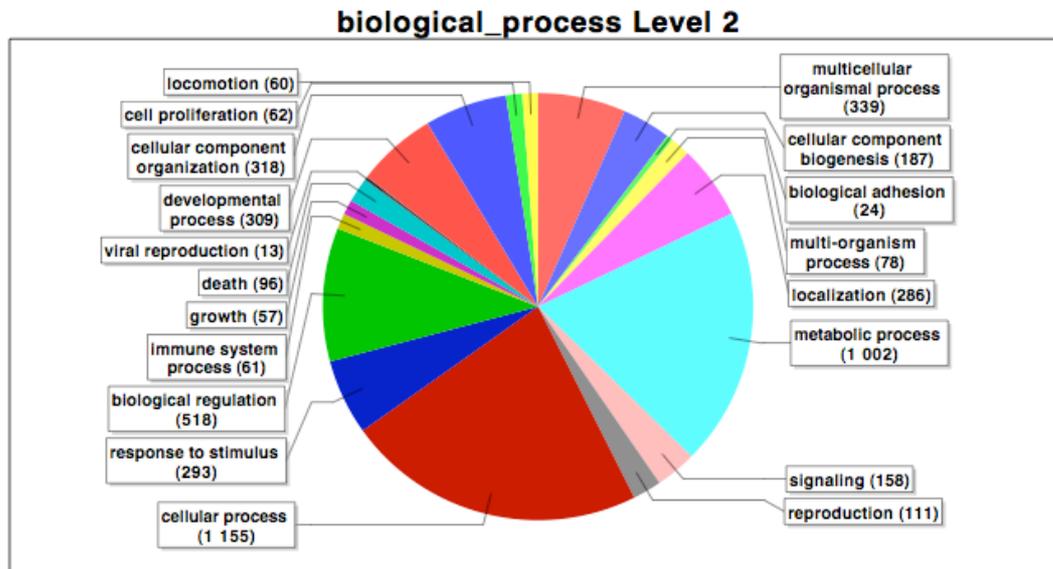
Annotation fonctionnelle du transcriptome de la population A : processus biologiques



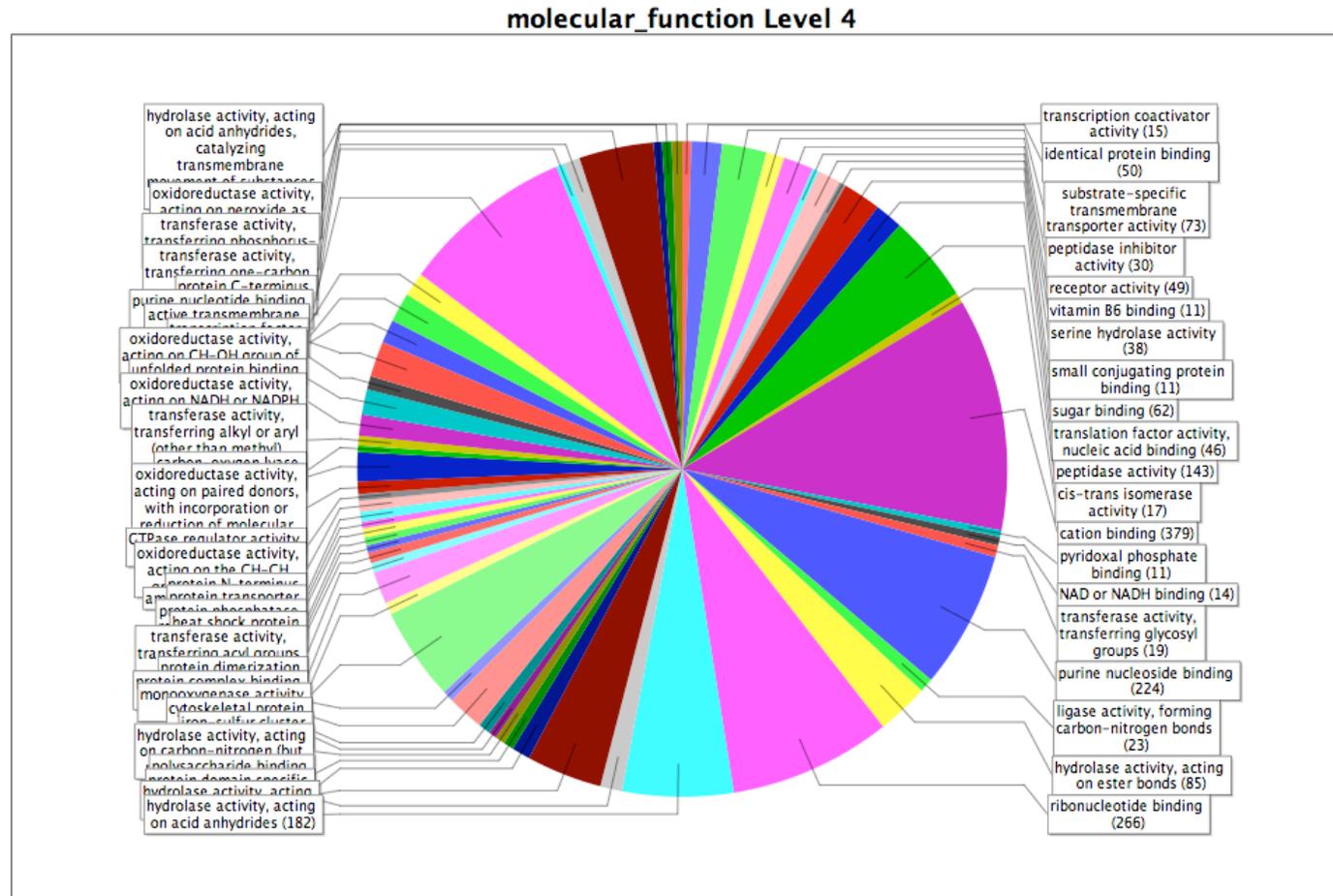
Annotation fonctionnelle du transcriptome de la population S : processus biologiques



Annotation fonctionnelle du transcriptome de la population G : processus biologiques

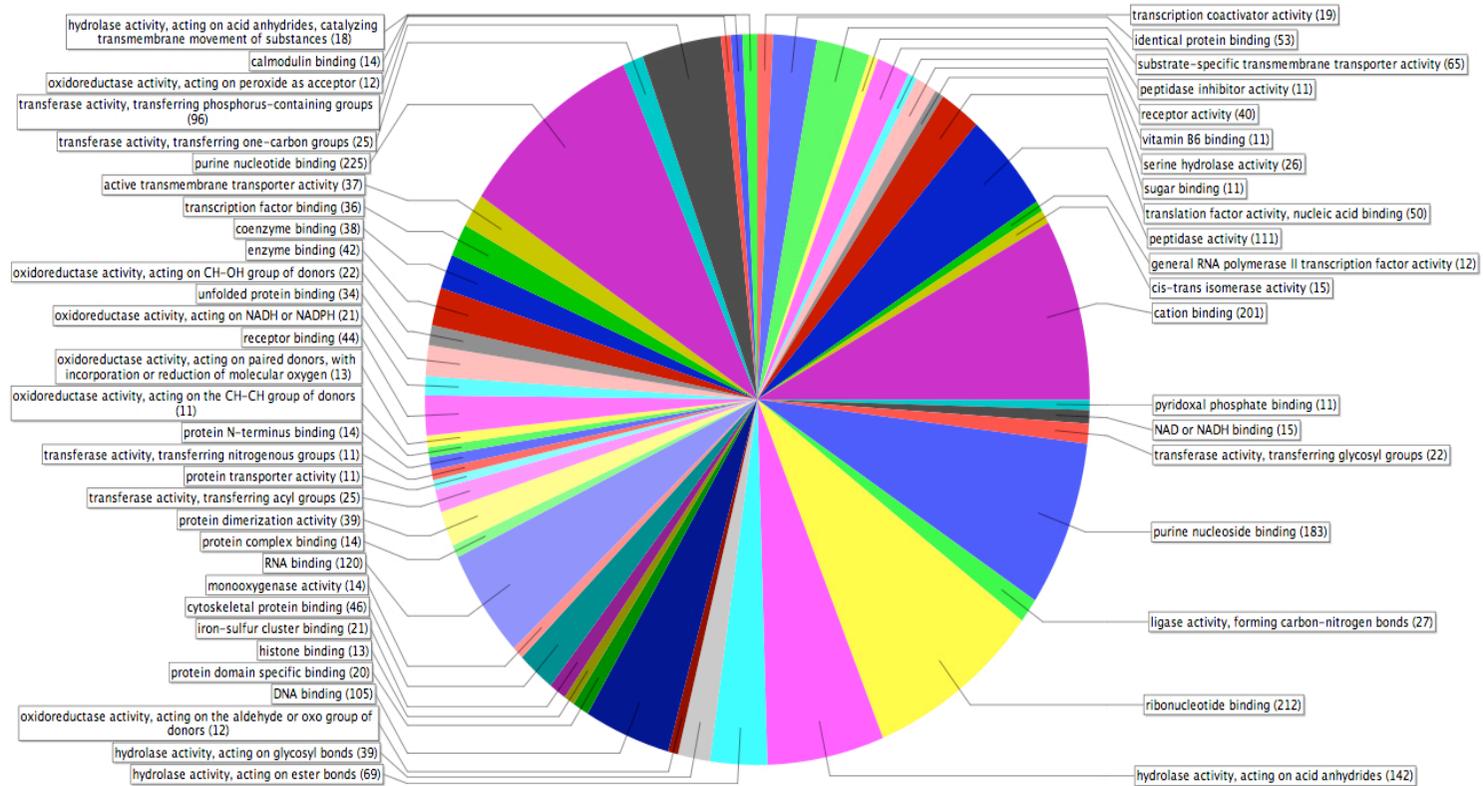


Annotation fonctionnelle du transcriptome de la population A : fonctions moléculaires



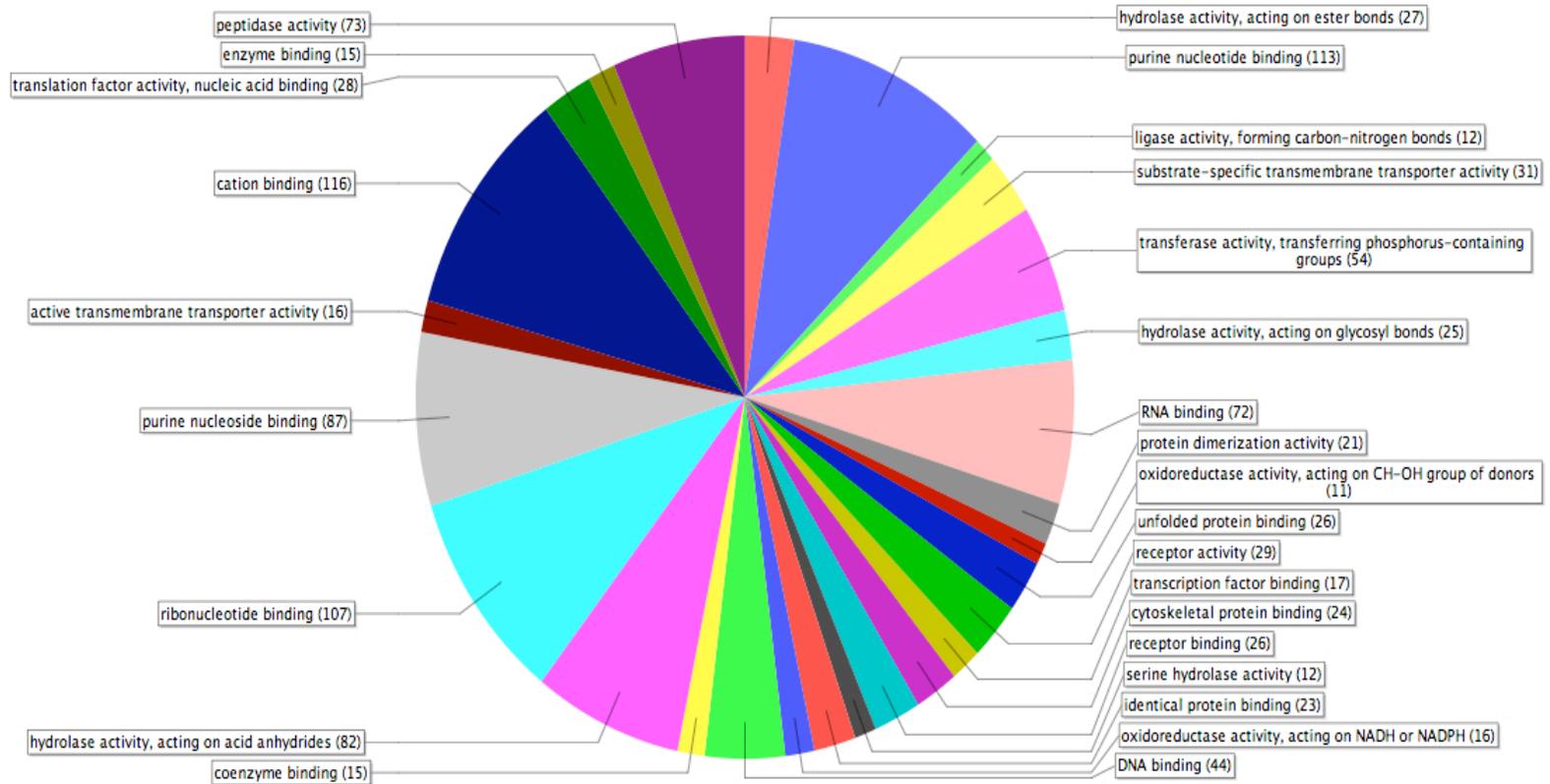
Annotation fonctionnelle du transcriptome de la population G : fonctions moléculaires

molecular_function Level 4



Annotation fonctionnelle du transcriptome de la population S : fonctions moléculaires

molecular_function Level 4



Liste de gènes candidats potentiellement impliqués dans la réponse cellulaire au stress

Implication hypothétique	Nom Contig	Nom Protéine
Mort cellulaire	AGSpops_A_c9551	apoptosis-linked gene 2
Mort cellulaire	AGSpops_G_c9348	cat eye syndrome critical region protein 2
Mort cellulaire	AGSpops_S_c113	defender against cell death 1
Mort cellulaire	AGSpops_G_c3085	eukaryotic translation elongation factor 1 epsilon 1
Mort cellulaire	AGSpops_A_c4060	histidine triad nucleotide binding protein 2
Mort cellulaire	AGSpops_S_c10994	inhibitor of apoptosis protein 1
Mort cellulaire	AGSpops_G_c3415	loc559360 protein
Mort cellulaire	AGSpops_S_c4542	phosphoprotein enriched in astrocytes 15
Mort cellulaire	AGSpops_A_c2631	ras homolog gene member ad
Mort cellulaire	AGSpops_S_c10745	luc7-like protein 3
Défense	AGSpops_G_c3010	comm domain-containing protein 7
Défense	AGSpops_S_c5178	defense protein
Défense	AGSpops_A_c4742	peptidoglycan binding domain-containing protein
Défense	AGSpops_G_c2009	peptidoglycan recognition protein
Défense	AGSpops_S_c565	peptidoglycan recognition protein 1
Défense	AGSpops_G_c1997	peptidoglycan recognition protein s1s-like
Défense	AGSpops_S_c442	peptidoglycan-binding domain 1 protein
Défense	AGSpops_S_c6889	peptidoglycan-recognition protein s2
Défense	AGSpops_S_c1810	protein
Défense	AGSpops_G_c7420	pyridoxine biosynthesis protein
Défense	AGSpops_G_c5119	pyridoxine biosynthesis protein pdx1
Défense	AGSpops_A_c8312	receptor protein-tyrosine kinase partial
Défense	AGSpops_S_c8616	selenium binding protein
Défense	AGSpops_G_c10526	selenium binding protein 1
Défense	AGSpops_G_c11096	slit2-a protein
Défense	AGSpops_G_c10152	stress-induced-phosphoprotein 1
Défense	AGSpops_S_c8098	transcription elongation factor 1 homolog
Défense	AGSpops_S_c12097	transcription elongation factor b polypeptide 1
Défense	AGSpops_A_c2777	transcription elongation factor b polypeptide 2
Défense	AGSpops_G_c7980	transcription elongation regulator 1 isoform 2
Défense	AGSpops_A_c3834	transcription factor ap-1
Défense	AGSpops_S_c183	transcription factor iib

Défense	AGSpops_S_c5035	transcription factor protein
Défense	AGSpops_G_c10285	transcription initiation factor iia gamma chain
Défense	AGSpops_G_c2399	transcription initiation factor iif alpha subunit
Défense	AGSpops_S_c9708	transcription intermediary factor 1-
Défense	AGSpops_G_c4769	uncoupling protein 2
Défense	AGSpops_A_c3814	universal stress protein
Défense	AGSpops_S_c807	versican core
Défense	AGSpops_A_c3874	zinc finger protein 225
Défense	AGSpops_A_c384	chondroitin sulfate proteoglycan 2 variant v3
Domage à l'ADN	AGSpops_G_c5121	4-aminobutyrate aminotransferase
Domage à l'ADN	AGSpops_A_c8341	aprataxin
Domage à l'ADN	AGSpops_G_c9345	brain and reproductive organ-expressed (tnfrsf1a modulator)
Domage à l'ADN	AGSpops_S_c7738	brca1 brca2-containing subunit 3
Domage à l'ADN	AGSpops_G_c3607	brca1-a complex subunit merit40
Domage à l'ADN	AGSpops_S_c31	btg member 2
Domage à l'ADN	AGSpops_G_c3095	cdk-activating kinase assembly factor mat1
Domage à l'ADN	AGSpops_G_c6191	deltex 3-like
Domage à l'ADN	AGSpops_G_c7013	denticleless homolog
Domage à l'ADN	AGSpops_A_c12003	dna replication licensing factor mcm2
Domage à l'ADN	AGSpops_A_c12559	eukaryotic initiation factor 4a
Domage à l'ADN	AGSpops_G_c5021	excision repair cross-complementing rodent repair complementation group 5 (xeroderma complementation group g (cockayn
Domage à l'ADN	AGSpops_A_c1046	guanine nucleotide-binding protein subunit beta-2-like 1
Domage à l'ADN	AGSpops_A_c9705	smell impaired 21f cg4710- isoform partial
Domage à l'ADN	AGSpops_G_c2859	uracil-dna glycosylase
Domage à l'ADN	AGSpops_S_c5290	chromosome 9 open reading frame 80
HSP	AGSpops_S_c227	10 kda heat shock mitochondrial-like
HSP	AGSpops_A_c10328	40 kda peptidyl-prolyl cis-trans isomerase
HSP	AGSpops_G_c6279	heat shock protein 22 isoform 1
HSP	AGSpops_G_c10642	heat shock protein 4
HSP	AGSpops_S_c8873	heat shock protein 40
HSP	AGSpops_S_c4505	heat shock protein 70
HSP	AGSpops_G_c7734	heat shock protein 9
HSP	AGSpops_S_c74	heat shock protein 90
HSP	AGSpops_S_c9007	heat-responsive protein 12
HSP	AGSpops_G_c6626	peptidylprolyl isomerase d
HSP	AGSpops_G_c2598	subfamily member 13

HSP	AGSpops_A_c2546	subfamily member 4
HSP	AGSpops_A_c4830	subfamily member 7
Protéases	AGSpops_G_c2876	26 serine protease
Protéases	AGSpops_G_c2650	26s protease regulatory subunit
Protéases	AGSpops_A_c2613	26s protease regulatory subunit 4
Protéases	AGSpops_G_c3746	26s protease regulatory subunit 6a
Protéases	AGSpops_A_c3109	26s protease regulatory subunit 6b
Protéases	AGSpops_A_c4279	26s protease regulatory subunit 7
Protéases	AGSpops_G_c5528	26s protease regulatory subunit 8
Protéases	AGSpops_S_c7456	26s protease regulatory subunit s10b
Protéases	AGSpops_A_c3959	26s proteasome non-atpase regulatory subunit 14
Protéases	AGSpops_G_c8543	26s proteasome non-atpase regulatory subunit 2
Protéases	AGSpops_G_c4714	26s proteasome non-atpase regulatory subunit 3
Protéases	AGSpops_A_c10776	26s proteasome non-atpase regulatory subunit 4
Protéases	AGSpops_S_c5032	26s proteasome non-atpase regulatory subunit 7
Protéases	AGSpops_A_c2775	26s proteasome non-atpase regulatory subunit 8
Protéases	AGSpops_A_c2795	26s proteasome subunit s9
Protéases	AGSpops_G_c14603	a chain elaborate manifold of short hydrogen bond arrays mediating binding of active site-directed serine protease inh
Protéases	AGSpops_S_c8706	a chain structural consequences of accommodation of four non- cognate amino-acid residues in the s1 pocket of bovine trypsin a
Protéases	AGSpops_G_c2947	alpha-aspartyl dipeptidase
Protéases	AGSpops_S_c5051	antistasin
Protéases	AGSpops_A_c3847	apical endosomal glycoprotein
Protéases	AGSpops_A_c5959	astacin-like squid metalloprotease type iii
Protéases	AGSpops_G_c5908	calcium binding egf domain containing protein
Protéases	AGSpops_S_c488	chymotrypsin b1
Protéases	AGSpops_G_s17605	chymotrypsin-2 (chymotrypsin ii)
Protéases	AGSpops_G_c402	chymotrypsin-like
Protéases	AGSpops_G_c1491	chymotrypsin-like elastase member 3b-like
Protéases	AGSpops_S_c1359	chymotrypsin-like protease ctrl-1
Protéases	AGSpops_A_c125	chymotrypsin-like protease ctrl-1 precursor
Protéases	AGSpops_G_c14843	chymotrypsinogen 2
Protéases	AGSpops_S_c5484	chymotrypsinogen b1
Protéases	AGSpops_A_c147	chymotrypsinogen b2
Protéases	AGSpops_G_c2449	complement factor b precursor
Protéases	AGSpops_A_c2150	ctrl protein
Protéases	AGSpops_S_c5255	cubilin (intrinsic factor-cobalamin receptor)

