



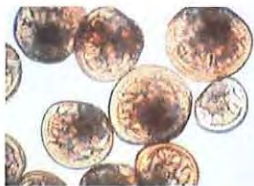
Institut Universitaire et Technologique
 Département Génie Biologique
 15, rue François Vaux de Foletier
 17026 LA ROCHELLE cedex 01



IFREMER
 Laboratoire Génétique
 et Pathologie
 Avenue du Mus de Loup
 Ronce les Bains
 17390 La Tremblade

Avril à Juin 2006

Génotypage moléculaire pour le développement de la carte génétique de la moule bleue : *Mytilus edulis*



69271
 E400-HEC.G



Par Céline HECQUET



IFREMER Bibliothèque de la Tremblade



OLR 03085



Institut Universitaire et Technologique
Département Génie Biologique
15, rue François Vaux de Foletier
17026 LA ROCHELLE cedex 01



IFREMER
Laboratoire Génétique
et Pathologie
Avenue du Mus de Loup
Ronce les Bains
17390 La Tremblade

Génotypage moléculaire pour le développement de la carte génétique de la moule bleue : *Mytilus edulis*

Pour la recherche de sites polymorphes, entre individus d'une famille préalablement sélectionnée.

Stage de fin d'étude validant un Diplôme Universitaire et Technologique en Génie Biologique, option Analyses Biologiques et Biochimiques.

Céline HECQUET

Année Universitaire 2005/2006

Stage effectué du 18 Avril au 24 Juin 2006.

Maîtres de stage : Madame Sylvie LAPEGUE et Mademoiselle Delphine LALLIAS
Tuteur Pédagogique : Monsieur Vincent RIDOUX

REMERCIEMENTS

Je tiens tout d'abord à remercier Monsieur Philippe GOULLETQUER, directeur du centre d'avoir accepté ma demande de stage, en me permettant ainsi de réaliser une partie de mon rêve, qui est de travailler dans un laboratoire de l'IFREMER.

Je remercie également Monsieur Tristan RENAULT, chef du laboratoire LGP de la station de La Tremblade.

Un grand merci à mes maîtres de stage Madame Sylvie LAPEGUE, pour son encadrement, sa gentillesse et sa bonne humeur, et plus particulièrement Mademoiselle Delphine LALLIAS, avec qui j'ai partagé la paillasse, les manipulations et plein d'autres choses. Son écoute, sa gentillesse, sa bonne humeur et son entrain permanent sont autant de stimulants, pour ces nombreuses heures passées au laboratoire.

Merci également à mon tuteur pédagogique, Monsieur Vincent RIDOUX, pour ses dialogues par mail et ses conseils.

Je tiens à remercier également tout le personnel de la station IFREMER de La Tremblade, que ce soit les techniciens, les cadres chercheurs, les stagiaires et les thésards, qui ont fait que mon stage se passe dans les meilleures conditions. De l'intégration, au travail dans le laboratoire (qu'il faut se partager), aux pauses café et repas, et autres bons moments passés ensemble. Merci pour tous les conseils, cette bonne humeur, même dans les moments où tout va mal et tout le reste. Salutations donc à Maeva, Christopher, Sylvie, Benjamin, Nicole FAURY (responsable des stagiaires), Pascaline, Stéphanie et tous les autres, y compris mes camarades de l'IUT.

Un petit clin d'œil à Jean-Luc et "Pasqual", pour la sortie en mer, qui restera longtemps gravée dans ma mémoire.

SOMMAIRE

INTRODUCTION

I) PRESENTATION DE L'IFREMER

- A) HISTORIQUE
- B) MISSIONS, MOYENS ET ORGANISATION
- C) LA STATION DE LA TREMBLADE
- D) LE LABORATOIRE GENETIQUE ET PATHOLOGIE

II) MATERIEL ET METHODES

A) PRESENTATION DU MODELE D'ETUDE : LA MOULE BLEUE *MYTILUS EDULIS*

- 1) TAXONOMIE
- 2) CARACTERISTIQUES GENERALES
 - HABITUDES ALIMENTAIRES
 - REPARTITION GEOGRAPHIQUE
 - REPRODUCTION ET PREMIERS STADES DE DEVELOPPEMENT
- 3) FAMILLE ETUDIEE

B) EXTRACTION ET PURIFICATION DE L'ADN

- 1) LYSE DES CELLULES
- 2) EXTRACTION DE L'ADN
- 3) PURIFICATION DE L'ADN
- 4) VERIFICATION DE L'ADN

C) LES AFLPs

- 1) PRESENTATION GENERALE
- 2) RESTRICTION ENZYMATIQUE-LIGATION
- 3) PCR
- 4) REVELATION DU POLYMORPHISME

D) ETABLISSEMENT DE LA CARTE GENETIQUE

- 1) SELECTION DES MARQUEURS

2) METHODES STATISTIQUES

3) CONSTRUCTION DE LA CARTE GENETIQUE

- Estimation de la distance entre marqueurs
- Fonctions de cartographie
- Logiciel MAPMAKER 3.0

III) RESULTATS

A) DISTORSION DE SEGREGATION

B) CARTES GENETIQUES SEXE-SPECIFIQUE

1) CARTE DE LIAISON GENETIQUE PRELIMINAIRE FEMELLE

2) CARTE DE LIAISON GENETIQUE PRELIMINAIRE MALE

IV) DISCUSSION

A) UTILISATION DES AFLPS

1) AVANTAGES ET INCONVENIENTS DES AFLPS

2) QUALITE DES AFLPS

B) DISTORSION DE SEGREGATION

C) CARTES GENETIQUES PRELIMINAIRES

D) VERS UNE CARTE MOYENNEE

CONCLUSION

BIBLIOGRAPHIE

ANNEXES

ANNEXE I : Organigramme de l'IFREMER

ANNEXE II : Protocole d'extraction de l'ADN

INTRODUCTION

La cartographie génétique, basée sur l'analyse de ségrégation, consiste à positionner les uns par rapport aux autres des loci, avec comme objectif la mise en évidence de liaisons avec des caractères phénotypiques.

Les premières cartes génétiques ont été construites au début du siècle chez la drosophile par Morgan. Ces cartes se basaient essentiellement sur la ségrégation de caractères phénotypiques disposés linéairement sur un chromosome.

La cartographie complète du génome quant à elle, n'a pu être réalisée qu'avec l'apport des techniques de génétique moléculaire qui ont permis d'exploiter une partie de la variabilité des génomes en augmentant considérablement le nombre de loci potentiellement polymorphes. En effet, par suite de la dégénérescence du code génétique (plusieurs codons correspondent au même acide aminé), le polymorphisme protéique, observé sur gel d'électrophorèse, ne reflète environ que le quart des mutations présentes au niveau de la séquence d'ADN. De plus, l'étude du polymorphisme protéique réduit le champ d'investigation aux seules séquences codantes (représentant 5 à 10 % seulement du génome), alors que l'étude du polymorphisme nucléotidique permet de positionner des marqueurs anonymes sur l'ensemble du génome. Ce réseau de marqueurs permettra ensuite de détecter des liaisons avec des gènes dont seul le phénotype est connu, et donc de les cartographier.

Parmi les marqueurs les plus utilisés, on peut citer les RFLPs et les microsatellites. D'autres techniques ont aussi été développées comme les techniques de RAPDs et d'AFLPs qui sera utilisée dans cette étude.

Les AFLPs sont des marqueurs moléculaires dominants qui ne nécessitent aucune connaissance préalable du génome. C'est une technique basée sur la mise en évidence conjointe de polymorphisme de site de restriction et de polymorphisme d'hybridation d'une amorce de séquence arbitraire.

La moule bleue, *Mytilus edulis*, est l'un des Bivalves les plus cultivés pour être consommé. Ainsi une surveillance étroite de ces animaux a été mise en place. Par ailleurs, de nombreuses études de génétique des populations ont été réalisées (Gosling, 1992).

L'établissement de la carte génétique des marqueurs AFLPs chez *M. edulis* permettra donc par la suite de mieux localiser et étudier des gènes impliqués dans l'expression de caractères d'intérêts, tels que des résistances à certaines maladies, la croissance ou autres. Autant de caractères qui permettront une meilleure production.

Mais comment avec la technique des AFLPs, peut-on réussir à établir une carte génétique chez la moule ?

Pour répondre à cette problématique, plusieurs points vont être abordés. Tout d'abord, la présentation de la structure d'accueil va être développée afin de mieux comprendre le cadre de l'étude. Ensuite dans la partie "matériel et méthodes" vont être abordées les différentes techniques utilisées, avant d'arriver à la présentation des résultats obtenus, qui débouchera sur une discussion finale.

I) PRESENTATION DE L'IFREMER

IFREMER ou Institut Français de Recherche pour l'Exploitation de la MER est un organisme public résultant de la fusion du Centre National d'Exploitation pour les Océans ou CNEXO et de l'ISPTM : Institut Scientifique et Technique des Pêches Maritimes, en juin 1984 et donc régit par le décret du 5 juin 1984, qui sera modifié par ceux du 18 février 1998 et 14 mars 2002.

C'est un établissement à caractère industriel et commercial, placé sous la tutelle conjointe du ministère de la Recherche, de l'Agriculture et de la Pêche, du ministère de l'Équipement et des Transports et du ministère de l'Écologie et du Développement durable. En effet le but de cet organisme est de développer l'économie et les technologies maritimes tout en permettant une exploitation durable des ressources.

De nombreux partenaires collaborent également activement à ses activités, comme la DDE, la DDASS, les agences de l'eau, le CEMAGREF, l'INRA et les professionnels de la conchyliculture.

A) HISTORIQUE

- 1913 : création de l'association d'Encouragement des Industries Ostréicoles et conchylicoles françaises (AEIO), avec un laboratoire de microbiologie à La Tremblade (Charente-Maritime).
- 1918 : Création de l'Office Scientifique et Technique des Pêches Maritimes (OSTPM).
- 1928 : Le laboratoire de l'AEIO de La Tremblade est rattaché à l'OSTPM. L'OSTPM présente les premières études relatives à l'appauvrissement des fonds marins, sur la cartographie sous-marine appliquée à la pêche, sur la reproduction de l'huître et sa mortalité.
- 1954 : Transformation de l'OSTPM en Institut Scientifique et Technique des Pêches Maritimes (ISTPM).
- 1967 : Création du CNEXO, établissement public à caractère industriel et commercial. Il se doit "de développer la connaissance des océans, les études et les recherches tendant à l'exploitation des ressources contenues à leur surface, dans leur masse, leur sol et leur sous-sol".
- 1972 : Participation du CNEXO à l'élaboration du Réseau National d'Observation (RNO), premier maillon de la chaîne de réseaux gérée par l'IFREMER.
- 1976 : Le CNEXO est chargé de coordonner les programmes de recherche et de développement, de l'ensemble des organismes publics concernés par l'océanologie.
- 1984 : Création de l'IFREMER : fusion du CNEXO et de l'ISTPM.
- 2001 : L'unité de microbiologie de La Tremblade a été accréditée par le COFRAC.

B) MISSIONS, MOYENS ET ORGANISATION

L'IFREMER a pour mission de connaître, évaluer, mettre en valeur les ressources des océans, permettre leur exploitation durable ainsi que d'améliorer les méthodes de surveillance, de protection et de mise en valeur du milieu marin et côtier.

Vingt six implantations sur le littoral métropolitain et dans les DOM-TOM, 7 navires, 1 submersible, 1 engin télé opéré pour grande profondeur, un ensemble de moyens d'élevage aquacole et d'expérimentation sont nécessaires pour ces objectifs.

L'institut est structuré en 5 centres : Boulogne-sur-Mer, Brest, Nantes, Toulon et Tahiti et en une vingtaine de stations rattachées à ces centres. Le siège social est situé à Paris (Issy-les-Moulineaux dans les Hauts de Seine).

L'IFREMER de La Tremblade est une des stations rattachées au centre de Nantes.

C) LA STATION DE LA TREMBLADE

L'IFREMER de La Tremblade est implanté dans le premier bassin conchylicole d'Europe, en Charente-Maritime (*Figure 1*).



Figure 1 : Photographie de la station de La Tremblade (vue du ciel).
Site Internet de l'IFREMER : www.ifremer.fr/latremblade/

La station de La Tremblade est structurée en deux laboratoires :

- Le Laboratoire Environnement et Ressources des Pertuis charentais : le LER/PC, qui remplit des missions de surveillance de l'environnement marin et d'évaluation des ressources conchylicoles. Le chef de Laboratoire est Monsieur Jean Prou.
- Le Laboratoire Génétique et Pathologie : LGP, qui étudie les caractères génétiques des mollusques bivalves dans le but d'améliorer la qualité des élevages. Il contrôle aussi les ressources conchylicoles, étudie les agents pathogènes et leurs impacts sur l'environnement. Le chef de laboratoire est Monsieur Tristan Renault.

