

Université des sciences de La Rochelle

Maîtrise de biologie des populations

mention environnement

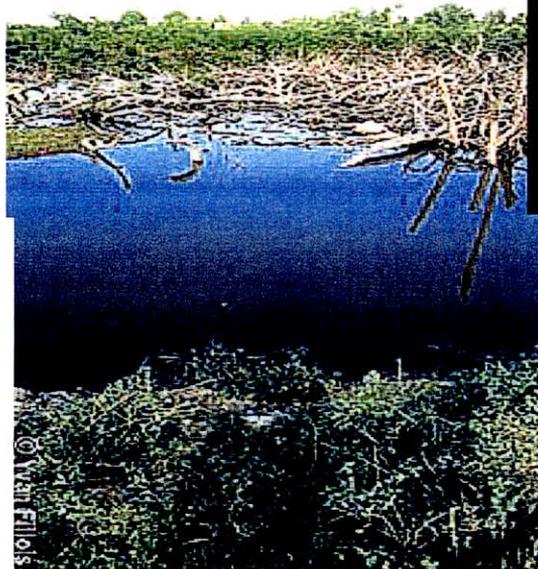


IFREMER
BIBLIOTHEQUE
LA TREMBLADE

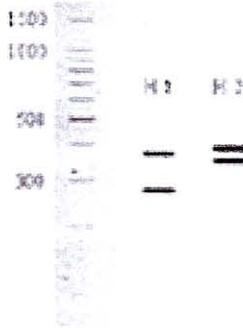
Rapport de stage par Dimitri Moreau

du 15 juin au 15 septembre 2001

Etude de génétique des populations sur l'huître creuse de mangrove, *Crassostrea gasar* (Adanson, 1757), par l'apport du marqueur ITS₂.



66530
E420
a2 MOR.E
042376



IFREMER
BIBLIOTHEQUE
LA TREMBLADE

Sous la direction de :

Pierre Boudry

IFREMER

Laboratoire de Génétique et de Pathologie

B.P.133

17390 La Tremblade

IFREMER Bibliothèque de la Tremblade



OLR 01376

Ifremer

SOMMAIRE

INTRODUCTION	2
MATERIEL ET METHODES	6
I/ Les populations étudiées	7
II/ Extraction de l'ADN	8
III/ Amplification de la séquence ITS₂	9
IV/ Révélation du polymorphisme	9
A. La RFLP	10
B. Le séquençage	10
V/ Analyse des données	11
RESULTATS	12
I/ Mise au point de l'amplification de la séquence ITS₂	13
II/ Détection du polymorphisme par la technique de RFLP	15
III/ Séquençage	22
A. Résultats du séquençage	22
B. Analyse de la séquence	23
1. Alignement des séquences	23
2. Alignement avec les bases de données	23
3. Analyse des sites de restriction	24
DISCUSSION CONCLUSION	26
BIBLIOGRAPHIE	30
ANNEXE 1 : Extraction d'ADN au phénol-chlorophorme	32
ANNEXE 2 : Protocole suivi pour la PCR	33
ANNEXE 3 : Protocole suivi pour la PCR- RFLP	34
ANNEXE 4 : Protocole suivi pour le séquençage	35

INTRODUCTION

Les huîtres sont des mollusques bivalves marins dont les nombreuses espèces ont colonisé la quasi totalité des milieux côtiers tempérés et tropicaux. Elles font l'objet, dans de nombreux cas, d'une culture intensive. Le genre *Crassostrea* regroupe plusieurs espèces d'huîtres creuses, il est le plus répandu et le mieux connu car il alimente un commerce important.

Les différentes espèces ont d'abord été décrites selon des critères morphologiques et de localisation géographique, ainsi que par des essais de croisements (Gaffney and Allen, 1994) et des études de caryotypes (Leitao *et al.*, 1999). Ces premières classifications ont pu être confirmées ou redéfinies grâce à l'apport de la biologie moléculaire et des marqueurs moléculaires (ex : *C. sikamea* Banks *et al.*, 1994). Les marqueurs sont des fragments d'ADN qui servent de repères pour suivre la transmission d'un segment de chromosome d'une génération à l'autre. Ils sont notamment utilisés en phylogénie moléculaire pour construire les arbres phylogénétiques (Ladron De Guevara *et al.*, 1996) ou, comme c'est le cas ici pour permettre d'observer l'évolution au sein d'une espèce et tenter de retracer les événements qui sont à l'origine de sa répartition actuelle. Grâce aux progrès récents de la biologie moléculaire, en particulier la découverte des séquences très variables (ITS, microsatellites, transposons) et la technique d'amplification des séquences d'ADN par PCR, de nombreux marqueurs génétiques sont maintenant disponibles (Klinbunga *et al.*, 2001). La fiabilité de ces méthodes permet d'écarter les incertitudes causées par des problèmes de plasticité morphologique et de répartition plus ou moins bien définie. Chez les huîtres, les marqueurs moléculaires, s'ils sont variables peuvent permettre d'observer une différenciation individuelle et donc de réaliser des analyses de parenté (Huvet *et al.*, in press), mais aussi d'obtenir des informations d'ordre historique sur l'évolution d'une espèce et la structuration des populations (Huvet *et al.*, 2000).

Ainsi dans le cas d'huîtres creuses peuplant les côtes de l'Amérique du sud et de l'Afrique, des études sur des marqueurs mitochondriaux 16S (Boutet, 1999) ont permis d'apporter des précisions sur la taxonomie et la phylogéographie des différentes espèces présentes. Ce travail a permis en particulier de différencier deux espèces sympatriques en Amérique du sud : *Crassostrea rizophorae* et *Crassostrea gasar* (Lapègue *et al.*, soumis).

C. gasar est une huître creuse vivant dans un milieu mésohalin, la mangrove. La mangrove est une formation végétale très particulière des zones côtières tropicales (cf : figure 1). Elle se développe sur des vases plus ou moins salées des zones balayées par les marées mais protégées des vagues. Ces formations sont importantes écologiquement et économiquement, par leur production primaire parmi les plus élevées du monde, par leur rôle protecteur vis-à-vis de la côte et peut-être surtout par le rôle qu'elles jouent dans la reproduction de

nombreuses espèces animales dont beaucoup sont exploitées comme des poissons et des crustacés (Ajana, 1979).

Décrite en 1757 sur les côtes Ouest africaines par Adanson, on sait donc aujourd'hui que *C. gasar* est également présente de l'autre côté de l'atlantique, en Amérique du sud. Cette espèce, parfaitement adaptée à ce milieu variable, semble pouvoir supporter d'importantes fluctuations de salinité (Gilles, 1991). En raison de ses excellentes qualités organoleptiques, cette huître est consommée, par les populations côtières et dans les grandes villes du Sénégal. Depuis de nombreuses années des essais d'aquaculture se sont développés, notamment à Joal et en Casamance. Mais, pour diverses raisons, dont la méconnaissance de la biologie de cette espèce, ces essais sont restés infructueux (Diadhiou, 1995).

La répartition de cette huître sur les côtes africaines et américaines a permis d'effectuer un échantillonnage sur plusieurs sites distincts. Cette répartition laisse supposer que les flux géniques peuvent être limités et pose la question de l'origine des populations. En effet, les transports marchands sont susceptibles de casser les barrières géographiques en permettant le déplacement « artificiel » d'individus sur de grandes distances (ex : *C. angulata*, Boudry *et al.*, 1998). De plus la nature ouverte de l'océan, la taille importante des populations et la phase larvaire planctonique sont autant de facteurs pouvant permettre des flux géniques importants entre populations.

Les premiers travaux effectués sur cette espèce avec le marqueur ITS₁ (InterTranscriptase sequence 1) ne permirent pas d'obtenir de résultats satisfaisants (Verdon, 2000). De plus, ne possédant aucune amorce pour des marqueurs moléculaires de type microsatellite chez cette espèce, une étude de génétique des populations a donc été réalisée avec un autre marqueur nucléaire variable : l'ITS₂ (InterTranscriptase sequence 2). Cette séquence sépare les gènes qui codent pour les ARN ribosomiques 5,8 S et 28 S. Cette séquence n'étant pas transcrite, elle n'est pas soumise aux pressions de sélection, ainsi les mutations ne sont pas délétères pour l'organisme et peuvent donc s'accumuler et être transmises à la descendance.

Ce fragment a été étudié grâce aux techniques « classiques » de la biologie moléculaire : amplification du fragment, à partir d'amorces mises au point chez d'autres espèces du genre *Crassostrea* (Xu *et al.*, 2001), révélation d'un polymorphisme de fragments de restriction (RFLP) et séquençage. La RFLP est une bonne technique pour l'étude de la divergence génétique entre les populations à large aire de répartition (Avisé *et al.*, 1987 ; Seutin *et al.*, 1993). Trois populations de chaque continent seront étudiées. Leur séparation géographique (cf : figure 1) laisse supposer une limitation des flux géniques naturels. Ce travail a donc pour but de rechercher des marqueurs nucléaires polymorphes susceptibles de mettre en évidence

la présence de barrières géographiques limitant les flux géniques. Le séquençage permettant d'affiner au maximum les informations recueillies pour chaque type observé devrait permettre de retracer l'évolution de la répartition de l'espèce.



Figure 1 : répartition mondiale des écosystèmes de type mangrove (Aubresil.net).

MATERIEL ET METHODES

La détection du polymorphisme entre les différentes populations d'huîtres *C. gasar* étudiées nécessite successivement plusieurs étapes. En premier lieu, l'ADN est extrait de chaque individu. On réalise ensuite une amplification du fragment ITS₂ qui sera le produit de base pour l'analyse du polymorphisme par les méthodes de RFLP et de séquençage.