Protéases	AGSpops_A_c7005	cysteine protease atg4a
Protéases	AGSpops_A_c74	elegans protein confirmed by transcript evidence
Protéases	AGSpops_A_c13499	elegans protein partially confirmed by transcript evidence
Protéases	AGSpops_G_c3646	flap endonuclease-1
Protéases	AGSpops_G_c9571	flap structure-specific endonuclease 1
Protéases	AGSpops_G_c2436	inter-alpha inhibitor h3
Protéases	AGSpops_G_c2148	inter-alpha-trypsin inhibitor heavy chain h3
Protéases	AGSpops_G_c14982	kallikrein plasma (fletcher factor) 1
Protéases	AGSpops_S_c476	kallikrein related-peptidase 6 precursor
Protéases	AGSpops_S_c9287	kuzbanian-like cg1964-pa
Protéases	AGSpops_A_c10672	leukocyte elastase inhibitor
Protéases	AGSpops_S_c5279	leukotriene a-4 hydrolase
Protéases	AGSpops_G_c10423	loc561562 protein
Protéases	AGSpops_G_c12143	math (meprin-associated traf homology) domain containing family member (math-33)-like
Protéases	AGSpops_G_c3606	matrix metalloproteinase
Protéases	AGSpops_A_c13848	membrane metallo-endoropeptidase-like 1
Protéases	AGSpops_S_c5064	neutral endopeptidase
Protéases	AGSpops_S_c11072	nuclear vcp-like protein
Protéases	AGSpops_G_c2781	periodic tryptophan protein 2 homolog
Protéases	AGSpops_G_c11197	presequence mitochondrial precursor
Protéases	AGSpops_G_c5892	prohormone convertase 1
Protéases	AGSpops_G_c5743	protease do
Protéases	AGSpops_A_c7504	proteasome (macropain) 26s 3
Protéases	AGSpops_A_c2766	proteasome (macropain) 26s 5
Protéases	AGSpops_S_c9780	proteasome (macropain) 26s 6
Protéases	AGSpops_S_c10461	proteasome (macropain) 26s atpase 2
Protéases	AGSpops_A_c2791	proteasome (macropain) 26s non- 1
Protéases	AGSpops_G_c3161	proteasome (macropain) 26s non- 10
Protéases	AGSpops_G_c3477	proteasome (macropain) 26s non- 11
Protéases	AGSpops_S_c3331	proteasome (macropain) 26s non- 12
Protéases	AGSpops_A_c3902	proteasome (macropain) 26s non- 13
Protéases	AGSpops_G_c10765	proteasome (macropain) 26s non- 2
Protéases	AGSpops_A_c3758	proteasome (macropain) 26s non- 3
Protéases	AGSpops_S_c5172	proteasome (macropain) 26s non- 6
Protéases	AGSpops_G_c2397	proteasome (macropain) 26s non- 7
Protéases	AGSpops_G_c7702	proteasome (macropain) 26s non- 9

Protéases	AGSpops_A_c6793	proteasome (macropain) activator subunit 3 (pa28 gamma ki)
Protéases	AGSpops_A_c2442	proteasome (macropain) alpha 3
Protéases	AGSpops_G_c3008	proteasome (macropain) alpha 5
Protéases	AGSpops_A_c13224	proteasome (macropain) alpha 6
Protéases	AGSpops_S_c1741	proteasome (macropain) beta 1
Protéases	AGSpops_S_c5022	proteasome (macropain) beta 7
Protéases	AGSpops_A_c7474	proteasome (macropain) beta type 4
Protéases	AGSpops_A_c6052	proteasome alpha
Protéases	AGSpops_A_c5138	proteasome assembly chaperone 2
Protéases	AGSpops_A_c8106	proteasome subunit alpha
Protéases	AGSpops_S_c4599	proteasome subunit alpha type-2
Protéases	AGSpops_G_c2433	proteasome subunit alpha type-3
Protéases	AGSpops_G_c8582	proteasome subunit alpha type-4
Protéases	AGSpops_A_c2343	proteasome subunit alpha type-7
Protéases	AGSpops_A_c2433	proteasome subunit beta type
Protéases	AGSpops_S_c7187	proteasome subunit beta type-2
Protéases	AGSpops_A_c2429	proteasome subunit beta type-3
Protéases	AGSpops_S_c8906	proteasome subunit beta type-4 precursor
Protéases	AGSpops_A_c9525	proteasome subunit beta type-6 precursor
Protéases	AGSpops_G_c2019	proteasome subunit beta type-7 precursor
Protéases	AGSpops_A_c1748	protein
Protéases	AGSpops_S_c8766	serine or cysteine proteinase inhibitor clade b member 1
Protéases	AGSpops_S_c10052	serine protease
Protéases	AGSpops_A_c13241	serine protease inhibitor dipetalogastin
Protéases	AGSpops_G_c7316	serine proteases and ovochymase regions
Protéases	AGSpops_G_c479	serine proteinase
Protéases	AGSpops_G_c8178	serine threonine kinase 17a
Protéases	AGSpops_A_c12632	serine threonine kinase 22b
Protéases	AGSpops_G_c11492	serine threonine kinase 38
Protéases	AGSpops_A_c11213	serine threonine kinase 38 like
Protéases	AGSpops_S_c9326	serine threonine protein kinase
Protéases	AGSpops_G_c8444	serine threonine protein phosphatase expressed
Protéases	AGSpops_G_c12019	serine threonine tyrosine-interacting protein
Protéases	AGSpops_A_c7010	serine threonine-protein kinase
Protéases	AGSpops_G_c11535	serine threonine-protein kinase 6
Protéases	AGSpops_S_c9360	serine threonine-protein kinase nek10-like

Protéases	AGSpops_A_c5617	serine threonine-protein kinase nim1
Protéases	AGSpops_S_c4671	serine threonine-protein kinase pim-3
Protéases	AGSpops_A_c11877	serine threonine-protein kinase plk1
Protéases	AGSpops_G_c10358	serine threonine-protein kinase rio3
Protéases	AGSpops_S_c7334	serine threonine-protein phosphatase 2a 55 kda regulatory subunit b alpha isoform
Protéases	AGSpops_A_c3259	serine threonine-protein phosphatase 2a 56 kda regulatory subunit epsilon
Protéases	AGSpops_G_c10994	serine threonine-protein phosphatase 2a 65 kda regulatory subunit a alpha isoform
Protéases	AGSpops_S_c2667	serine threonine-protein phosphatase pp1
Protéases	AGSpops_S_c4891	serine threonine-protein phosphatase pp1-gamma catalytic subunit
Protéases	AGSpops_G_c5777	signal peptide peptidase
Protéases	AGSpops_S_c344	thioester-containing protein
Protéases	AGSpops_S_c5390	transmembrane serine 9
Protéases	AGSpops_A_c178	trya4_luccu_ame: full=trypsin alpha-4 flags: precursor
Protéases	AGSpops_S_c557	trydg_drome_ame: full=trypsin delta gamma flags: precursor
Protéases	AGSpops_A_c8322	venom c-type lectin mannose binding isoform 2 variant 1
Protéases	AGSpops_A_c3520	venom c-type lectin mannose binding isoform 5 variant 1
Protéases	AGSpops_A_c2642	venom serine protease
Protéases	AGSpops_S_c9172	vitamin k-dependent protein c
Protéases	AGSpops_S_c5673	zinc carboxypeptidase
Protéases	AGSpops_S_c4952	zinc carboxypeptidase a 1
Stress oxydatif	AGSpops_S_c6399	15-hydroxyprostaglandin dehydrogenase
Stress oxydatif	AGSpops_G_c3478	2-oxoglutarate and iron-dependent oxygenase domain-containing protein 1
Stress oxydatif	AGSpops_S_c4838	3-oxoacyl-acyl-carrier-protein reductase
Stress oxydatif	AGSpops_A_c3618	aldo-keto reductase
Stress oxydatif	AGSpops_G_c5563	aldo-keto reductase family 1 member b1
Stress oxydatif	AGSpops_S_c8450	alkylation repair homolog 6 (coli)
Stress oxydatif	AGSpops_G_c3727	apex nuclease (multifunctional dna repair enzyme) 1
Stress oxydatif	AGSpops_G_c6879	coenzyme q6 monoxygenase (cerevisiae)
Stress oxydatif	AGSpops_G_c4614	dehydrogenase reductase (sdr family) member 11
Stress oxydatif	AGSpops_G_c4072	dehydrogenase reductase (sdr family) member 7b
Stress oxydatif	AGSpops_S_c10212	dehydrogenase reductase sdr family member 1
Stress oxydatif	AGSpops_S_c5420	dehydrogenase reductase sdr family member 11 precursor
Stress oxydatif	AGSpops_S_c5326	dsba oxidoreductase
Stress oxydatif	AGSpops_G_c6443	dual oxidase 1
Stress oxydatif	AGSpops_G_c9385	fad-linked oxidoreductase
Stress oxydatif	AGSpops_G_c9468	ferric reductase domain-containing protein

Stress oxydatif	AGSpops_G_c3061	ferric-chelate reductase 1
Stress oxydatif	AGSpops_A_c1205	ferritin
Stress oxydatif	AGSpops_G_c2898	ferritin-like protein
Stress oxydatif	AGSpops_A_c11413	glutaredoxin-1
Stress oxydatif	AGSpops_G_c14963	glutathione peroxidase
Stress oxydatif	AGSpops_A_c13176	glutathione peroxidase 2
Stress oxydatif	AGSpops_A_c4184	glutathione peroxidase 6
Stress oxydatif	AGSpops_G_c3718	glutathione s- theta 3
Stress oxydatif	AGSpops_A_c1928	glutathione s-transferase
Stress oxydatif	AGSpops_G_c14312	glutathione s-transferase pi-class
Stress oxydatif	AGSpops_S_c6914	glutathione s-transferase theta 1
Stress oxydatif	AGSpops_S_c518	microsomal glutathione s-transferase 1
Stress oxydatif	AGSpops_G_c6552	microsomal glutathione s-transferase 2
Stress oxydatif	AGSpops_S_c8871	microsomal glutathione s-transferase 3
Stress oxydatif	AGSpops_S_c6347	nucleoredoxin
Stress oxydatif	AGSpops_G_c3183	nucleoredoxin 1
Stress oxydatif	AGSpops_S_c7123	peroxiredoxin 5
Stress oxydatif	AGSpops_S_c11919	peroxiredoxin 6
Stress oxydatif	AGSpops_A_c11757	peroxiredoxin- mitochondrial precursor
Stress oxydatif	AGSpops_S_c323	pi-class glutathione s-transferase
Stress oxydatif	AGSpops_S_c8321	pyrroline-5-carboxylate reductase 1
Stress oxydatif	AGSpops_S_c6092	pyrroline-5-carboxylate reductase-like protein c14orf148
Stress oxydatif	AGSpops_A_c5786	ras-related c3 botulinum toxin substrate 2 (rho small gtp binding protein rac2)
Stress oxydatif	AGSpops_G_c2634	sarcosine oxidase
Stress oxydatif	AGSpops_G_c7004	sarcosine oxidase
Stress oxydatif	AGSpops_A_c11837	selenium-dependent glutathione peroxidase
Stress oxydatif	AGSpops_G_c11704	selenoprotein 1
Stress oxydatif	AGSpops_A_c4513	sh3 domain-binding glutamic acid-rich-like protein
Stress oxydatif	AGSpops_A_c14088	short-chain dehydrogenase reductase
Stress oxydatif	AGSpops_S_c259	succinate dehydrogenase
Stress oxydatif	AGSpops_A_c2960	succinate dehydrogenase cytochrome b560 mitochondrial precursor
Stress oxydatif	AGSpops_G_c2272	succinate dehydrogenase iron-sulfur mitochondrial precursor
Stress oxydatif	AGSpops_S_c11039	succinate dehydrogenase subunit flavoprotein
Stress oxydatif	AGSpops_A_c2414	superoxide dismutase
Stress oxydatif	AGSpops_G_c4312	thioredoxin
Stress oxydatif	AGSpops_A_c6170	thioredoxin domain containing 3

Stress oxydatif	AGSpops_G_c10363	thioredoxin domain containing 6
Stress oxydatif	AGSpops_G_c10931	thioredoxin domain-containing protein 12
Stress oxydatif	AGSpops_S_c516	thioredoxin peroxidase
Stress oxydatif	AGSpops_A_c5094	thioredoxin reductase cytoplasmic isoform 2
Stress oxydatif	AGSpops_G_c4933	thioredoxin-like protein
Stress oxydatif	AGSpops_A_c5198	trans- -dihydrobenzene- -diol dehydrogenase
Stress oxydatif	AGSpops_G_c7703	tyrocidine synthetase 1
Stress oxydatif	AGSpops_G_c5591	tyrosine aminotransferase
Stress oxydatif	AGSpops_A_c6326	yme1-like 1 (cerevisiae)
Stress oxydatif	AGSpops_S_c12400	zeta (quinone reductase)-like 1
Stress oxydatif	AGSpops_A_c1164	zgc:101788 protein
Stress oxydatif	AGSpops_S_c4610	zinc binding alcohol domain containing 2
Stress oxydatif	AGSpops_A_c3724	zinc-binding dehydrogenase family protein
Stress oxydatif	AGSpops_G_c8689	biotin sulfoxide reductase
Réponse à un agent bactérien	AGSpops_G_c2080	histone h2b
Réponse à un agent bactérien	AGSpops_A_c1012	lysozyme
Réponse à un agent bactérien	AGSpops_S_c5346	ovary-specific c1q-like factor
Réponse à un composé chimique	AGSpops_S_c4724	apolipoprotein d
Réponse à un composé chimique	AGSpops_G_c5415	cation efflux system protein
Réponse à un composé chimique	AGSpops_G_c6288	innexin family protein
Réponse à un composé chimique	AGSpops_G_c3816	phosphoserine aminotransferase
Réponse à un composé chimique	AGSpops_A_c4571	signal recognition particle 72kda
Réponse à un composé chimique	AGSpops_G_c5863	sulfotransferase family member 1
Réponse à un composé chimique	AGSpops_S_c6731	sulfotransferase member 2
Réponse à un composé chimique	AGSpops_S_c5845	sulfotransferase member 3
Réponse à un composé chimique	AGSpops_S_c617	thaumatin-like protein
Réponse à un composé chimique	AGSpops_A_c2780	inhibitor of dna binding dominant negative helix-loop-helix protein
Réponse immunitaire	AGSpops_S_c12716	thrombospondin 1
Réponse immunitaire	AGSpops_G_c8415	thrombospondin 2
Réponse immunitaire	AGSpops_S_c492	thrombospondin-related anonymous protein
Réponse immunitaire	AGSpops_G_c10699	fibrinogen-like 1
Réponse immunitaire	AGSpops_A_c11297	c type lectin receptor c
Réponse immunitaire	AGSpops_S_c5848	c-type lectin
Réponse immunitaire	AGSpops_A_c2758	c-type lectin 2
Réponse immunitaire	AGSpops_G_c5027	c-type lectin 2 like protein
Réponse immunitaire	AGSpops_A_c9753	c-type lectin a

Réponse immunitaire	AGSpops_S_c8917	c-type lectin domain family 10 member a-like isoform 2
Réponse immunitaire	AGSpops_A_c10966	c-type lectin domain family 4 member e
Réponse immunitaire	AGSpops_G_c442	c-type lectin domain family member c
Réponse immunitaire	AGSpops_A_c3482	collectin sub-family member 12
Réponse immunitaire	AGSpops_S_c10325	fc fragment of low affinity receptor for
Réponse immunitaire	AGSpops_G_c9346	ficolin 4
Réponse immunitaire	AGSpops_A_c10446	immune- lectin-like receptor 1
Réponse immunitaire	AGSpops_S_c987	immune- lectin-like receptor 3
Réponse immunitaire	AGSpops_A_c17447	immune- lectin-like receptor 4
Réponse immunitaire	AGSpops_S_c764	immunity adhesion receptor cd209l2
Réponse immunitaire	AGSpops_A_c3036	interleukin enhancer binding factor 2
Réponse immunitaire	AGSpops_S_c5154	macrophage-inducible c-type lectin
Réponse immunitaire	AGSpops_G_c5726	preli domain containing 1
Réponse immunitaire	AGSpops_A_c1011	sialic acid binding lectin
Réponse immunitaire	AGSpops_A_c11238	tumor necrosis factor member 11
Réponse immunitaire	AGSpops_A_c12933	tumor necrosis factor receptor member 19
Réponse immunitaire	AGSpops_A_c3645	tumor rejection antigen 1
Réponse immunitaire	AGSpops_G_c4845	tumor suppressing subtransferable candidate 1
Réponse immunitaire	AGSpops_A_c12085	tumor suppressor candidate 3
Réponse immunitaire	AGSpops_G_c2525	ccat enhancer binding protein
Réponse immunitaire	AGSpops_S_c4854	ccat enhancer binding protein (c ebp) gamma
Réponse immunitaire	AGSpops_A_c10723	macrophage galactose n-acetyl-galactosamine specific lectin 2
Réponse immunitaire	AGSpops_G_c4063	macrophage migration inhibitory factor
Réponse immunitaire	AGSpops_G_c8980	membrane protein
Réponse immunitaire	AGSpops_A_s21033	salivary c-type lectin
Réponse inflammatoire	AGSpops_S_c4467	acidic mammalian chitinase precursor
Réponse inflammatoire	AGSpops_A_c6873	allograft inflammatory factor 1
Réponse inflammatoire	AGSpops_A_c3882	annexin a1
Réponse inflammatoire	AGSpops_G_c3087	annexin a5
Réponse inflammatoire	AGSpops_S_c6978	annexin a6
Réponse inflammatoire	AGSpops_S_s18131	cathepsin a
Réponse inflammatoire	AGSpops_S_c294	cathepsin b
Réponse inflammatoire	AGSpops_A_c4705	cathepsin c
Réponse inflammatoire	AGSpops_S_c220	cathepsin d
Réponse inflammatoire	AGSpops_S_c9923	cathepsin d
Réponse inflammatoire	AGSpops_A_c5522	cathepsin f

Réponse inflammatoire	AGSpops_A_c5455	cathepsin k
Réponse inflammatoire	AGSpops_S_c41	cathepsin l
Réponse inflammatoire	AGSpops_G_s17258	cathepsin l precursor
Réponse inflammatoire	AGSpops_A_c3466	cathepsin l-like cysteine proteinase
Réponse inflammatoire	AGSpops_G_c3774	cathepsin l-like precursor
Réponse inflammatoire	AGSpops_G_c7851	cathepsin l1
Réponse inflammatoire	AGSpops_A_s19803	cathepsin l1 precursor
Réponse inflammatoire	AGSpops_G_c9969	cathepsin l2
Réponse inflammatoire	AGSpops_G_c2694	cathepsin s
Réponse inflammatoire	AGSpops_G_c262	cathepsin z
Réponse inflammatoire	AGSpops_S_c4497	cathepsin z precursor
Réponse inflammatoire	AGSpops_A_c2389	cathepsin z precursor
Réponse inflammatoire	AGSpops_G_c11214	fatty acid-binding heart
Réponse inflammatoire	AGSpops_A_c2173	plasma kallikrein precursor
Réponse inflammatoire	AGSpops_G_c4310	ras-related c3 botulinum toxin substrate 1 (rho small gtp binding protein rac1)
Réponse à une agression virale	AGSpops_A_c4002	baculoviral iap repeat-containing 2
Réponse à une agression virale	AGSpops_S_c10987	barrier-to-autointegration factor
Réponse à une agression virale	AGSpops_S_c5693	barrier-to-autointegration factor
Réponse à une agression virale	AGSpops_A_c13892	cd209 antigen-like protein a
Réponse à une agression virale	AGSpops_S_c8374	cd209 protein
Réponse à une agression virale	AGSpops_G_c4862	cd209-like protein
Réponse à une agression virale	AGSpops_G_c4997	cd209b antigen
Réponse à une agression virale	AGSpops_A_c3485	cd209l1 protein
Réponse à une agression virale	AGSpops_G_c4725	hepatitis b virus x interacting protein
Réponse à une agression virale	AGSpops_G_c2458	interferon gamma inducible protein 47
Ubiquitine	AGSpops_G_c2085	ubiquitin
Ubiquitine	AGSpops_S_c8776	ubiquitin carboxyl-terminal
Ubiquitine	AGSpops_A_c11357	ubiquitin carboxyl-terminal hydrolase 5
Ubiquitine	AGSpops_S_c4600	ubiquitin conjugating enzyme e2
Ubiquitine	AGSpops_G_c7832	ubiquitin fusion degradation protein 1 homolog
Ubiquitine	AGSpops_S_c7607	ubiquitin fusion degradaton protein
Ubiquitine	AGSpops_G_c6138	ubiquitin like
Ubiquitine	AGSpops_A_c4223	ubiquitin protein
Ubiquitine	AGSpops_A_c9487	ubiquitin protein ligase
Ubiquitine	AGSpops_A_c2463	ubiquitin protein ligase e3 component n-recognin 7
Ubiquitine	AGSpops_A_s21106	ubiquitin ribosomal protein s27a fusion protein

Ubiquitine	AGSpops_G_c10813	ubiquitin specific peptidase 10-like
Ubiquitine	AGSpops_G_c6020	ubiquitin specific peptidase 14 (trna-guanine transglycosylase)
Ubiquitine	AGSpops_G_c10027	ubiquitin specific peptidase 16
Ubiquitine	AGSpops_G_c7553	ubiquitin specific peptidase 2
Ubiquitine	AGSpops_G_c6725	ubiquitin specific peptidase 3
Ubiquitine	AGSpops_A_c9651	ubiquitin specific peptidase x chromosome
Ubiquitine	AGSpops_G_c6010	ubiquitin- containing phd and ring finger 1
Ubiquitine	AGSpops_G_c2522	ubiquitin-conjugating enzyme
Ubiquitine	AGSpops_A_c7441	ubiquitin-conjugating enzyme 2
Ubiquitine	AGSpops_G_c7288	ubiquitin-conjugating enzyme e2
Ubiquitine	AGSpops_S_c9986	ubiquitin-conjugating enzyme e2 g
Ubiquitine	AGSpops_A_c4392	ubiquitin-conjugating enzyme e2 i
Ubiquitine	AGSpops_A_c3792	ubiquitin-conjugating enzyme e2d 3 (ubc4 5 yeast) isoform cra_a
Ubiquitine	AGSpops_A_c8163	ubiquitin-conjugating enzyme e2n
Ubiquitine	AGSpops_G_c2519	ubiquitin-conjugating enzyme e2r 2
Ubiquitine	AGSpops_G_c2339	ubiquitin-conjugating enzyme e2s
Ubiquitine	AGSpops_G_c7747	ubiquitin-conjugating enzyme e2t
Ubiquitine	AGSpops_G_c5903	ubiquitin-conjugating enzyme e2u
Ubiquitine	AGSpops_A_c8021	ubiquitin-like domain containing ctd phosphatase 1
Ubiquitine	AGSpops_A_c5780	ubiquitin-protein ligase