La station comprend une éclosérie et un marais expérimental. Une soixantaine de salariés travaillent dans cette structure, dont le Directeur est Monsieur Philippe Gouletquer.

D) LE LABORATOIRE GENETIQUE ET PATHOLOGIE

Spécialisé dans les domaines de la génétique et de la pathologie des invertébrés marins et plus spécifiquement des mollusques bivalves, le *Laboratoire Génétique et Pathologie (LGP)* dépend du **Département Amélioration Génétique, Santé Animale et Environnement**.

Les principaux objectifs du laboratoire visent essentiellement à développer des recherches chez les mollusques bivalves marins dans les domaines de :

- **La pathologie** : surveillance des ressources conchylicoles, identification des agents pathogènes, description de leur cycle de développement, mise au point de techniques de reproduction expérimentale des maladies, développement d'outils performants de diagnostic utilisables à des fins de recherche ou de contrôle, étude de l'impact de ces maladies et de leur évolution géographique et temporelle.

- **La génétique** : étude des ressources génétiques, testage de nouvelles espèces, de nouvelles populations et d'hybrides pour limiter les risques liés à la monoculture ; obtention de souches résistantes ou tolérantes aux maladies pour essayer d'apporter des réponses aux épizooties qui remettent en cause les productions ; création de souches ou de lignées présentant de meilleures performances de croissance, de qualité de chair, une meilleure adaptation aux conditions de milieu d'élevage ou éventuellement de faibles besoins métaboliques, pour améliorer la productivité des entreprises.

II) MATERIEL ET METHODES

A) PRESENTATION DU MODELE D'ETUDE : LA MOULE BLEUE MYTILUS EDULIS

1) TAXONOMIE

D'après la taxonomie linnéenne, la moule appartient au règne animal et à l'embranchement des **Mollusques**.

En effet les mollusques (du latin *mollis*, mou) sont des métazoaires, triblastiques à symétrie bilatérale. Leur tégument est mou et contient de nombreuses glandes. Bien que les mollusques ne soient généralement pas métamérisés, certaines espèces présentent des traces de métamérisation. Leur corps se subdivise en trois parties :

- * le pied est un organe musculueux destiné à la locomotion.
- * la masse viscérale, comme son nom l'indique, contient les viscères. Elle est contenue dans une mince tunique qu'on appelle le manteau. C'est le manteau qui sécrète la coquille des mollusques, qui leur sert de protection et/ou de squelette.
- * le bourrelet palléal constitue une cavité qui protège les organes respiratoires et où débouchent les organes excréteurs, l'intestin et les conduits génitaux.

Les **bivalves** sont des mollusques dont la coquille est constituée de deux parties distinctes et attachées, plus ou moins symétriques, pouvant s'ouvrir ou se refermer, tout comme celle de la moule. Les bivalves sont tous exclusivement aquatiques.

La classe se nomme également Pelecypoda ou Lamellibranchia car chaque branchie en lamelles possède un feuillet direct et un feuillet réfléchi reliés entre eux par des septa branchiaux transverses (*Figure 2*).

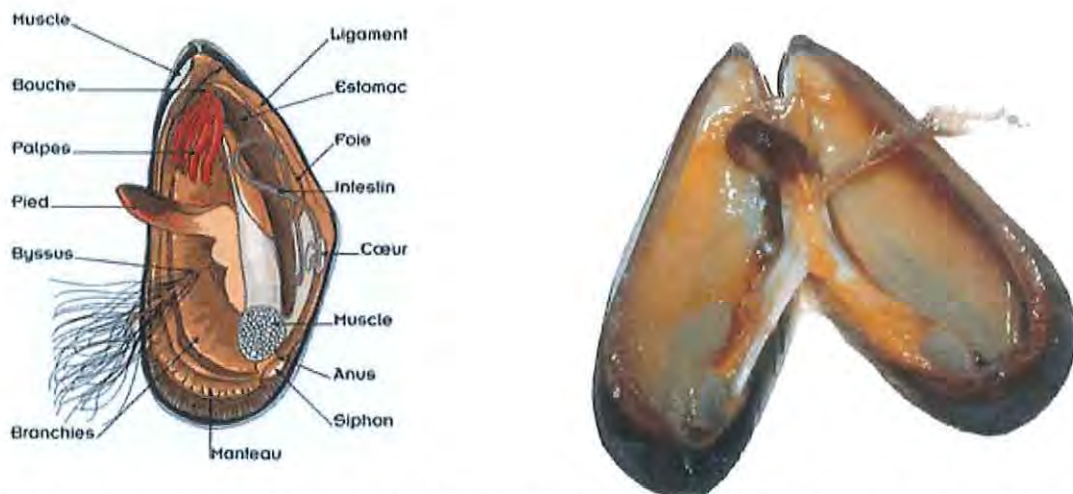


Figure 2 : *Principaux éléments anatomiques d'une moule, sous forme de schéma et en vue interne.*

La moule appartient aux **Mytiloïdes** (ordre des **Mytiloïda**), qui sont des mollusques bivalves, équivalves mais pas symétriques, dépourvues de dents de charnière. Les crochets se trouvent à l'extrémité antérieure. Le ligament est développé mais les muscles adducteurs sont vestigiaux.

La moule bleue (*Mytilus edulis*) adulte est un animal fixé à son support par le byssus. Elle ne peut se déplacer que très lentement; le byssus sécrète un liquide qui se polymérise avec l'eau formant des filaments très résistants qui se lient au support sur lequel la moule évolue.

L'extérieur de la coquille est bleuâtre ou noir bleuâtre et luisant, mais parfois brun pâle et strié, tandis que l'intérieur est généralement pourpre (*Figure 3*). Dans la coquille, la chair de la moule est le plus souvent de couleur rose ou orangée (femelle) ou blanchâtre (mâle).



Figure 3 : Chair et coquille de moule bleue *Mytilus edulis*

2) CARACTERISTIQUES GÉNÉRALES DE LA MOULE

• Habitudes alimentaires

Les moules se nourrissent en suspension et, comparativement à d'autres mollusques, elles sont les plus efficaces à cette activité. Elles se nourrissent en filtrant activement les particules de l'eau qui entre et qui sort de la cavité du manteau en passant par les siphons dentelés. La respiration se fait quand un courant d'eau passe sur les branchies de la moule (*Figure 4*). Les cellules phytoplanctoniques, vivantes et mortes, constituent la principale source de nourriture, mais d'autres sources de carbone, comme des macrophytes en décomposition ou des débris en suspension peuvent aussi compléter le régime alimentaire.

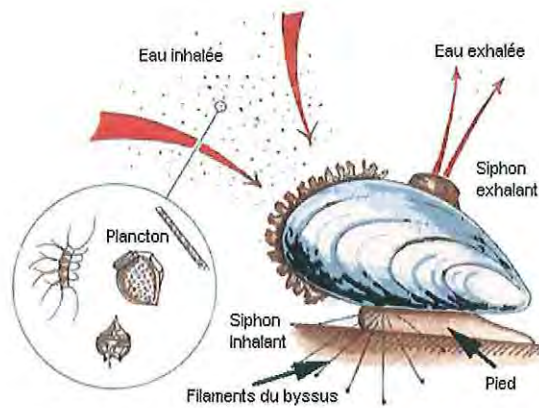


Figure 4 : Respiration de la moule

• Répartition géographique

M. edulis est présente partout dans le monde, dans la plupart des eaux polaires et tempérées de l'hémisphère Nord (de la mer blanche aux côtes Atlantique du sud de la France) (Figure 5). Elle peut vivre dans des estuaires peu profonds et légèrement saumâtres jusque dans des milieux très profonds à forte salinité en haute mer, mais elle privilégie les baies et les estuaires qui ont un taux élevé de nutriments en raison du ruissellement, ce qui favorise un accroissement du phytoplancton. Elle peut vivre dans une eau dont la gamme de salinité étendue varie entre 0 et 31 ppm (26 ppm étant la valeur optimale), et dont la température est comprise entre 0°C et 25°C. Il existe d'autres espèces de moules présentes dans le monde entier.

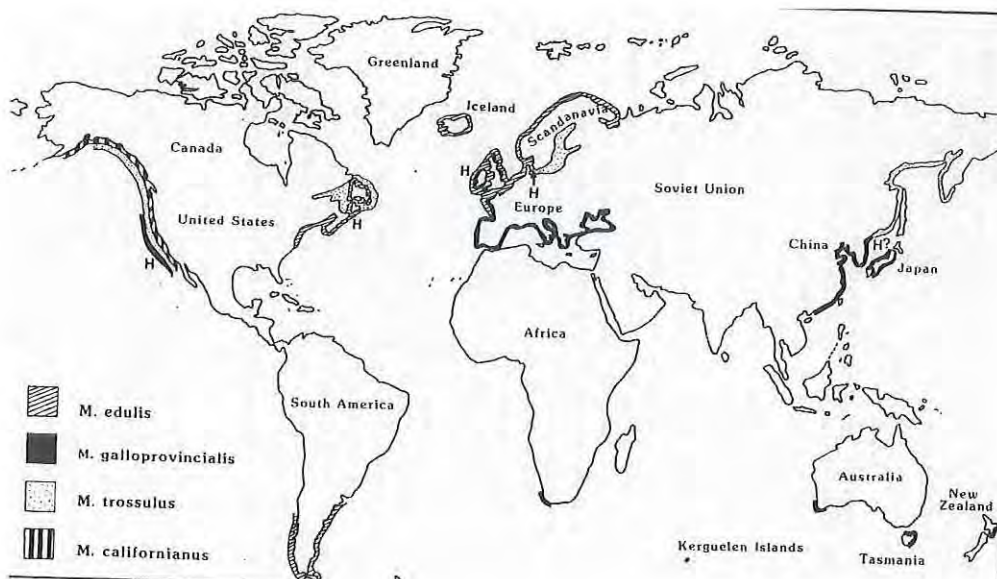


Figure 5 : Distribution géographique de la moule

• Reproduction et premiers stades de développement

Les sexes sont habituellement distincts (dioïques), mais certains individus ont parfois les deux sexes (hermaphrodites).

Les moules atteignent généralement la maturité sexuelle à la fin du printemps ou au début de l'été, la formation de sperme ou d'œufs (gamétogenèse) se produisant dans de nombreux conduits situés principalement dans les lobes du manteau.

Après l'émission de gamètes (sperme et ovocytes), la fécondation se fait dans la colonne d'eau.

Il faut seulement cinq heures pour que l'embryon commence à nager, à se nourrir de petites cellules de phytoplancton et à grossir pour atteindre le dernier stade larvaire, celui des larves véligères. Les larves nagent librement pendant trois à quatre semaines, jusqu'à la métamorphose, alors qu'elles cherchent à se fixer.

À la métamorphose, la moule allonge son pied, rétrécit son velum et secrète les filaments du byssus pour se fixer sur des substrats solides comme des collecteurs, des pieux de quai, des algues filamenteuses ou des rochers. On parle alors de la chute du naissain, qui a lieu entre la mi-juin et la mi-juillet. À ce stade, les jeunes moules peuvent facilement se détacher et changer d'endroit, soit en utilisant leur pied pour ramper véritablement ou en se laissant flotter dans la colonne d'eau. Si elles sont soumises à un stress, les moules peuvent, par conséquent, continuellement s'accrocher à un nouveau substrat jusqu'à ce qu'elles se fixent en permanence. Cependant, leurs mouvements deviennent plus limités à mesure qu'elles grossissent puis finissent par se fixer définitivement à un substrat solide et approprié où elles s'engraissent jusqu'à ce qu'elles deviennent des moules adultes (*Figure 6, page ci-après*).

3) FAMILLE ETUDIÉE

Plusieurs croisements biparentaux ont été réalisés en juillet 2004 avec des géniteurs en provenance du Pays de Galles.