I/ Les populations étudiées :

Pour cette étude, 6 populations provenant d'Afrique et d'Amérique ont été choisies parmi, un nombre important d'échantillons de diverse provenance. La répartition géographique de ces populations est représentée sur la figure 2.

Le Tableau 1 présente le lieu de prélèvement de chacune des populations et le nombre d'individus étudiés. Pour chaque individu un morceau de tissu, pouvant être de la branchie, du manteau ou du muscle a été prélevé et conservé dans l'éthanol absolue à température ambiante.

Tableau 1 : Liste des différentes populations étudiées, site de prélèvement et nombre d'individus disponibles.

Continent	Population	Lieu de prélèvement	Nombre d'individus
AFRIQUE	DAM	Nigéria - Delta du niger	29
	NOB	Sénégal- - Kafountine, Casamance	10
	ALM	Sénégal- Pointe des Almadies	22
AMERIQUE DU SUD	SIN	Guyane - Mont Sinery	44
	PAR	Brésil - Baie de paranagua	27
	CAN	Brésil - Cananéia	50

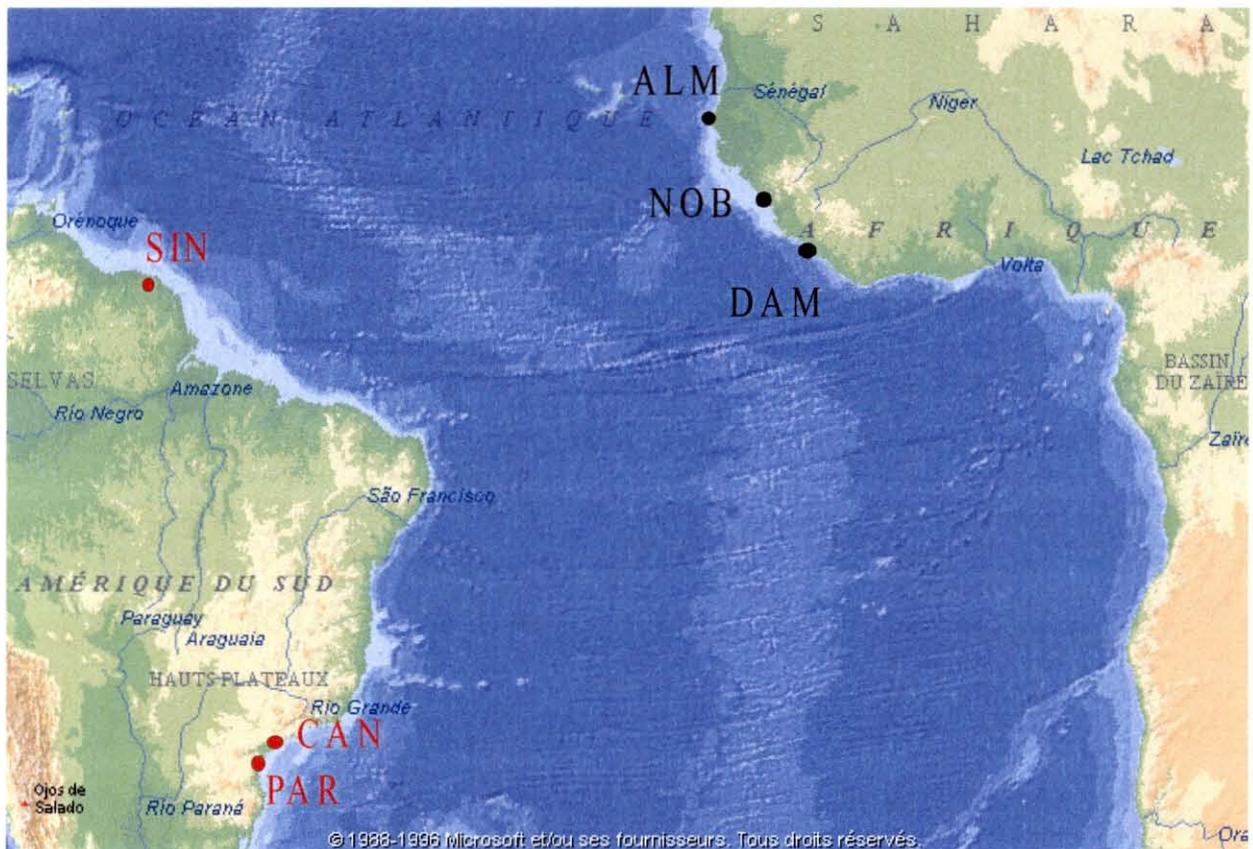


Figure 1 : Carte situant les différentes populations étudiées de chaque côté de l’océan Atlantique.

II/ Extraction de l’ADN :

L’ADN est extrait par la méthode dite du phénol-chloroforme comme décrit par Desmarais *et al.* (1995).

Trois étapes se succèdent (cf : protocole détaillé et composition des tampons en annexe 1).

- 1- Dissociation des tissus et destruction des membranes cellulaires par l’action d’un tampon dénaturant contenant : du SDS (Sodium Dodécil Sulfate) qui est un détergent ionique se fixant uniformément sur les protéines et les dénaturant en les chargeant négativement, de l’EDTA (éthylène diamine tétraacétique) qui est un agent chélateur complexant les ions Ca^{2+} indispensables à la cohésion cellulaire (composition du tampon en annexe 1) et par l’action de la protéinase K active à $55^{\circ}C$, qui dégrade les protéines par rupture des liaisons esters entre les acides aminés.
- 2- Extraction des acides nucléiques par adjonction successive de phénol aqueux, qui est un déprotéinisant, au tampon d’extraction, puis de chloroforme au surnageant aqueux contenant l’ADN pour éliminer les traces de phénol résiduelles.

3- Précipitation de l'ADN en présence d'éthanol absolu et stockage 2h à -20°C . Après centrifugation les culots d'ADN sont lavés avec de l'éthanol à 70° . Enfin les culots secs sont repris avec 75 à 100 μl de TE (Tris-EDTA) 1X.

III/ Amplification de la séquence ITS₂ :

Pour permettre l'étude du polymorphisme, il faut un grand nombre de copies du fragment d'ADN ITS₂ pour appliquer les différentes techniques et permettre une révélation des résultats par électrophorèse. Ce grand nombre de copies est obtenu grâce à la réaction de polymérisation en chaîne (Saiki R. K. *et al.*, 1985).

Cette technique permet l'amplification *in vitro* d'ADN spécifique. Une polymérase thermostable (à 95°C) de la bactérie *Thermus aquaticus*, la Taq polymérase est utilisée pour catalyser l'élongation simultanée de deux oligonucléotides (amorces) complémentaires aux régions bordant la séquence à amplifier. Lorsque la réplication du segment compris entre les deux amorces est terminée (un cycle), les deux duplex formés sont dénaturés par la chaleur pour produire des matrices monocaténaïres et un second cycle de réplication est effectué en abaissant la température en présence de tous les composants requis pour la polymérisation. La répétition des cycles de synthèse et de dénaturation entraîne un accroissement exponentiel du nombre de fragments (ici ITS₂). Ainsi de très petites quantités d'ADN sont requises comme matrice et une quantité suffisante d'ADN est amplifiée pour qu'une bande visible sur gel soit obtenue.

Les amorces utilisées ont été mises au point lors d'une étude sur les huîtres creuses *C. virginica* et *C. gigas* (Xu Z. *et al.*, 2001), et sont supposées utilisable pour les différentes espèces du genre. Les deux amorces nommées ITS₂ R et F se composent de 19 et 20 paires de bases (les séquences sont données en annexe 2). Le fragment ITS₂ amplifiée comprend environ 750 paires de bases (cf : séquence en annexe 5)

Le protocole et la séquence des amorces sont détaillés en annexe 2.

IV/ Révélation du polymorphisme :

Lors de cette étude deux étapes se sont succédées pour permettre la mise en évidence de polymorphisme génétique entre les différentes populations. Les deux techniques utilisées sont : la RFLP (Restriction Fragment Length Polymorphism) et le séquençage.

A. La RFLP :

La RFLP utilise les propriétés d'enzymes bactériennes dites de restrictions, synthétisées par ces organismes dans le but de casser l'ADN étranger pouvant pénétrer dans la cellule. Ces enzymes reconnaissent une séquence spécifique de 4 à 10 paires de bases et coupent la molécule d'ADN en un site très précis. Ainsi des différences entre les échantillons dans la séquence amplifiée peuvent être mises en évidence par reconnaissance ou non de sites de restriction par l'enzyme. Les différences des profils de restriction révélées par électrophorèse permettent donc de conclure à la présence d'un polymorphisme génétique. Les enzymes sont mises en présence du produit de PCR brut et incubées au bain marie à leur température optimale d'activité. Les résultats de la digestion sont révélés sur gel d'acrylamide. Les deux protocoles sont détaillés en annexe 3.