B- Résultats de la Partie II,
Chapitre 3

*Adaptation différentielle selon
les contraintes environnementales*

Comparaison des transcriptomes des populations A et G par test exact de Fisher

Gene Ontology	Activité peptidique	FDR	FWER	p	Nb séquences test (popA)	Nb séquences référent (popG)	Nb séquences test non annotées	Nb séquences référent non annotées	Niveau d'expression
GO:0008270	zinc ion binding	2.38807E-7	5.97018E-8	7.86939E-10	137	39	2212	1791	over
GO:0030246	carbohydrate binding	1.55844E-6	7.79221E-7	9.83152E-9	85	18	2264	1812	over
GO:0005622	intracellular	7.19135E-6	5.3935E-6	6.32695E-8	71	14	2278	1816	over
GO:0005529	sugar binding	1.10235E-5	1.10235E-5	1.28186E-7	62	11	2287	1819	over
GO:0043169	cation binding	5.7968E-5	8.6013E-5	9.01508E-7	379	201	1970	1629	over
GO:0046914	transition metal ion binding	5.7968E-5	8.69483E-5	9.1833E-7	188	80	2161	1750	over
GO:0043167	ion binding	6.54981E-5	1.14615E-4	1.1846E-6	379	202	1970	1628	over
GO:0046872	metal ion binding	2.10514E-4	4.20939E-4	3.7672E-6	365	197	1984	1633	over
GO:0005488	binding	0.0047144	0.0105514	7.93974E-5	1606	1148	743	682	over
GO:0044424	intracellular part	0.0074624	0.00186386	1.22824E-5	1263	1104	1086	726	under
GO:0008061	chitin binding	0.100149	0.221511	0.00131522	15	1	2334	1829	over
GO:0044444	cytoplasmic part	0.104085	0.0507126	2.57079E-4	683	625	1666	1205	under
GO:0004252	serine-type endopeptidase activity	0.194014	0.431231	0.00290445	36	11	2313	1819	over
GO:0005509	calcium ion binding	0.194014	0.441326	0.00298283	75	33	2274	1797	over
GO:0003824	catalytic activity	0.209357	0.14532	7.37E-4	982	856	1367	974	under
GO:0016410	N-acyltransferase activity	0.233313	0.208111	0.00104359	4	16	2345	1814	under
GO:0006030	chitin metabolic process	0.280984	0.626127	0.00490611	20	4	2329	1826	over
GO:0044464	cell part	0.325617	0.334408	0.00167068	1556	1291	793	539	under
GO:0008080	N-acetyltransferase activity	0.327897	0.388553	0.00193702	4	15	2345	1815	under
GO:0001871	pattern binding	0.397637	0.796376	0.00654846	17	3	2332	1827	over
GO:0030247	polysaccharide binding	0.397637	0.796376	0.00654846	17	3	2332	1827	over
GO:0009108	coenzyme biosynthetic process	0.444835	0.660406	0.00399411	7	18	2342	1812	under
GO:0051716	cellular response to stimulus	0.444835	0.675772	0.0043755	63	77	2286	1753	under
GO:0050896	response to stimulus	0.444835	0.725845	0.00481156	287	275	2062	1555	under
GO:0051704	multi-organism process	0.444835	0.728582	0.00486654	48	62	2301	1768	under
GO:0051188	cofactor biosynthetic process	0.444835	0.734011	0.00500589	13	25	2336	1805	under
GO:0005829	cytosol	0.444835	0.736862	0.00515753	150	156	2199	1674	under
GO:0006022	aminoglycan metabolic process	0.451238	0.853266	0.0083046	25	7	2324	1823	over
GO:0007156	homophilic cell adhesion	0.572968	0.924275	0.00991321	8	0	2341	1830	over
GO:0030414	peptidase inhibitor activity	0.659219	0.958909	0.0117735	24	7	2325	1823	over

GO:0004857	enzyme inhibitor activity	0.659219	0.963105	0.0125404	35	13	2314	1817	over
GO:0004402	histone acetyltransferase activity	0.667077	0.913405	0.00738033	1	8	2348	1822	under
GO:0004468	lysine N-acetyltransferase activity	0.667077	0.913405	0.00738033	1	8	2348	1822	under
GO:0033554	cellular response to stress	0.667077	0.920567	0.00801639	54	66	2295	1764	under
GO:0043227	membrane-bounded organelle	0.667077	0.949264	0.00966312	441	398	1908	1432	under
GO:0043231	intracellular membrane-bounded organelle	0.667077	0.949264	0.00966312	441	398	1908	1432	under
GO:0032991	macromolecular complex	0.667077	0.95043	0.00998006	638	558	1711	1272	under
GO:0006974	response to DNA damage stimulus	0.710057	0.965817	0.0117227	40	51	2309	1779	under
GO:0016407	acetyltransferase activity	0.733783	0.97672	0.0124989	8	17	2341	1813	under
GO:0051707	response to other organism	0.733783	0.97886	0.0134153	16	26	2333	1804	under
GO:0044462	external encapsulating structure part	0.860421	0.991712	0.0151595	1	7	2348	1823	under
GO:0044446	intracellular organelle part	0.860421	0.992947	0.0159208	624	542	1725	1288	under
GO:0009279	cell outer membrane	0.878003	0.997678	0.0160531	0	5	2349	1825	under
GO:0035257	nuclear hormone receptor binding	0.878003	0.998498	0.0162187	3	10	2346	1820	under
GO:0043226	organelle	0.878003	0.9986	0.0171226	676	583	1673	1247	under
GO:0043229	intracellular organelle	0.878003	0.998636	0.0174422	675	582	1674	1248	under
GO:0042221	response to chemical stimulus	0.878003	0.998961	0.0194117	136	136	2213	1694	under
GO:0016746	transferase activity, transferring acyl groups	0.878003	0.998986	0.0195455	16	25	2333	1805	under
GO:0006766	vitamin metabolic process	0.878003	0.999275	0.0201167	4	11	2345	1819	under
GO:0009987	cellular process	0.878003	0.999284	0.0203286	1267	1046	1082	784	under
GO:0044422	organelle part	0.878003	0.999396	0.021193	634	547	1715	1283	under
GO:0006810	transport	0.878003	0.999545	0.02191	253	235	2096	1595	under
GO:0071495	cellular response to endogenous stimulus	0.878003	0.999695	0.022684	2	8	2347	1822	under
GO:0016903	oxidoreductase activity, acting on the aldehyde or oxo group of donors	0.878003	0.999766	0.0235605	5	12	2344	1818	under
GO:0044281	small molecule metabolic process	0.878003	0.999806	0.0254086	201	190	2148	1640	under
GO:0008415	acyltransferase activity	0.878003	0.999881	0.0275623	15	23	2334	1807	under
GO:0003702	RNA polymerase II transcription factor activity	0.878003	0.999894	0.0281099	16	24	2333	1806	under
GO:0051234	establishment of localization	0.878003	0.999903	0.0286413	258	237	2091	1593	under
GO:0035107	appendage morphogenesis	0.878003	0.999923	0.0291424	7	14	2342	1816	under
GO:0042180	cellular ketone metabolic process	0.878003	0.999968	0.0306947	80	84	2269	1746	under
GO:0001889	liver development	0.878003	0.999988	0.0307247	1	6	2348	1824	under
GO:0045785	positive regulation of cell adhesion	0.878003	0.999988	0.0307247	1	6	2348	1824	under
GO:0071375	cellular response to peptide hormone	0.878003	1.0	0.036704	0	4	2349	1826	under

	stimulus								
GO:0032869	cellular response to insulin stimulus	0.878003	1.0	0.036704	0	4	2349	1826	under
GO:0016840	carbon-nitrogen lyase activity	0.878003	1.0	0.036704	0	4	2349	1826	under
GO:0015144	carbohydrate transmembrane transporter activity	0.878003	1.0	0.036704	0	4	2349	1826	under
GO:0019861	flagellum	0.878003	1.0	0.0372406	12	19	2337	1811	under
GO:0016769	transferase activity, transferring nitrogenous groups	0.878003	1.0	0.0391421	5	11	2344	1819	under
GO:0008483	transaminase activity	0.878003	1.0	0.0391421	5	11	2344	1819	under
GO:0016020	membrane	0.878003	1.0	0.0396141	300	269	2049	1561	under
GO:0016747	transferase activity, transferring acyl groups other than amino-acyl groups	0.878003	1.0	0.0398801	16	23	2333	1807	under
GO:0042439	ethanolamine and derivative metabolic process	0.878003	1.0	0.0423865	2	7	2347	1823	under
GO:0042133	neurotransmitter metabolic process	0.878003	1.0	0.0423865	2	7	2347	1823	under
GO:0008484	sulfuric ester hydrolase activity	0.878003	1.0	0.0423865	2	7	2347	1823	under
GO:0032870	cellular response to hormone stimulus	0.878003	1.0	0.0423865	2	7	2347	1823	under
GO:0016251	general RNA polymerase II transcription factor activity	0.878003	1.0	0.0428553	6	12	2343	1818	under
GO:0005578	proteinaceous extracellular matrix	0.878003	1.0	0.0428553	6	12	2343	1818	under
GO:0007610	behavior	0.878003	1.0	0.046545	35	41	2314	1789	under
GO:0016740	transferase activity	0.878003	1.0	0.0474947	220	201	2129	1629	under
GO:0004725	protein tyrosine phosphatase activity	0.969594	0.995444	0.0176592	7	0	2342	1830	over
GO:0004222	metalloendopeptidase activity	0.969594	0.997414	0.0209378	12	2	2337	1828	over
GO:0022610	biological adhesion	0.969594	0.999286	0.026601	34	14	2315	1816	over
GO:0007155	cell adhesion	0.969594	0.999286	0.026601	34	14	2315	1816	over
GO:0019915	lipid storage	0.969594	0.999614	0.0275391	9	1	2340	1829	over
GO:0008237	metallopeptidase activity	0.969594	0.999666	0.0286192	25	9	2324	1821	over
GO:0005856	cytoskeleton	0.969594	0.999694	0.0290236	42	19	2307	1811	over
GO:0031589	cell-substrate adhesion	0.969594	0.999962	0.031452	6	0	2343	1830	over
GO:0006508	proteolysis	0.969594	0.999994	0.0372561	161	100	2188	1730	over
GO:0005576	extracellular region	0.969594	0.999997	0.0411977	73	40	2276	1790	over

Comparaison des transcriptomes des populations G et S par test exact de Fisher

Gene Ontology	Activité peptidique	FDR	FWER	p	Nb séquences test (popG)	Nb séquences référent (popS)	Nb séquences test non annotées	Nb séquences référent non annotées	Niveau d'expression
GO:0044424	intracellular part	8.87726E-9	4.36114E-9	2.12961E-12	1104	610	726	762	over
GO:0044464	cell part	8.87726E-9	4.43863E-9	3.57427E-12	1291	768	539	604	over
GO:0043226	organelle	1.11332E-6	9.30559E-7	1.50355E-8	583	315	1247	1057	over
GO:0043229	intracellular organelle	1.11332E-6	1.11332E-6	1.81012E-8	582	315	1248	1057	over
GO:0044422	organelle part	1.5199E-6	1.89988E-6	2.85038E-8	547	293	1283	1079	over
GO:0044446	intracellular organelle part	1.69591E-6	2.54386E-6	4.14731E-8	542	291	1288	1081	over
GO:0005515	protein binding	2.12066E-6	3.71115E-6	5.58617E-8	650	366	1180	1006	over
GO:0044428	nuclear part	8.27884E-6	1.65575E-5	2.35209E-7	225	95	1605	1277	over
GO:0065007	biological regulation	1.93865E-5	4.36187E-5	6.07445E-7	515	283	1315	1089	over
GO:0050789	regulation of biological process	3.10769E-5	7.76892E-5	1.04785E-6	422	223	1408	1149	over
GO:0043231	intracellular membrane-bounded organelle	4.29481E-5	1.28836E-4	1.69069E-6	398	209	1432	1163	over
GO:0043227	membrane-bounded organelle	4.29481E-5	1.28836E-4	1.69069E-6	398	209	1432	1163	over
GO:0032991	macromolecular complex	1.5972E-4	5.18956E-4	5.8002E-6	558	322	1272	1050	over
GO:0050794	regulation of cellular process	2.37319E-4	9.31617E-4	1.06662E-5	363	193	1467	1179	over
GO:0019222	regulation of metabolic process	2.37319E-4	9.60501E-4	1.0859E-5	231	109	1599	1263	over
GO:0044444	cytoplasmic part	2.37319E-4	0.00100087	1.15097E-5	625	372	1205	1000	over
GO:0060255	regulation of macromolecule metabolic process	2.37319E-4	0.0010081	1.16158E-5	211	97	1619	1275	over
GO:0080090	regulation of primary metabolic process	6.30684E-4	0.00299127	3.01809E-5	208	98	1622	1274	over
GO:0005829	cytosol	0.00135703	0.00676223	6.78532E-5	156	69	1674	1303	over

GO:0043234	protein complex	0.0023083	0.0120456	1.18688E-4	360	201	1470	1171	over
GO:0016043	cellular component organization	0.00257333	0.0140538	1.26457E-4	256	133	1574	1239	over
GO:0016070	RNA metabolic process	0.00271604	0.0158077	1.40857E-4	134	58	1696	1314	over
GO:0009889	regulation of biosynthetic process	0.00271604	0.017759	1.54284E-4	137	60	1693	1312	over
GO:0010556	regulation of macromolecule biosynthetic process	0.00271604	0.0180604	1.60292E-4	130	56	1700	1316	over
GO:0022402	cell cycle process	0.00271604	0.0181665	1.64887E-4	109	44	1721	1328	over
GO:0051704	multi-organism process	0.00290759	0.0201477	1.8377E-4	62	19	1768	1353	over
GO:0010468	regulation of gene expression	0.00507591	0.0361323	3.18699E-4	147	68	1683	1304	over
GO:0044451	nucleoplasm part	0.00523595	0.0385096	3.3942E-4	54	16	1776	1356	over
GO:0006260	DNA replication	0.00657175	0.0301229	1.86083E-12	23	107	1807	1265	under
GO:0016779	nucleotidyltransferase activity	0.00657175	0.0301229	3.64871E-12	13	105	1817	1267	under
GO:0006259	DNA metabolic process	0.00657175	0.0301229	4.73757E-12	73	167	1757	1205	under
GO:0003676	nucleic acid binding	0.00657175	0.0301229	4.7905E-12	287	361	1543	1011	under
GO:0003964	RNA-directed DNA polymerase activity	0.00657175	0.0301229	5.12327E-12	3	94	1827	1278	under
GO:0034061	DNA polymerase activity	0.00657175	0.0301229	5.16728E-12	3	99	1827	1273	under
GO:0008270	zinc ion binding	0.00657175	0.0301229	5.45975E-12	39	99	1791	1273	under
GO:0046914	transition metal ion binding	0.00657175	0.030123	9.0099E-10	80	135	1750	1237	under
GO:0016772	transferase activity, transferring phosphorus-containing groups	0.00657175	0.0301231	3.05839E-9	96	149	1734	1223	under
GO:0003723	RNA binding	0.00657175	0.0301241	1.56319E-8	120	169	1710	1203	under
GO:0034645	cellular macromolecule biosynthetic process	0.00657175	0.0301352	1.38738E-7	173	213	1657	1159	under
GO:0009059	macromolecule biosynthetic process	0.00657175	0.0301392	1.83008E-7	174	213	1656	1159	under
GO:0090304	nucleic acid metabolic process	0.00657175	0.0305712	4.00295E-6	204	229	1626	1143	under
GO:0043170	macromolecule metabolic process	0.00657175	0.0311394	9.0362E-6	561	521	1269	851	under
GO:0043169	cation binding	0.00657175	0.0311613	9.32894E-	201	223	1629	1149	under

				6					
GO:0043167	ion binding	0.00657175	0.0313967	1.16127E-5	202	223	1628	1149	under
GO:0044238	primary metabolic process	0.00657175	0.0324791	1.79197E-5	676	607	1154	765	under
GO:0046872	metal ion binding	0.00657175	0.0327727	2.1918E-5	197	216	1633	1156	under
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.00681509	0.0355934	4.21855E-5	257	265	1573	1107	under
GO:0005634	nucleus	0.00691958	0.0522159	4.00995E-4	207	106	1623	1266	over
GO:0006974	response to DNA damage stimulus	0.00713121	0.0586693	4.62337E-4	51	15	1779	1357	over
GO:0032502	developmental process	0.00713121	0.0603255	4.94468E-4	263	143	1567	1229	over
GO:0048519	negative regulation of biological process	0.00713121	0.060451	4.98692E-4	182	91	1648	1281	over
GO:0005737	cytoplasm	0.00713121	0.0604936	4.99742E-4	201	103	1629	1269	over
GO:0006807	nitrogen compound metabolic process	0.0072657	0.0396182	6.96956E-5	305	303	1525	1069	under
GO:0051716	cellular response to stimulus	0.00785577	0.0682635	5.89362E-4	77	29	1753	1343	over
GO:0005488	binding	0.00876847	0.0496091	1.41017E-4	1148	946	682	426	under
GO:0030246	carbohydrate binding	0.0096271	0.0565633	1.98436E-4	18	37	1812	1335	under
GO:0005622	intracellular	0.0102094	0.0622505	2.07292E-4	14	32	1816	1340	under
GO:0034641	cellular nitrogen compound metabolic process	0.0156673	0.0972439	4.55474E-4	294	284	1536	1088	under
GO:0006810	transport	0.0162102	0.139256	0.00107732	235	128	1595	1244	over
GO:0030247	polysaccharide binding	0.0172382	0.114342	5.12722E-4	3	15	1827	1357	under
GO:0001871	pattern binding	0.0172382	0.114342	5.12722E-4	3	15	1827	1357	under
GO:0016740	transferase activity	0.0172382	0.117895	5.49519E-4	201	205	1629	1167	under
GO:0051234	establishment of localization	0.0191508	0.166358	0.00126209	237	130	1593	1242	over
GO:0048869	cellular developmental process	0.0193362	0.17184	0.0013642	93	40	1737	1332	over
GO:0006355	regulation of transcription, DNA-dependent	0.0193542	0.175983	0.00140701	70	27	1760	1345	over
GO:0005694	chromosome	0.0196305	0.182985	0.0014656	21	3	1809	1369	over

				6					
GO:0000166	nucleotide binding	0.0196305	0.189744	0.0015073 6	278	158	1552	1214	over
GO:0016020	membrane	0.0196305	0.19027	0.0015104 9	269	152	1561	1220	over
GO:0033554	cellular response to stress	0.0197185	0.195015	0.0015480 6	66	25	1764	1347	over
GO:0048523	negative regulation of cellular process	0.0231825	0.229598	0.0017436 4	159	81	1671	1291	over
GO:0008152	metabolic process	0.0235846	0.16252	8.99749E- 4	908	758	922	614	under
GO:0010604	positive regulation of macromolecule metabolic process	0.0274286	0.27474	0.0021010 9	79	33	1751	1339	over
GO:0044427	chromosomal part	0.0274286	0.275551	0.0021039	63	24	1767	1348	over
GO:0016071	mRNA metabolic process	0.0276721	0.282604	0.0021853 4	46	15	1784	1357	over
GO:0051649	establishment of localization in cell	0.0277537	0.295321	0.0022676 8	89	39	1741	1333	over
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.0277537	0.302703	0.0022996 9	114	54	1716	1318	over
GO:0051252	regulation of RNA metabolic process	0.0277537	0.305529	0.0023402 4	82	35	1748	1337	over
GO:0044281	small molecule metabolic process	0.0277537	0.306103	0.0023465 5	190	102	1640	1270	over
GO:0048856	anatomical structure development	0.0277537	0.307753	0.0024194	130	64	1700	1308	over
GO:0010605	negative regulation of macromolecule metabolic process	0.0328662	0.361942	0.0028194 3	88	39	1742	1333	over
GO:0009893	positive regulation of metabolic process	0.0328662	0.368368	0.0029322 1	81	35	1749	1337	over
GO:0051171	regulation of nitrogen compound metabolic process	0.0328662	0.368869	0.0029371 9	116	56	1714	1316	over
GO:0045449	regulation of transcription	0.0356975	0.398802	0.0032905 5	94	43	1736	1329	over
GO:0031325	positive regulation of cellular metabolic process	0.0374736	0.424172	0.0036611 5	80	35	1750	1337	over
GO:0048518	positive regulation of biological process	0.0374736	0.424723	0.0036696 5	184	100	1646	1272	over
GO:0051246	regulation of protein metabolic process	0.0438698	0.482251	0.0037709 4	95	44	1735	1328	over
GO:0048583	regulation of response to stimulus	0.0462355	0.506061	0.0041407 9	32	9	1798	1363	over
GO:0044454	nuclear chromosome part	0.0467947	0.515956	0.0041691 8	16	2	1814	1370	over
GO:0006519	cellular amino acid and derivative metabolic	0.0484454	0.533877	0.0042520	62	25	1768	1347	over

	process			2					
GO:0046907	intracellular transport	0.050805	0.565829	0.0045551 5	79	35	1751	1337	over
GO:0022607	cellular component assembly	0.050805	0.566172	0.0045982 7	128	65	1702	1307	over
GO:0042180	cellular ketone metabolic process	0.050805	0.567695	0.0046164 5	84	38	1746	1334	over
GO:0051188	cofactor biosynthetic process	0.0533393	0.590911	0.0052002 1	25	6	1805	1366	over
GO:0004888	transmembrane receptor activity	0.0536196	0.340373	0.0020258	8	20	1822	1352	under
GO:0009108	coenzyme biosynthetic process	0.0539593	0.602747	0.0053719	18	3	1812	1369	over
GO:0003682	chromatin binding	0.0539593	0.612002	0.0054891	27	7	1803	1365	over
GO:0051186	cofactor metabolic process	0.0539593	0.617127	0.0055531	52	20	1778	1352	over
GO:0032268	regulation of cellular protein metabolic process	0.0539593	0.627199	0.0056753 4	88	41	1742	1331	over
GO:0015031	protein transport	0.0539593	0.629416	0.0057158	66	28	1764	1344	over
GO:0050896	response to stimulus	0.0539593	0.6373	0.0057994 7	275	163	1555	1209	over
GO:0043228	non-membrane-bounded organelle	0.0539593	0.64056	0.0059230 6	208	118	1622	1254	over
GO:0043232	intracellular non-membrane-bounded organelle	0.0539593	0.64056	0.0059230 6	208	118	1622	1254	over
GO:0009892	negative regulation of metabolic process	0.0539593	0.641472	0.0060273 7	91	43	1739	1329	over
GO:0008237	metallopeptidase activity	0.0549089	0.355841	0.0023437 4	9	21	1821	1351	under
GO:0019899	enzyme binding	0.0671964	0.725912	0.0069461	42	15	1788	1357	over
GO:0006397	mRNA processing	0.0731796	0.773147	0.0076472 8	36	12	1794	1360	over
GO:0023034	intracellular signaling pathway	0.0731796	0.77517	0.0077482 2	56	23	1774	1349	over
GO:0005815	microtubule organizing center	0.0731796	0.77843	0.0077739 6	26	7	1804	1365	over
GO:0008380	RNA splicing	0.0731796	0.781442	0.0078151 2	34	11	1796	1361	over
GO:0065009	regulation of molecular function	0.0731796	0.781608	0.0078430 5	101	50	1729	1322	over
GO:0044106	cellular amine metabolic process	0.0731796	0.787655	0.0079526 3	47	18	1783	1354	over
GO:0005681	spliceosomal complex	0.0731796	0.791242	0.0079754 3	30	9	1800	1363	over
GO:0001882	nucleoside binding	0.0731796	0.802609	0.0081824 4	180	101	1650	1271	over