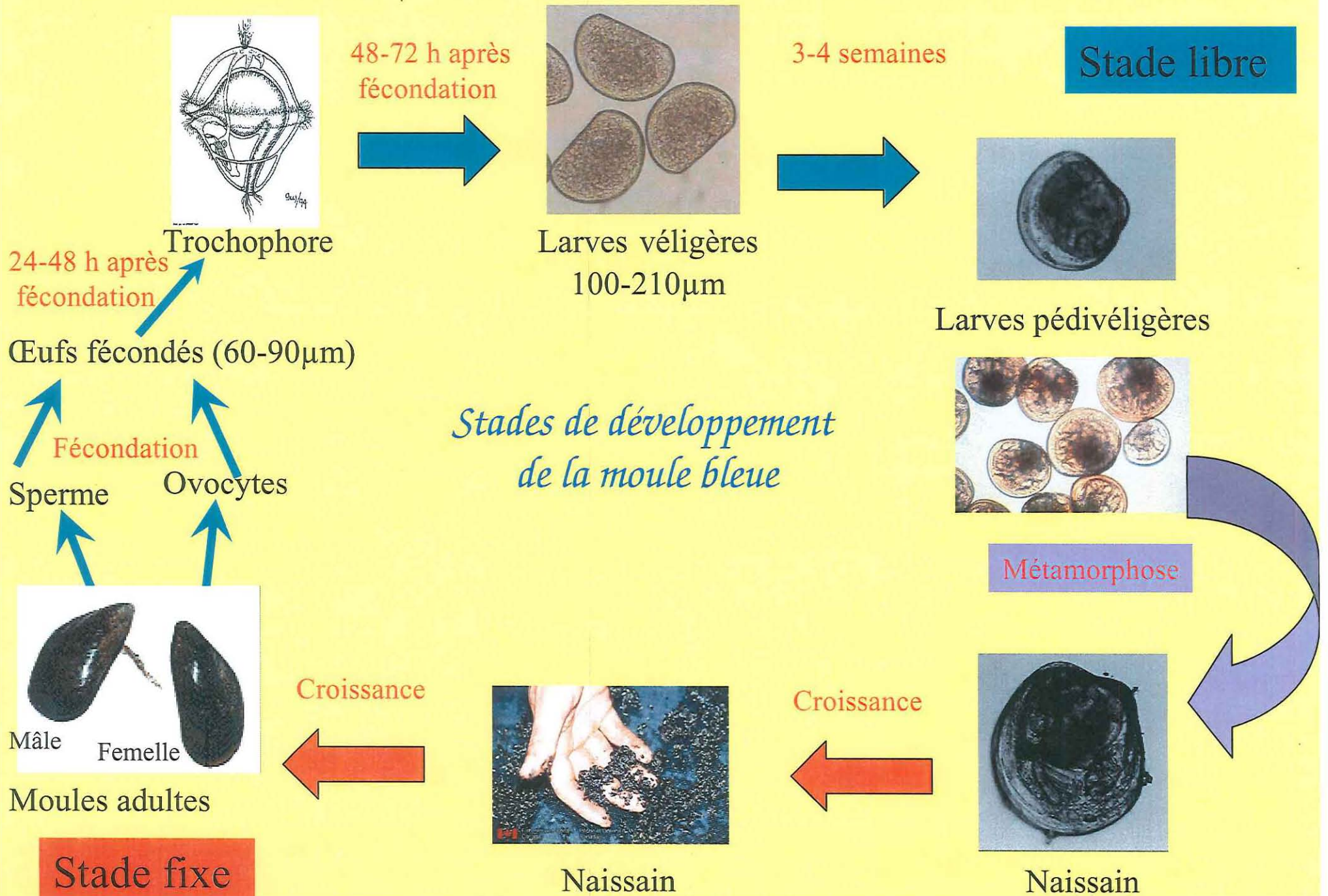
Une soixantaine de moules ont été maintenues en béccher individuel, après injection de KCl 0,5M dans la cavité palléale pour stimuler la ponte des gamètes.

La fécondation a été réalisée in vitro par le mélange des ovocytes d'une femelle et du sperme d'un mâle (10 spermatozoïdes pour un ovocyte). Après 72 heures les larves D obtenues ont été transférées dans des bécchers de 2 litres pour l'élevage larvaire de trois à quatre semaines.

À la suite de la métamorphose, les naissains se sont fixés sur des tamis et ont été conservés à l'écloserie de La Tremblade, pour grossissement. Chaque famille (issue d'un croisement biparental impliquant 2 géniteurs différents) est élevée dans des récipients différents pour éviter les mélanges de lots.

La famille étudiée ici est la famille 10, composée de 92 F1 et des deux géniteurs (mâle et femelle).

Figure 6 : Stades de développement de la moule bleue *Mytilus edulis*



B) EXTRACTION ET PURIFICATION DE L'ADN

L'extraction d'un ADN de grande qualité implique plusieurs étapes :

1) LYSE DES CELLULES

Dans un premier temps, les 94 individus de la famille sélectionnée (92 F1 et 2 parents) ont été sacrifiés pour prélever un fragment de branchies ($0,5\text{mm}^2$). Chaque fragment a été placé dans un tube Eppendorf de 1,5mL contenant 400 μl de tampon de lyse, le 2% CTAB (Hexadecylthimethyl ammonium bromide : 100 mM Tris-HCl, 20 mM NaCl pH = 8, 20 mM EDTA, eau bidistillée, 2% CTAB).

20 μl de protéinase K (10mg/mL) ont été ajoutés à chaque tube avant incubation au bain-marie à 60°C pendant toute une nuit. Cette enzyme permet la digestion des protéines des parois cellulaires pour rendre liquides ces dernières.

2) EXTRACTION DE L'ADN

L'extraction de l'ADN se fait au chloroforme : isoamyl alcohol (24 : 1). L'ajout de ce produit chimique permet la précipitation des protéines, mais pas celle des acides nucléiques, étant donné qu'ils ne sont pas solubles dans celui-ci.

Deux extractions successives par ce procédé sont nécessaires, celles-ci se faisant en trois étapes :

- Ajout de 400 μl de chloroforme, mixe par retournement
- Centrifugation à 13500 tours/min pendant 10 minutes
- Prélèvement de la phase aqueuse contenant l'ADN, en veillant à ne pas prélever la galette de protéines qui s'est formée à l'interface des 2 phases

Seule la phase aqueuse contenant l'ADN est récupérée et conservée, pour la suite de l'extraction.

3) PURIFICATION DE L'ADN

Pour obtenir un ADN de grande qualité, la solution d'ADN récupérée suite à l'extraction au chloroforme : isoamyl alcohol (24 : 1) a été purifiée.

La purification de l'ADN s'est faite grâce au kit "Wizard DNA Clean Up System" (Promega). C'est une purification sur minicolonne (*Annexe II*).

L'ADN retenu sur la membrane est alors élué avec 50 μl de TE (0,01M Tris-HCl, 0,5mM EDTA) chauffé à 70°C.

4) VERIFICATION DE L'ADN

Après la purification, la mesure de la concentration d'ADN ($\mu\text{g/ml}$) dans chaque échantillon est mesurée au spectrophotomètre Eppendorf Bio Photometer, de même que le rapport de densité optique, DO à 260nm/DO à 280nm. Ce rapport permet de déterminer si les échantillons sont contaminés ou non par des protéines (rapport compris entre 1,7 et 1,8 pour un échantillon pur et supérieur à 2 pour un échantillon contaminé par les protéines).

La gamme de concentration obtenue pour les échantillons testés est comprise entre 140 et 750 $\mu\text{g/ml}$, avec un rapport toujours compris entre 1,7 et 1,8.

Des dilutions des échantillons ont été effectuées pour obtenir une concentration finale d'ADN de 100 $\mu\text{g/ml}$

Afin de confirmer visuellement la bonne qualité de l'ADN extrait, une électrophorèse sur gel d'agarose 2% a été effectuée pour des échantillons de concentrations différentes. La présence d'une bande de haut poids moléculaire, de forte intensité avec un faible smear, confirme l'absence de contamination protéique et de dégradation de l'ADN (*Figure 7*).

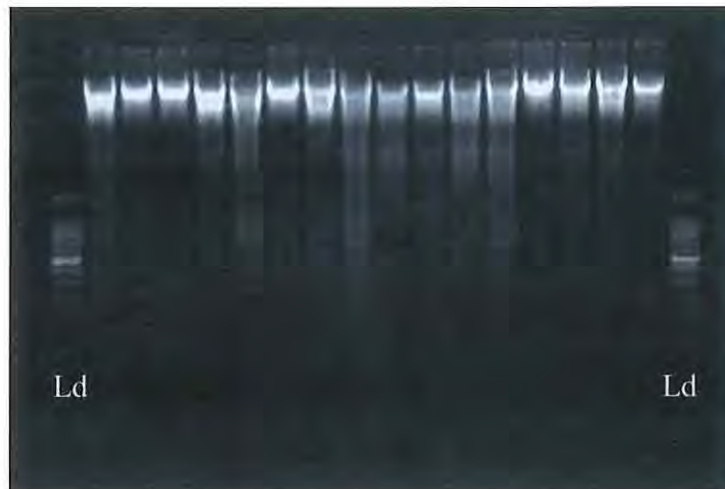


Figure 7 : *Electrophorèse sur gel d'agarose d'ADN de moules après extraction au chloroforme : isoamyl alcohol (24:1) et purification au kit DNA CleanUp System (Promega).*

Ld : Marqueur de taille 100 à 1500 pb.

C) LES AFLPs

1) PRESENTATION GENERALE

Les AFLPs (Amplified Fragment Length Polymorphism) sont des marqueurs moléculaires, utilisés pour l'étude et la pratique de la génétique, qui peuvent servir dans la cartographie des génomes, comme ce sera le cas ici. Ces marqueurs ne nécessitent

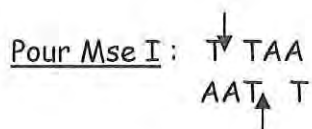
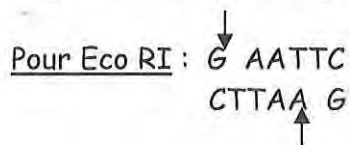
aucune connaissance préalable du génome : ils sont donc utilisés comme alternative chez des espèces où peu de marqueurs locus spécifiques (microsatellites) sont disponibles, ce qui est le cas chez *M. edulis*.

C'est une technique basée sur la mise en évidence conjointe de polymorphisme de site de restriction et de polymorphisme d'hybridation d'une amorce de séquence arbitraire.

Le polymorphisme étant le caractère de ce qui se présente sous différentes formes, il peut être relié à des différences de séquences d'ADN entre individus. Les AFLPs sont des parties polymorphes du génome révélées aléatoirement, après coupure de l'ADN par des enzymes de restriction, ligature d'adaptateurs spécifiques aux sites de restriction et amplification sélective de certains de ces fragments de restriction avec deux cycles successifs de PCR.

2) RESTRICTION ENZYMATIQUE - LIGATION

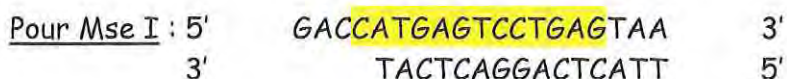
L'ADN est digéré par deux enzymes de restriction (endonucléases coupant de manière définie et reproductible l'ADN double brin en des sites spécifiques) : Eco RI et Mse I dans notre cas, mais d'autres enzymes peuvent être utilisées. Leurs sites de restriction respectifs sont les suivants :



Dans le même temps deux adaptateurs (spécifiques de chacun des sites de restriction) sont ajoutés aux extrémités à l'aide de la T4 DNA ligase, chaque adaptateur étant différent pour chaque site de restriction.

Les adaptateurs sont ajoutés en concentration supérieure à celle des fragments d'ADN pour optimiser la ligation.

Les différents adaptateurs utilisés ici sont :



Pour la restriction enzymatique, 5µl d'ADN (100µg/ml) est mélangé à 45µl de Mix (1X NE Buffer, 1X BSA, Eco RI 5U, Mse I 5U, 5pmol Adaptateur Eco RI, 50pmol Adaptateur Mse I, T4 DNA ligase NEB 200U, 1X ligase Buffer et eau bidistillée). Ce

mélange est placé dans un thermocycleur à 16°C pendant 16 heures pour que la restriction enzymatique et ligation puissent se faire dans des conditions optimales.

3) PCR

Le substrat obtenu par la restriction enzymatique est dilué au 1/10^e dans du TE (20µl d'ADN et 180µl de tampon 0.1X TE).

Les fragments obtenus à l'issue de la restriction enzymatique doivent être amplifiés par PCR avec des oligonucléotides spécifiques, afin de révéler le polymorphisme et de réduire le nombre de fragments visualisés.

La PCR (Polymerase Chain Reaction) est une technique permettant l'amplification sélective in vitro d'une séquence d'ADN donnée, par la répétition de cycles de réplication d'ADN. Chaque cycle est constitué de trois étapes, nécessitant chacune une température différente : dénaturation ; hybridation avec les oligonucléotides de synthèse, spécifiques de la séquence cible et qui servent d'amorce; élongation de la chaîne nucléotidique.

Les amorces d'AFLPs sont composées de trois parties : dans l'ordre, une **partie centrale** permettant d'augmenter la spécificité de la réaction en la stabilisant; la séquence spécifique de l'enzyme de restriction et une **extension sélective**.