B. Le séquençage :

Le séquençage consiste à révéler directement la séquence des acides nucléiques (cf : annexe 4). Les produits de PCR purifiés (cf : annexe 4) sont analysés par la méthode de séquençage multiplex d'ADN double brin avec des primers marqués et l'enzyme sequiterm excell II (épicycle) préconisée pour l'utilisation du séquenceur disponible au laboratoire (Licor). Le principe de cette technique repose 1) sur le séquençage enzymatique décrit par Sanger en 1977, 2) sur la technique de PCR asymétrique :

1) Elongation à partir d'une amorce d'un ADN complémentaire à une matrice par une ADN polymérase ; terminaison aléatoire de la chaîne néosynthétisée par des didéoxyribonucléotides. Ici deux amorces (sens et antisens) encadrant le fragment à séquencer sont utilisées en même temps. L'une des amorces est marquée par un fluorochrome absorbant à 700 nm, l'autre à 800 nm. Le laser du séquenceur active simultanément les deux fluorochromes et permet ainsi la lecture de la séquence à partir des deux amorces soit dans les deux sens.

2) Une série de 30 cycles de dénaturation, hybridation des amorces et élongation permet d'amplifier le signal.

En pratique, 1 à 6µl de produit de PCR purifié (selon la concentration d'ADN évaluée par comparaison à un ladder 40 et 250 fmole) sont ajoutés à un mélange réactionnel contenant 1 picomole de chaque amorce, le tampon de l'enzyme et 5 unités d'Excell II DNA polymérase. Ce mélange est réparti dans 4 tubes contenant les didéoxyribonucléotides ddATP, ddGTP, ddCTP, ddTTP. Le mélange est incubé 2 minutes à 92°C suivi de 30 cycles de 30 secondes à 92°C, 30 secondes à 54°C et 1 minute à 70 °C. A la fin de la réaction, une solution d'arrêt est

ajoutée aux échantillons. Après une dénaturation de 3 minutes à 92°C, les échantillons sont déposés dans un gel d'acrylamide-urée (composition du gel cf :annexe 4) dénaturant à 5.5% pour la séparation des fragments nucléotide par nucléotide.

V/ Analyse des données :

Les données compilées sur tableur sont analysées avec le logiciel Statgraph. Plusieurs tests de comparaisons non paramétriques de Kruskal-wallis comparer les proportions des différents types observés ont été réalisés.

Les séquences obtenues sont comparées par le programme d'alignement multiple *Clustal W* disponible sur Internet (site www.infobiogen) et grâce au logiciel DNA man .

Les cartes de restriction ont été réalisées avec le logiciel d'analyse nucléique du site infobiogen.

Les séquences ont été comparées avec la base de donné genbank grâce au logiciel Blast disponible sur internet.

RESULTATS

I/ Mise au point de l'amplification de la séquence ITS₂ :

L'utilisation des amorces ITS₂ « Forward » et « Reverse » avec le protocole de PCR détaillé plus avant donne de bons résultats et une amplification quasi constante des ADN_s extraits utilisés à faible concentration. Cependant, il est vite apparu, lors d'essais de RFLP et d'électrophorèse des produits de PCR sur gel d'acrylamide, que des bandes parasites, plus ou moins nettes, résultantes d'hybridations aspécifiques des amorces apparaissaient. Ce problème complique la lecture des profils de restriction et surtout fausse le séquençage et donc entraîne des erreurs d'interprétation des résultats. Pour résoudre cela, la première démarche fut d'augmenter la stringence (donc la spécificité) du protocole de la PCR. On a donc tout d'abord augmenté la température d'hybridation des amorces (jusqu'à 10°C de plus que pour le protocole initial) et on a réalisé une gamme de concentration en MgCl₂ (la diminution de la concentration en MgCl₂ augmentant la stringence). Ces premiers essais de réglage ne permirent pas d'éliminer les bandes aspécifiques observées et de plus diminuaient la qualité de la RFLP par manque de copies de la séquence ITS₂ (séquence qui sépare les gènes qui codent pour les ARN ribosomiques 5,8 S et 28 S) dans le produit de PCR.

Une nouvelle mise au point fut alors effectuée en recherchant la température d'hybridation la mieux adaptée et en modifiant les concentrations d'ADN matrice. Pour cela nous avons réalisé une gamme combinant la température d'hybridation, qui s'étale de 44 à 66°C avec un intervalle de 2°C et la dilution de l'ADN extrait, qui va de 1 à 50% (cf : figures 3 à 5). Cette gamme a donc permis de déterminer la température optimale d'hybridation chez *C. gasar*, qui fut établie à 54°C. En revanche pour la concentration de l'ADN, il est apparu que la trop grande disparité de la qualité des extractions ne permettait pas de fixer une concentration idéale pour tous les échantillons. Il a donc fallu adapter la concentration d'ADN pour chaque individu, lors de la PCR. En effet les différents tissus ne sont pas tous de la même qualité, la conservation étant de plus ou moins bonne qualité à cause de la quantité des tissus dans l'épendorf et de problème d'évaporation de l'alcool. Malgré tout il est difficile de prévoir la qualité de l'extraction d'ADN et de l'amplification à la vue du tissu.

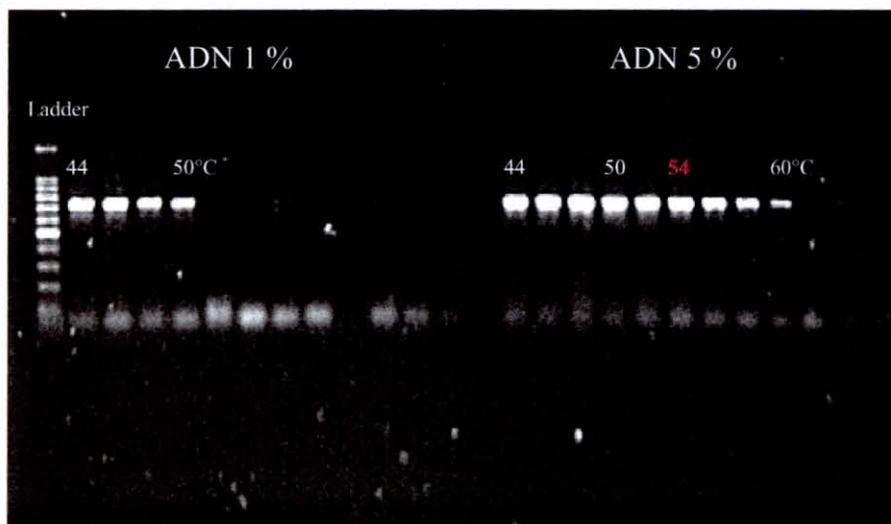


Figure 3 : Gamme de température pour la mise au point de la PCR avec ici des ADN matrices diluées à 1 et 5 %.

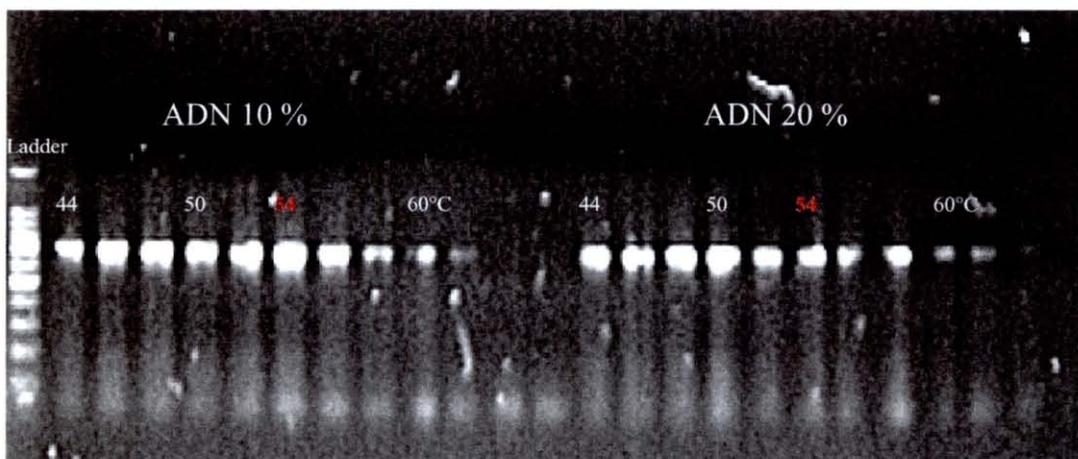


Figure 4 : Gamme de température pour la mise au point de la PCR avec des ADN matrices diluées à 10 et 20 %.

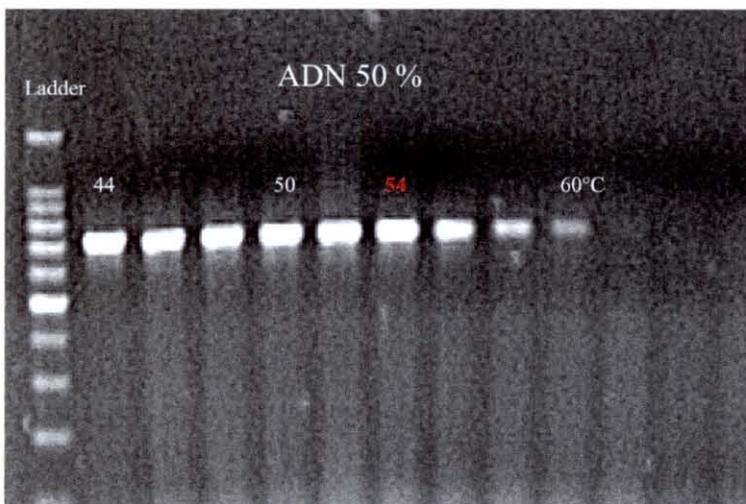


Figure 5 : Gamme de température pour la mise au point de la PCR avec des ADN matrices diluées à 50 %.

II/ Détection du polymorphisme par la technique de RFLP :

Dans cette étude, l'utilisation des d'amorces (ITS₂ R et F) a permis d'amplifier le fragment d'ADN ITS₂. Ce fragment s'est révélé être un bon marqueur moléculaire pour mettre en évidence le polymorphisme entre les différentes populations de *C. gasar* étudiées.