GO:0001883	purine nucleoside binding	0.0731796	0.802609	0.0081824 4	180	101	1650	1271	over
GO:0045184	establishment of protein localization	0.0731796	0.806445	0.0084105 4	66	29	1764	1343	over
GO:0032559	adenyl ribonucleotide binding	0.0731796	0.80835	0.0085740 8	166	92	1664	1280	over
GO:0048522	positive regulation of cellular process	0.0731796	0.808912	0.0086688 7	146	79	1684	1293	over
GO:0032555	purine ribonucleotide binding	0.0731796	0.825631	0.0093714 1	212	123	1618	1249	over
GO:0032553	ribonucleotide binding	0.0731796	0.825631	0.0093714 1	212	123	1618	1249	over
GO:0030554	adenyl nucleotide binding	0.0731796	0.825753	0.0093973	179	101	1651	1271	over
GO:0017076	purine nucleotide binding	0.0731796	0.829608	0.0098049 9	225	132	1605	1240	over
GO:0005524	ATP binding	0.0731796	0.830878	0.0098886 3	165	92	1665	1280	over
GO:0031327	negative regulation of cellular biosynthetic process	0.0731796	0.833767	0.0099117 1	37	13	1793	1359	over
GO:0009890	negative regulation of biosynthetic process	0.0731796	0.833767	0.0099117 1	37	13	1793	1359	over
GO:0022403	cell cycle phase	0.0731796	0.833767	0.0099117 1	37	13	1793	1359	over
GO:0010629	negative regulation of gene expression	0.0731796	0.833767	0.0099117 1	37	13	1793	1359	over
GO:0010558	negative regulation of macromolecule biosynthetic process	0.0747547	0.843038	0.0102544	35	12	1795	1360	over
GO:0006082	organic acid metabolic process	0.0776303	0.859549	0.0105237	75	35	1755	1337	over
GO:0051254	positive regulation of RNA metabolic process	0.0776303	0.865039	0.0106664	23	6	1807	1366	over
GO:0045893	positive regulation of transcription, DNA-dependent	0.0776303	0.865039	0.0106664	23	6	1807	1366	over
GO:0031090	organelle membrane	0.0776303	0.865353	0.0107191	88	43	1742	1329	over
GO:0006732	coenzyme metabolic process	0.0776303	0.86928	0.010848	44	17	1786	1355	over
GO:0010628	positive regulation of gene expression	0.0776303	0.872436	0.0109345	25	7	1805	1365	over
GO:0045941	positive regulation of transcription	0.0776303	0.872436	0.0109345	25	7	1805	1365	over
GO:0071495	cellular response to endogenous stimulus	0.0882932	0.909708	0.011308	8	0	1822	1372	over
GO:0071103	DNA conformation change	0.0882932	0.909708	0.011308	8	0	1822	1372	over
GO:0006325	chromatin organization	0.0882932	0.911289	0.0116423	56	24	1774	1348	over
GO:0005739	mitochondrion	0.0882932	0.912014	0.011925	105	54	1725	1318	over
GO:0008061	chitin binding	0.0884054	0.51809	0.0029533 4	1	9	1829	1363	under
GO:0000902	cell morphogenesis	0.0953403	0.92925	0.0123669	16	3	1814	1369	over

GO:0006520	cellular amino acid metabolic process	0.0954387	0.931647	0.0125639	38	14	1792	1358	over
GO:0010557	positive regulation of macromolecule biosynthetic process	0.0954387	0.934185	0.0131257	36	13	1794	1359	over
GO:0031324	negative regulation of cellular metabolic process	0.0954387	0.935347	0.0132647	82	40	1748	1332	over
GO:0005886	plasma membrane	0.0954387	0.935882	0.0135684	77	37	1753	1335	over
GO:0005509	calcium ion binding	0.102127	0.580552	0.0038319 7	33	46	1797	1326	under
GO:0043412	macromolecule modification	0.103132	0.949944	0.014918	127	69	1703	1303	over
GO:0016481	negative regulation of transcription	0.10349	0.951731	0.0149641	28	9	1802	1363	over
GO:0042802	identical protein binding	0.105606	0.957674	0.0155952	53	23	1777	1349	over
GO:0043436	oxoacid metabolic process	0.105606	0.958098	0.0156359	73	35	1757	1337	over
GO:0019752	carboxylic acid metabolic process	0.105606	0.958098	0.0156359	73	35	1757	1337	over
GO:0016874	ligase activity	0.112137	0.966595	0.0169477	44	18	1786	1354	over
GO:0009605	response to external stimulus	0.112137	0.968173	0.0172734	35	13	1795	1359	over
GO:0043687	post-translational protein modification	0.112137	0.968821	0.0177619	90	46	1740	1326	over
GO:0009653	anatomical structure morphogenesis	0.112137	0.969236	0.0178482	82	41	1748	1331	over
GO:0016568	chromatin modification	0.112601	0.970521	0.0180921	33	12	1797	1360	over
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	0.114273	0.973584	0.0185502	15	3	1815	1369	over
GO:0005813	centrosome	0.114273	0.973584	0.0185502	15	3	1815	1369	over
GO:0051726	regulation of cell cycle	0.115622	0.975419	0.0190887	40	16	1790	1356	over
GO:0000123	histone acetyltransferase complex	0.119529	0.981316	0.0194913	10	1	1820	1371	over
GO:0048806	genitalia development	0.119529	0.981316	0.0194913	10	1	1820	1371	over
GO:0008022	protein C-terminus binding	0.119529	0.981316	0.0194913	10	1	1820	1371	over
GO:0040035	hermaphrodite genitalia development	0.119529	0.981316	0.0194913	10	1	1820	1371	over
GO:0006281	DNA repair	0.119529	0.982182	0.0196234	29	10	1801	1362	over
GO:0048731	system development	0.119529	0.982182	0.0196234	29	10	1801	1362	over
GO:0048037	cofactor binding	0.119529	0.98242	0.0196869	57	26	1773	1346	over
GO:0006366	transcription from RNA polymerase II promoter	0.131814	0.990303	0.0198184	7	0	1823	1372	over
GO:0032870	cellular response to hormone stimulus	0.131814	0.990303	0.0198184	7	0	1823	1372	over
GO:0000228	nuclear chromosome	0.131814	0.990303	0.0198184	7	0	1823	1372	over
GO:0008484	sulfuric ester hydrolase activity	0.131814	0.990303	0.0198184	7	0	1823	1372	over
GO:0050657	nucleic acid transport	0.131814	0.990957	0.0201144	17	4	1813	1368	over
GO:0050658	RNA transport	0.131814	0.990957	0.0201144	17	4	1813	1368	over
GO:0006733	oxidoreduction coenzyme metabolic process	0.131814	0.990957	0.0201144	17	4	1813	1368	over
GO:0065008	regulation of biological quality	0.131814	0.991432	0.0206712	173	101	1657	1271	over

GO:0016746	transferase activity, transferring acyl groups	0.131814	0.991821	0.0208113	25	8	1805	1364	over
GO:0043933	macromolecular complex subunit organization	0.131814	0.991835	0.0208444	114	62	1716	1310	over
GO:0016773	phosphotransferase activity, alcohol group as acceptor	0.131814	0.992688	0.0211056	55	25	1775	1347	over
GO:0032989	cellular component morphogenesis	0.131814	0.992934	0.0211568	23	7	1807	1365	over
GO:0044419	interspecies interaction between organisms	0.131814	0.992934	0.0211568	23	7	1807	1365	over
GO:0016747	transferase activity, transferring acyl groups other than amino-acyl groups	0.131814	0.992934	0.0211568	23	7	1807	1365	over
GO:0008415	acyltransferase activity	0.131814	0.992934	0.0211568	23	7	1807	1365	over
GO:0045934	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.132515	0.99352	0.0213902	36	14	1794	1358	over
GO:0009850	auxin metabolic process	0.132515	0.993772	0.0220139	84	43	1746	1329	over
GO:0009851	auxin biosynthetic process	0.132515	0.993772	0.0220139	84	43	1746	1329	over
GO:0004672	protein kinase activity	0.136313	0.994801	0.0242564	51	23	1779	1349	over
GO:0048513	organ development	0.136313	0.994974	0.0245762	69	34	1761	1338	over
GO:0007399	nervous system development	0.141888	0.996095	0.0246119	12	2	1818	1370	over
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.142018	0.996271	0.0249961	30	11	1800	1361	over
GO:0009891	positive regulation of biosynthetic process	0.142018	0.996632	0.026018	37	15	1793	1357	over
GO:0006357	regulation of transcription from RNA polymerase II promoter	0.142018	0.996632	0.026018	37	15	1793	1357	over
GO:0031328	positive regulation of cellular biosynthetic process	0.142018	0.996632	0.026018	37	15	1793	1357	over
GO:0051028	mRNA transport	0.148161	0.997557	0.0275845	14	3	1816	1369	over
GO:0047485	protein N-terminus binding	0.148161	0.997557	0.0275845	14	3	1816	1369	over
GO:0003702	RNA polymerase II transcription factor activity	0.150006	0.997818	0.0282708	24	8	1806	1364	over
GO:0016410	N-acyltransferase activity	0.154635	0.998283	0.0291146	16	4	1814	1368	over
GO:0006886	intracellular protein transport	0.154635	0.99833	0.0291658	40	17	1790	1355	over
GO:0042446	hormone biosynthetic process	0.154976	0.998416	0.0293673	87	46	1743	1326	over
GO:0044248	cellular catabolic process	0.160668	0.9988	0.0305859	119	67	1711	1305	over
GO:0004674	protein serine/threonine kinase activity	0.160894	0.998859	0.0311537	38	16	1792	1356	over
GO:0004181	metallocarboxypeptidase activity	0.169414	0.773184	0.00615004	0	6	1830	1366	under
GO:0005876	spindle microtubule	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0034637	cellular carbohydrate biosynthetic process	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0006368	RNA elongation from RNA polymerase II promoter	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0006405	RNA export from nucleus	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0000793	condensed chromosome	0.169829	0.999468	0.0313963	9	1	1821	1371	over

GO:0045216	cell-cell junction organization	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0006354	RNA elongation	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0042375	quinone cofactor metabolic process	0.169829	0.999468	0.0313963	9	1	1821	1371	over
GO:0023052	signaling	0.169829	0.999477	0.0316062	102	56	1728	1316	over
GO:0016192	vesicle-mediated transport	0.169829	0.999499	0.0320292	43	19	1787	1353	over
GO:0032501	multicellular organismal process	0.169829	0.999528	0.0329854	126	72	1704	1300	over
GO:0051172	negative regulation of nitrogen compound metabolic process	0.169829	0.999544	0.0332592	36	15	1794	1357	over
GO:0044092	negative regulation of molecular function	0.169829	0.999551	0.0337353	56	27	1774	1345	over
GO:0048468	cell development	0.170824	0.999606	0.0346347	27	10	1803	1362	over
GO:0005875	microtubule associated complex	0.170824	0.999606	0.0346347	27	10	1803	1362	over
GO:0000271	polysaccharide biosynthetic process	0.191996	0.9999	0.0347257	6	0	1824	1372	over
GO:0007043	cell-cell junction assembly	0.191996	0.9999	0.0347257	6	0	1824	1372	over
GO:0050790	regulation of catalytic activity	0.191996	0.999901	0.0348277	75	39	1755	1333	over
GO:0065003	macromolecular complex assembly	0.191996	0.99991	0.0363078	98	54	1732	1318	over
GO:0016301	kinase activity	0.191996	0.999911	0.0363754	81	43	1749	1329	over
GO:0003712	transcription cofactor activity	0.191996	0.999917	0.0363988	25	9	1805	1363	over
GO:0005625	soluble fraction	0.191996	0.999917	0.0363988	25	9	1805	1363	over
GO:0016563	transcription activator activity	0.191996	0.999917	0.0363988	25	9	1805	1363	over
GO:0016853	isomerase activity	0.191996	0.999921	0.0367898	39	17	1791	1355	over
GO:0051247	positive regulation of protein metabolic process	0.191996	0.999924	0.0369517	49	23	1781	1349	over
GO:0032270	positive regulation of cellular protein metabolic process	0.191996	0.999924	0.0369517	49	23	1781	1349	over
GO:0043086	negative regulation of catalytic activity	0.191996	0.999925	0.0371843	44	20	1786	1352	over
GO:0006950	response to stress	0.191996	0.999926	0.0372298	143	84	1687	1288	over
GO:0044265	cellular macromolecule catabolic process	0.191996	0.999927	0.0374377	73	38	1757	1334	over
GO:0006396	RNA processing	0.191996	0.999928	0.0375968	87	47	1743	1325	over
GO:0004175	endopeptidase activity	0.193727	0.825365	0.0070587 4	56	66	1774	1306	under
GO:0004190	aspartic-type endopeptidase activity	0.194257	0.8423	0.0072428 7	3	11	1827	1361	under
GO:0070001	aspartic-type peptidase activity	0.194257	0.8423	0.0072428 7	3	11	1827	1361	under
GO:0046982	protein heterodimerization activity	0.19912	0.999959	0.0376002	11	2	1819	1370	over
GO:0006367	transcription initiation from RNA polymerase II promoter	0.19912	0.999959	0.0376002	11	2	1819	1370	over
GO:0000904	cell morphogenesis involved in differentiation	0.19912	0.999959	0.0376002	11	2	1819	1370	over
GO:0071310	cellular response to organic substance	0.19912	0.999959	0.0376002	11	2	1819	1370	over

GO:0044260	cellular macromolecule metabolic process	0.199627	0.857459	0.0077125	494	425	1336	947	under
GO:0045892	negative regulation of transcription, DNA-dependent	0.201302	0.999965	0.0380854	23	8	1807	1364	over
GO:0005789	endoplasmic reticulum membrane	0.202799	0.999971	0.0396153	21	7	1809	1365	over
GO:0006730	one-carbon metabolic process	0.202799	0.999971	0.0396153	21	7	1809	1365	over
GO:0044432	endoplasmic reticulum part	0.202799	0.999972	0.0399443	42	19	1788	1353	over
GO:0010564	regulation of cell cycle process	0.204289	0.999981	0.0406261	13	3	1817	1369	over
GO:0006352	transcription initiation	0.204289	0.999981	0.0406261	13	3	1817	1369	over
GO:0032774	RNA biosynthetic process	0.204289	0.999981	0.0406261	13	3	1817	1369	over
GO:0080134	regulation of response to stress	0.204289	0.999982	0.0408703	19	6	1811	1366	over
GO:0003713	transcription coactivator activity	0.204289	0.999982	0.0408703	19	6	1811	1366	over
GO:0019861	flagellum	0.204289	0.999982	0.0408703	19	6	1811	1366	over
GO:0030528	transcription regulator activity	0.204289	0.999983	0.0411659	74	39	1756	1333	over
GO:0006996	organelle organization	0.204289	0.999983	0.0413234	91	50	1739	1322	over
GO:0006022	aminoglycan metabolic process	0.204559	0.870943	0.0086437 4	7	16	1823	1356	under
GO:0051236	establishment of RNA localization	0.207396	0.999986	0.0416706	17	5	1813	1367	over
GO:0008080	N-acetyltransferase activity	0.209464	0.999989	0.0417376	15	4	1815	1368	over
GO:0031399	regulation of protein modification process	0.209464	0.999989	0.0420221	50	24	1780	1348	over
GO:0000267	cell fraction	0.209464	0.999989	0.0421864	66	34	1764	1338	over
GO:0006030	chitin metabolic process	0.210633	0.884798	0.0093084 8	4	12	1826	1360	under
GO:0051301	cell division	0.211692	0.999991	0.0428335	28	11	1802	1361	over
GO:0016887	ATPase activity	0.212864	0.999992	0.0450998	75	40	1755	1332	over
GO:0051707	response to other organism	0.212864	0.999993	0.045472	26	10	1804	1362	over
GO:0003777	microtubule motor activity	0.212864	0.999993	0.045472	26	10	1804	1362	over
GO:0005654	nucleoplasm	0.213402	0.999994	0.0461112	38	17	1792	1355	over
GO:0004386	helicase activity	0.21567	0.999995	0.0481628	24	9	1806	1363	over
GO:0051173	positive regulation of nitrogen compound metabolic process	0.21567	0.999995	0.0484247	31	13	1799	1359	over
GO:0004222	metalloendopeptidase activity	0.227081	0.908076	0.0100406	2	9	1828	1363	under
GO:0004252	serine-type endopeptidase activity	0.251081	0.932941	0.0120064	11	20	1819	1352	under
GO:0008235	metalloexopeptidase activity	0.288118	0.958134	0.0131676	1	7	1829	1365	under
GO:0006026	aminoglycan catabolic process	0.301743	0.966596	0.0135593	3	10	1827	1362	under
GO:0007156	homophilic cell adhesion	0.372602	0.988641	0.0143831	0	5	1830	1367	under
GO:0003887	DNA-directed DNA polymerase activity	0.372602	0.988641	0.0143831	0	5	1830	1367	under
GO:0005044	scavenger receptor activity	0.372602	0.988641	0.0143831	0	5	1830	1367	under

GO:0044249	cellular biosynthetic process	0.38539	0.991158	0.0152544	341	299	1489	1073	under
GO:0009058	biosynthetic process	0.386792	0.992112	0.0163379	348	304	1482	1068	under
GO:0005507	copper ion binding	0.422632	0.995474	0.0193958	5	12	1825	1360	under
GO:0032982	myosin filament	0.527314	0.998961	0.0249273	3	9	1827	1363	under
GO:0000272	polysaccharide catabolic process	0.628215	0.999797	0.0293503	4	10	1826	1362	under
GO:0007186	G-protein coupled receptor protein signaling pathway	0.628215	0.999797	0.0293503	4	10	1826	1362	under
GO:0003677	DNA binding	0.653568	0.999878	0.0318849	105	102	1725	1270	under
GO:0008745	N-acetylmuramoyl-L-alanine amidase activity	0.839326	0.999996	0.0336236	0	4	1830	1368	under
GO:0006027	glycosaminoglycan catabolic process	0.839326	0.999996	0.0336236	0	4	1830	1368	under
GO:0009253	peptidoglycan catabolic process	0.839326	0.999996	0.0336236	0	4	1830	1368	under
GO:0000270	peptidoglycan metabolic process	0.839326	0.999996	0.0336236	0	4	1830	1368	under
GO:0006508	proteolysis	0.998298	1.0	0.0435359	100	96	1730	1276	under

Comparaison des transcriptomes des populations S et A par test exact de Fisher

Gene Ontology	Activité peptidique	FDR	FWER	p	Nb séquences test (popS)	Nb séquences référent (popA)	Nb séquences test non annotées	Nb séquences référent non annotées	Niveau d'expression
GO:0006260	DNA replication	4.86072E-9	8.28695E-9	0.0	107	27	1265	2322	over
GO:0003676	nucleic acid binding	4.86072E-9	8.28695E-9	0.0	361	365	1011	1984	over
GO:0003964	RNA-directed DNA polymerase activity	4.86072E-9	8.28695E-9	0.0	94	4	1278	2345	over
GO:0034061	DNA polymerase activity	4.86072E-9	8.28695E-9	0.0	99	6	1273	2343	over
GO:0006259	DNA metabolic process	4.86072E-9	8.28695E-9	0.0	167	80	1205	2269	over
GO:0006278	RNA-dependent DNA replication	4.86072E-9	8.50626E-9	7.3218E-13	93	6	1279	2343	over
GO:0016779	nucleotidyltransferase activity	4.86072E-9	8.50626E-9	1.02554E-12	105	18	1267	2331	over
GO:0016772	transferase activity, transferring phosphorus-containing groups	1.43081E-8	2.86162E-8	2.3163E-10	149	122	1223	2227	over
GO:0003723	RNA binding	5.20961E-8	1.17216E-7	1.23421E-9	169	152	1203	2197	over
GO:0034645	cellular macromolecule biosynthetic process	5.1718E-7	1.29295E-6	1.31705E-8	213	219	1159	2130	over
GO:0009059	macromolecule biosynthetic process	6.03158E-7	1.65868E-6	1.6891E-8	213	220	1159	2129	over
GO:0015074	DNA integration	8.06661E-7	2.41998E-6	2.58882E-8	32	8	1340	2341	over
GO:0090304	nucleic acid metabolic process	1.47332E-6	4.78828E-6	4.11124E-8	229	246	1143	2103	over
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	6.54316E-6	2.29008E-5	1.99834E-7	265	305	1107	2044	over
GO:0016740	transferase activity	7.15026E-6	2.68131E-5	2.35442E-7	205	220	1167	2129	over
GO:0034641	cellular nitrogen compound metabolic process	2.71476E-5	1.08584E-4	8.99195E-7	284	341	1088	2008	over
GO:0006807	nitrogen compound metabolic process	4.26741E-5	1.81349E-4	1.39962E-6	303	372	1069	1977	over
GO:0044464	cell part	8.47211E-5	6.20521E-5	3.13979E-10	768	1556	604	793	under
GO:0005515	protein binding	8.47211E-5	6.28846E-5	1.53755E-8	366	833	1006	1516	under
GO:0044424	intracellular part	8.47211E-5	6.35407E-5	2.54483E-8	610	1263	762	1086	under
GO:0065007	biological regulation	3.07707E-4	3.07661E-4	3.24094E-6	283	640	1089	1709	under
GO:0050789	regulation of biological process	3.66402E-4	4.57899E-4	4.83162E-6	223	523	1149	1826	under
GO:0050794	regulation of cellular process	8.4825E-4	0.00127157	1.45442E-5	193	457	1179	1892	under
GO:0008152	metabolic process	0.00107767	0.00510585	3.55252E-5	758	1138	614	1211	over
GO:0060255	regulation of macromolecule metabolic	0.00131445	0.00267188	3.00082E-5	97	260	1275	2089	under