Pour Eco RI : 5' **GACTGCGTACC** AATTCNNN 3'
Pour Mse I : 5' **CATGAGTCCTGAG** TAA NNN 3'

Les amorces ci-dessus sont utilisées pour amplifier par PCR les AFLPs, ceci nécessitant deux étapes, qui sont la PCR pré-sélective et la PCR sélective.

La PCR pré-sélective consiste à lier sur l'adaptateur les amorces possédant un seul nucléotide spécifique, afin de sélectionner un sous-échantillon de tous les fragments de produits de la restriction.

Pour ce faire le mix de la préselective-PCR se compose de tampon 1X, 1,5 mM de MgCl₂, 0,1 mM de dNTP, 25pmol d'amorces contenant un nucléotide sélectif, 1U de Taq polymérase (Promega) et d'eau bidistillée. La réaction de PCR se fait dans un volume final de 50µl, contenant 45µl de mix et 5µl de produit de restriction-ligation dilué au 1/10^e.

Le programme d'amplification se compose des étapes suivantes :

- 5 minutes à 94°C (dénaturation initiale)
- 20 cycles successifs : 30 secondes à 94°C (dénaturation), 1 minute à 56°C (appariement des amorces), 1 minute à 72°C (élongation)
- 7 minutes à 72°C (élongation finale).

A l'issue la PCR pré-sélective, le substrat est à nouveau dilué au 1/10^e dans du TE, avant l'étape de PCR sélective.

Lors de la PCR sélective, ce sont trois nucléotides sélectifs qui sont ajoutés à l'amorce, pour sélectionner un nombre encore plus restreint de fragments. Certains de

ces nucléotides sont marqués par des fluorochromes de couleurs différentes suivants les paires d'AFLPs utilisées (Tableau 1). Seule l'étape de PCR sélective utilise les amorces marquées, 36 paires d'AFLPs ont été utilisées au cours de cette étude.

Tableau 1 : Paires d'amorces d'AFLPs utilisées pour le génotypage, avec les abréviations employées.

	Eco + CAG FAM (bleu)	Eco + ACG HEX (vert)	Eco + ACT NED (jaune)
Mse + CGA	A1	B1	E1
Mse + CAA	A2	B2	E2
Mse + CTG	A3	B3	E3
Mse + CAT	A4	B4	E4
Mse + CTT	A5	B5	E5
Mse + ATC	A6	B6	E6
Mse + AGT	A7	B7	E7
Mse + CTC	A8	B8	E8
Mse + CTA	A9	B9	E9
Mse + CAC	A10	B10	E10
Mse + CAG	A11	B11	E11
Mse + CCT	A12	B12	E12

Pour la PCR sélective, le mix se compose de tampon 1X, 1,5mM de MgCl₂, 0,1 mM de dNTP, 5 ng d'amorces contenant 3 nucléotides sélectifs, 0,4U de Taq polymérase (Promega) et d'eau bidistillée.

Le volume final étant de 20µl, 5µl d'ADN est mélangé à 15µl de mix avant le début du programme de la sélective-PCR, qui se décompose de la manière suivante :

- 5 minutes à 94°C (dénaturation initiale)
- 13 cycles successifs de type touchdown : 30 secondes à 94°C, 30 secondes à 65°C (avec la température diminuant de 0,7°C à chaque cycle), et 1 minute à 72°C
- 23 cycles successifs : 30 secondes à 94°C, 30 secondes à 56°C, 1 minute à 72°C
- 30 minutes à 72°C (élongation finale).

4) REVELATION DU POLYMORPHISME

L'utilisation de trois fluorochromes différents pour le marquage des amorces Eco+3 permet le multiplexage des produits de PCR lors du passage des échantillons sur le séquenceur qui sert alors de génotypeur. Par exemple l'analyse des paires A1, B1, E1 se fera simultanément.

Un mix contenant 0,5 µl de marqueur de taille de couleur rouge GeneScan™-500 ROX™ STANDARD (Applied Biosystems) et 10 µl de formamide est mélangé aux 3 µl de produits de PCR (1 µl de la paire A1, 1µl de la paire B1 et 1 µl de la paire E1), avant une dénaturation de 5 minutes à 95°C. Suite à la dénaturation, les échantillons subissent une électrophorèse par capillaires dans un séquenceur ABI 3100-Avant. La révélation du

polymorphisme se fait par analyse des profils électrophorétiques obtenus à l'aide du logiciel GeneMapper 3.7 (Applied Biosystems).

D) ETABLISSEMENT DE LA CARTE GENETIQUE

Pour établir une carte génétique, il faut :

- Disposer d'une descendance en ségrégation pour de nombreux marqueurs ;
- Déterminer le génotype des individus d'une ou plusieurs descendance pour un maximum de locus marqueurs ;
- Eliminer des loci qui présentent de fortes distorsions de ségrégation c'est-à-dire ne respectant pas les règles de ségrégation Mendélienne. Pour cela un test statistique est établi (test du χ^2).

1) SELECTION DES MARQUEURS

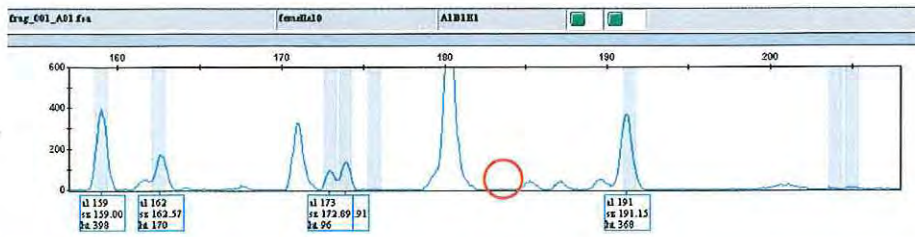
Lors de l'analyse des AFLPs, seuls les marqueurs présents chez l'un des parents et qui ségrègent chez les enfants sont retenus. Comme les AFLPs sont des marqueurs dominants, sélectionner les marqueurs ségrégeant en F1 qui sont spécifiques de l'un des parents permet de connaître l'origine parentale des allèles dans la descendance, ainsi que leur génotype complet. Ainsi la présence du pic sera associée à un génotype Aa et l'absence par un génotype aa.

Par exemple, l'allèle 176 du marqueur A1 est présent chez le mâle et chez les individus F1 1 et 4, tandis qu'il est absent chez la femelle et chez les individus F1 2 et 3. On peut donc conclure que cet allèle est spécifique du mâle (*Figure 8*).

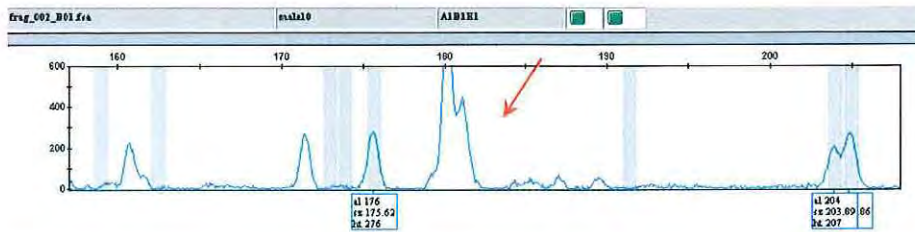
Ces deux jeux de données (mâle et femelle) seront donc analysés indépendamment.

Les marqueurs AFLPs remplissant ces critères sont sélectionnés, puis vérifiés une seconde fois à la main, pour corriger les erreurs du logiciel. Celles-ci correspondent le plus souvent à une non-détection des pics se trouvant en dessous du seuil fixé par le logiciel, ils sont donc rajoutés s'ils sont visibles à l'œil nu.

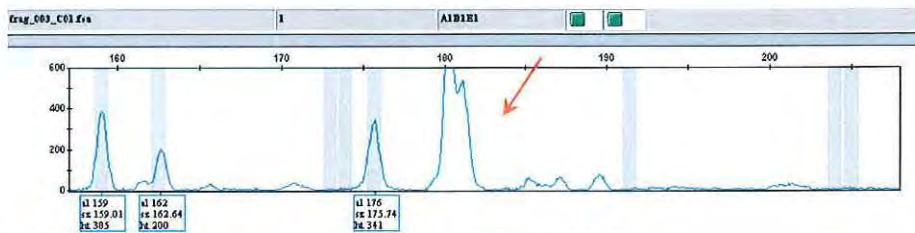
Après cette vérification, seuls les marqueurs non distordus sont gardés pour l'établissement de la carte génétique.



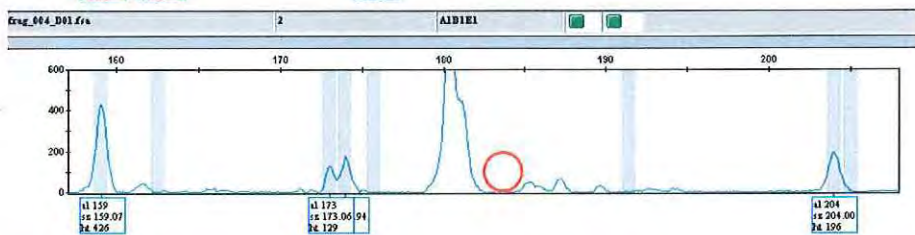
Femelle



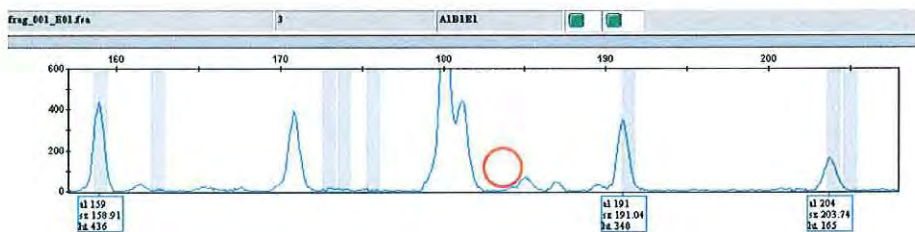
Mâle



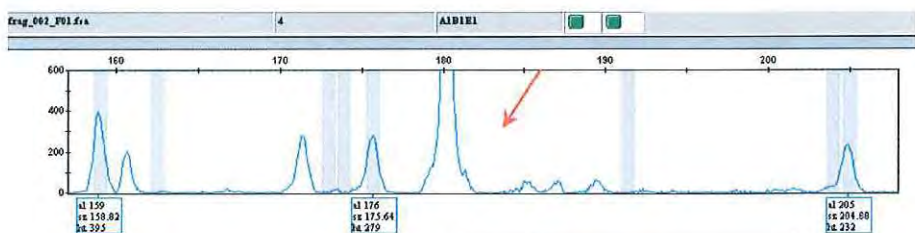
Individu F1 1



Individu F1 2



Individu F1 3



Individu F1 4

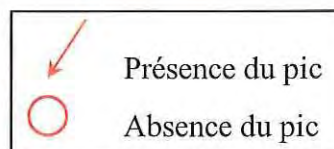


Figure 8 : Capture d'écran du logiciel GeneMapper 3.7 : profils électrophorétiques montrant le type d'allèles retenus pour l'analyse.

2) METHODES STATISTIQUES

Afin d'éliminer les locus qui présentent de fortes distorsions de ségrégation, le test statistique du χ^2 a été réalisé sur tous les allèles comme suit :

$$\chi^2 = \frac{(n_{pic-obs} - n_{pic-TH})^2}{n_{pic-TH}} + \frac{(n_{abs-obs} - n_{abs-TH})^2}{n_{abs-TH}}$$

où $n_{pic-obs}$ et n_{pic-TH} correspondent au nombre d'individus F1 observé et attendu ayant le pic; $n_{abs-obs}$ et n_{abs-TH} correspondent au nombre d'individus F1 observé et attendu sans le pic.

Pour les AFLPs, seulement deux classes sont possibles : présence ou absence du pic représentant l'allèle du marqueur, si bien que le nombre de degrés de liberté est un nombre de classe 1, on notera : ddl=1.

Pour chaque valeur de χ^2 trouvée, une interprétation s'impose :

- $\chi^2 < 3,841$ NS le test est non significatif, le marqueur ségrège selon les attendus Mendéliens, on le considère donc comme non distordu, il servira pour la carte génétique.

- $3,841 < \chi^2 < 6,635$ $p < 0,05$ le test est significatif au seuil de test statistique $\alpha=5\%$, le marqueur est considéré comme distordu, il ne sera donc pas conservé pour établir la carte.

- $6,635 < \chi^2 < 10,828$ $p < 0,01$ le test est très significatif au seuil de test $\alpha=1\%$, le marqueur est considéré comme fortement distordu, il ne sera donc pas conservé pour établir la carte.

- $\chi^2 > 10,828$ $p < 0,001$ le test est significatif au seuil $\alpha=1\%$, le marqueur est alors hyper distordu, il ne sera donc pas conservé pour établir la carte.