La première étape de l'étude du polymorphisme fut de tester plusieurs enzymes de restrictions en RFLP. Les différents types d'enzymes et les résultats obtenus sont regroupés dans le tableau 2.

Tableau 2 : Différentes enzymes testées et résultats obtenus (protocole RFLP pour chaque enzyme cf : annexe 3) :

Enzymes testées	Résultats
Hha I	Polymorphisme, 2 profils
Taq I	Polymorphisme, 2 profils
Mse I	monomorphe
Xho I	monomorphe
Bgl I	monomorphe (pas de site de restriction)
Acc I	monomorphe (pas de site de restriction)
Bsp 143 I	monomorphe

L'enzyme Hha I a permis de mettre en évidence des différences dans les profils de restriction entre les individus des populations africaines. En effet deux profils apparaissent : Un, H 1, avec deux bandes, assez éloignées l'une de l'autre, une en dessous et une au dessus de 300 paires de base et un autre, H 2, avec deux bandes de tailles très proches (cf : figures 7 à 11). On remarque de plus que dans ces populations africaines le profil H 1 semble à première vu plus présent que le deuxième. En revanche pour les populations américaines ce polymorphisme n'est pas observé, seul le profil H 2 est présent (cf : figures 12 à 16).

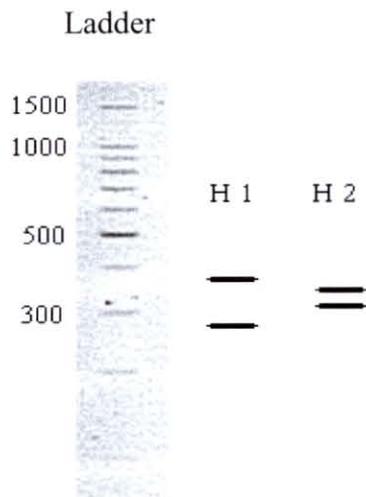


Figure 7 : Schéma représentatif des deux profils.

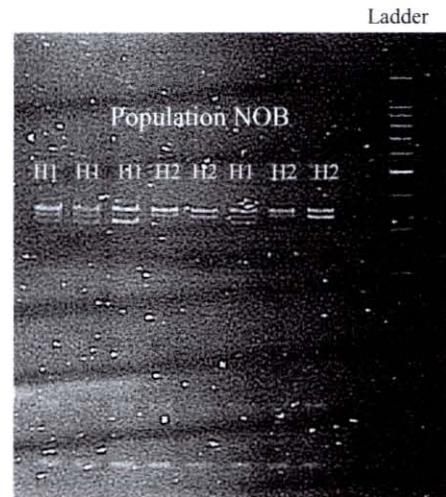


Figure 8 : Résultats de la digestion des produits de PCR des individus NOB par l'enzyme Hha I sur gel d'acrylamide à 8%. On voit bien apparaître ici les deux profils différents décrits plus avant.

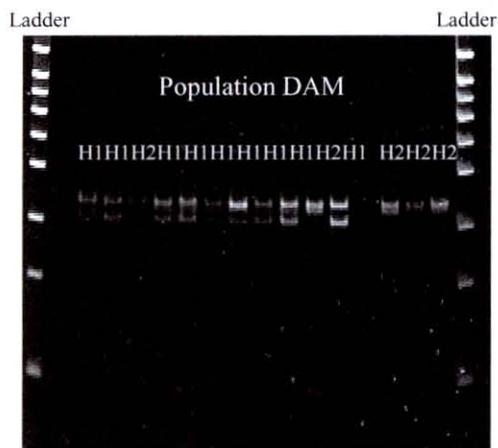
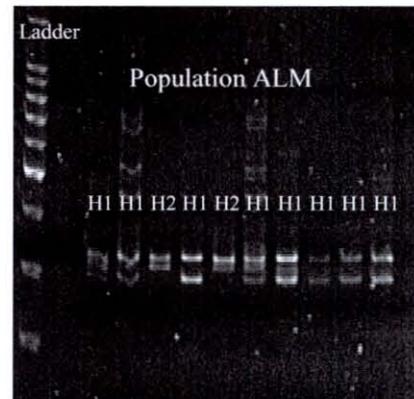
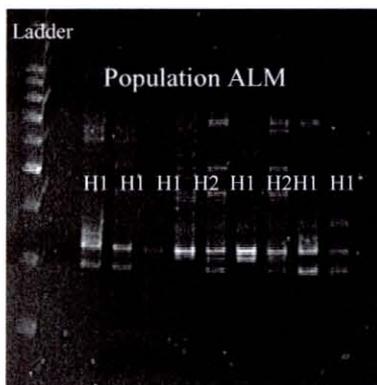
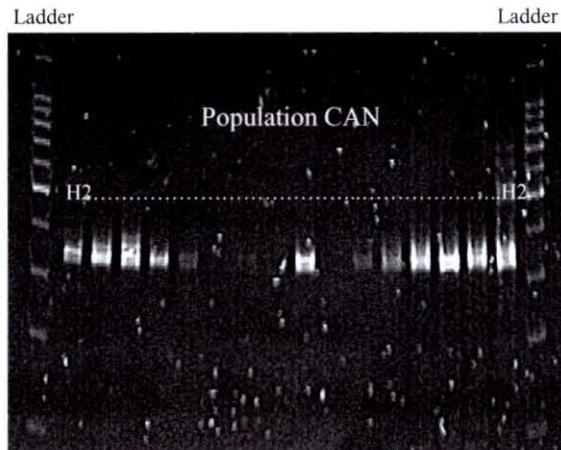
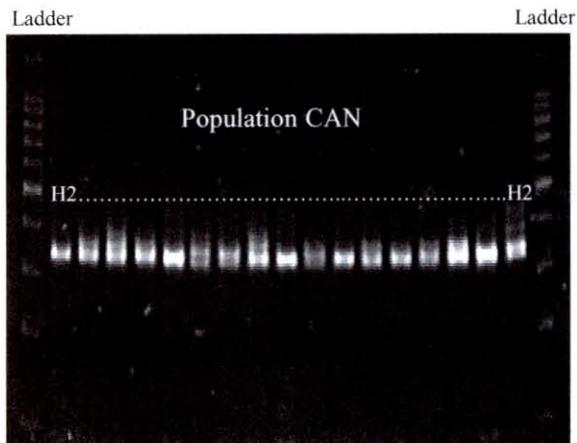


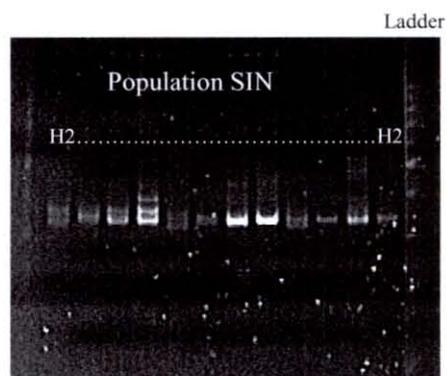
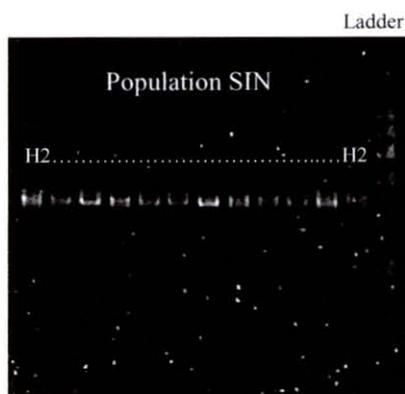
Figure 9 : Résultats de la digestion des produits de PCR des individus DAM par l'enzyme Hha I sur gel d'acrylamide à 8%. On voit bien apparaître ici encore les deux profils différents décrits plus avant (H 1, H 2).



Figures 10 et 11 : Résultats de la digestion des produits de PCR des individus ALM par l'enzyme Hha I sur gel d'acrylamide à 8%. On voit bien apparaître ici aussi les deux profils différents décrits plus avant (H 1, H 2).



Figures 12 et 13 : Résultats de la digestion des produits de PCR des individus CAN par l'enzyme Hha I sur gel d'acrylamide à 8%. Dans ce cas seul le profil H 2 apparaît.



Figures 14 et 15 : Résultats de la digestion des produits de PCR des individus SIN par l'enzyme Hha I sur gel d'acrylamide à 8%. Dans ce cas aussi seul le profil H 2 apparaît.

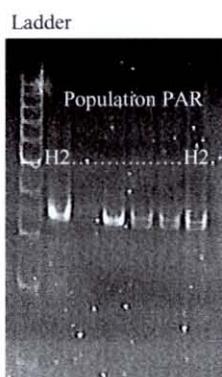


Figure 16 : Résultats de la digestion des produits de PCR des individus PAR par l'enzyme Hha I sur gel d'acrylamide à 8%. Dans ce cas encore seul le profil H 2 apparaît.

L'enzyme Taq I met en évidence une différence génétique entre les populations africaines et américaines. En effet deux profils différents sont obtenus par RFLP : T 1 et T 2 (cf : figure 17 représentant les deux profils). Chaque profil semble inféodé à un continent le profil T 1 pour les populations africaines (cf : figures 18 à 22) et le profil T 2 pour les populations américaines (cf : figures 23 à 27). Cependant le T 2 semble apparaître aussi chez 4 individus de la population ALM qui se situe en Afrique (cf : figure 22). Les gels d'acrylamides sont difficiles à lire en raison du nombre important de fragments de restriction et de la petite taille des fragments polymorphes.