	process								
GO:0019222	regulation of metabolic process	0.00131445	0.00317308	3.4096E-5	109	284	1263	2065	under
GO:0044428	nuclear part	0.00131445	0.00317946	3.44739E-5	95	255	1277	2094	under
GO:0031323	regulation of cellular metabolic process	0.00131445	0.00328075	3.6703E-5	100	265	1272	2084	under
GO:0043226	organelle	0.00187557	0.00514457	5.65369E-5	315	676	1057	1673	under
GO:0043229	intracellular organelle	0.00205204	0.00636267	6.32454E-5	315	675	1057	1674	under
GO:0044422	organelle part	0.00205204	0.00664698	6.63343E-5	293	634	1079	1715	under
GO:0010556	regulation of macromolecule biosynthetic process	0.00225425	0.00801967	8.02406E-5	56	167	1316	2182	under
GO:0080090	regulation of primary metabolic process	0.00225425	0.00897654	9.39491E-5	98	255	1274	2094	under
GO:0044238	primary metabolic process	0.00247349	0.0122913	7.28905E-5	607	889	765	1460	over
GO:0031326	regulation of cellular biosynthetic process	0.00276055	0.0116639	1.10985E-4	60	174	1312	2175	under
GO:0044446	intracellular organelle part	0.00299201	0.013374	1.33987E-4	291	624	1081	1725	under
GO:0016070	RNA metabolic process	0.0124402	0.0573813	5.21374E-4	58	161	1314	2188	under
GO:0022402	cell cycle process	0.0142372	0.0687145	5.9906E-4	44	130	1328	2219	under
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.0156947	0.079098	6.7196E-4	54	151	1318	2198	under
GO:0000166	nucleotide binding	0.0184121	0.0963142	8.04888E-4	158	358	1214	1991	under
GO:0051171	regulation of nitrogen compound metabolic process	0.0184175	0.103156	8.33547E-4	56	154	1316	2195	under
GO:0044260	cellular macromolecule metabolic process	0.0195925	0.0977514	5.42153E-4	425	609	947	1740	over
GO:0044451	nucleoplasm part	0.0208378	0.122122	9.93706E-4	16	63	1356	2286	under
GO:0044249	cellular biosynthetic process	0.0265946	0.136084	7.99884E-4	299	411	1073	1938	over
GO:0019899	enzyme binding	0.0317914	0.218407	0.00186198	15	58	1357	2291	under
GO:0046982	protein heterodimerization activity	0.0473564	0.320556	0.00250114	2	21	1370	2328	under
GO:0016043	cellular component organization	0.0473564	0.32347	0.00256234	133	300	1239	2049	under
GO:0009058	biosynthetic process	0.0483083	0.242552	0.00126037	304	423	1068	1926	over
GO:0048518	positive regulation of biological process	0.0536637	0.384043	0.00322254	100	234	1272	2115	under
GO:0043227	membrane-bounded organelle	0.0536637	0.387921	0.00328725	209	441	1163	1908	under
GO:0043231	intracellular membrane-bounded organelle	0.0536637	0.387921	0.00328725	209	441	1163	1908	under
GO:0048523	negative regulation of cellular process	0.0536637	0.391361	0.00339825	81	196	1291	2153	under
GO:0006397	mRNA processing	0.0540852	0.401878	0.00353313	12	48	1360	2301	under
GO:0005654	nucleoplasm	0.0552683	0.416687	0.00366593	17	60	1355	2289	under
GO:0005681	spliceosomal complex	0.0565568	0.432067	0.00388102	9	40	1363	2309	under
GO:0003824	catalytic activity	0.0570979	0.290099	0.00150626	643	982	729	1367	over
GO:0006366	transcription from RNA polymerase II	0.0610296	0.468977	0.00396432	0	12	1372	2337	under

	promoter								
GO:0008380	RNA splicing	0.0610296	0.473253	0.00400889	11	45	1361	2304	under
GO:0043232	intracellular non-membrane-bounded organelle	0.0610811	0.489383	0.00410604	118	267	1254	2082	under
GO:0043228	non-membrane-bounded organelle	0.0610811	0.489383	0.00410604	118	267	1254	2082	under
GO:0070001	aspartic-type peptidase activity	0.0614272	0.329234	0.00178609	11	3	1361	2346	over
GO:0004190	aspartic-type endopeptidase activity	0.0614272	0.329234	0.00178609	11	3	1361	2346	over
GO:0010628	positive regulation of gene expression	0.0654118	0.535127	0.00474324	7	34	1365	2315	under
GO:0009893	positive regulation of metabolic process	0.0654118	0.53602	0.00482251	35	99	1337	2250	under
GO:0045449	regulation of transcription	0.0654118	0.536479	0.00483251	43	116	1329	2233	under
GO:0044454	nuclear chromosome part	0.0698049	0.57493	0.00531653	2	19	1370	2330	under
GO:0007059	chromosome segregation	0.0698049	0.57493	0.00531653	2	19	1370	2330	under
GO:0051246	regulation of protein metabolic process	0.0751257	0.609193	0.00577673	44	117	1328	2232	under
GO:0032502	developmental process	0.0758916	0.620204	0.00618871	143	311	1229	2038	under
GO:0045941	positive regulation of transcription	0.0831731	0.6615	0.00633017	7	33	1365	2316	under
GO:0010557	positive regulation of macromolecule biosynthetic process	0.0831731	0.668015	0.0064911	13	48	1359	2301	under
GO:0010604	positive regulation of macromolecule metabolic process	0.0866839	0.689917	0.00651485	33	93	1339	2256	under
GO:0032991	macromolecular complex	0.0888073	0.706536	0.00708502	322	638	1050	1711	under
GO:0010605	negative regulation of macromolecule metabolic process	0.0888073	0.711794	0.00739384	39	105	1333	2244	under
GO:0032774	RNA biosynthetic process	0.095986	0.760166	0.0080651	3	21	1369	2328	under
GO:0031325	positive regulation of cellular metabolic process	0.095986	0.760812	0.00811147	35	96	1337	2253	under
GO:0009890	negative regulation of biosynthetic process	0.095986	0.763873	0.00823736	13	47	1359	2302	under
GO:0051649	establishment of localization in cell	0.095986	0.773657	0.00870484	39	104	1333	2245	under
GO:0043234	protein complex	0.095986	0.782188	0.00922317	201	415	1171	1934	under
GO:0065008	regulation of biological quality	0.095986	0.787201	0.00933678	101	227	1271	2122	under
GO:0010558	negative regulation of macromolecule biosynthetic process	0.095986	0.79204	0.00951433	12	44	1360	2305	under
GO:0009891	positive regulation of biosynthetic process	0.095986	0.794695	0.00970358	15	51	1357	2298	under
GO:0006996	organelle organization	0.095986	0.795062	0.00970716	50	126	1322	2223	under
GO:0023034	intracellular signaling pathway	0.100839	0.815571	0.00987502	23	69	1349	2280	under
GO:0005200	structural constituent of cytoskeleton	0.116564	0.862398	0.00998065	0	10	1372	2339	under
GO:0032268	regulation of cellular protein metabolic process	0.117175	0.869452	0.0104791	41	107	1331	2242	under

GO:0048522	positive regulation of cellular process	0.117175	0.871585	0.0108361	79	183	1293	2166	under
GO:0010627	regulation of intracellular protein kinase cascade	0.117507	0.876034	0.0109242	4	23	1368	2326	under
GO:0046907	intracellular transport	0.125005	0.894852	0.0113285	35	94	1337	2255	under
GO:0006351	transcription, DNA-dependent	0.125713	0.899404	0.0113356	3	20	1369	2329	under
GO:0031328	positive regulation of cellular biosynthetic process	0.126896	0.905866	0.012106	15	50	1357	2299	under
GO:0009850	auxin metabolic process	0.126896	0.909513	0.0124576	43	110	1329	2239	under
GO:0048468	cell development	0.126896	0.910514	0.0126567	10	38	1362	2311	under
GO:0044237	cellular metabolic process	0.128496	0.580044	0.00367166	583	892	789	1457	over
GO:0005634	nucleus	0.129198	0.91905	0.0138252	106	233	1266	2116	under
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.129198	0.921421	0.0139764	11	40	1361	2309	under
GO:0009851	auxin biosynthetic process	0.129198	0.924758	0.0144712	43	109	1329	2240	under
GO:0009892	negative regulation of metabolic process	0.129198	0.924758	0.0144712	43	109	1329	2240	under
GO:0008022	protein C-terminus binding	0.134355	0.934392	0.014513	1	13	1371	2336	under
GO:0043687	post-translational protein modification	0.134606	0.936891	0.0147151	46	115	1326	2234	under
GO:0032555	purine ribonucleotide binding	0.137349	0.94432	0.0156979	123	264	1249	2085	under
GO:0032553	ribonucleotide binding	0.137349	0.94432	0.0156979	123	264	1249	2085	under
GO:0071103	DNA conformation change	0.152678	0.961185	0.0158326	0	9	1372	2340	under
GO:0051254	positive regulation of RNA metabolic process	0.160162	0.969957	0.0165292	6	27	1366	2322	under
GO:0051052	regulation of DNA metabolic process	0.160162	0.969957	0.0165292	6	27	1366	2322	under
GO:0017076	purine nucleotide binding	0.160162	0.972055	0.0171629	132	280	1240	2069	under
GO:0032559	adenyl ribonucleotide binding	0.160162	0.972383	0.0175384	92	204	1280	2145	under
GO:0005737	cytoplasm	0.160162	0.972926	0.0175697	103	225	1269	2124	under
GO:0001882	nucleoside binding	0.160502	0.974639	0.0179235	101	221	1271	2128	under
GO:0010647	positive regulation of cell communication	0.160502	0.976285	0.0185037	5	24	1367	2325	under
GO:0022607	cellular component assembly	0.160502	0.977017	0.0189261	65	151	1307	2198	under
GO:0005524	ATP binding	0.160502	0.978428	0.0195382	92	203	1280	2146	under
GO:0042446	hormone biosynthetic process	0.160502	0.978502	0.019564	46	113	1326	2236	under
GO:0007165	signal transduction	0.160502	0.9789	0.0196135	16	50	1356	2299	under
GO:0010629	negative regulation of gene expression	0.1613	0.981356	0.0205039	13	43	1359	2306	under
GO:0022403	cell cycle phase	0.1613	0.981356	0.0205039	13	43	1359	2306	under
GO:0006886	intracellular protein transport	0.1613	0.981668	0.0205448	17	52	1355	2297	under
GO:0004713	protein tyrosine kinase activity	0.171693	0.987652	0.0215617	1	12	1371	2337	under
GO:0045934	negative regulation of nucleobase,	0.171693	0.988342	0.0217679	14	45	1358	2304	under

	nucleoside, nucleotide and nucleic acid metabolic process								
GO:0080134	regulation of response to stress	0.171693	0.988869	0.0218347	6	26	1366	2323	under
GO:0045893	positive regulation of transcription, DNA-dependent	0.171693	0.988869	0.0218347	6	26	1366	2323	under
GO:0001883	purine nucleoside binding	0.171693	0.989063	0.0220332	101	219	1271	2130	under
GO:0030554	adenyl nucleotide binding	0.171693	0.989063	0.0220332	101	219	1271	2130	under
GO:0035468	positive regulation of signaling pathway	0.173642	0.990596	0.0220352	3	18	1369	2331	under
GO:0060249	anatomical structure homeostasis	0.173642	0.990596	0.0220352	3	18	1369	2331	under
GO:0023046	signaling process	0.173642	0.990883	0.022298	26	71	1346	2278	under
GO:0051172	negative regulation of nitrogen compound metabolic process	0.180214	0.992767	0.0229456	15	47	1357	2302	under
GO:0004386	helicase activity	0.180214	0.993339	0.0239295	9	33	1363	2316	under
GO:0006350	transcription	0.180214	0.993339	0.0239295	9	33	1363	2316	under
GO:0016604	nuclear body	0.180752	0.994164	0.0247209	5	23	1367	2326	under
GO:0030154	cell differentiation	0.180752	0.994274	0.0249061	23	64	1349	2285	under
GO:0023060	signal transmission	0.180752	0.994274	0.0249061	23	64	1349	2285	under
GO:0003677	DNA binding	0.196393	0.747257	0.0052421	102	124	1270	2225	over
GO:0048477	oogenesis	0.2075	0.99764	0.0251116	0	8	1372	2341	under
GO:0023052	signaling	0.2075	0.997649	0.0252176	56	131	1316	2218	under
GO:0042445	hormone metabolic process	0.2075	0.99772	0.0255457	52	123	1320	2226	under
GO:0019904	protein domain specific binding	0.215677	0.998405	0.0278042	8	30	1364	2319	under
GO:0002682	regulation of immune system process	0.215677	0.998602	0.0286774	6	25	1366	2324	under
GO:0031324	negative regulation of cellular metabolic process	0.215677	0.998712	0.0294604	40	98	1332	2251	under
GO:0016301	kinase activity	0.215677	0.998728	0.0295358	43	104	1329	2245	under
GO:0004672	protein kinase activity	0.215677	0.998749	0.0296612	23	63	1349	2286	under
GO:0016481	negative regulation of transcription	0.215677	0.998863	0.0304317	9	32	1363	2317	under
GO:0005625	soluble fraction	0.215677	0.998863	0.0304317	9	32	1363	2317	under
GO:0005694	chromosome	0.215677	0.999003	0.0304478	3	17	1369	2332	under
GO:0006275	regulation of DNA replication	0.215677	0.999003	0.0304478	3	17	1369	2332	under
GO:0016773	phosphotransferase activity, alcohol group as acceptor	0.215677	0.999022	0.030722	25	67	1347	2282	under
GO:0006464	protein modification process	0.215677	0.99903	0.0309465	63	143	1309	2206	under
GO:0051173	positive regulation of nitrogen compound metabolic process	0.215677	0.999065	0.0315042	13	41	1359	2308	under
GO:0005876	spindle microtubule	0.226823	0.999417	0.0318807	1	11	1371	2338	under
GO:0015031	protein transport	0.226823	0.999418	0.0319578	28	73	1344	2276	under

GO:0007264	small GTPase mediated signal transduction	0.226823	0.999451	0.0322836	7	27	1365	2322	under
GO:0032101	regulation of response to external stimulus	0.229806	0.999599	0.0323508	2	14	1370	2335	under
GO:0022604	regulation of cell morphogenesis	0.229806	0.999599	0.0323508	2	14	1370	2335	under
GO:0007600	sensory perception	0.229806	0.999599	0.0323508	2	14	1370	2335	under
GO:0010817	regulation of hormone levels	0.229806	0.999605	0.0327947	53	123	1319	2226	under
GO:0030054	cell junction	0.230808	0.99964	0.0328258	5	22	1367	2327	under
GO:0051301	cell division	0.237239	0.999728	0.03501	11	36	1361	2313	under
GO:0043412	macromolecule modification	0.239978	0.999779	0.037043	69	153	1303	2196	under
GO:0006355	regulation of transcription, DNA-dependent	0.239978	0.999781	0.037073	27	70	1345	2279	under
GO:0045184	establishment of protein localization	0.243558	0.999821	0.0377146	29	74	1343	2275	under
GO:0023033	signaling pathway	0.243558	0.999829	0.0381469	47	110	1325	2239	under
GO:0003712	transcription cofactor activity	0.243558	0.999843	0.0384884	9	31	1363	2318	under
GO:0051252	regulation of RNA metabolic process	0.243558	0.999849	0.038738	35	86	1337	2263	under
GO:0000018	regulation of DNA recombination	0.288787	0.999976	0.0398226	0	7	1372	2342	under
GO:0045132	meiotic chromosome segregation	0.288787	0.999976	0.0398226	0	7	1372	2342	under
GO:0051128	regulation of cellular component organization	0.288787	0.999976	0.0403481	21	57	1351	2292	under
GO:0033043	regulation of organelle organization	0.290387	0.999979	0.041541	7	26	1365	2323	under
GO:0047485	protein N-terminus binding	0.291349	0.999984	0.0417915	3	16	1369	2333	under
GO:0005813	centrosome	0.291349	0.999984	0.0417915	3	16	1369	2333	under
GO:0001932	regulation of protein amino acid phosphorylation	0.291349	0.999984	0.0417915	3	16	1369	2333	under
GO:0031399	regulation of protein modification process	0.291349	0.999985	0.0421435	24	63	1348	2286	under
GO:0007399	nervous system development	0.308207	0.999994	0.0456188	2	13	1370	2336	under
GO:0070161	anchoring junction	0.308207	0.999994	0.0456188	2	13	1370	2336	under
GO:0016192	vesicle-mediated transport	0.308207	0.999994	0.0462806	19	52	1353	2297	under
GO:0001894	tissue homeostasis	0.321022	0.999997	0.0468803	1	10	1371	2339	under
GO:0007292	female gamete generation	0.321022	0.999997	0.0468803	1	10	1371	2339	under
GO:0007017	microtubule-based process	0.324572	0.999998	0.0486896	48	110	1324	2239	under
GO:0006396	RNA processing	0.324572	0.999998	0.0489757	47	108	1325	2241	under
GO:0016607	nuclear speck	0.328202	0.999998	0.0498718	4	18	1368	2331	under
GO:0010212	response to ionizing radiation	0.99835	0.999936	0.0282249	5	1	1367	2348	over
GO:0046914	transition metal ion binding	0.99835	0.999973	0.0321885	135	188	1237	2161	over
GO:0006414	translational elongation	0.99835	0.999999	0.0361724	46	54	1326	2295	over
GO:0007389	pattern specification process	0.99835	0.999999	0.042021	24	24	1348	2325	over

GO:0004888	transmembrane receptor activity	0.99835	1.0	0.045625	20	19	1352	2330	over
GO:0015002	heme-copper terminal oxidase activity	0.99835	1.0	0.0463932	12	9	1360	2340	over
GO:0004129	cytochrome-c oxidase activity	0.99835	1.0	0.0463932	12	9	1360	2340	over
GO:0016676	oxidoreductase activity, acting on heme group of donors, oxygen as acceptor	0.99835	1.0	0.0463932	12	9	1360	2340	over
GO:0016675	oxidoreductase activity, acting on heme group of donors	0.99835	1.0	0.0463932	12	9	1360	2340	over
GO:0005507	copper ion binding	0.99835	1.0	0.0463932	12	9	1360	2340	over

*C- Article soumis à la revue
Conservation Genetics*

GLACIAL REFUGIUM VERSUS RANGE LIMIT: CONSERVATION GENETICS OF *MACOMA BALTHICA*, A KEY SPECIES IN THE BAY OF BISCAY.

V. Becquet^{1,3}, B. Simon-Bouhet^{1,4}, H. Hummel^{2,5} and P. Garcia^{1,6}.

Corresponding author: V. Becquet

Tel: 033546507637

Fax: 033546458264

¹ Littoral Environnement et Sociétés, UMR 6250, 2 rue Olympe de Gouges, 17000 La Rochelle, France

² Centre for Estuarine and Marine Ecology, Netherlands Institute of Ecology, Koringaweg 7, 4401 NT Yerseke, The Netherlands

³ E-mail: vbecquet@univ-lr.fr

⁴ E-mail: besibo@gmail.com

⁵ E-mail: H.Hummel@nioo.knaw.nl

⁶ E-mail: pgarciam@univ-lr.fr

Abstract: The bivalve *Macoma balthica* (L.) is a key-species of the intertidal mudflats in Europe and moreover in France. It is thus of importance to consider the conservation implication of this species in the trophic web. Yet, its natural range experienced a shift towards the northeast along the European coastline during the past 5 decades. This southern boundary shift seems to be correlated with the increased sea surface temperature in the Bay of Biscay (France), a major glacial refugium during LGM (Last Glacial Maximum, 18000 years ago). Both mitochondrial and nuclear genetic markers revealed some highly divergent populations consistent with populations close to a glacial refugium. Moreover, populations in the Bay of Biscay are structured with a gene flow rupture near Brittany. Despite a high genetic diversity, southernmost populations exhibit a low tolerance to elevated temperature. Then, considering (i) the impact of seasonal variations of temperature on bivalve recruitment and (ii) the sensibility of *M. balthica* to an elevated temperature, the southernmost *M. balthica* populations should be monitored in the context of climate warming.

Key words: Refuge zone, Global warming, Boundary shift, *Macoma balthica*, Microsatellites, Mitochondrial DNA.

1. INTRODUCTION

For any given species, populations located near the boundary of the distribution area should be carefully monitored in the context of conservation biology. Such populations theoretically present structural and genetic characteristics that may limit their adaptive capacity (Eckert et al. 2008). Indeed, near the species range limit, populations are not at their optimum fitness and present most of the time a reduced growth rate (Brown et al. 1995; Thomas and Kunin 1999). As a consequence they are often small fragmented populations. Such fragmented structures facilitate the action of genetic drift and expose populations to the Allee effect and reduced mutation rate (Courchamp et al. 1999; Alleaume-Benharira et al. 2006). Moreover, asymmetric gene flow from central populations to the periphery inhibits any opportunity to develop local adaptations. This process known as the migration-selection balance, causes chronic maladaptation (Ronce and Kirkpatrick 2001; Bolnick et al. 2008). Furthermore environmental changes associated to supposedly low genetic diversity in populations close to the species' range limit, might increase the risk of extinction for such populations (Lawton 1993; Vucetich and Wayte 2003). The dynamic equilibrium between extinction and recolonization in these areas is more fragile, and renewal of populations more difficult (Lande et al. 2003). Under these conditions, the rate of environmental changes is of direct importance because it is the main limitation to the development of an adaptive process.

In recent decades, the anthropogenic pressures applied on natural populations have dramatically increased (Vitousek et al. 1997; Walker and Kendrick 1998) accelerating habitat fragmentation, change of species ranges and maladaptations. Moreover, for the last 30 years, these effects have been accelerated by global warming (Bell and Collins 2008) and are thus too recent and too fast for adaptation to have occurred. Indeed, stressors related to human activities (pollution, loss and fragmentation of habitat, release of invasive species...) that lead to an interruption of connectivity within and between populations come in synergy with the increase in temperature (Gordon 1998; Schroter et al. 2005). Thus, between 1906 and 2005, the mean temperature of the earth's surface increased by 0.74 °C, partially due to the

greenhouse effect (GIEC 2007) and in the marine environment, the average temperature of surface waters increased by 0.23°C per decade between 1977 and 2007 in the North East Atlantic (elevation of the surface water temperature in the Bay of Biscay in France, Goikoetxea et al. 2009). For more than 1500 species, Parmesan and Yohe (2003) showed that this phenomenon resulted in switches of biotic trends coherent with the expected response to global warming: shifts of species range towards the pole, higher altitudes for terrestrial species or higher depths for aquatic species (see also Dulvy et al. 2008; Wethey and Woodin 2008). The response of populations located near their range limit to such environmental pressures remain mostly unknown (Walther et al. 2002) and consequences of species range shifts for the functioning and structure of ecosystems might be heavy, particularly when key-species are concerned.

As an example, the Baltic clam, *Macoma balthica* (L.) (Bivalvia: Tellinidae), is an infaunal tellinid bivalve commonly present in marine and estuarine soft-bottom habitats of the northern hemisphere. As an important prey for migratory birds (Piersma and Beukema 1993), macro-invertebrates (Edjung and Bonsdorff 1992) and fish (Mattila and Bonsdorff 1998), *M. balthica* holds an important place in trophic webs of the European coasts (Philippart et al. 2003). It usually occupies tidal and subtidal sandy and muddy bottoms to a maximal depth of 30 m (Scarlatto 1981; Hummel 1985). *Macoma balthica* is an eurythermal species and can thus thrive along a large temperature gradient. In Europe, it is currently largely distributed from the eastern Pechora Sea (northern range limit) in Russia (Hummel et al. 1997b) to the Gironde estuary (southern range limit) in France (Bachelet 1980). Nevertheless, during the past five decades, the natural range of *Macoma balthica* strongly shifted towards the northeast of the European coasts (Hummel et al. 2000). Abundant along the Atlantic coasts of the Iberian Peninsula more than 40 years ago (Otero and Milan 1970) it has now completely disappeared in this area.

On the evolutionary scale, several studies have already been performed on *Macoma balthica*. Väinölä (2003) showed with allozymes the existence of two groups of populations, one large NE Atlantic homogeneous

population (including the Norwegian sea, the North sea and the coasts of the British Islands), and one structured meta-population with a high genetic diversity and a high number of alleles in the northern part of the natural range of the species (Pacific ocean, Baltic and White sea). As stated by Väinölä (2003) these two groups are probably two subspecies of *Macoma balthica* coexisting in Europe, the Pacific lineage *Macoma balthica balthica* and the Atlantic one *Macoma balthica rubra*. He designates the Baltic Sea population as a hybrid swarm resulting of a mix of two genomes after a secondary invasion from the Pacific Ocean during Pliocene. Luttikhuisen et al. (2003) and Nikula et al. (2008) have confirmed these results using mitochondrial markers. However, Luttikhuisen et al. (2003) also revealed a very specific mitochondrial pattern in the southernmost population of *M. balthica* (i.e. Gironde estuary, France), with haplotypes quite divergent from the other European and Pacific populations. On the contrary, Hummel et al (1995, 1997b) and Strelkov et al. (2007) having studied French Atlantic populations (including the Gironde and the Loire estuary) did not reveal any isolation from other populations with allozymes. These populations were indeed similar to the Norwegian Sea and western Barents Sea populations.