Exemple : pour un marqueur génotypé sur 86 individus, 16 possédant le pic, 70 ne le possédant pas, d'où $\chi^2 = [(16-43)^2/43] + [(70-43)^2/43] = 33,9$ ($p < 0.001$).

Une fois les marqueurs sélectionnés pour la carte génétique, il faut essayer d'assigner chacun des marqueurs à un groupe de liaison.

Pour cela on utilise une méthode probabiliste qu'est le LODscore (Logarithm of the ODD ratio, ou logarithme décimal du rapport de vraisemblance de 2 hypothèses) (Morton, 1955).

On considère alors deux hypothèses :

- Hypothèse 1 : il y a une liaison génétique entre 2 loci ($r = \theta < 0,5$) ; la vraisemblance de cette hypothèse est notée e^L ($r = \theta$).

- Hypothèse 2 : il n'y a pas de liaison génétique entre les 2 loci ($r = 0,5$) ; la vraisemblance de cette hypothèse est notée e^L ($r = 0,5$).

On note ainsi le LODscore : $\text{LODscore} = \log_{10} (e^L (r = \theta) / e^L (r = 0,5))$

Un LODscore positif et supérieur à une valeur seuil correspond à une liaison génétique entre les 2 loci. Dans notre étude, un LOD de 3 a été retenu comme valeur seuil.

Exemple : LODscore = 3 : l'hypothèse de liaison est 1000 fois plus vraisemblable que l'hypothèse d'indépendance entre 2 locus, tout comme un LODscore = -1, correspond à une hypothèse de liaison étant 10 fois moins vraisemblable qu'une hypothèse d'indépendance entre 2 loci.

3) CONSTRUCTION DE LA CARTE GENETIQUE

• ESTIMATION DE LA DISTANCE ENTRE MARQUEURS

Si l'on considère deux loci distincts A et B, 4 gamètes haploïdes différents peuvent se former lors de la reproduction entre deux individus : 2 de type parental et 2 de type recombiné. La proportion de gamètes recombinés r correspond au taux de recombinaison, ce dernier rendant compte de la distance entre les deux locus. En effet, plus deux marqueurs sont distants, plus l'occurrence d'un crossing-over est probable, donc plus il y aura de recombinaison entre eux. La fréquence de gamètes recombinés sera alors plus grande.

Le taux de recombinaison ne peut être utilisé directement comme mesure ou unité de distance. En raison des recombinaisons multiples, et du phénomène d'interférence (empêchement de l'occurrence d'un second crossing-over à proximité d'un premier), cette quantité n'est pas additive. C'est pourquoi la proportion de recombinaison doit être transformée en une variable additive, en utilisant une fonction de cartographie. Ce phénomène est d'autant plus marqué que la distance entre deux marqueurs est faible.

• FONCTIONS DE CARTOGRAPHIE

L'unité de mesure de la distance génétique entre deux loci est le centimorgan (cM), unité utilisée pour la première fois par Morgan et Sturtevant en 1913. Le centimorgan correspond à la distance qui sépare deux loci entre lesquels l'espérance du taux de recombinaison est de 0,01, soit un crossing-over pour 100 méioses.

Les fonctions de cartographie sont au nombre de deux :

- La distance d'Haldane, qui prend en compte les crossing-over multiples.
 $D_H = -50 \ln (1-2r)$
- La distance de Kosambi, qui prend en compte les crossing-over multiples et l'existence d'interférence. C'est la distance utilisée dans notre cas.
 $D_K = 25 \ln [(1+2r) / (1-2r)]$

Suivant la fonction utilisée, on peut se rendre compte de l'influence sur la fréquence de recombinaison en fonction de la distance génétique (*Figure 9*).

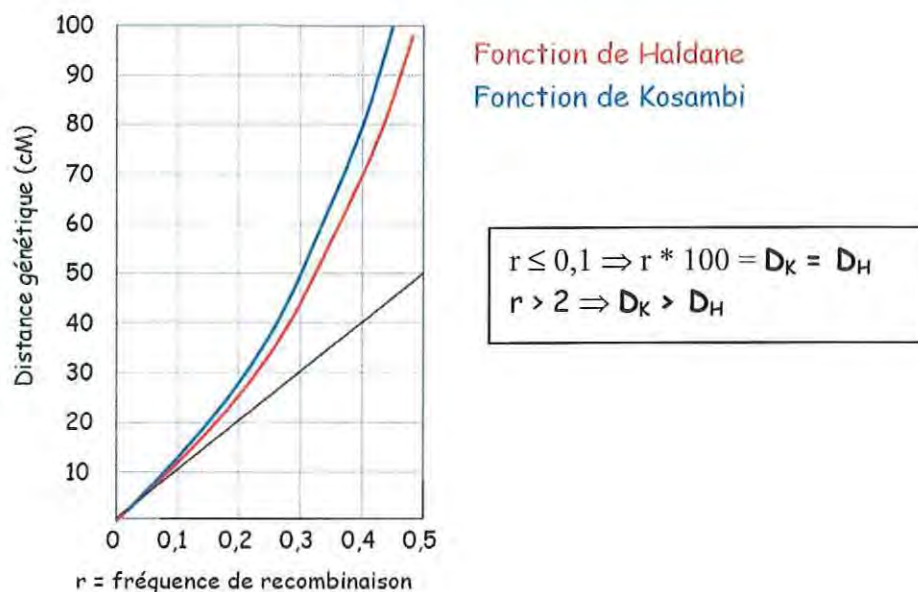


Figure 9 : Fréquence de recombinaison en fonction de la distance génétique, suivant les deux fonctions de cartographie.

• LOGICIEL MAPMAKER 3.0.

De nombreux logiciels qui permettent de construire des cartes génétiques existent, mais celui utilisé ici est le logiciel MAPMAKER 3.0. (Lander, Green *et al.* 1987).

Après codage des données toutes les analyses sont faites à partir de ce logiciel.

Dans notre cas le codage des données a été effectué de la façon suivante : les génotypes Aa sont codés "H", les génotypes aa sont codés "A" et les données manquantes sont remplacées par un signe "-".

Le travail du logiciel se fait de manière séquentielle. En effet, dans un premier temps, il calcule les taux de recombinaison entre chaque paire de marqueurs associée au maximum de vraisemblance. Seuls les marqueurs associés à un taux de recombinaison θ avec un LODscore supérieur à 3 sont considérés comme liés. Le logiciel établit ainsi les groupes de liaisons à l'aide de la commande "GROUP".

L'étape suivante permet de déterminer l'ordre des différents marqueurs à l'intérieur de chaque groupe de liaison, et de calculer les LODscore associés à chaque séquence possible, pour déterminer la plus vraisemblable. Plusieurs commandes sont possibles, en fonction du nombre de marqueurs présents dans les groupes de liaisons :

- Pour les groupes de liaisons comptant de 5 à 9 marqueurs (groupes de petite taille), la fonction "COMPARE" est directement utilisée. Le logiciel calcule le LODscore

associé à chaque ordre possible et nous donne alors les 20 séquences les plus probables. La première séquence sera retenue si la deuxième plus vraisemblable à un LODscore de deux ordres plus petits (10^2 fois moins vraisemblable).

- Pour les groupes de liaisons comptant plus de 10 marqueurs, il faut utiliser les fonctions "THREE POINT" puis "ORDER". Trois cas de figures sont alors possibles :

- Premier cas, le logiciel arrive seul à ordonner les marqueurs et donne la distance entre chaque paire de marqueurs du groupe.
- Deuxième cas, le logiciel souhaite plus d'informations, il y a donc nécessité de choisir les 5 marqueurs les plus informatifs dans le groupe, grâce aux commandes "LIST LOCI" puis "LOD TABLE", les ordonner avec la fonction "COMPARE", puis ajouter les autres marqueurs un par un avec la commande "TRY".
- Dernier cas, le groupe contient trop de marqueurs, il faut alors partir d'un sous-groupe de marqueurs puis rajouter les autres avec la fonction "TRY".

Une fois que tous les groupes de liaison sont ordonnés, la fonction "MAP" donne la distance entre les marqueurs des différentes séquences trouvées (uniquement après les fonctions "COMPARE" et "TRY").

Le logiciel MAPCHART 2.0 est alors utilisé pour dessiner la carte.

III) RESULTATS

A) DISTORSION DE SEGREGATION

Après analyse et vérification des données avec le logiciel GeneMapper 3.7, 636 marqueurs AFLPs ont été retenus. Les tests du χ^2 pour chaque marqueur révèlent que 71% des marqueurs (452) ségrègent selon les attendus Mendéliens (Figure 10). Ces marqueurs sont ensuite repartis en deux groupes : les marqueurs qui ségrègent chez la femelle et ceux qui ségrègent chez le mâle, pour qu'ensuite une carte génétique soit effectuée pour chacun des deux sexes. Ainsi, le nombre de marqueurs ségrégeant est de 341 chez la femelle et de 295 chez le mâle. L'analyse par test du χ^2 permet de déterminer que la distorsion de ségrégation est similaire chez les deux sexes : 71 % des marqueurs respectent les attendus Mendéliens (Figure 11). Seuls les AFLPs non distordus sont gardés pour l'établissement de la carte génétique, à savoir 209 marqueurs chez le mâle et 243 chez la femelle. De plus 6 individus sont exclus des analyses ultérieures car leurs profils électrophorétiques différents révèlent qu'ils proviennent d'une autre famille.

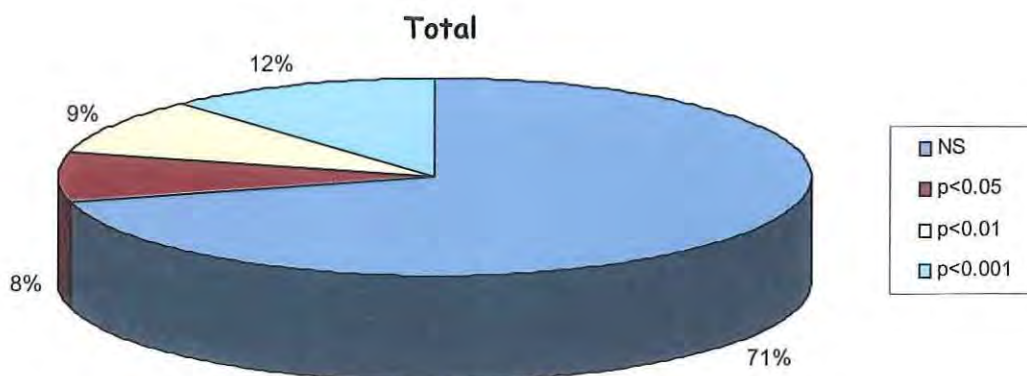


Figure 10 : Diagramme représentant la distorsion de ségrégation des AFLPs (n=636) chez *Mytilus edulis*, Famille 10.



Figure 11: Diagrammes représentant la distorsion de ségrégation des AFLPs (n=636) chez *Mytilus edulis* Famille 10 : comparaison entre femelle et mâle.

Il existe une grande hétérogénéité dans le nombre de marqueurs AFLPs selon la paire d'amorces utilisée (*Tableau 2*). En effet, le nombre total de marqueurs varie de 4 à 46 ; chez la femelle de 2 à 24 et chez le mâle de 2 à 22. De même, il existe une grande hétérogénéité dans le pourcentage de marqueurs distordus, que ce soit dans le jeu de données total (de 0 à 42.1%) ou sexe-spécifique (de 0 à 75% chez la femelle et de 0 à 100% chez le mâle).

Tableau 2 : Répartition du nombre de marqueurs AFLPs en fonction des paires et du sexe. Nombre et pourcentage de marqueurs gardés et distordus en fonction du sexe.