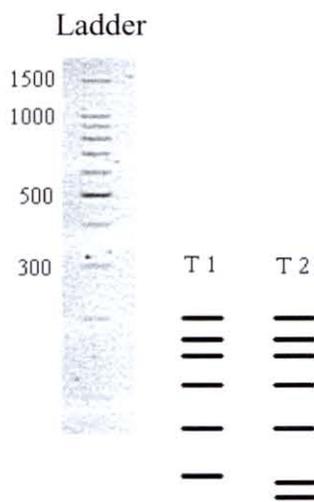


Figure 17 : Schéma représentant les profils T 1 et T 2.

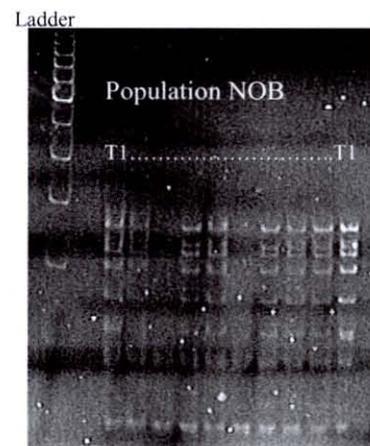
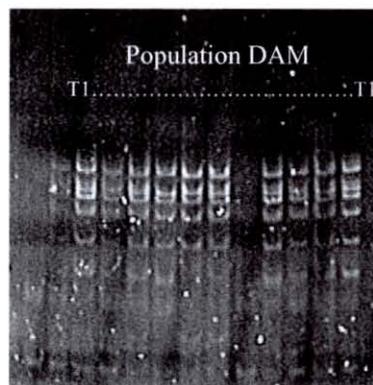
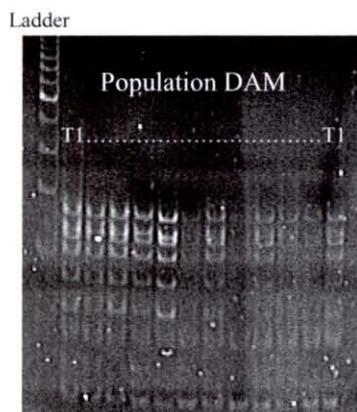
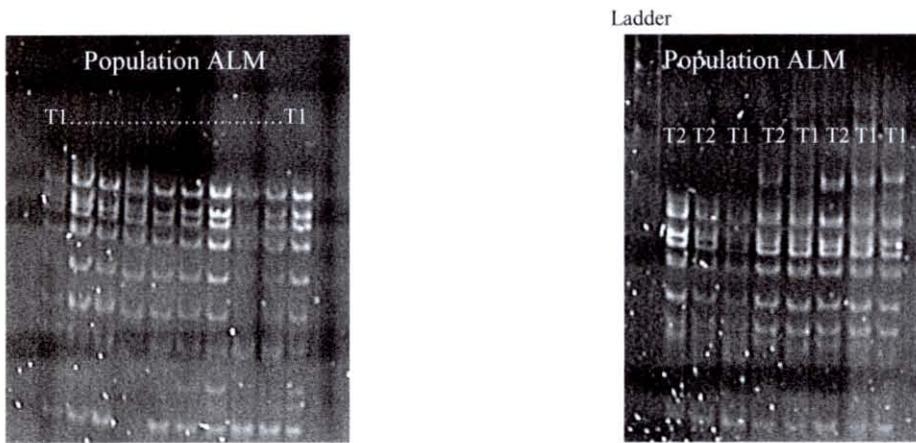


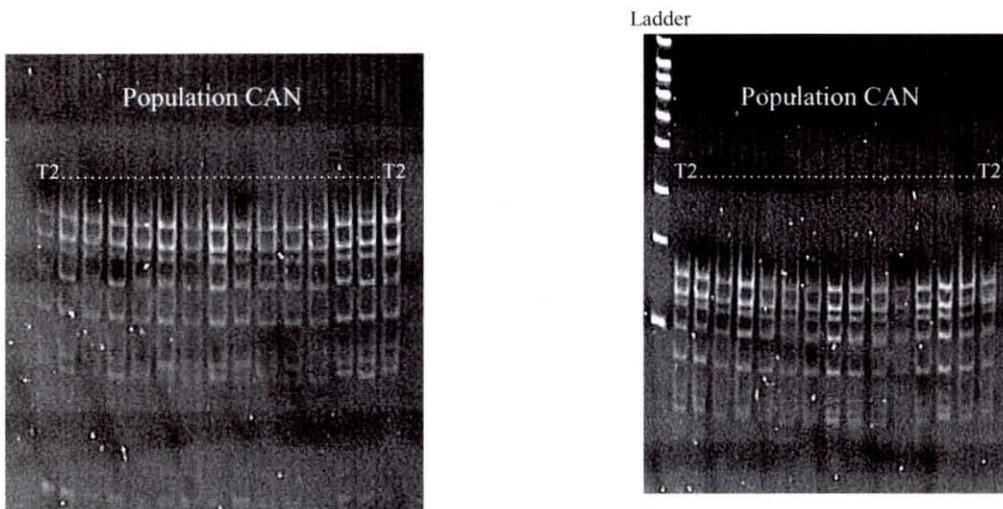
Figure 18 : Résultats de la digestion des produits de PCR des individus NOB par l'enzyme Taq I sur gel d'acrylamide à 8%. Dans ce cas seul le profil T 1 apparaît.



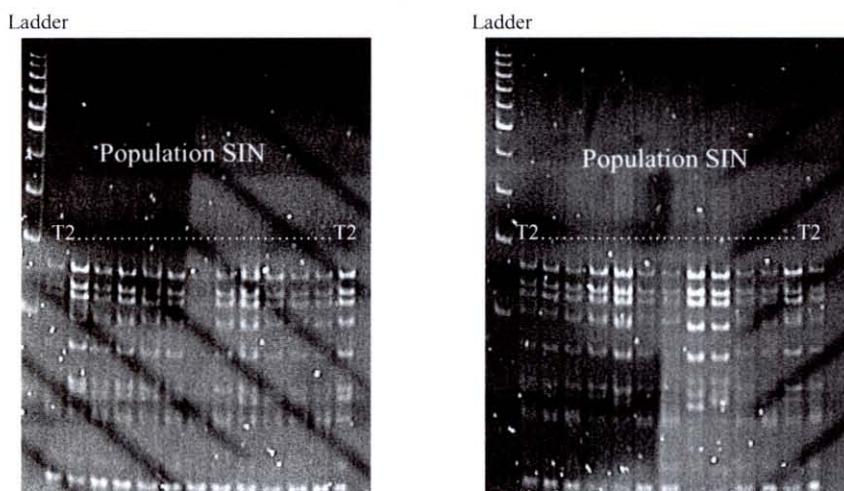
Figures 19 et 20 : Résultats de la digestion des produits de PCR des individus DAM par l'enzyme Taq I sur gel d'acrylamide à 8%. Dans ce cas aussi seul le profil T 1 apparaît.



Figures 21 et 22 : Résultats de la digestion des produits de PCR des individus ALM par l'enzyme Taq I sur gel d'acrylamide à 8%. Dans ce cas deux profils apparaissent : T 1 et T 2.



Figures 23 et 24 : Résultats de la digestion des produits de PCR des individus CAN par l'enzyme Taq I sur gel d'acrylamide à 8%. Dans ce cas seul le profil T 2 apparaît.



Figures 25 et 26 : Résultats de la digestion des produits de PCR des individus SIN par l'enzyme Taq I sur gel d'acrylamide à 8%. Dans ce cas aussi seul le profil T 2 apparaît.

Ladder

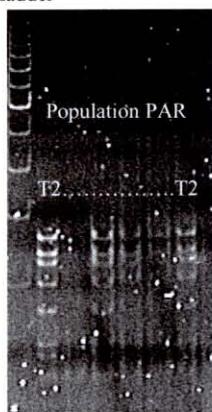


Figure 27 : Résultats de la digestion des produits de PCR des individus SIN par l'enzyme Taq I sur gel d'acrylamide à 8%. Dans ce cas aussi seul le profil T 2 apparaît.

Ces différentes observations ont fait l'objet d'une analyse statistique. Les fréquences de chaque profil sont répertoriées dans le tableau : 3, et celles des différents génotypes observés sont répertoriés dans le tableau : 4. La figure 28 représente sous forme d'un histogramme la répartition des profils H 1 et H 2 révélés par l'enzyme Hha I.

Tableau 3 : Fréquence des différents profils observés pour chaque population.