The attention paid on *Macoma balthica* populations living near the southern limit of the range remains limited to only two populations previously studied in the Bay of Biscay (Gironde and/or Loire estuary for Hummel 1997b, Luttikhuisen et al. 2003 and Strelkov et al. 2007). The Bay of Biscay is known to have been one of the four major pliocene refuge in Europe during the Last Glacial Maximum (LGM) 18000 years ago (Gomez and Lunt 2007). As the ice retreated, species recolonized northward by successive founder-events (Hewitt 1999). Thereby genetic diversity is expected to decrease from the refuge zone to newly established populations (Petit et al. 2003). In an extensive genetic study based on isoenzymes for several French stations compared to central and northern populations, Hummel et al. (1997b) showed however an increasing diversity (measured as observed and expected heterozygosity and number of alleles) from south to north, the lowest diversity being found in the Gironde population.

For *Macoma balthica* populations located near the southern range limit of the species, we have thus two sets of opposite theoretical expectations: on the one hand, these populations are expected to harbor a low level of genetic diversity since they are located at the edge of the species natural range, but on the other hand, these populations should harbor a high level of genetic diversity since they are located in a recent glacial refugium. From a conservation point of view and given the theoretical vulnerability of these populations, it is therefore of importance to have a proper description of their precise distribution and genetic diversity and structure to infer their potential for adaptation under shifting selective environments (e.g., climate change, anthropogenic pressure, population collapse, Moritz 1994; Crandall et al. 2000). In this context, we conducted a mitochondrial analysis of 12 populations sampled along the French coastline and 6 populations sampled in the North Sea, the English Channel and the North Atlantic Ocean to assess historical processes. Then contemporary features such as recent demographic history and gene flow were investigated using a set of newly developed microsatellites markers (Becquet et al. 2009) with an emphasis on populations located at the edge of the current natural range of *Macoma balthica*.

2. MATERIALS AND METHODS

2.1 Sampling

Adult specimens of *Macoma balthica* have been collected between 2003 and 2007 in 18 locations ranging from the Gironde estuary (known southern limit) to the Barents Sea (Murman population, see fig. 1 and table 1 for precise locations and sample sizes). Specimens were deep frozen or stored in 95% ethanol until DNA extractions.

2.2 Mitochondrial COI sequencing of PCR products

Total DNA was extracted from less than 15 mg of muscle using the DneasyTM Tissue Kit (Qiagen, Germany). Amplifications of a COI gene fragment (Cytochrome Oxidase Subunit 1) were performed using the set of primers described in Luttikhuizen et al. (2003). Amplifications were carried

out in a total volume of 50 μ L consisting of 1X PCR buffer, 1.85 mM MgCl₂, 125 μ M dNTPs, 0.25 μ M of each primer, 1.6 U of Taq DNA polymerase (Red Hot ® *TAQ* DNA Polymerase, Abgène) and about 10 ng of template DNA. A MJResearch PTC 100 Thermal Cycler was used with the following cycling profile: 3 min of initial denaturation at 95 °C, followed by 35 cycles of denaturation at 95 °C for 45 s, annealing at 60 °C for 60 s and extension at 72 °C for 90 s. A final extension step was carried out for 5 min at 72 °C. Double-stranded PCR products were cleaned using MultiScreen-PCR MANU03010 plates (Millipore). The sequencing was performed by Genoscreen corporation (Campus Pasteur - 1 rue du Professeur Calmette - 59 000 Lille - France), using an ABI PRISM® 3730 XL automated DNA Sequencer (Perkin-Elmer Applied Biosystems, Foster City, CA). Sequence data were aligned using ClustalX (Thompson et al., 1994) and ambiguities were manually checked comparing each sequence with its complementary fragment using BioEdit (Hall 1999).

2.3 Microsatellite amplification and genotyping

PCR amplifications of 8 microsatellite loci were conducted using the primers described in Becquet et al. (2009). The protocol for DNA amplification consisted in a touch-down PCR procedure: 3 min of an initial denaturation step at 95°C followed by 10 cycles of denaturation at 95°C for 45 s, annealing at temperature decreasing by 1°C per cycle from T_{initial} to $T_{\text{annealing}}$ (T_{initial} varied between 68°C and 62°C depending on the primers, Becquet et al. 2009) for 45s and extension at 72°C for 45s, then followed by 30 cycles at 95°C for 45s, annealing at final $T_{\text{annealing}}$ for 45s and extension at 72°C for 45s. A final extension step was carried out for 7 min at 72°C. PCR products were screened on 6,5% polyacrylamide gels using a Li-Cor NEN Global IR2 DNA sequencer. Allele sizes were determined using a known DNA sequence with the SAGA-GT software (v3.1: Automated microsatellite Analysis Software, LI-COR Biosciences).

2.4 COI sequence analyses

Haplotype number H , haplotype diversity H_d (Nei 1987), number of polymorphic sites S , and nucleotide diversity π (i.e. average number of

nucleotide differences between pairs of sequences) were calculated for each population using the software DnaSP v.5 (Librado and Rozas 2009). We examined the overall and regional population genetic structure by performing an analysis of molecular variance (AMOVA, Excoffier et al. 1992) as implemented in ARLEQUIN v.3.11. (Excoffier et al. 2005). Fixation indices (Wright 1951) analogous to F_{st} , F_{ct} and F_{sc} parameters, were calculated to analyze genetic differentiation between populations over the whole study area (Φ_{st}), among groups of populations within the species range (Φ_{ct}) and among populations within groups (Φ_{sc}) as defined in Excoffier et al. (1992). Φ -statistics take into account both haplotype frequencies and the molecular distances between haplotypes. Statistical significance of the Φ indices was tested using a non-parametric permutation procedure implemented in ARLEQUIN v.3.11. Phylogenetic analyses at the intra-specific level were performed with the software Network (Bandelt et al. 1999) that builds haplotypic networks based on the median-joining algorithm (Cassens et al. 2003). This process combines both minimum spanning trees and maximum parsimony approaches to simplify the complex branching pattern and represent all the most parsimonious intra-specific phylogenies. Such networks are picturing the evolutionary relationships (based on mutational events) between the different haplotypes observed over the whole data set.

Sets of populations were defined based on mitochondrial similarities observed at the population level in order to test for genetic differences between groups of populations. One sided tests to compare means were performed in R v 2.11.1 (R Development Core Team, 2010) using either Student's t -test or Wilcoxon test when parametric assumptions were not met.

2.5 Microsatellites analyses

2.5.1 Genetic diversity within localities

Pairwise linkage disequilibrium among loci was tested using GENEPOP version 4.0.10 (Rousset 2008), and significance levels were evaluated using a Markov-chain randomization procedure (MCMC) with 10000 dememorization steps, 5000 batches and 10000 iterations per batch.

The number of alleles per population and locus (Nall), the observed (H_o) and expected (H_e) heterozygotes under Hardy-Weinberg equilibrium were calculated using Genetix 4.05.2. (Belkhir et al. 1996-2004). Allelic richness (Rall) was computed using FSTAT 2.9.3.2 (Goudet 2002) to account for the differences in sampling sizes across populations. Mean and overall Nall, Rall, H_o and H_e were also computed for each set of populations (BayBiscay and Core-populations). As for the mitochondrial dataset, one-sided tests to compare mean values for each set of populations were performed following the procedure described earlier.

2.5.2 Genetic differentiation and relationships among localities

Pairwise genetic differentiation between populations (F_{st}) and departure from Hardy-Weinberg equilibrium (F_{is}) were computed in Genetix 4.05.2 (Belkhir et al. 1996-2004). To estimate relationship between population at the European scale, Cavalli-Sforza and Edwards (1967) genetic distances were calculated from allelic frequencies using PHYLIP (Felsenstein 2004). A consensus unrooted neighbor-joining (NJ) population tree (Saitou and Nei 1987) was then constructed using CONSENSE (Margush and McMorris 1981), with 1000 bootstrap iterations computed using Seqboot (Felsenstein 1985), as implemented in PHYLIP 3.69. Because the assumption that populations are in mutation-drift equilibrium may be violated for the computation of F_{st} statistics in some populations that have suffered drastic demographic declines, we further tested the population structure via a Bayesian clustering method implemented in the R package Geneland (Guillot et al. 2005). This package avoids the manual grouping of populations and takes into account both the geographical coordinates of sampled individuals and their genotypes at multiple loci to identify groups at Hardy-Weinberg equilibrium. For each run, the number of groups or classes (K) at Hardy-Weinberg equilibrium is assessed using maximum-likelihood estimations. Then, K is fixed at its most likely value and the posterior probability to belong to each class is computed for each individual of the dataset. Finally, maps of posterior probabilities are generated. A total of 30 independent runs have been computed in order to check for the convergence of MCMC

computations with the following parameters: length of MCMC = 150 000, thinning = 10, maximum rate of Poisson process = 550, maximum seed number for the Poisson-Voronoi tessellation = 1650.

3. RESULTS

3.1 Mitochondrial structure

A 313-bp fragment of the COI gene was sequenced for 424 individuals sampled in 17 locations (Table 1). We identified 20 polymorphic sites defining 19 haplotypes (Genbank accession numbers: HM756170 to HM756189). The genealogical relationships between these 19 haplotypes are depicted in the minimum spanning network presented in fig. 2. The network is structured in four major clades organized around the 4 most frequent haplotypes (H1 to H4 frequencies sum up to 94.7%). Three of these four haplotypes (H1 to H3) are separated by a single mutation while the more distant haplotype (H4) was found 3 mutations apart. The other 15 haplotypes were found at low frequencies and are equally spread around the 4 major clades. The overall structure of the network is consistent with the geographic locations of sampling sites: haplotypes H1 and H2 were found within the Bay of Biscay populations only, haplotype H3 was found mostly in the English Channel and North Sea, while the most distant haplotype (i.e. H4) is characteristic of the Ireland and Murman populations (fig.1 and fig. 3).

3.2 Molecular diversity

Based on the mitochondrial structure, 2 populations sets have been identified to test for potential molecular diversity differences between populations presenting contrasting haplotypic patterns. The Bay of Biscay set includes the 9 southernmost locations (populations 1 to 9, table 1), while the Core set includes the remaining 9 sampling sites (populations 10 to 18, table 1).

Overall (except for the monomorphic population of Murman) the number of haplotypes H and segregating sites S ranged from 2 to 5 (mean = 3.35, sd =1.32) and from 2 to 7 (mean = 3.71, sd =1.65) respectively (Table 2). When comparing mitochondrial diversity between the two population sets, the mean haplotype number (see Table 2) was found significantly higher

in the Bay of Biscay set than in the Core set (one sided Wilcoxon test, $W = 55$, $p = 0.033$). The same result was obtained for the nucleotide diversity (one sided Wilcoxon test, $W = 57$, $p = 0.023$) with an even more striking difference between both population sets: mean nucleotide diversity within the Bay of Biscay set ($\pi = 0.03153 \pm 0.02646$) was found to be more than ten fold greater than within the Core set ($\pi = 0.00275 \pm 0.00192$, Table 2).

At nuclear level, the multilocus genotype was obtained for 542 individuals collected in 18 sampling sites (Table 1). No linkage disequilibrium was detected (pairwise tests p values ranged from 0.2308 to 1.0000). Overall, the number of alleles (Nall) ranged from 6,62 (Bonne Anse) to 11.25 (Aytre) with a total of 20.87 over the whole dataset (Table 2). Most of the differences observed across populations were due to varying sample sizes since all allelic richness values (Rall) were found to be similar, ranging from 4.0295 (11) to 4.7716 (6) (population overall mean and standard deviation = 4.39 ± 0.23). However, a highly significant difference (one sided two sample Student t -test, $t = 5.933$, $df = 16$, $p = 1.051e^{-5}$) was found when comparing mean Rall values for both population sets, with a greater mean Rall value within the Bay of Biscay set (Rall = 4.58 ± 0.15) than within the Core set (Rall = 4.21 ± 0.12 , see also fig. 4a).

The same pattern was observed for both the observed and the expected heterozygosities (see fig. 4b). While H_o and H_s showed little variation across the whole dataset (mean and standard deviation across the 18 populations equals $0.366 \pm 0,0425$ and $0.6325 \pm 0,0347$ for H_o and H_s respectively, Table 2), both indices were found to be significantly greater in the Bay of Biscay set than in the Core set ($t = 2.513$, $df = 16$, $p = 0.012$ for H_o and $t = 2.246$, $df = 16$, $p = 0.020$ for H_s). Heterozygote deficiency was quite high for all populations as expected for a marine mollusk (mean F_{is} value and standard deviation equals 0.424 ± 0.054). The same levels of heterozygote deficiency were found in both sets (mean F_{is} and standard deviation = 0.4187 ± 0.0491 and 0.4305 ± 0.0603 for the Bay of Biscay set and for the Core set respectively) and a two-sided Student's t test showed no difference between the sets means ($t = 0.457$, $df = 16$, $p = 0.654$).

3.3 Population structure

The mitochondrial pairwise F_{st} estimate presented in Table 3 showed no significant difference between populations from the Bay of Biscay set (populations 1 to 9) but these populations were clearly differentiated from (i) the Ireland and Murman populations and (ii) the other populations from the Core set. This clustering is statistically significant according to the AMOVA results: the overall high genetic structure ($\Phi_{st} = 0,613, p = 0$) was mostly due to a large genetic differentiation between these 3 groups ($\Phi_{ct} = 0,587, p = 0$) while the genetic differentiation among populations within these groups appeared to be low ($\Phi_{sc} = 0,0613, p = 0$).

The multilocus nuclear structure was found to be coherent but slightly different from the structure observed with mitochondrial data. The overall genetic structure ($F_{it} = 0,4425$) was essentially due to a strong heterozygote deficiency at the population level ($F_{is} = 0,4189$) as populations were not strongly differentiated ($F_{st} = 0,0407$). However, a population tree constructed with PHYLIP software (Fig 5) revealed that the Bay of Biscay set was mostly (57,7 % of the 1000 simulated trees) isolated from other populations. Within this group, a fine scale structure was revealed: a subgroup called Marennes-Oléron (MO, sampling sites 1 to 6) was found to be distinct from a second subgroup called Loire (sampling sites 7 to 9). These two subgroups were observed in respectively 73,6 % and 94 % of the 1000 simulated trees. This fine scale structure was not detected by mitochondrial data. To overcome the manual (and somewhat arbitrary) grouping of classical F-statistics computations, we used a probabilistic approach implemented in the software Geneland (Guillot et al. 2005). The 30 replicated simulations all produced the same results (see Fig. 6): a strong barrier to gene flow was revealed within the Bay of Biscay between populations 6 and 7. Thus, 2 populations at Hardy-Weinberg equilibrium were identified: the Marennes-Oléron group (i.e. sampling sites 1 to 6) were clearly isolated from populations located further North (i.e. sampling sites 7 to 18) all clustered in a second group.

4. DISCUSSION

The biogeographical distribution of natural populations of *Macoma balthica* has been studied for several years using various types of

markers. However no study has yet considered the conservation implications for this key species of the rapid Northward shift of its natural range. Our work has therefore focused on the potentially endangered populations located near the Southern range boundary of the species.

4.1 Genetic diversity within peripheral populations

In the classical abundant-center model, species are expected to achieve lower abundance at the edge of the range than at the geographical centre, with populations becoming progressively smaller and more fragmented (Lawton 1993, Vucetich & Waite 2003). Genetic organization is thus affected, in terms of diversity and population structure. As an example, several studies showed that peripheral populations exhibit low genetic diversity, greater genetic differentiation (structured metapopulations), small effective population sizes and geographic isolation as compared to more central populations (El Mousadik and Petit 1996, Durka 1999, Bouzat and Johnson 2004).

Near the current southern range limit of *Macoma balthica*, populations seem regarding mitochondrial and nuclear markers to differ from this biogeographic model. Populations from the Bay of Biscay display a significantly higher genetic diversity than in the Core-population set ($p = 0,0122$ for heterozygosity and $p = 1,051.10^{-5}$ for allelic richness). Furthermore, despite a high global heterozygote deficiency ($F_{is} = 0,4305$) observed in *Macoma balthica* populations and consistent with strong F_{is} values observed for many marine mollusks, the Bay of Biscay populations do not seem to be affected by inbreeding as confirmed by the high levels of genetic diversity in these populations. This peculiar pattern might be linked to historical processes. Indeed, *Macoma balthica* colonized the Atlantic Ocean down to the Iberic coasts after the opening of the Bering Strait during Pleistocene. During the Last Glacial Maximum (LGM, 18000 years ago) all the northern European coasts except the Bay of Biscay were covered by ice, the Spanish region being one of the four major temperate refuge zones (Gomez and Lunt, 2007). Populations in such areas are expected to harbor a large fraction of the intraspecific diversity of temperate biotas (Petit et al 2003) and to become the source of recolonization of northward sites after

LGM (see for instance Austin et al. (2004) or Howes et al. (2006)). Hewitt (1999) clearly established that a feature of temperate species (*e.g.* alder, beech, grasshoppers and newts) in southern Europe was the presence of several distinct geographical genomes with a variety of alleles. This pattern is present in southern populations of *Macoma balthica* which exhibit 2 major haplotypes (H1 and H2) found exclusively within the Bay of Biscay. Likewise, the higher nuclear diversity observed in the Bay of Biscay populations is mainly due to the high number of private alleles found in these populations. Regarding the loci mac40, mac10 and mac84, respectively 44%, 47% and 69% of alleles were found only in the southernmost populations. These populations exhibit a total of 42 rare and specific alleles across 8 loci while Core-populations exhibit only 19.

These results are congruent with a rapid northward expansion after the LGM from a refuge zone located either in Spain or in the Bay of Biscay.

4.2 Structure in rear edge

Both mitochondrial and nuclear markers presented a significant genetic structure within our dataset. However, the structure inferred based on both types of markers differed slightly: the mtDNA structure revealed a clear geographic barrier between the Bay of Biscay populations (*i.e.* populations 1 to 9) and populations located up North (see fig 1), while nuclear results highlighted an additional barrier to gene flow located within the Bay of Biscay (*i.e.* between populations 6 and 7, fig 5). For both types of analyses, all populations belonging to a given group shared similar characteristics. For instance, mitochondrial patterns for all populations from the Bay of Biscay set are similar: two of the four most frequent haplotypes (namely H1 and H2) are present in each of the Bay of Biscay populations but were never found in populations from the Core set. On the other hand, the haplotypes H3 and H4 were rarely found in the Bay of Biscay populations but were the most frequent in every population from the Core set (except for the strictly monomorphic population of Murman, fig. 1). This mitochondrial pattern is coherent with a post-glacial recolonization of the northern european coasts along two distinct routes: (i) a South to North route from an Iberian refugium

and (ii) a North to South route from the Pacific Ocean.

For nuclear results, the neighbor joining tree (fig. 5) as well as bayesian simulations (fig. 6) showed also a strong and well supported genetic break within the Bay of Biscay (NJ trees bootstrap values equals 73% and 94% for the Marennes-Oléron group and the Loire group respectively). Previous allozymic studies (Hummel 1997b) probably failed to reveal these results because of the limited resolution of such markers.

The genetic structures observed raise two questions. First, how to explain such a strong structure for a long-distance disperser? And second, how to explain the discrepancy between the positions of the barrier to gene flow inferred by both types of markers?

4.2.1 Long distance dispersal and geographic barrier

Using both types of markers we identified a strong and significant genetic structure located somewhere around Brittany (*i.e.* the westernmost French coasts marking the southern end of the English channel). Such a structure observed between geographically close populations of *Macoma balthica* is paradoxical considering the life history traits of this species. Indeed, its first stage of development is a 2 to 5 weeks period of larval pelagic life during which larvae can be carried over long distances by ocean currents (Beukema and De Vlas 1989) before their settlement as benthic adults (Caddy 1967; Drent 2002). In the absence of a physical or genetic barrier, this pelagic larval stage is expected to maintain the gene flow between southern and northern Brittany populations. Two main kinds of geographic barriers could explain the global genetic structure observed (*i.e.* southernmost *M. balthica* populations isolated from the core area of the natural range).

First, Brittany is known to be a major marine biogeographic barrier. As underlined by Leppäkoski et al. (2002), the effects of the Gulf Stream along the European coastline result in a strong biogeographic barrier around Brittany for a variety of marine phyla, limiting the species exchange between the northern and southern areas.

Second, local hydrodynamics during the larval stage might

explain why southern populations remain isolated from the core area of *M. balthica*'s natural range. A larval dispersion model tested on the marine polychete *Pectinaria koreni* confirmed that a strong larval retention in a bay led to self-recruitment (Ellien et al. 2000). Such a post-zygotic isolation mechanism might occur within the semi-closed Marennes-Oléron Bay, and on a wider scale, within the Pertuis Charentais area (*i.e.* populations 1 to 6) where two 30 km-long islands only a few kilometers apart each other and from the main land delineate a 1300 km² semi-closed area.

4.2.2 Discrepancy between markers and genetic barriers

Apart from the geographic barrier presented above, genetic barriers to gene flow could also be an explanation to the strong genetic structure observed and to the contrasting patterns observed using both types of markers (Lemaire et al. 2005).

The position of the barrier to gene flow inferred with both types of markers differ for a few hundred kilometers. As underlined earlier, Brittany is a major marine biogeographic barrier. From that perspective, the 3 populations of the Loire group (*i.e.* populations 7 - 9) that were assigned to different regions depending on the molecular marker used (*i.e.* Bay of Biscay set based on mitochondrial analyzes or Core set based on nuclear markers) can be considered to be located in a zone where genes do not move freely. The post glacial recolonization from two isolated *M. balthica* stocks could have been slowed near this biogeographic barrier, explaining why mitochondrial and nuclear patterns appear so different on both sides of Brittany. Moreover, migration can be a cause of cyto-nuclear disequilibria, especially when disequilibria exist in the migrant pool (Asmusen and Arnold 1991; Asmussen and Orive 2000), which is likely in areas where migrants are derived from two (or more) genetically distinct pools (Latta et al. 2001). Then, since *M. balthica* populations from the core area and populations located near the southern range limit are divergent populations, it is not surprising to observe a discrepancy between mitochondrial and nuclear signatures in a zone where the mixture of both types of populations is expected to occur (transition zone as described in Lemaire et al. 2005).

Besides, sex biased gene flow is known to induce prezygotic isolation (Crochet et al. 2003). However, mitochondrial and nuclear gene flows (*i.e.* number of migrants per generation, N_m) between the two regions (Marennes-Oléron and Loire), were found to be similar using the private allele method (Slatkin, 1985) and the F_{st} method (Wright 1951).