Marqueurs	Nombre total de marqueurs	Nombre chez la femelle	Nombre chez le mâle	Nombre gardé chez la femelle (% de distordus)	Nombre gardé chez le mâle (% de distordus)	% total de distordus
A1	24	15	9	13 (13%)	9 (0%)	8,30%
B1	20	15	5	13 (13%)	1 (80%)	30%
E1	27	12	15	9 (25%)	7 (53,33%)	40,74%
A3	21	12	9	8 (33,33%)	9 (0%)	19,04%
B3	16	10	6	7 (30%)	5 (16,66%)	25%
E3	19	13	6	11 (15,38%)	4 (33,33%)	21,05%
A4	29	12	17	6 (50%)	12 (29,41%)	37,93%
B4	20	11	9	8 (37,5%)	8 (12,5%)	20%
E4	24	14	10	8 (75%)	6 (66,6%)	41,60%
A5	8	3	5	2 (50%)	5 (0%)	12,50%
B5	11	7	4	7 (0%)	0 (100%)	36,36%
E5	6	4	2	3 (25%)	1 (50%)	30%
A6	19	5	14	3 (40%)	9 (55,55%)	36,84%
B6	14	6	8	4 (33,33%)	6 (25%)	28,57%
E6	4	2	2	2 (0%)	2 (0%)	0%
A7	46	24	22	16 (50%)	15(46,6%)	32,60%
B7	15	7	8	2 (71,43%)	7 (12,5%)	40%
E7	17	8	9	6 (33,33%)	7 (28,57%)	23,50%
A8	18	10	8	8 (20%)	5 (37,5%)	27,77%
B8	11	7	4	5 (28,57%)	4 (0%)	18,18%
E8	14	9	5	9 (0%)	3 (40%)	14,28%
A9	15	9	6	7 (22,22%)	4 (33,33%)	26,66%
B9	15	8	7	7 (12,5%)	4 (42,86%)	26,66%
E9	19	9	10	5 (44,4%)	7 (30%)	36,84%
A10	27	14	13	10 (28,6%)	10 (23,08%)	25,93%
B10	15	4	11	3 (25%)	7 (36,36%)	33,33%
E10	27	19	8	12 (26,32%)	7 (12,5%)	22,22%
A11	25	14	11	10 (28,57%)	6 (45,45%)	36%
B11	21	11	10	9 (18,18%)	9 (10%)	14,28%
E11	25	14	11	8 (42,85%)	8 (27,27%)	36%
A12	19	7	12	4 (42,85%)	7 (41,66%)	42,10%
B12	19	11	8	8 (27,27%)	7 (12,5%)	21,05%
E12	27	15	12	10 (33,33%)	7 (41,66%)	37,04%

B) CARTES GENETIQUES SEXE-SPECIFIQUE

1) CARTE DE LIAISON GENETIQUE PRELIMINAIRE FEMELLE

Tout d'abord, la fonction GROUP a déterminé 26 groupes de liaison : 11 groupes comprenant 5 à 9 marqueurs, 9 avec plus de 10 marqueurs, 2 tétramères, 2 trimères et 2 dimères. Sur les 243 AFLPs spécifiques de la femelle, seulement 12 ne sont liés à aucun autre marqueur, ce qui représente un pourcentage de 5% de non-assignement.

L'analyse a d'abord porté sur les groupes de liaison contenant plus de 10 marqueurs.

L'analyse du groupe LG1 contenant 34 marqueurs n'a pu être réalisée sur la totalité de ces marqueurs du fait de leur trop grand nombre. Une liste préliminaire de 21 AFLPs a permis d'ordonner 10 d'entre eux avec les fonctions THREE POINT et ORDER avec une longueur de 110cM. Puis, la fonction TRY a permis d'en cartographier 4 supplémentaires, aboutissant ainsi à une longueur "finale" de 116cM.

Les marqueurs des groupes LG2 à LG4 ont pu être ordonnés et cartographiés à l'aide des fonctions THREE POINT et ORDER. La taille de ces groupes varie de 110cM pour LG2 à 65cM pour LG4.

En revanche, 5 de ces 9 groupes n'ont pu être ordonnés par ces fonctions. En effet, pour l'un de ces groupes (LG5) 5 marqueurs conservatifs ont été sélectionnés puis ordonnés avec la fonction COMPARE. L'ajout des autres marqueurs du groupe avec la fonction TRY nous a permis d'obtenir un groupe d'environ 58cM (*Figure 12-a*).

L'analyse a ensuite porté sur les groupes contenant moins de 10 marqueurs.

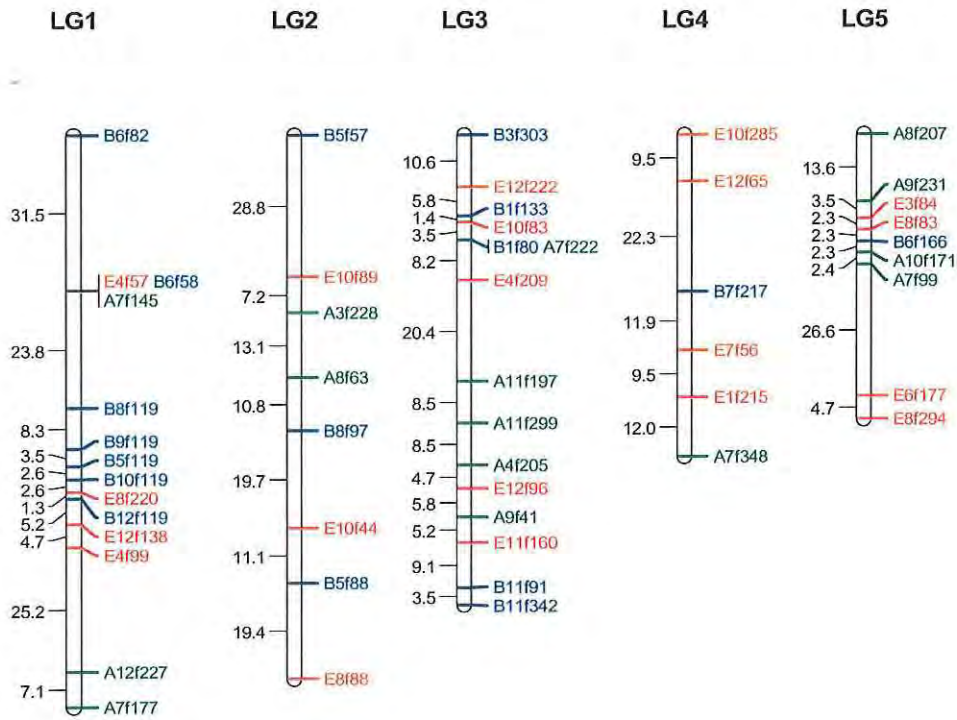
Les marqueurs des groupes LG6 à LG9 ont été ordonnés avec la fonction COMPARE puis cartographiés avec la fonction MAP. Les tailles ainsi obtenues varient de 85cM pour LG6 à 13cM pour LG8 et LG9.

Ces dernières fonctions n'ont pas permis d'aboutir à un ordre plus vraisemblable pour 7 de ces groupes. En effet, certains marqueurs ont une position fixe, alors que d'autres changent de position sans influencer sur la valeur du LOD score. D'où l'impossibilité de choisir l'ordre le plus vraisemblable (*Figure 12-b*).

Les marqueurs issus des 3 amorces marquées utilisées (Eco+CAG, Eco+ACG, Eco+ACT) ont été représentés sur la *Figure 12* par des couleurs différentes. Cette visualisation nous permet de constater que la répartition des AFLPs semble aléatoire sur chacun des groupes de liaison.

Enfin, quelques clusters (impliquant 2 ou 3 AFLPs) ont été mis en évidence. Par exemple, les marqueurs E4f57 et B6f58 sont localisés au même endroit sur LG1.

a)



b)

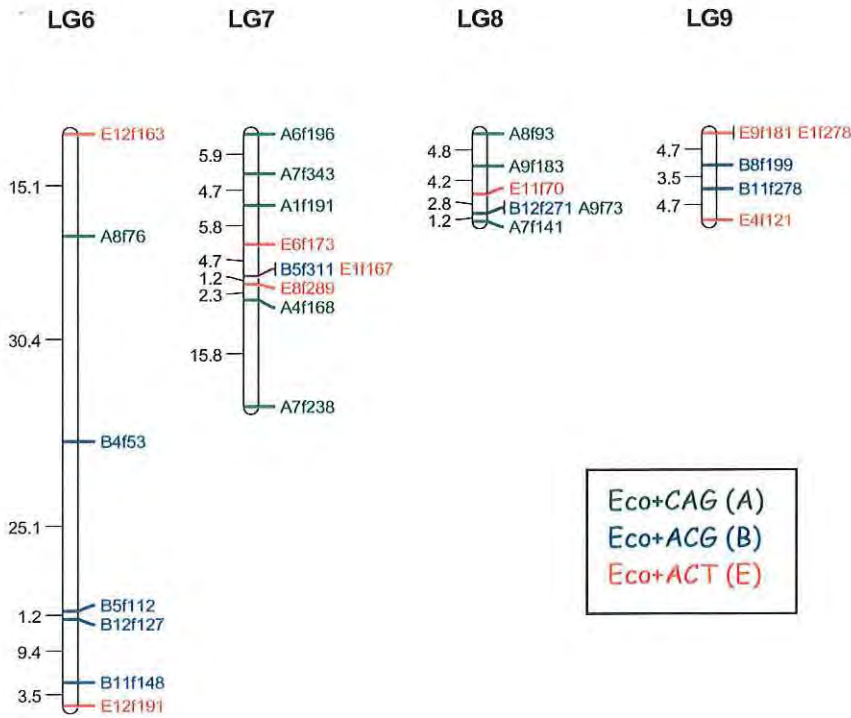


Figure 12 : Carte génétique préliminaire femelle.

a) Groupes avec plus de 10 marqueurs (THREE POINT, ORDER, TRY)

b) Groupes avec moins de 10 marqueurs (COMPARE et MAP)

Pour chacun des 9 groupes cartographiés, l'intervalle le plus long et l'intervalle moyen ont été déterminés. Ils sont respectivement compris entre 15.8 et 31.5cM; et entre 3.3 et 15.7cM (*Tableau 3*).

Tableau 3 : Nombre de marqueurs total (N total) et cartographié (N carte), taille, intervalle le plus long et moyen pour les groupes de liaison de la carte femelle.

Groupe	N total	N carte	Taille (cM)	Intervalle le plus long (cM)	Intervalle moyen (cM)
LG1	34	14	115.8	31.5	10.1
LG2	14	8	110.1	28.8	15.7
LG3	17	15	95.2	20.4	6.7
LG4	10	6	65.3	22.3	13.0
LG5	17	9	57.7	26.6	6.7
LG6	7	7	84.7	30.4	14.1
LG7	9	9	40.4	15.8	5.8
LG8	6	6	13.0	4.8	3.3
LG9	5	5	12.9	4.7	4.3

2) CARTE DE LIAISON GÉNÉTIQUE PRÉLIMINAIRE MÂLE

Tout d'abord, la fonction GROUP a déterminé 28 groupes de liaison : 17 groupes comprenant 5 à 9 marqueurs, 5 avec plus de 10 marqueurs, 3 tétramères, 1 trimère et 2 dimères. Sur les 209 AFLPs spécifiques du mâle, seulement 8 ne sont liés à aucun autre marqueur, ce qui représente un pourcentage de 3.8% de non-assignement.

L'analyse a d'abord porté sur les groupes de liaison contenant plus de 10 marqueurs.

Les marqueurs des groupes LG1 et LG2 ont pu être ordonnés et cartographiés à l'aide des fonctions THREE POINT et ORDER. La taille de ces groupes est respectivement de 73cM et de 31cM.

En revanche, 3 de ces 5 groupes n'ont pu être ordonnés par ces fonctions (*Figure 13-a*). Pour ces groupes, il serait nécessaire de sélectionner 5 marqueurs conservatifs comme base de départ pour ensuite rajouter les autres marqueurs un à un avec la fonction TRY.

L'analyse a ensuite porté sur les groupes contenant moins de 10 marqueurs.

Les marqueurs des groupes LG3 à LG11 ont été ordonnés avec la fonction COMPARE puis cartographiés avec la fonction MAP. Les tailles ainsi obtenues varient de 91cM pour LG3 à 20cM pour LG11.

Ces dernières fonctions n'ont pas permis d'aboutir à un ordre plus vraisemblable pour 11 de ces groupes (*Figure 13-b*).

Les marqueurs issus des 3 amorces marquées utilisées (Eco+CAG, Eco+ACG, Eco+ACT) ont été représentés sur la *Figure 13* par des couleurs différentes. Cette visualisation nous permet là encore de constater que la répartition des AFLPs semble aléatoire sur chacun des groupes de liaison, et de noter la présence de clusters.