Continents	Populations	Effectif	H 1	H 2	T 1	T 2
Afrique	NOB	8	0.500	0.500	1.000	0.000
	DAM	14	0.786	0.214	1.000	0.000
	ALM	19	0.789	0.211	0.789	0.211
Amérique	CAN	29	0.000	1.000	0.000	1.000
	SIN	24	0.000	1.000	0.000	1.000
	PAR	5	0.000	1.000	0.000	1.000

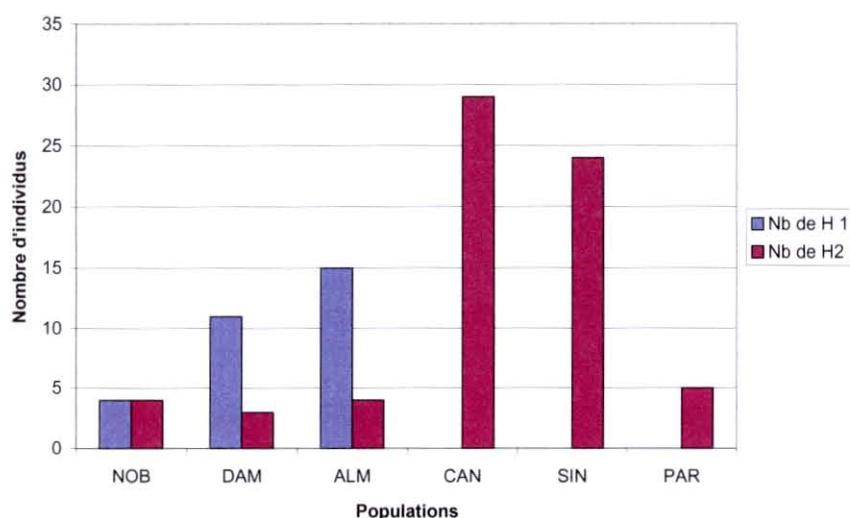


Figure 28 : Histogramme regroupant les résultats des RFLP avec l'enzyme Hha I. Ce schéma indique le nombre de chacun des profils pour chaque population.

Tableau 4 : Fréquences des quatre génotypes observés.

Continents	Populations	H 1/T 1	H 1/T 2	H 2/T 1	H 2/T 2
Afrique	NOB	0.500	0.000	0.500	0.000
	DAM	0.786	0.000	0.214	0.000
	ALM	0.632	0.105	0.158	0.105
Amérique	CAN	0.000	0.000	0.000	1.000
	SIN	0.000	0.000	0.000	1.000
	PAR	0.000	0.000	0.000	1.000

Deux tests de Kruskal-Wallis sont donc effectués. Le premier démontre une différence significative sur la présence des profils en fonction des continents ($p < 0.01$) (cf : tableau 5), le deuxième concernant le continent africain (seul possédant deux profils distincts) montre qu'il n'y a pas de différence significative entre les populations pour le facteur profil ($p = 0,265$) (cf : tableau 6). Aucun test n'a pu être effectué spécifiquement sur le continent américain, celui-ci ne présentant qu'un seul profil.

Tableau 5 : Test non paramétrique de Kruskal-Wallis sur le facteur profil en fonction des continents.

Kruskal-Wallis Test for morphe by continent

continent	Sample Size	Average Rank
Afrique	41	28,7805
Amerique	58	65,0

Test statistic = 60,2757 P-Value = 0,0

Tableau 6 : Test non paramétrique de Kruskal-Wallis sur le facteur profil en fonction des populations du continent africain.

Kruskal-Wallis Test for morphe by population

population	Sample Size	Average Rank
ALM	19	19,8158
DAM	14	19,8929
NOB	8	25,75

Test statistic = 2,65254 P-Value = 0,265466

III/ Séquençage :

A. Résultats du séquençage :

Les différentes tentatives de séquençage ont été réalisées avec le séquenceur automatique Licor. Le séquençage de huit individus de 5 populations (CAN 1 et 3, NOB 10 et 11, DAM 19 et 21, PAR 3 et SIN 17) a été réalisé en testant trois concentrations d'ADN (1, 3 et 6µl de produit de PCR). Ces quantités ont été choisies aux vues de la concentration en ADN assez faible des produits de PCR. Les séquences des deux brins réverses sont obtenues grâce aux deux lasers sensibles aux longueurs d'ondes 700 et 800 nm excitant les deux fluorochromes complexés aux amorces sens et antisens. Malgré un réel manque de temps pour la mise au point du protocole, il a été possible d'obtenir une partie de la séquence ITS₂ pour deux individus appartenants aux populations CAN et DAM. En effet nous avons pu lire environ 600 pb pour CAN 1 et 450 pb pour DAM 21. Pour les autres individus seulement 100 à 300 pb sont lues avec trop d'incertitudes et ne permettent pas d'effectuer une bonne analyse. Il semble donc qu'il y ait de grosses disparités entre les échantillons en raison de la qualité des produits de PCR. Il faut donc envisager une meilleure détermination de la concentration en ADN des produits amplifiés pour ajuster la quantité de ces produits à utiliser pour la réaction de séquence. Il faudrait de plus tenter d'obtenir des fragments plus uniformes grâce à une

meilleure spécificité des amorces. La figure 29 représente une partie du gel de séquence révélé par le laser émettant à 700 nm.

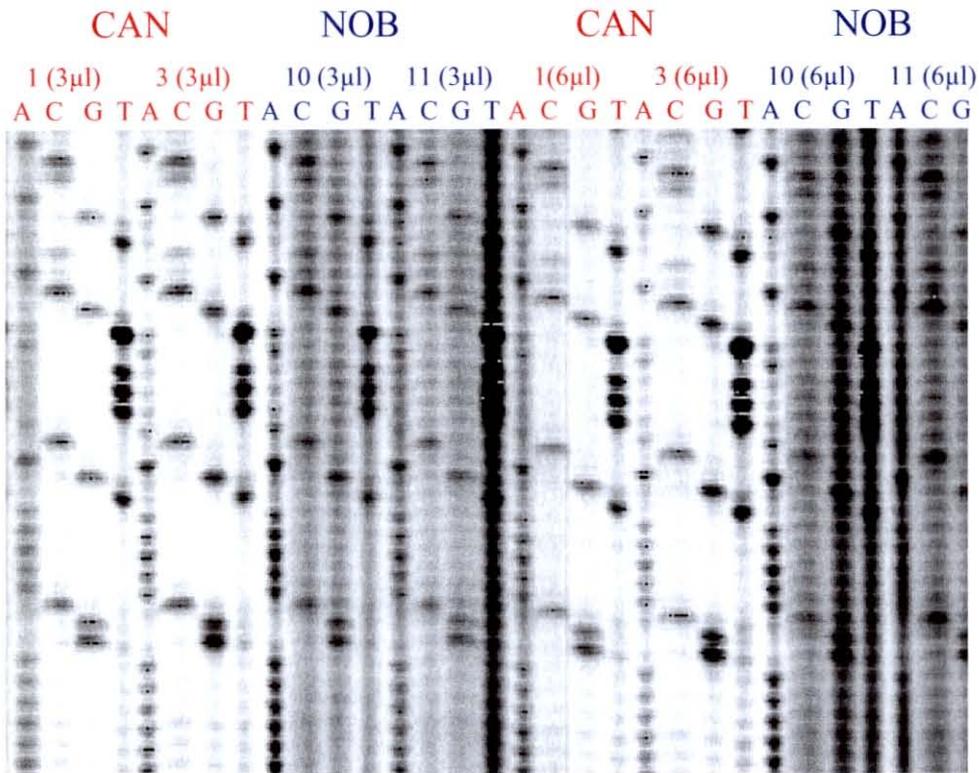


Figure 29 : Extrait de la photo du gel de séquence, obtenue avec le laser 700 nm. Sont indiqués au-dessus : le nom des populations, le numéro de l'échantillon (la quantité d'ADN placée dans le mix), les quatre déoxynucléotides. On peut voir sur cette photo que le gel est plus foncé pour les individus NOB ce qui a rendu impossible la lecture du gel.

B. Analyse de la séquence :

1. Alignement des séquences :

Nous avons pu aligner les séquences CAN 1 et DAM 21 sur environ 400 pb avec le logiciel clustlaw. Nous obtenons 97% d'homologie (cf : figure 30). Un alignement avec la séquence de *Crassostrea gigas* nous donne 56% d'homologie. Seules 400 bases ont été retenues car le reste présentait trop d'ambiguïtés.

2. Alignement avec les bases de données :

Le logiciel blast (Altschul, 1990) comparant la séquence soumise à la base de données genbank met en évidence un recouvrement avec la séquence ITS₂ de *C. gigas*.

3. Analyse des sites de restriction :

Le logiciel d'analyse nucléique du site d'infobiogen nous a permis d'obtenir la liste des enzymes de restriction les plus courantes coupant la séquence de 400 paires de bases. Il a ainsi été possible de sélectionner les enzymes divergents entre les deux séquences. Une carte de restriction a donc été réalisée avec ces enzymes (cf : figure 30). Celle-ci nous permet de voir que plusieurs enzymes susceptibles de révéler du polymorphisme entre les deux continents peuvent encore être testées pour compléter l'étude.

- Neufs enzymes ont ainsi pu être sélectionnées :
- 4 coupant sur la séquence africaine, Bsl I, Mbo II, Taq I et Bsa XI.
 - 5 coupant sur la séquence américaine, Bcg I, Bsa II, Sty I, Acl I et Tsc I.

Toutes les enzymes reconnaissent des sites de coupure présentant des différences entre les deux séquences. Ainsi elles ne coupent qu'un seul type et permettent la mise en évidence du polymorphisme génétique.

Un site polymorphe pour Taq I a bien été trouvé, sur une région présentant des délétions, ce qui confirme les résultats obtenus par RFLP. En revanche sur cette portion du fragment aucun site polymorphe pour Hha I n'a été mis en évidence.