4.3 Conservation implications

Even if both types of markers did not allow us to precisely identify the position of a barrier to gene flow, it remains clear that such a barrier is present and isolates in a certain extent the southern populations of the Bay of Biscay from the central area of the natural range of the species. Another strong result of this study is the characterization of higher levels of mitochondrial and nuclear diversities in the southernmost populations of *Macoma balthica* as compared to populations from the core area. These results might have important implications for the conservation of this bivalve.

Indeed, *M. balthica's* natural range experienced a strong northward shift during the last 50 years. Such shifts are observed in many other species (Koutsikopoulos et al. 1998 ; Parmesan and Yohe 2003). As an example, in the last 25 years in the North Sea, an entire community of 28 demersal fish species was found to occur deeper and deeper, up to 3 to 6 m per decade, where temperatures are cooler (Dulvy et al. 2008). In European intertidal zones, the southern range limit of the cirriped *Semibalanus balanoides* shifted 300 km towards North between 1872 and 2006, with a rate of displacement of 50 km per decade over the last thirty years (Wetthey and Woodin 2008). Other studies on rocky shores have confirmed the same trends: latitudinal or abundance shifts from south to north (Barry et al. 1995).

Many anthropogenic pressures might accelerate these boundary shifts. Pollution and habitat fragmentation induced by littoral industrialization are two possible factors. Nevertheless many studies have been conducted on the performance of *Macoma balthica* in highly polluted sites (Szefer and Skwarzec 1988; Pempkowiak and Szefer 1992; Pempkowiak et al. 1999) and it seemed to remain a major component of the macrozoobenthic communities in heavily polluted environments while other

species were declining or disappeared (Szaniawska et al. 1996). Likewise, gene flow disturbances induced by potential habitat fragmentation along the French littoral might be partially counteracted by the long pelagic larval stage of *Macoma balthica*.

On the opposite, an increase in temperature is known to greatly affect the physiological capacities of individuals, especially for organisms living at the low-latitude margin of species natural ranges (Hugues 2000). As an illustration, a recent study on the metabolic and respiratory rates of *M. balthica* (Jansen et al. 2007) revealed a zero-plasticity to temperature for populations close to the southern range limit. The author translocated a Gironde population to the Bidasoa Bay (where *M. balthica* was present 50 years ago) and observed a rapid mortality of animals. These results reinforced those published by Hummel et al. (1998) showing limited ecophysiological performances for *M. balthica* populations at the edge of its natural range.

The negative consequences of the potential disappearance of *Macoma balthica* along the Atlantic French coasts could be far worse than the loss of any other single species. As the favorite prey for many shorebirds, fishes and macro-invertebrates, *Macoma balthica* is a key-species for intertidal communities. Its disappearance could thus have a dramatic and lasting impact on the whole ecosystem. As an illustration, the decreasing bivalve recruitment observed in the Wadden Sea lead to a massive emigration, starvation and mortality of bivalve-eating birds (Camphuysen et al. 2002). The Pertuis Charentais area and southern Brittany (where populations 1 to 9 have been sampled) are known to be a major wintering area for the Black Tailed Godwit *Limosa limosa* (Gill et al. 2007) a species feeding almost exclusively on *M. balthica* stocks on intertidal mudflats (Bocher P, Robin F, personal communication). Indeed, Zwarts & Blomert (1992) highlighted the energetic importance of *M. balthica* intake among a 5 species-mix of bivalves in the diet of another shorebird *Calidris canutus*.

5. CONCLUSION

We showed that *M. balthica* populations located near the southern range limit of the species exhibited a high genetic diversity probably

because of their location in, or near, a major glacial refugium. An elevated level of genetic diversity especially highlighted with neutral markers, is probably not enough in itself to guarantee the maintenance of populations as suggested by the disappearance of the southern French and Spanish populations during the last 50 years. The strong barrier to gene flow between the populations from the Bay of Biscay and the central area of the species range probably results in the isolation of these populations making them even more vulnerable to environmental changes. Given that seasonal variations of temperature is known to greatly impact bivalves recruitment (Beukema et al. 2009) and considering the low tolerance of *M. balthica* individuals to an elevated temperature, a rapid rise of temperature (as expected by the GIEC projections, GIEC 2007) could rapidly put the southernmost *M. balthica* populations at great risk. Thus only the current experimental genomic approach on genetically divergent populations (e.g. MO and Northward Brittany populations), would provide some answers to tolerance to elevated temperature of *Macoma balthica* in a context of global warming.

Aknowledgements

This research was supported by the European Committee (Research Directorate General, Environmental Program Marine Ecosystems) through the BIOCOMBE-project (contrat EVK3-2001-00146). We wish to thank Sandra Kube, Jeroen Jansen and Frederic Robin for their fieldwork and Alice Saunier for the laboratory work.

Alleaume-Benharira M, Pen I.R, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *J Evol Biol* 19 : 203–215.

Asmussen M.A, Arnold J (1991) The Effects of Admixture and Population Subdivision on Cytonuclear Disequilibrium. *Theor Popul Biol* 39 : 273-300.

Asmussen M.A, Orive M.E (2000) The Effect of Pollen and Seed Migration on Nuclear-dicytoplasmic Systems. I. Nonrandom Associations and Equilibrium Structure with Both Maternal and Paternal Cytoplasmic Inheritance. *Genetics* 155 : 813-831.

Austin J.D, Loughheed S.C. Boag P.T (2004) Discordant temporal and geographic patterns in maternal lineages of eastern north American frogs, *Rana catesbeiana* (Ranidae) and *Pseudacris crucifer* (Hylidae). *Mol Phylo Evol* 32 : 799–816.

Bachelet G (1980) Growth and recruitment of the tellinid bivalve *Macoma balthica* at the southern limit of its geographical distribution, the Gironde Estuary (SW France). *Mar Biol* 59 : 105–117.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16 : 37-48.

Barry J.P. Baxter C.H, Sagarin R.H, Gilman S.E (1995) Climate-related longterm faunal changes in a California rocky intertidal community. *Sciences* 267 : 672-675.

Belkhir K, Borsa P, Chikhi P, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous windows pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Universités de Montpellier II, Montpellier, France.

Becquet V, Lanneluc I, Simon-Bouhet B, Garcia P (2009) Microsatellite markers for the Baltic clam, *Macoma balthica* (Linné, 1758), a key species concerned by changing southern limit, in exploited littoral

ecosystems. *Cons Genet Ressources* 1 : 265-269.

Bell G, Collins S (2008) Adaptation, extinction and global change. *Evol Appl* 1 : 3-16.

Beukema J.J, De Vlas, J (1989) Tidal-current transport of thread-drifting postlarval juveniles of the bivalve *Macoma balthica* from the Wadden Sea to the North Sea. *Mar Ecol Prog Ser* 52 : 193-200.

Beukema J.J, Dekker R, Jansen J.M, (2009) Some like it cold: populations of the tellinid bivalve *Macoma balthica* (L.) suffer in various ways from a warming climate. *Mar Ecol Prog Ser* 384: 135–145.

Bolnick D.I, Caldera E.J, Matthews B (2008) Evidence for asymmetric migration load in a pair of ecologically divergent stickleback populations. *Biol J Linn Soc* 94 : 273–287.

Bouzat J.L, Johnson K (2004) Genetic structure among closely spaced leks in a peripheral population of lesser prairie-chickens. *Mol Ecol* 13 : 499-505.

Brown J.H, Mehlman D.W, Stevens G.C (1995) Spatial variation in abundance. *Ecology* 76 : 2028–2043.

Caddy J.F (1967) Development of mantle organs, feeding and locomotion in postlarval *Macoma balthica* (L.) (Lamellibranchiata). *Can J Zool* 47 : 609-617.

Camphuysen C.J, Berrevoets C.M, Cremers H.J.W.M, Dekinga A, Dekker R, Ens B.J, Van der Have T.M, Kats R.K.H, Kuiken T, Leopold M.F, Van der Meer J, Piersma T (2002) Mass mortality of common eiders (*Somateria mollissima*) in the Dutch Wadden Sea, winter 1999/2000 : Starvation in a commercially exploited wetland of international importance. *Biol Cons* 106 : 303–317.

Cassens I, Van Waerebeek K, Best P.B, Crespo E.A, Reyes J, Milinkovitch M.C (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Mol Ecol* 12 : 1781-1792.

Cavalli-Sforza L.L, Edwards A.W.F (1967) Phylogenetic

analysis. Models and estimation procedures. *Am J Hum Genet* 19 : 233–257.

Courchamp P, Clutton-Brock T, Grenfell B (1999) Inverse density dependence and the Allee effect. *Trends Ecol Evol* 14 : 405–410.

Crandall K.A, Bininda-Emonds O.R.P, Mace G.M, Wayne R.K (2000) Considering evolutionary processes in conservation biology. *Trends in Ecol Evol* 15 : 290-295.

Crochet P.A, Chen J.Z, Pons J.M, Lebreton J.D, Hebert P.D, Bonhomme F (2003) Genetic Differentiation At Nuclear And Mitochondrial Loci Among Large White-Headed Gulls : sex-Biased Interspecific Gene Flow ? *Evolution* 57 : 2865–2878.

Drent J (2002) Temperature responses in larvae of *Macoma balthica* from a northerly and southerly population of the European distribution range. *J Exp Mar Biol Ecology* 275 : 117-129.

Dulvy N.K, Rogers S.I, Jennings S, Stelzenmüller V, Dye S.R, Skoldal H.R (2008) Climate change and deepening of the North Sea fish assemblage: a biotic indicator of warming seas. *J Appl Ecol* 45 : 1029-1039.

Durka W (1999) Genetic diversity in peripheral and subcentral populations of *Corrigiola litoralis* L. (Illecebracea). *Heredity* 83 : 476-484.

Ecker, C.G, Samis K.E, Loughheed S.C (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol Ecol* 17 : 1170-1188.

Edjung G, Bonsdorff E (1992) Predation on the bivalve *Macoma balthica* by the isopod *Saduria entomon*: laboratory and field experiments. *Mar Ecol Prog Ser* 88 : 207-214.

Ellien C, Thiebaut E, Barnay A.S, Dauvin J.C, Gentil F, Salomon J.C (2000) The influence of variability in larval dispersal on the dynamics of a marine metapopulation in the eastern Channel. *Oceanologica Acta*. 23 : 423-442.

El Mousadik A, Petit R.J (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theor Appl Genet* 92 :

832-839.

Excoffier L, Smouse P.E, Quattro J.M (1992) Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* 131 : 479-491.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0) : An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1 : 47–50.

Felsenstein J (1985) Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* 39 : 783-791.

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6.

GIEC (2007). Bilan des changements climatiques: Contribution des groupes de travail I, II et III au quatrième rapport d'évaluation du Groupe d'Experts Intergouvernemental sur l'évolution du climat (Equipe de rédaction principale : Pachauri R.K., Resinger A.) GIEC. Genève, Suisse. 103pp.

Gill J.A, Langston R.H.W, Alves J.A, Atkinson P.W, Bocher P, Cidraes Vieira N, Crockford N.J, Gélinaud G, Groen N, Gunnarsson T.G, Hayhow B, Hooijmeijer J, Kentie R, Kleijn D, Lourenço P.M, Masero J.A, Meunier F, Potts P.M, Roodbergen M, Schekkerman H, Schröder J, Wymenga E, Piersma T (2007) Contrasting trends in two Black-tailed Godwit populations: a review of causes and recommendations. *Wader Study Group Bull* 114: 43–50.

Goikoetxea N, Borja A, Fontán A, González M, Valencia V (2009) Trends and anomalies in sea-surface temperature, observed over the last 60 years, within the southeastern Bay of Biscay. *Cont Shelf Res* 29 : 1060–1069.

Gómez A, Lunt A.H, (2007) Refugia within refugia : patterns of phylogeographic concordance in the Iberian Peninsula, in : Wein S., Ferrand N., (eds), *Phylogeography of Southern European Refugia*, 155-188.

Gordon D.R (1998) Effect of invasive, non-indigenous plant

species on ecosystem processes: lessons from Florida. *Ecol Appl* 8 : 975–989.

Goudet J (2002) FSTAT version 2.9.3.2. Available from Jerome.goudet@ie.zea.unil.ch, via email. Institute of Ecology, UNIL, CH-1015, Lausanne, Switzerland.

Guillot G, Mortier F, Estoup F (2005) Geneland : a program for landscape genetics. *Mol Ecol Notes* 5 : 712-715.

Hall T.A (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nuc Ac Symp Ser* 41 : 95-98.

Hewitt G.M (1999) Post-glacial re-colonization of European biota. *Biol J Linn Soc* 68 : 87-112.

Howes B.J, Lindsay B, Loughheed S.C (2006) Phylogeography of a temperate lizard. *Mol Phylo Evol* 40 :183-194.

Hugues L (2000) Biological consequences of global warming: is the signal already. *Trends Ecol Evol* 15 : 56-61.

Hummel H (1985) Food intake of *Macoma balthica* (Mollusca) in relation to seasonal changes in its potential food on a tidal flat in the Dutch Wadden Sea *Neth J Sea Res* 19 : 52–76.

Hummel H, Bogaards R.H, Amiard-Triquet C, Bachelet G, Desprez M, Marchand J, Rybarczyk H, Sylvand B, de Wit Y, de Wolf L (1995) Uniform variation in genetic traits of a marine bivalve related to starvation, pollution and geographic clines. *J exp mar Biol Ecol* 191 : 133-150.

Hummel H, Amiard-Triquet C, Bachelet G, Desprez M, Marchand J, Sylvand B, Amiard J.C, Rybarczyk H, Bogaards R.H, de Wolf L (1997a) Comparison of ecophysiological, biochemical and genetic traits in the estuarine bivalve *Macoma balthica* from areas between the Netherlands and its southern limits (Gironde): geographic clines parallel effects of starvation and copper exposure. L.E. Hawkins & S. Hutchinson (eds.). *Proceeding 30th European Marine Biological Symposium*, University of

Southampton, pp 15-20.

Hummel H, Bogaards R, Bek T, Polishchuk L, Amiard-Triquet C, Bachelet G, Desprez M, Strelkov P, Sukhotin A, Naumov A, Dahle S, Denisenko S, Gantsevich M, Sokolov K, de Wolf L (1997b) Sensitivity to stress in the bivalve *Macoma balthica* from the most northern (Arctic) to the most southern (French) populations: low sensitivity in Arctic populations because of genetic adaptations? *Hydrobiologia* 355 : 127–138.

Hummel H, Bogaards R, Bek T, Polishchuk L, Sokolov K, Amiard-Triquet C, Bachelet G, Desprez M, Naumov A, Strelkov P (1998) Growth in the bivalve *Macoma balthica* from its northern to its southern distribution limit: a discontinuity in North Europe because of genetic adaptations in Arctic populations? *Comp Biochem Physiol A* 120 : 133-141.

Hummel H, Bogaards R.H, Bachelet G, Caron F, Sola J.C, Amiard-Triquet C (2000) The respiratory performance and survival of the bivalve *Macoma balthica* (L.) at the southern limit of its distribution area: a translocation experiment. *J exp Mar Biol Ecol* 251 : 85-102

Jansen J.M, Pronker A.E, Kube S, Sokolowski A, Carlos Sola J, Marquiegui M.A, Schiedek D, Sjoerd Wendelaar Bonga, Wolowicz M, Hummel, H (2007) Geographic and seasonal patterns and limits on the adaptive response to temperature of European *Mytilus* spp. and *Macoma balthica* populations. *Oecologia* 154 : 23-34.

Koutsikopoulos C, Beillois P, Leroy C, Taillefer F (1998) Temporal trends and spatial structures of the sea surface temperature in the Bay of Biscay. *Oceanologica Acta* 21 : 335–344.

Lande R, Engen S, Sæther B.E (2003) *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, New York.

Latta R.G, Linhart Y.B, Mitton J.B (2001) Cytonuclear Disequilibrium and Genetic Drift in a Natural Population of *Ponderosa* Pine. *Genetics* 158 : 843-850.

Lawton J.H (1993) Range, population abundance and conservation. *Trends Ecol. Evol.* 8 : 409-413.

Lemaire C, Versini J.J, Bonhomme F (2005) Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. *J Evol Biol* 18 : 70-80.

Leppäkoski E, Gollasch S, Olenin S (2002) (Eds.), *Invasive Aquatic Species of Europe Distribution, Impacts and Management*. Hardcover, p600.

Librado P, Rozas J (2009) DNASP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25 : 1451-1452.

Luttikhuisen P.C, Drent J, Baker A.J (2003) Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. *Mol Ecol* 12 : 2215–2229.

Margush T, McMorris F.R (1981) Consensus-trees. *Bulletin of Mathematical Biology* 43 : 234-244.

Mattila J, Bonsdorff E (1998) Predation by juvenile flounder (*Platichthys flesus* L.): a test of prey vulnerability, predator preference, switching behaviour and functional response. *J Exp Mar Biol Ecol* 227 : 221-236.

Moritz C (1994) Defining ‘Evolutionarily Significant Units’ for conservation. *Trends Ecol Evol* 9 : 373-375.

Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New- York.

Nikula R, Strelkov P, Väinölä R (2008) A broad transition zone between an inner Baltic hybrid swarm and a pure North Sea subspecies of *Macoma balthica* (Mollusca, Bivalvia). *Mol Ecol* 17 : 1505–1522.

Otero J.H, Milan F.J (1970) Distribucion de los moluscos: gasteropodos y pelecipodos, marinos, de las costas de Galicia. *Cuad Biol* 1 : 79–93.

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421 : 37-42.

Pempkowiak J, Szefer P (1992) Origin, sources and concentrations of selected heavy metals in the southern Baltic biota. Bulletin of the Sea, Rybackiego Institute. 125 : 29-32.

Pempkowiak J, Sikora A, Biernacka E (1999) Speciation of heavy metals in marine sediments vs their bioaccumulation by mussels. Chemosphere. 39 : 313-321.

Petit R.J, Aguinagalde I, De Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin G.G (2003) Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. Science 300 : 1563 –1565.

Philippart C, Van Aken H.M, Beukema J.J, Bos O.G, Cadee G.C, Dekker R (2003) Climate-related changes in recruitment of the bivalve *Macoma balthica* .Limnol Oceanogr 48 : 2171–2185.

Piersma T, Beukema J.J (1993) Tropic interactions between shorebirds and their invertebrate prey. Neth J Sea Res 31 : 299–512.

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Ronce O, Kirckpatrick M (2001) When sources become sinks: migrational meltdown in heterogeneous habitats. Evolution 55 : 1520-1531.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Ressources 8 : 103-106.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4 : 406-425.

Scarlatto O.A (1981) The Bivalve Molluscs of the Boreal Western Pacific. Nauka, Leningrad.

Schroter D, Cramer W, Leemans R, Prentice I.C, Araújo M.B, Arnell N.W, Bondeau A, Bugmann H, Carter T.R, Gracia C.A, De la Vega

Leinert A.C, Erhard M, Ewert F, Glendining M, House J.I. Kankaanpää S, Klein R.J.T, Lavorel S, Lindner M, Metzger M.J, Meyer J, Mitchell T.D, Reginster I, Rounsevell M, Sabaté S, Sitch S, Smith B, Smith J, Smith P, Sykes M.T, Thonicke K, Thuiller W, Tuck G, Zaehle S, Zierl B (2005) Ecosystem Service Supply and Vulnerability to Global Change in Europe. *Science* 310: 1333-1337.

Slatkin M (1985) Rare Alleles as Indicators of Gene Flow. *Evolution* 39: 53-65.

Strelkov P, Nikula R, Väinölä R (2007) *Macoma balthica* in the White and Barents Seas: properties of a widespread marine hybrid swarm (Mollusca: Bivalvia). *Mol Ecol* 16 : 4110–4127.

Szaniawska A, Janas U, Normant M (1996) Changes in Macrozoobenthos Communities Induced by Anthropogenic Eutrophication of the Gulf of Gdansk, in : Gray J.S., et al., (eds), *Biogeochemical Cycling and Sediment Ecology* 147-152 .

Szefer P, Skwarzec B (1988) Distribution and Possible Sources of Some Elements in the Sediment Cores of the Southern Baltic. *Marine Chemistry* 23 : 109-129.

Thomas C.D, Kunin W.E (1999) The spatial structure of populations. *J Anim Ecol* 68 : 647–657.

Thompson J.D, Higgins D.G, Gibson T.J (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nuc Ac Res* 22 : 4673-4680.

Väinölä R (2003) Repeated trans-Arctic invasions in littoral bivalves: molecular zoogeography of the *Macoma balthica* complex. *Mar Biol* 143 : 935–946.

Vitousek P, Mooney H, Lubchenco J, Mellilo J, (1997) Human domination of Earth's ecosystems. *Sciences* 277 : 494-499.

Vucetich J.A, Waite T.A (2003) Spatial patterns of demography and genetic processes across the species' range: null hypotheses for

landscape conservation genetics. *Cons Gen* 4 : 639–645.

Walker D, Kendrick G (1998) Threats to macroalgal diversity: marine habitat destruction and fragmentation, pollution and introduced species. *Botanica Marina*. 41 (1), 105-112.

Walther G.R, Post E, Convey P, Menzel A, Parmesan C, Beebee T.J.C, Fromentin JM, Hoegh-Guldberg O, Bairlein F (2002) Ecological responses to recent climate change. *Nature* 416 : 389-395.

Wethey D.S, Woodin S.A (2008) Ecological hindcasting of biogeographic responses to climate change in the European intertidal zone. *Hydrobiologia* 606 : 139-151.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* 15 : 323-354.

Zwarts L, Blomert A.M (1992) Why knot *Calidris canutus* take medium-sized *Macoma balthica* when six prey species are available. *Mar Ecol Prog Ser* 83 : 113-128.

Fig. 1: Haplotypes distribution of *Macoma balthica* populations in Europe and in the southern range (Bay of Biscay).

Fig. 2: Minimum spanning network among haplotypes. H1 and H2 are typical haplotypes of the Bay of Biscay and H3 and H4 are characteristic of the Core-populations. Each point between haplotype consists in one mutation.

Fig. 3: Distribution of major haplotype frequencies along the European coasts. Except for Ireland, sampling sites are ordered by increasing latitude along the x-axis.

Fig. 4a: Comparative allelic richness in Bay of Biscay and Core-populations.

Fig. 4b: Comparatives heterozygoties in Bay of Biscay and Core-populations.