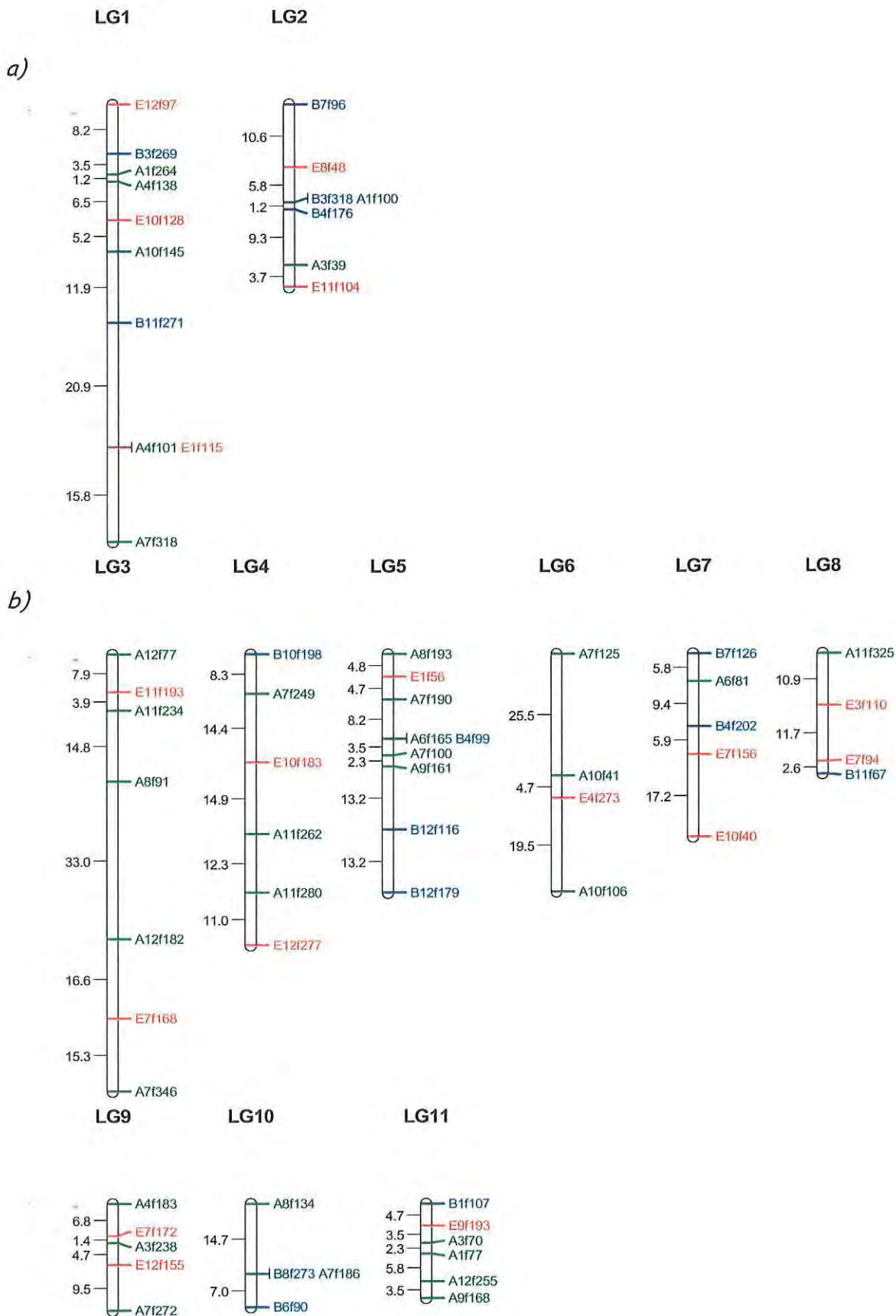


Figure 13 : Carte génétique préliminaire mâle.

a) Groupes avec plus de 10 marqueurs (THREE POINT, ORDER, TRY)

b) Groupes avec moins de 10 marqueurs (COMPARE et MAP)

Pour chacun des 11 groupes cartographiés, l'intervalle le plus long et l'intervalle moyen ont également été déterminés. Ils sont respectivement compris entre 15.8 et 31.5cM; et entre 3.3 et 15.7cM (*Tableau 4*).

Tableau 4 : Nombre de marqueurs total (N total) et cartographié (N carte), taille, intervalle le plus long et moyen pour les groupes de liaison de la carte mâle.

Groupe	N total	N carte	Taille (cM)	Intervalle le plus long (cM)	Intervalle moyen (cM)
LG1	13	10	73.2	20.9	9.2
LG2	10	8	30.7	10.6	6.1
LG3	7	7	91.1	33.0	15.3
LG4	6	6	60.9	14.9	12.2
LG5	9	9	49.9	13.2	7.1
LG6	4	4	49.7	25.5	16.6
LG7	5	5	38.3	17.2	9.6
LG8	4	4	25.2	11.7	8.4
LG9	5	5	22.4	9.5	5.6
LG10	4	4	21.7	14.7	10.9
LG11	6	6	19.8	5.8	4.0

IV) DISCUSSION

A) UTILISATION DES AFLPs

1) AVANTAGES ET INCONVENIENTS DES AFLPs

Les AFLPs sont des marqueurs dominants (Vos, Hogers *et al.* 1995), ce qui nous oblige à utiliser des types de croisements adaptés afin d'obtenir des jeux de données informatifs.

En effet, afin de connaître le génotype complet des parents et des F1, seuls les marqueurs présents chez l'un des parents et qui ségrègent en F1 ont été conservés dans notre étude. Ainsi notre jeu de données est pleinement informatif, puisque la présence du pic est associée à un génotype hétérozygote "Aa", alors que son absence se caractérise par le génotype "aa".

Malgré ces limitations d'utilisation des AFLPs, du fait de leur nature dominante, ces marqueurs moléculaires restent un outil de choix chez des espèces où peu de marqueurs codominants sont disponibles et dont le génome est très peu connu.

2) QUALITE DES AFLPs

Suivant les paires d'amorces utilisées, les profils électrophorétiques sont différents du point de vue de la qualité et du nombre d'AFLPs sélectionnés. En effet, seuls 4 marqueurs ont été retenus pour la paire E6, alors que pour la paire A7, 46 marqueurs ségrégeants ont été gardés.

Pourtant lors des cycles de PCR, des cycles de type touchdown (Hecker and Roux 1996) ont été utilisés, pour que, quelle que soit l'amorce utilisée, les cycles soient identiques et les réactions de PCR toujours spécifiques. Selon les paires, les pics sont plus ou moins rapprochés, ce qui compliquait l'analyse des données. Ceci explique en partie la différence dans le nombre de marqueurs retenus en fonction des paires (Tableau 2). Un problème technique passager n'est probablement pas à l'origine de ces différences de qualité. En effet, les paires A8, B8 et E8, ont été amplifiées une seconde fois pour vérification. Néanmoins, les profils électrophorétiques étaient identiques, et la paire B8, toujours difficilement analysable, avec seulement 11 marqueurs retenus.

Dans la majorité des cas, les paires Eco+ACT (E) présentaient un signal plus faible, résultant en une analyse et une vérification des données plus laborieuses. Ceci peut être relié à une qualité moindre des amorces marquées avec le fluorochrome NED. Ce même phénomène a été mis en évidence par (Li and Guo 2004).

B) DISTORSION DE SEGREGATION

La distorsion de ségrégation est une caractéristique courante chez les bivalves. La proportion de loci distordus observés dans cette étude, 29%, est similaire à celle rapportée dans d'autres études ((Li and Guo 2004) : 27% chez *Crassostrea gigas*). Ces distorsions de ségrégation pourraient être causées par des artefacts techniques lors du génotypage. En effet, une efficacité non totale dans la réaction de restriction enzymatique ou lors des amplifications sélectives, peut résulter dans un déficit de génotype homozygote "aa". Cependant des écarts dans les deux sens (déficits ou excès d'homozygotes) ont été observés dans notre étude. De ce fait, l'hypothèse d'artefacts techniques est peu probable. Les déficits en homozygotes observés pourraient être la résultante du fort fardeau génétique, fréquemment rapporté chez les bivalves marins (Launey and Hedgecock 2001).

C) CARTES GENETIQUES PRELIMINAIRES

Les cartes génétiques mâle et femelle obtenues sont préliminaires, du fait du manque de temps. En effet, la fonction GROUP du logiciel MAPMAKER 3.0 nous a donné 28 et 26 groupes de liaisons respectivement. Or les études de cytogénétiques ont démontré l'existence de 14 paires de chromosomes chez *Mytilus edulis* (Thiriot-Quievreux 1984). De ce fait nous aurions dû obtenir tout au plus 14 groupes de liaison.

Lors de cette étude, un LODscore de 3 a été choisi pour grouper les marqueurs afin d'être très conservatif; de ce fait on obtient beaucoup de groupes avec un petit nombre de marqueurs. On peut alors penser que, dans une deuxième approche, l'utilisation d'un LODscore plus faible, par exemple 2, permettra de fusionner certains de ces groupes, abaissant ainsi le nombre total de groupes de liaison.

Tous les marqueurs de chaque groupe n'ont pas été cartographiés, il faudrait donc, avec la fonction TRY rajouter ces marqueurs pour ainsi essayer de combler les intervalles de trop grande taille, et diminuer l'intervalle moyen entre chaque marqueur; tout comme la fusion de plusieurs groupes de petites tailles.

Trois cartes de liaison génétiques ont été établies chez les huîtres :

- une utilisant les AFLPs chez *C. gigas* (Li and Guo 2004)
- une utilisant les microsatellites chez *C. gigas* (Hubert and Hedgecock 2004)
- une utilisant les AFLPs et les Microsatellites et EST chez *Crassostrea virginica* (Yu and Guo 2003).

Dans ces différentes cartes, il existe également une grande hétérogénéité dans les intervalles trouvés entre chaque paire de marqueurs, en effet dans presque chaque groupe de liaison, la gamme d'intervalle s'étend de quelques cM à 20 voir 40cM. Dans notre étude, même préliminaire, les résultats ne sont pas en décalage avec les cartes déjà publiées, de ce point de vue là.

Des clusters ont pu être observés dans nos groupes de liaison, ce qui est quasiment systématique dans les cartes uniquement basées sur les AFLPs (Yu and Guo 2003; Li and Guo 2004).

D) VERS UNE CARTE MOYENNEE

L'ajout d'autres types de marqueurs moléculaires, comme les microsatellites, serait un atout, car ils permettraient de combiner les cartes mâle et femelle, pour obtenir ainsi une carte moyennée. En effet les microsatellites, qui sont des marqueurs codominants, pourraient être informatifs chez les deux parents simultanément, permettant donc de les cartographier chez les deux sexes. Les microsatellites servent de points d'ancrage puisqu'ils permettent ensuite de cartographier les uns par rapport aux autres des AFLPs présents chez des parents différents.

Chez *M. edulis*, seuls 7 microsatellites sont disponibles à l'heure actuelle (Presa, Pérez et al. 2002). Au cours de cette étude nous avons commencé l'optimisation de ces marqueurs, seulement 5 ont pu être amplifiés correctement sur gel d'agarose (*Mgμ1*, *Mgμ3*, *Mgμ4*, *Mgμ6* et *Mgμ7*). Afin de vérifier que les amplicons contiennent bien le microsatellite, nous avons séquencé les produits de PCR. Le séquençage de ces produits de PCR a confirmé l'amplification du locus attendu (*Figures 14 et 15*).

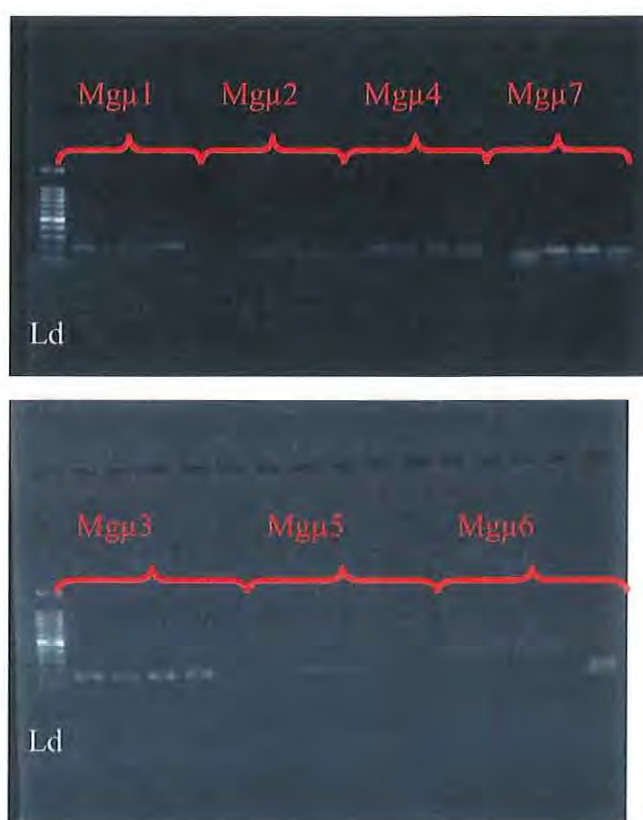


Figure 14 : *Electrophorèses sur gel d'agarose des microsatellites de moule amplifiés*

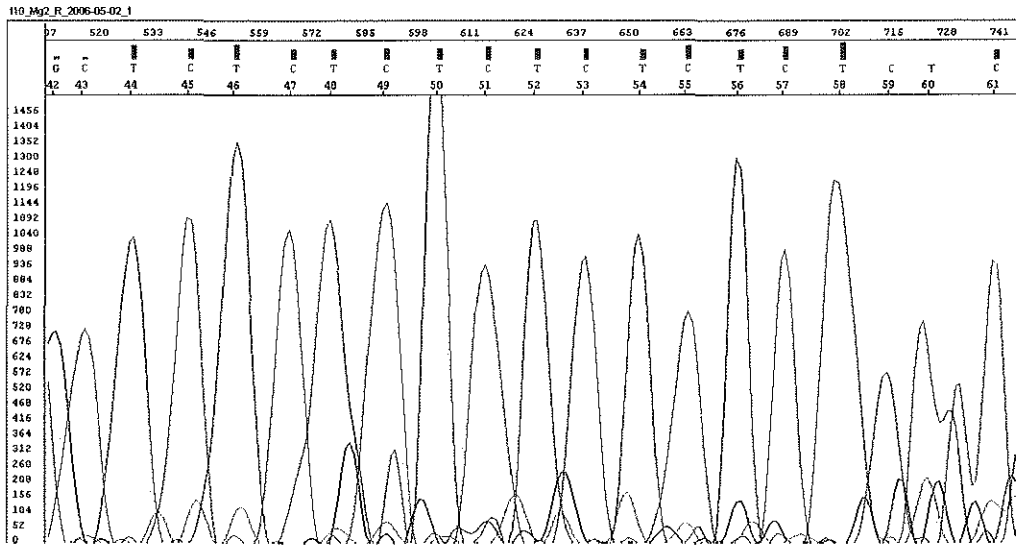


Figure 15 : *Electrophorégramme d'un microsatellite séquencé de moule (de motif CT n fois).*

CONCLUSION

Les marqueurs moléculaires du type AFLPs sont simple d'utilisation et leur protocole est facilement transférable d'une espèce à une autre. Ils peuvent en effet, être appliqués à de nombreuses espèces, car il n'est pas nécessaire de connaître le génome de celles-ci pour s'en servir. Cette caractéristique constitue un énorme atout pour les généticiens.