Upper line : DAM21, from 1 to 421
 Lower line : CAN1, from 1 to 430
 identity= 97%

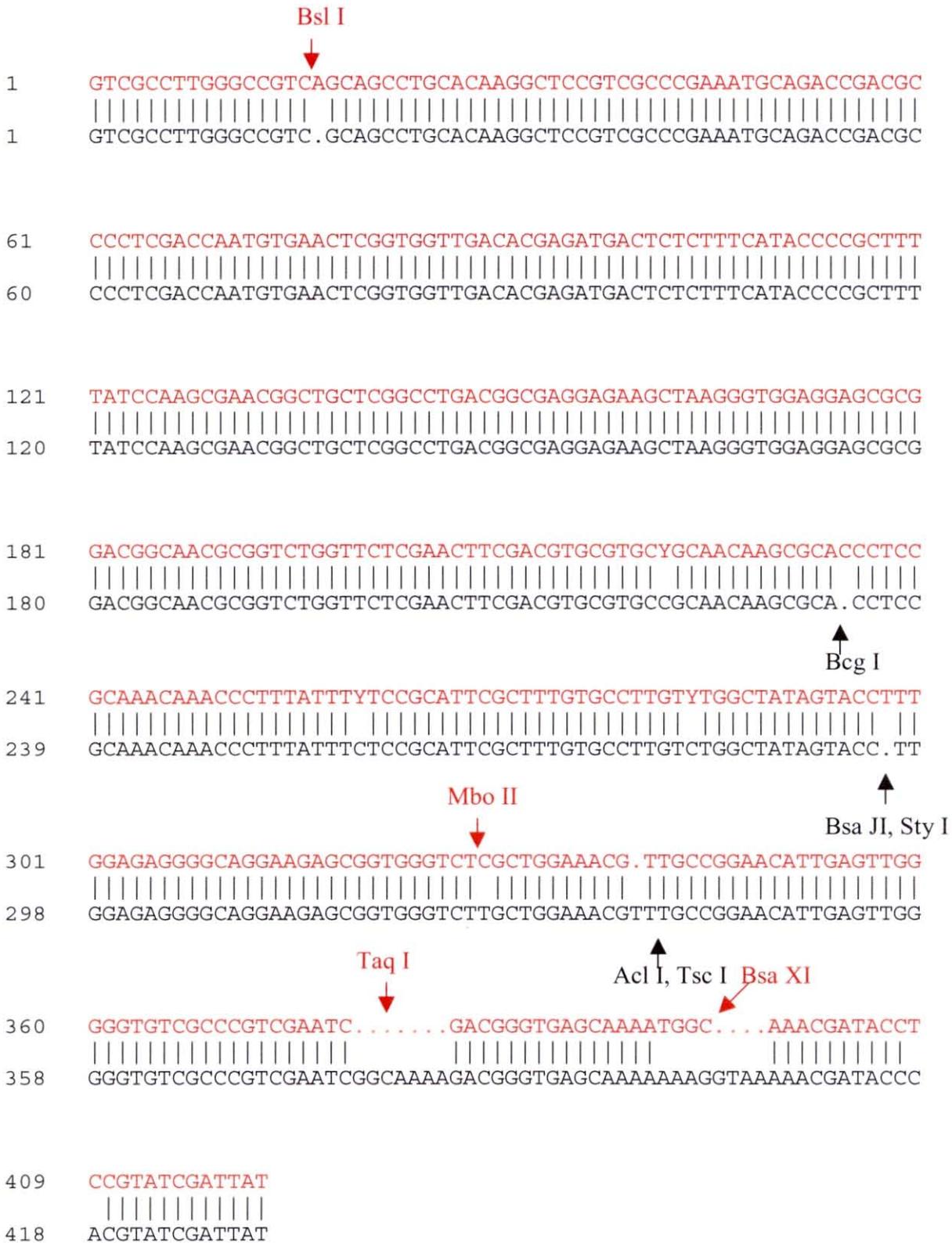


Figure 30 : Alignement des deux séquences obtenues. Les différentes enzymes retenues après l'analyse sont positionnées sur la figure.

DISCUSSION
CONCLUSION

La découverte de la présence de l'espèce *C. gasar* sur les côtes américaines (Lapègue *et al.*, soumis) initialement décrites sur les côtes africaines (Adanson, 1757) a inévitablement suscité de nombreuses interrogations.

En effet la présence d'une espèce dans deux continents séparés par un océan laisse penser :

- certaines populations sont plus récentes que les autres.
- l'éloignement géographique implique une limitation des flux géniques et ainsi la mise en place de divergences génétiques.

C'est par une étude de génétique des populations sur les séquences variables du génome qu'il est possible de mettre en évidence des différences entre des individus morphologiquement identiques et génétiquement très proches.

Notre étude réalisée sur le fragment ITS₂, séparant les gènes 5,8 et 28 S, a tout d'abord permis de démontrer que celui-ci était un bon marqueur moléculaire dont la variabilité permet de mettre en évidence des différences entre les individus vivant dans des sites différents.

Le premier objectif de cette étude était donc de détecter une différenciation entre populations. Pour cela les fragments ITS₂, amplifiés par PCR, de tous les individus disponibles de chaque population ont été comparés par RFLP. Après une analyse statistique des résultats il est possible de confirmer l'hypothèse de la présence de populations génétiquement distinctes. En effet les résultats, obtenus grâce à l'étude réalisée avec l'enzyme de restriction Hha I, mettent en évidence une différence nette entre les populations africaines et américaines. En effet, on peut observer sur le continent africain un polymorphisme intrapopulation (présence de deux profils de restriction H 1 et H 2) totalement absent sur le continent américain (présence uniquement du profil H 2). En revanche, ces premiers résultats n'ont pas permis de différencier les populations d'un même continent. Les pourcentages de chaque profil pour les populations africaines ne sont pas significativement différents et pour le continent américain il n'y a aucun polymorphisme mis en évidence par cette première étude. Une des hypothèses au vu de ces résultats serait donc que la différence génétique entre les deux continents est très nette, ce qui implique une absence de flux géniques récents de l'Afrique vers l'Amérique.

L'étude réalisée avec l'enzyme Taq I met elle aussi en évidence un polymorphisme intercontinent qui confirme les premiers résultats. Chaque continent présente un profil différent, majoritaire ou unique dans chaque population. Seule la population ALM du Sénégal présenterait les deux profils. Ces résultats sont à confirmer étant donné la difficulté de lecture

des profils de restriction. En revanche aucun polymorphisme interpopulation n'est observé sur le continent américain.

Cela pourrait s'expliquer par le fait que la colonisation américaine est trop récente pour que des mutations différenciant les populations aient pu s'installer et que la colonisation a été faite à partir d'une base génétique restreinte.

Le deuxième objectif est de commencer à retracer l'histoire évolutive de la répartition de l'espèce sur les différentes régions de l'océan Atlantique. Les résultats obtenus avec l'enzyme Hha I montrent un polymorphisme intrapopulation uniquement en Afrique. Ceci laisse donc supposer que les populations de ce continent sont plus anciennes. En effet l'étude du polymorphisme est basée sur l'observation de l'accumulation des mutations au cours du temps. Il est donc logique de penser que le nombre de mutations augmente au cours du temps et qu'une population est d'autant plus ancienne que le polymorphisme interindividuel est important. Les populations américaines n'ont semble t'il pas eu le temps d'accumuler des mutations sur le site de restriction de l'enzyme Hha I et sont donc plus récentes. Ces résultats nous permettent donc de penser que les populations américaines ont une origine africaine.

Ces premières conclusions semblent pouvoir être affinées avec les résultats moins probables avec l'enzyme Taq I. En effet celle-ci révèle deux profils différents dont chacun semble spécifique d'un continent. Seule la population ALM du Sénégal semble « faire le lien » entre les deux continents en possédant les deux profils, deux individus possédant le génotype américain. Mais ces résultats sont encore trop incertains et ne nous permettent pas pour le moment de conclure à un effet fondateur ayant pour origine des individus de la population sénégalaise de la pointe des Almadies.

Le séquençage nous a permis d'obtenir une séquence pour deux individus vivants sur les deux continents. L'analyse de ces séquences révèle une forte homologie (97%) entre les individus des deux continents alors que l'homologie avec *C. gigas* est beaucoup plus faible (56%). Ceci nous laisse penser que la divergence bien que claire d'après les résultats de RFLP ne semble pas ancienne ce que confirme les résultats obtenus lors de l'étude du gène mitochondriale 16S (100% d'homologie). Le logiciel blast nous confirme que le fragment séquencé est bien une partie de l'ITS₂ en corrélant une petite partie de notre séquence à celle de l'ITS₂ de *C. gigas*. Enfin l'analyse des sites de restrictions permet d'envisager une poursuite du travail plus rapide avec les enzymes appropriées à la mise en évidence du polymorphisme. Il faudra malgré tout très certainement tester plusieurs enzymes parmi celles retenues car l'étude n'a été réalisée que sur une partie plus restreinte du fragment amplifié par PCR.

Ce travail a donc permis d'apporter des réponses aux questions, restées en suspens, sur l'histoire évolutive de cette espèce et sur les possibilités de flux géniques. Il reste malgré tout à confirmer ces résultats par une étude sur un nombre plus important d'individus et de populations. De plus le séquençage réalisé laisse envisager de bonnes possibilités pour la suite de l'étude, après une mise au point plus fine du protocole, il sera possible de séquencer rapidement tous les individus. Cette nouvelle base de donnée ainsi obtenue permettra sans nulle doute de retrouver la ou les populations africaines originaires des individus américains et peut être de déterminer les circonstances de la colonisation américaine.