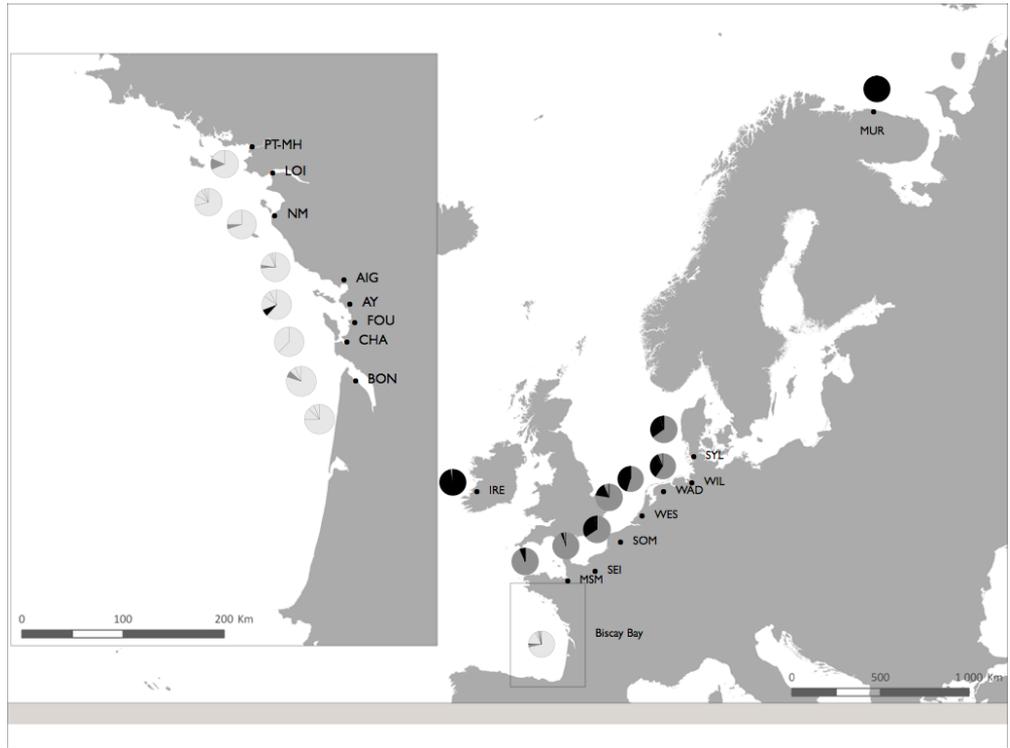
Fig. 5: Neighbour-joining consensus tree. Genetic distances were calculated from microsatellites allelic frequencies (1000 bootstrap iterations).

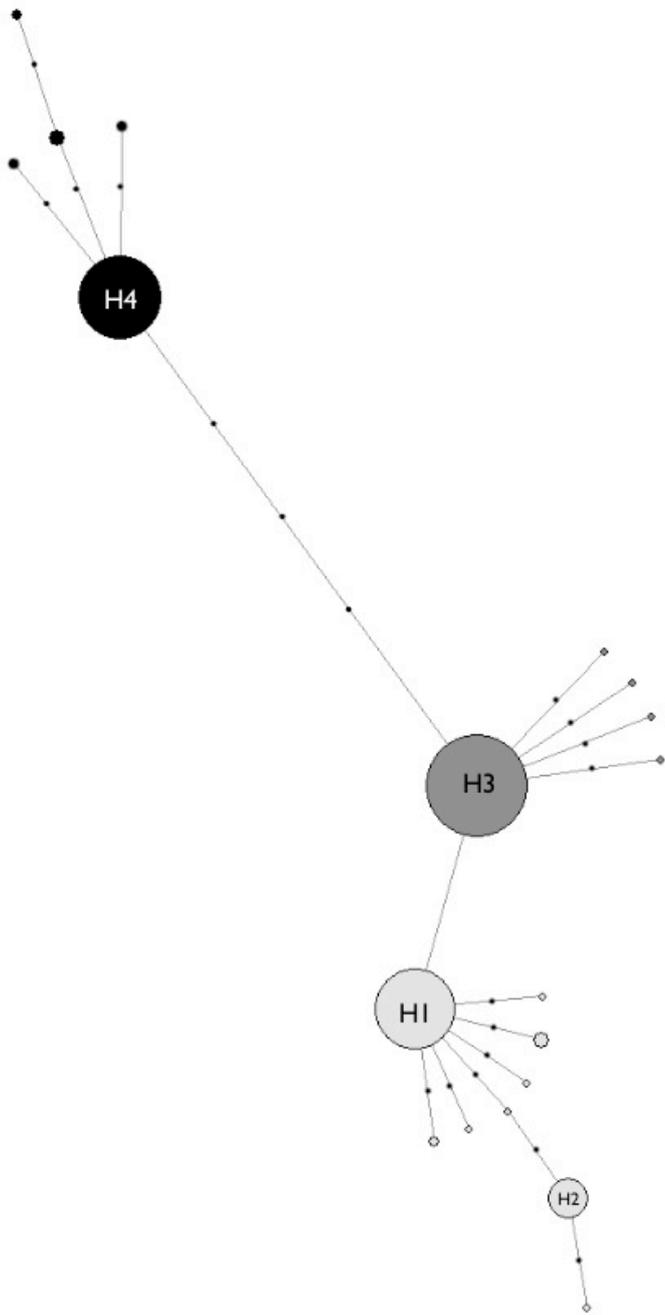
Fig. 6: Map of posterior probability to belong to the first population at Hardy-Weinberg equilibrium as inferred by Geneland. Light zones correspond to the lowest probabilities and dark zones to the highest. Isolines of probability are pictured as white solid lines, coastlines as black solid lines and sampling sites as grey points.

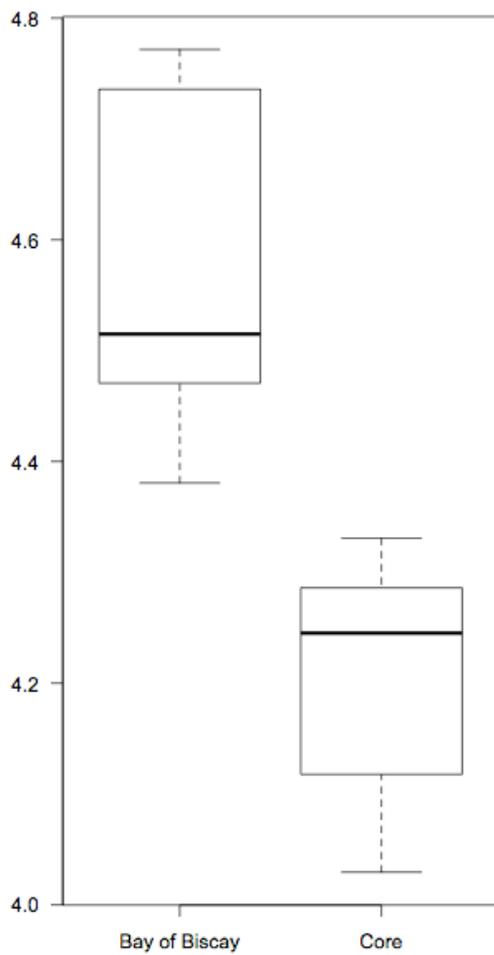
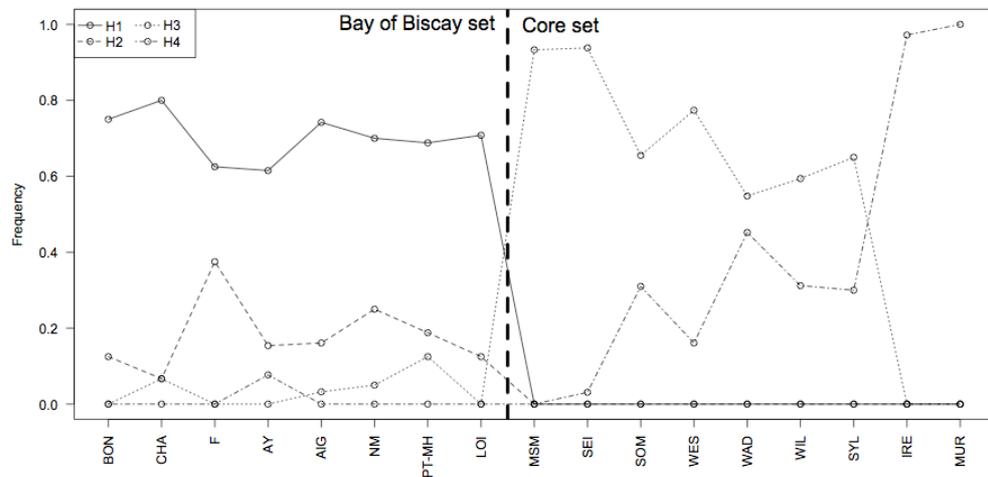
Table. 1: Geographic location of the eighteen European populations of *Macoma balthica*.

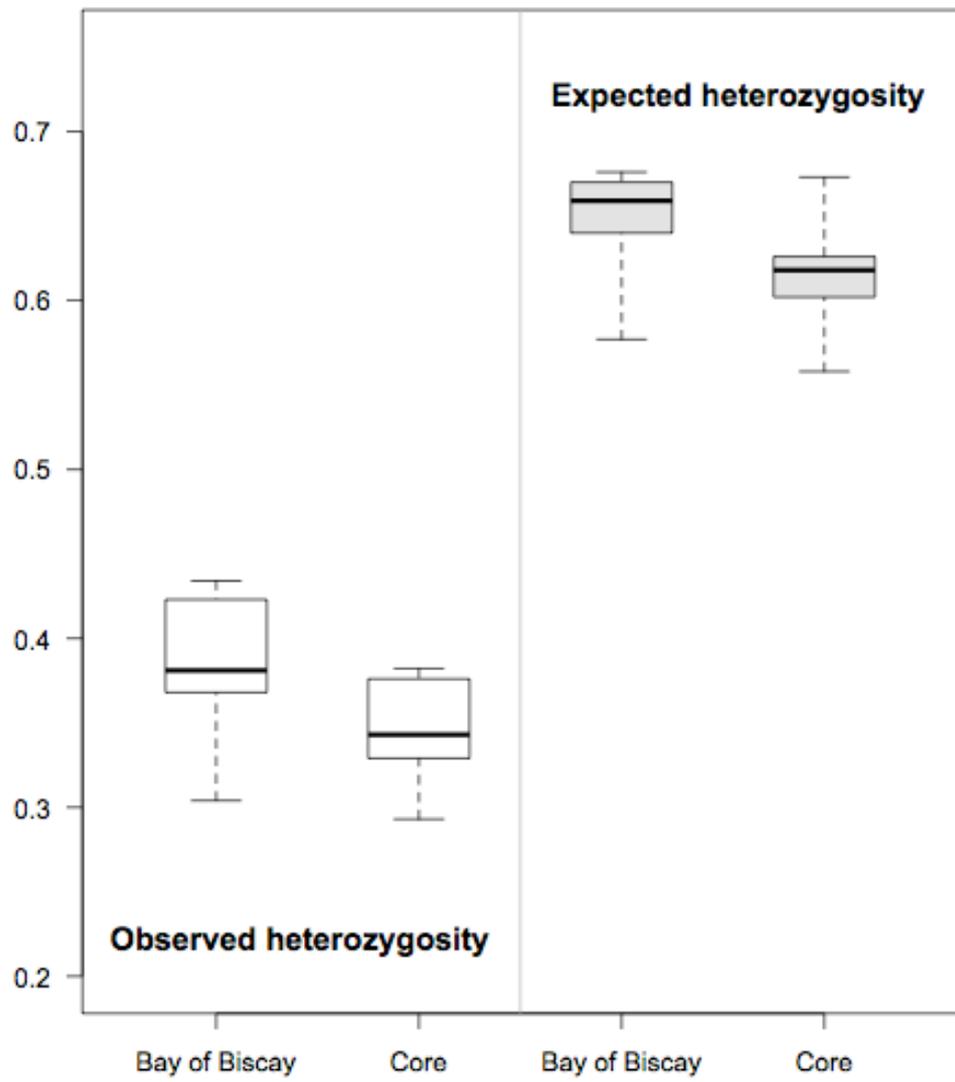
Table. 2: Mitochondrial and nuclear diversity for each population and each set of population (Bay and Biscay and Core-populations).

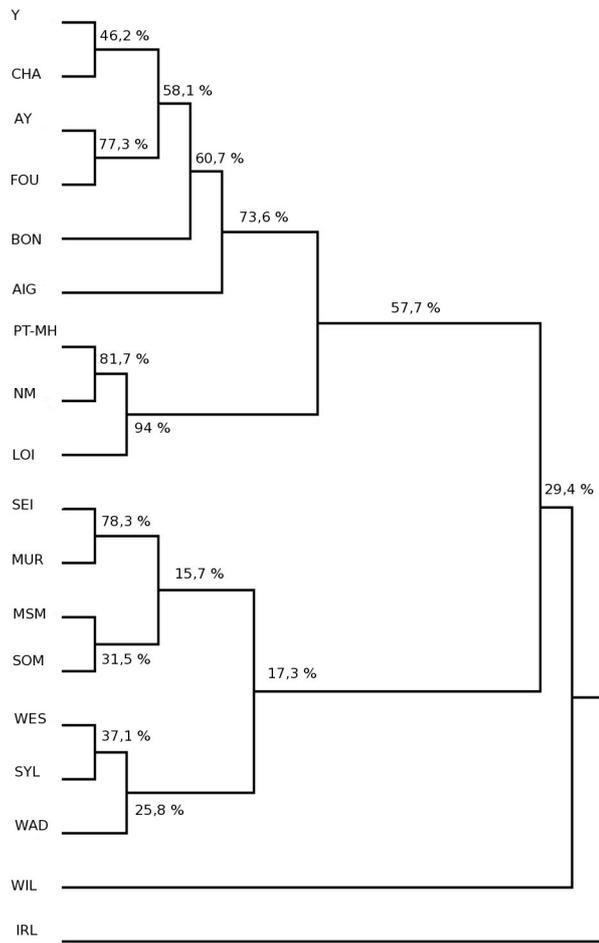
Table. 3: Pairwise F_{st} values..

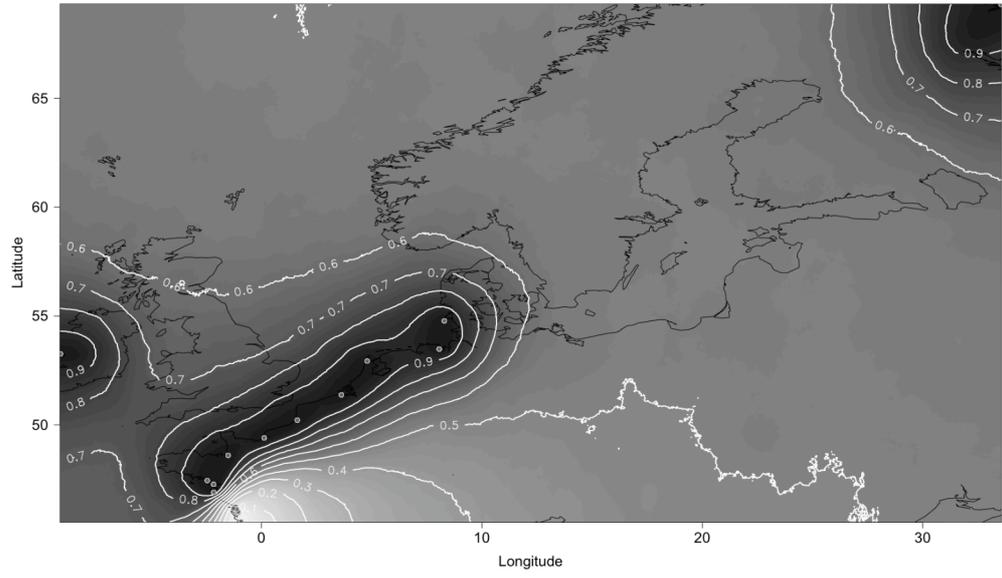












Sampling Site	Code	GPS Coordinate N	GPS Coordinate W	N Mitochondrial	N Microsatellites
(1) Bonne Anse	BON	45,5523	-0,9294	16	13
(2) Charente	CHA	45,9583	-1,0501	15	25
(3) Fouras	F	45,984	-1,0925	8	36
(4) Yves	Y	46,0116	-1,0561	0	39
(5) Aytre	AY	46,126	-1,1306	13	50
(6) Aiguillon	AIG	46,1617	-1,1298	31	16
(7) Noirmoutiers	NM	46,9035	-2,1671	20	40
(8) Loire	LOI	47,2675	-2,1715	24	47
(9) Pont-Mahe	PT-MH	47,4442	-2,4608	16	23
(10) Mont Saint Michel	MSM	48,438	-1,5153	30	30
(11) Seine	SEI	49,4029	0,1205	32	29
(12) Somme	SOM	50,2146	1,6227	29	28
(13) Paulina Westerschelde	WES	51,3792	3,6272	31	27
(14) Balgsand	WAD	52,9301	4,7953	31	28
(15) Irlande	IRE	53,257	-9,1201	36	31
(16) Wilhemshaven	WIL	53,4802	8,0641	32	28
(17) Sylt	SYL	54,7789	8,2954	40	40
(18) Murman	MUR	69,3073	33,556	20	15
Total				424	573

Mitochondrial						Microsats						
Population	N	S	H	Hd	π	Population	N	Nall	Rall	Ho	Hs	Fis
(1) Bonne Anse	16	4	4	0,442	0,00269	(1) Bonne Anse	13	6,625	4,380625	0,304	0,577	0,5027
(2) Charentes	15	3	5	0,476	0,00188	(2) Charentes	25	8,526	4,514875	0,368	0,63	0,4325
(3) Yves	–	–	–	–	–	(3) Yves	39	11	4,73575	0,375	0,661	0,4437
(4) Fouras	8	2	2	0,536	0,05357	(4) Fouras	36	1,75	4,656	0,368	0,64	0,4342
(5) Aytre	13	7	5	0,628	0,07179	(5) Aytre	50	11,25	4,744	0,434	0,659	0,3339
(6) Aiguillon	31	4	5	0,434	0,00219	(6) Aiguillon	16	9,625	4,771625	0,408	0,67	0,4207
(7) Noirmoutiers	20	2	3	0,468	0,04184	(7) Noirmoutiers	37	9,125	4,470625	0,43	0,676	0,3764
(8) Loire	24	5	5	0,493	0,03913	(8) Loire	47	10,25	4,5025	0,381	0,674	0,4444
(9) PontMahé	16	2	3	0,508	0,03917	(9) PontMahé	23	7,375	4,444625	0,423	0,654	0,3794
BayBiscay group	143	11	11	0,464	0,0411	BayBiscay group	286		4,58	0,395	0,668	0,4305
Mean	17,875	3,625	4	0,498	0,03153	Mean	31,77777778	8,391777778	4,580069444	0,387888889	0,649	0,418655556
(S.D.)	7,08	1,768	1,195	0,062	0,02646	(S.D.)	13,16033603	2,931941924	0,147452258	0,041199043	0,031004032	0,0491
(10) MontSaintMichel	30	4	2	0,129	0,00154	(10) MontSaintMichel	30	8,375	4,245125	0,376	0,618	0,40871
(11) Seine	32	4	3	0,123	0,00127	(11) Seine	29	8,125	4,0295	0,333	0,599	0,45983
(12) Somme	29	4	3	0,49	0,00441	(12) Somme	28	8,375	4,286	0,329	0,602	0,46807
(13) Westerschelde	31	5	4	0,385	0,0029	(13) Westerschelde	27	8,125	4,330625	0,303	0,626	0,52984
(14) Waddensee	31	3	2	0,512	0,0046	(14) Waddensee	28	8,75	4,330375	0,357	0,639	0,45703
(15) Wilhemshaven	32	6	5	0,565	0,00491	(15) Wilhemshaven	28	8,125	4,21125	0,382	0,607	0,38639
(16) Sylt	40	4	3	0,497	0,00448	(16) Sylt	40	9,5	4,281375	0,343	0,622	0,46129
(17) Irlande	36	4	2	0,056	0,00067	(17) Irlande	31	7,625	4,11775	0,379	0,558	0,33544
(18) Murman	20	0	1	0	0	(18) Murman	15	7,75	4,05225	0,293	0,673	0,36806
Core-populations	281	11	10	0,523	0,07684	Core-populations	256		4,211	0,358	0,628	0,44345
Mean	31,222	3,777	2,777777778	0,306	0,00275	Mean	28,44444444	8,305555556	4,209361111	0,343888889	0,616	0,430517778
(S.D.)	5,403	1,641	1,201850425	0,225	0,00192	(S.D.)	6,38574802	0,559405066	0,115714061	0,032578538	0,03132092	0,060361106
Total	424	18	19	0,72	0,08984	Total	542	20,875		0,374	0,65	
Mean	24,941	3,706	3,353	0,396	0,0163	Mean	30,11	8,305	4,395	0,366	0,6325	0,42459
(S.D.)	9,148	1,649	1,32	0,192	0,023	(S.D.)	10,18	0,165	0,23	0,0425	0,0347	0,05373

Site	(1)	(2)	(3)	(5)	(6)	(7)	(9)	(8)	(10)	(11)	(12)	(13)	(14)	(16)	(17)	(15)	(18)
(1)	0.00000																
(2)	-0.04643	0.00000															
(3)	-0.01751	0.04082	0.00000														
(5)	-0.04713	-0.03454	-0.01860	0.00000													
(6)	-0.03726	-0.03688	0.00213	-0.03231	0.00000												
(7)	-0.02924	0.00591	-0.07508	-0.02946	-0.01480	0.00000											
(9)	-0.03590	-0.00940	-0.05986	-0.03964	-0.02623	-0.05377	0.00000										
(8)	-0.03545	-0.03609	0.05326	-0.03002	-0.01648	0.02345	0.01098	0.00000									
(10)	0.49164*	0.60124*	0.37380*	0.36808*	0.52032*	0.44447*	0.43141*	0.54737*	0.00000								
(11)	0.50683*	0.62230*	0.38592*	0.37857*	0.53774*	0.46000*	0.44743*	0.56403*	-0.02119	0.00000							
(12)	0.50943*	0.62698*	0.38690*	0.38169*	0.54101*	0.46224*	0.44961*	0.56734*	-0.02419	-0.03175	0.00000						
(13)	0.44193*	0.52758*	0.34152*	0.32544*	0.46259*	0.39782*	0.38489*	0.48971*	0.01225*	0.01427	0.02304	0.00000					
(14)	0.50894*	0.56882*	0.44648*	0.39461*	0.52523*	0.48536*	0.48005*	0.54034*	0.28847*	0.30179*	0.31818*	0.14634*	0.00000				
(16)	0.45560*	0.51721*	0.38265*	0.34400*	0.47114*	0.42503*	0.41825*	0.49072*	0.19349*	0.20404*	0.21761*	0.07179	-0.01916	0.00000			
(17)	0.46125*	0.52555*	0.38837*	0.34703*	0.47740*	0.43066*	0.42259*	0.49705*	0.17974*	0.19038*	0.20276*	0.05886	-0.00664	-0.02627	0.00000		
(15)	0.86212*	0.90905*	0.83602*	0.78202*	0.88013*	0.86403*	0.86836*	0.87779*	0.86940*	0.88430*	0.89094*	0.76685*	0.47212*	0.52894*	0.57068*	0.00000	
(18)	0.89091*	0.93810*	0.86607*	0.81333*	0.90933*	0.89407*	0.89892*	0.90609*	0.91018*	0.92520*	0.93103*	0.81250*	0.53333*	0.58372*	0.62564*	0.00000	0.00000