Une grande hétérogénéité entre chaque paire d'AFLPs a pu être observée, que ce soit au niveau de la qualité du profil électrophorétique, au niveau du nombre total de marqueurs ou encore du pourcentage de distorsion. Cependant le fort polymorphisme de la moule a permis d'obtenir un nombre moyen de marqueurs AFLPs d'environ 300 pour chacun des sexes. Le pourcentage de distorsion de ces marqueurs étant relativement faible, 29%, environ 200 marqueurs non distordus par sexe ont été retenus.

Le logiciel de cartographie utilisé, MAPMAKER 3.0, est adapté au schéma de croisement réalisé (analyse en F1). De plus le codage des données n'est pas très complexe.

Un certain nombre de fonctions peuvent être utilisées pour construire les groupes de liaison et ordonner les marqueurs entre eux. L'obtention des résultats est assez rapide.

En effet, des cartes sexe-spécifique préliminaires ont pu être établies, bien que par manque de temps les résultats n'aient pu être approfondis. Malgré tout, 9 groupes de liaisons ont pu être cartographiés pour la femelle, et 11 pour le mâle, ce qui constitue une bonne base de départ. Même si le nombre de groupes de liaisons est plus faible que 14 (nombre haploïde de chromosome chez la moule bleue), certains groupes cartographiés ont une longueur totale tout à fait satisfaisante. De plus ils possèdent un intervalle moyen, autour de 10 cM, comparable aux cartes publiées chez des espèces proches.

De ce point de vue, les cartes préliminaires obtenues se révèlent très encourageantes.

Malgré la courte durée du stage, qui ne m'a permis de ne faire qu'un travail préliminaire, cette étude m'a apportée beaucoup sur le plan professionnel. En effet, le travail en recherche est une question d'organisation perpétuelle, d'anticipation quant aux résultats qui peuvent ne pas être ceux attendus, de partage des appareils qui servent à tout le laboratoire. Le temps défile à une vitesse incroyable, et l'arrivée des dates butoirs effraie, quand le travail n'est pas tout à fait abouti.

Sur le plan relationnel, le travail en équipe, le soutien entre stagiaires et les autres moments passés en dehors du laboratoire apporte beaucoup à l'ambiance générale, et au moral.

Avec ce type de stage on se rend compte du travail à fournir en permanence et du poids de nos résultats. Le travail personnel n'est pas à proscrire quand on désire avancer. A la vue de tous ces points, ce genre de carrière ne me déplairait pas, bien au contraire.

BIBLIOGRAPHIE

PUBLICATIONS

- Hecker, K. H. and K. H. Roux (1996). "High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR." Biotechniques 20(3): 478-85.
- Hubert, S. and D. Hedgecock (2004). "Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*." Genetics 168(1): 351-62.
- Lander, E. S., P. Green, et al. (1987). "MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations." Genomics 1(2): 174-81.
- Launey, S. and D. Hedgecock (2001). "High genetic load in the Pacific oyster *Crassostrea gigas*." Genetics 159(1): 255-65.
- Li, L. and X. Guo (2004). "AFLP-based genetic linkage maps of the Pacific oyster *Crassostrea gigas* Thunberg." Mar Biotechnol (NY) 6(1): 26-36.
- Morton, N. E. (1955). "Sequential tests for the detection of linkage." Am J Hum Genet 7(3): 277-318.
- Presá, P., M. Pérez, et al. (2002). "Polymorphic microsatellite markers for blue mussels (*Mytilus* spp.)." Conservation genetics 3: 441-443.
- Thiriou-Quievreux, C. (1984). "Chromosome analysis of three species of *Mytilus* (Bivalvia : Mytilidae)." Marine Biology Letter 5: 265-273.
- Vos, P., R. Hogers, et al. (1995). "AFLP: a new technique for DNA fingerprinting." Nucleic Acids Research 23(21): 4407-14.
- Yu, Z. and X. Guo (2003). "Genetic linkage map of the eastern oyster *Crassostrea virginica* Gmelin." Biol Bull 204(3): 327-38.
- Gosling, E. (1992). "Genetics of *Mytilus*". In *Developments in aquaculture and fisheries science, The mussels Mytilus : ecology, physiology, genetics and culture*. Volume 25, chapitre 7, p309-381.

SITES INTERNET

Site de l'IFREMER : www.ifremer.fr

Site de Promega : www.promega.com

Site sur la mytiliculture : www.mytiliculture.com

Site canadien sur la mytiliculture : www.glf.dfo-mpo.gc.ca/pe-pe/es-se/mussel-moule/mussel-moule-f.html

AUTRES DOCUMENTS

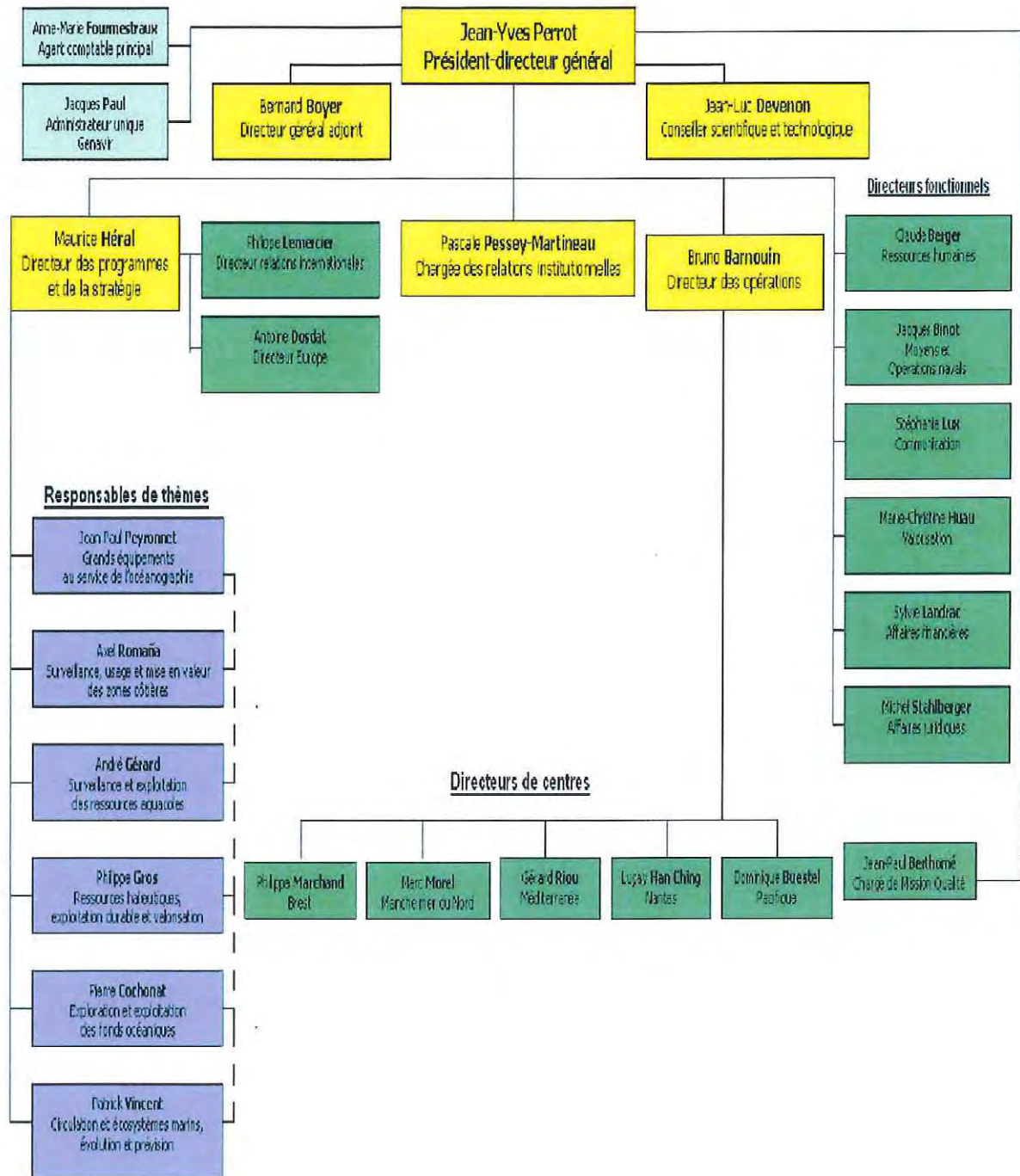
Thèse de Melle Sophie LAUNEY : "Marqueurs microsatellites chez l'huître plate *Ostrea edulis* L. : Caractérisation et applications à un programme de sélection pour une résistance au parasite *Bonamia ostreae* et à l'étude de populations naturelles".

Rapport de stage de Melle Delphine LALLIAS : "Analyse du succès reproducteur chez les populations naturelles d'huître plate *Ostrea edulis*"

Cours de Mme QUILLET, université de Lille : "Cartographie génétique des génomes eucaryotes".

ANNEXES

ANNEXE I : Organigramme de l'IFREMER



ANNEXE II : Protocole d'extraction d'ADN

Extraction ADN (Kit Promega - Wizard DNA Clean Up)

1. Lyse des cellules :

- Mettre dans un tube eppendorf de 2ml :
- 400 μ l de tampon d'extraction CTAB (20 ml TrisHCl 1 M pH 8.0 ; 16.36 g NaCl ; 8 mL - EDTA 0.5 M ; 4 g CTAB ; qsp 200 mL eau bidistillée).
- Fragment de branchie (4 à 5 mm²).
- 10 μ l de protéinase K à 10 mg/mL.
- Laisser au bain marie (65°C) pendant au moins 3 heures et vortexer régulièrement.

2. Lavage au Chloroforme

- Ajouter 400 μ l de Chloroforme.
- Mélanger par inversion pendant 2 minutes.
- Centrifuger à vitesse maximale pendant 5 minutes.
- Prélever le surnageant et ajouter 400 μ l de chloroforme et mélanger par inversion pendant 2 minutes.
- Répéter la centrifugation puis prélever le surnageant.

3. Extraction de l'ADN

- Homogénéiser le flacon contenant la résine.
- Ajouter 1mL de résine au surnageant puis mélanger le tout avec la pipette.
- Monter la micro-colonne sur un eppendorf de 2 mL et placer la seringue, sans le piston, sur la micro-colonne.
- Prélever le mélange résine-surnageant et placer dans la seringue. Placer le piston puis presser doucement et régulièrement.
- Retirer la seringue avec son piston puis enlever le piston.
- Injecter 2 mL d'isopropanol à 80% dans la seringue, placer le piston, puis injecter doucement.
- Retirer la seringue et la jeter.
- Centrifuger l'eppendorf surmonté de la micro-colonne pendant 2 minutes à vitesse maximale.
- Placer la micro-colonne sur un eppendorf neuf. Déposer 50 μ l de TE préalablement chauffé à 65°C.
- Centrifuger 20 secondes à vitesse maximale pour éluer l'ADN