BIBLIOGRAPHIE :

- Ajana A.M.**, 1979. Preliminary investigation into some factors affecting the settlement of the larvae of the mangrove oyster *Crassostrea gasar* (Adanson) in the Lagos lagoon. *Malacologia*, 18(1-2), 271-275.
- Allen S.K., Gaffney P.M.**, 1993. Genetic confirmation of hybridization between *Crassostrea gigas* (Thunberg) and *Crassostrea rivularis* (Gould). *Aquaculture* 113 : 291- 300.
- Avisé J.C., Arnold J., Martin Ball R., Bermingham E., Lamb T., Neigel J. E., Reeb C.A., Saunders N.C.**, 1987. Interspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Revue ; Ecolology System* 18 : 489-522.
- Banks M.A., McGoldrick D.J., Borgeson W., Hedgecock D.**, 1994. Gametic incompatibility and genetics divergence of Pacific and Kumamoto oysters, *Crassostrea gigas* and *C. sikamea* *Marine Biology* 121(1): 127-135
- Boutet I.**, 1999. Taxonomie et phylogéographie des huîtres creuses dans l'Atlantique sud. Rapport de stage de maîtrise, Université de La Rochelle, effectué au LGP.
- Boudry P., Heurtebise S., Collet B., Cornette F., Gérard A.**, 1998. Differentiation between populations of the Portuguese oyster, *Crassostrea angulata* (Lamarck) and the Pacific oyster, *Crassostrea gigas* (Thunberg), revealed by mtDNA RFLP analysis. *Journal of Experimentation Marine Biology and Ecology* 226: 279-291.
- Desmarais E., Vigneron S., Buresi C., Roizes G.**, 1995. Détection du polymorphisme dans l'ADN : application en biologie et médecine diagnostique, épidémiologique et pronostique. Edition INSERM, Paris.
- Diadhiou H.**, 1995. Biologie de l'huître de palétuvier *Crassostrea gasar* dans l'estuaire de la casamance (Sénégal) : reproduction, larves et captage du naissain. Thèse, Université de Bretagne Occidentale.
- Gilles S.**, 1991. Settlement and growth of the West African mangrove oyster, *Crassostrea gasar*, in the Casamance, Senegal. *Revue Hydrobiology Tropical* 24 (3) :197-207.
- Huvet A., Balabaud K., Bierne N., Boudry P.**, 2001. Microsatellite analysis of 6-hour-old embryos reveals no preferential intra-specific fertilization between cupped oysters *Crassostrea gigas* or *Crassostrea angulata*. *Marine Biotechnology* (in press).
- Huvet A., Lapègue S., Magoulas A., Boudry P.**, 2000. Mithochondrial and nuclear DNA phylogeography of *Crassostrea angulata*, th portuguese oyster endangered in Europe. *Conservation Genetique* 1 : 251-262.
- Klinbunga S., Ampayup P., Tassanakajon A., Jarayabhand P., Yoosukh W.**, 2001. Genetic diversity and molecular markers of cupped oysters (genera *Crassostrea*, *Saccostrea*, and *Striostrea*) in Thailand revealed by randomly amplified polymorphic DNA analysis. *Marine Biotechnology* 3 (2) : 133-144.

Ladron De Guevara B., Winkler F., Rodriguez-Romero F., Palma-Rojas C., 1996. Comparative karyology of four american oyster species. *The Veliger* 39 (3) : 260- 266.

Lapègue S., Boutet I., Leitao A., Heutebise S., Garcia P., Thiriot-Quévieux, Boudry P. The African mangrove oyster, *Crassostrea gasar* (Adanson, 1757), is present along the Atlantic coast of south America. *Soumis à Biology Bulletin*.

Leitao A., Boudry P., Labat J.-P., Thiriot-Quévieux C., 1999. Comparative karyological study of cupped oyster species. *Malacologia* 41 (1) : 175-186.

Saiki R. K., Scharf S., Faloona F. et al, 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230 : 1350-1354.

Seutin G., Brawn J., Ricklefs R.E., Bermingham E., 1993. Genetic divergence among populations of a tropical passerine, the streaked saltator (*Saltator albicollis*). *Auk*. 110 : 117-116.

Verdon B., 2000. Phylogéographie des huîtres creuses des mangroves de l'Atlantique sud par l'apport des marqueurs moléculaire. Rapport de stage de maîtrise , Université de La Rochelle, effectué au LGP.

Xu Z., Guo X., Gaffney P. M., Pierce J. C., 2001. Chromosomal location of the major ribosomal RNA genes in *C. virginica* and *C. gigas*. *The Veliger* 44 (1) : 79-83.

ANNEXE 1 : EXTRACTION D'ADN AU PHENOL-CHLOROPHORME

Protocole

400 μ L de tampon d'extraction* + échantillon de tissu + 10 μ L de protéinase K
(10 mg/ mL)

Bain- marie 55°C pendant une nuit

Ajouter 500 μ L de phénol, homogénéiser

Centrifuger 12000 tr/ min, 10 min

- la phase aqueuse (surnageant) contient les acides nucléiques

- la phase intermédiaire contient les protéines

- la phase organique contient les lipides

Prélever le surnageant. Ajouter 500 μ L de chloroforme (ou chloroforme/ alcool isoamylique 24: 1). Homogénéiser

Centrifuger 12000 tr/ min, 10 min

Prélever le surnageant. Ajouter 500 μ L d'éthanol 100 % glacé (-20°C). Homogénéiser

Précipiter au froid (-20°C pendant 1 h)

Centrifuger 12000 tr/ min, 10 min

Enlever l'éthanol. Rincer par 150 μ L d'éthanol 70 % glacé (-20°C)

Centrifuger 8000 tr/ min, 5 min

Enlever l'éthanol. Faire sécher sur la paillasse

Reprendre dans de l'eau ou du TE 1X

* Tampon d'extraction

<i>EDTA</i>	<i>9.3 g</i>
<i>tris</i>	<i>1.5 g</i>
<i>SDS</i>	<i>2.5 g</i>
<i>eau bidistillée</i>	<i>qsp 250 mL</i>

Remarques

- Toutes les étapes du protocole doivent être faites sous la hotte.

- Les déchets liquides de phénol et de chloroforme sont collectés dans un récipient prévu à cet effet.

ANNEXE 2 : PROTOCOLE SUIVI POUR LA PCR

Préparation du mélange pour PCR ITS₂ :

<i>ADN</i>	2 μ L
<i>tampon pro- HA</i>	5 μ L
<i>MgCl₂</i>	3 μ L
<i>dNTP (2mM de chaque)</i>	5 μ L
<i>amorces ITS₂ R</i>	1 μ L
<i>amorces ITS₂ F</i>	1 μ L
<i>Taq pro- HA</i>	0.1 μ L
<i>eau bidistillée</i>	32.9 μ L
<i>huile minérale</i>	2 gouttes avant de lancer la réaction

Remarque : l'huile permet d'éviter l'évaporation de l'échantillon pendant les différentes variations de température.

Séquence des amorces ITS₂ :

Amorce ITS₂ R : 5'- GCTCTTCCCGCTTCACTCG

Amorce ITS₂ F : 5'- GGGTCGATGAAGAACGCAGC

Caractéristiques de la Taq polymérase

<i>masse moléculaire</i>	94 kD
<i>température optimale de synthèse</i>	70 à 75°C
<i>demi- vie</i>	120 minutes à 92°C
<i>MgCl₂ (200- 300 mM dNTPs)</i>	1 à 4 mM
<i>KCl</i>	<50 mM
<i>exo 3'- 5'</i>	-
<i>exo 5'- 3'</i>	+

ANNEXE 3 : PROTOCOLE SUIVI POUR LA PCR- RFLP

Préparation des milieux réactionnels pour chaque enzyme

Volumes en μL	Taq I	Mse I	Bgl I	Hha I	Acc I	Bsp 143 I	Xho I
Produits de PCR	10	10	10	10	10	10	10
Enzymes	0.5	0.4	0.4	0.4	0.4	0.4	0.4
Tampon	2	2	2	2	2	2	2
eau bidistillée	8.5	8.6	8.5	8.7	8.5	8.6	8.6
Température de digestion (°C)	65	37	37	37	37	37	37
Bleu de charge en fin de réaction	2	2	2	2	2	2	2

Le temps de réaction est de 1h30 à 2h pour toutes les enzymes.

Préparation du TBE 10X pH 8.8

<i>tris</i>	<i>162 g</i>
<i>acide borique</i>	<i>27.5 g</i>
<i>EDTA 0.5 M</i>	<i>50 mL</i>
<i>eau bidistillée</i>	<i>qsp 1L</i>

Préparation d'un gel d'acrylamide 8%

<i>TBE 10X</i>	<i>3.75 mL</i>
<i>acrylamide /bis- acrylamide 40% (37.5 : 1 Biofit)</i>	<i>7.5 mL</i>
<i>eau bidistillée</i>	<i>26.25 mL</i>
<i>TEMED</i>	<i>30 μL</i>
<i>ammonium persulfate 10 %</i>	<i>180 μL</i>

ANNEXE 4 : PROTOCOLE SUIVI POUR LE SEQUENÇAGE.

Composition du gel de séquence :

<i>Solution 5,5% gel matrix Licor</i>	<i>50ml</i>
<i>TMED</i>	<i>33μl</i>
<i>APS</i>	<i>330μl</i>

Purification des produits de PCR :

30 μ l d'acétate de sodium sont additionnés à 30 μ l de produits de PCR.

Le mélange est vortexé.

60 μ l d'éthanol 100% à -20°C sont ajoutés.

Le mélange est vortexé et repose 5 min sur la paillasse.

Centrifugation de 15 min à 10 000 rpm à 4°C.

200 μ l d'éthanol 100% à -20°C sont ajoutés.

Centrifugation de 10 min à 10 000 rpm à 4°C.

Après élimination de l'alcool par simple renversement des tubes, ceux sont séchés 2h au speedvack.