École doctorale : Gènes, Génomes, Cellules

Structure génétique des populations de trois espèces de poissons de récifs cubains : Stegastes partitus, Haemulon flavolineatum et Acanthurus tractus.

Soutenue le 1 juin 2012 devant le jury constitué de :

M. Pierre Capy          Professeur Paris-Sud 11  Examineur
M. Jorge Angulo Valdés  Professeur UH          Examineur
M. Marc Girondot        Professeur MNHN         Rapporteur
M. Serge Planes         DR CNRS EPHE            Rapporteur
M. Erik García Machado  Chercheur CIM          Directeur de thèse
M. Didier Casane        Professeur Paris Diderot Directeur de thèse
# Table des matières

Remerciements ................................................................................................. 7  
Agradecimientos ................................................................................................ 9  

I. Introduction ..................................................................................................15  

II. Informations Complémentaires .................................................................21  
   II.1 Marqueurs moléculaires............................................................................21  
      II.1.1 ADN microsatellite..............................................................................21  
      II.1.2 ADN mitochondrial .............................................................................23  
      II.1.2.1 Région Non Codante de l’ADN mitochondrial ....................................25  
      II.2 Structure génétique des populations de poissons de récifs ......................25  
      II.3 Modèles d´étude .....................................................................................29  

III. *Stegastes partitus* .....................................................................................33  
   III. 1 Caractéristiques générales de l’espèce .....................................................33  
   III. 2 Résultats ..............................................................................................35  
      III.2.1 ADN microsatellite.............................................................................35  
       III.2.1.1 Structure populationnelle .............................................................36  
       III.2.1.2 Diversité génétique ......................................................................39  
       III.2.1.3 Analyse démographique ..............................................................40  
      III.2.2 ADN mitochondrial ............................................................................42  
       III.2.2.1 Diversité génétique et différenciation des populations ......................42  
       III.2.2.2 Démographie ..............................................................................44  

IV. *Haemulon flavolineatum* ...........................................................................49  
   IV. 1 Caractéristiques générales de l’espèce ......................................................49  
   IV. 2 Résultats ..............................................................................................51  
      IV.2.1 ADN microsatellite .............................................................................51  
       IV.2.1.1 Structure populationnelle .............................................................52  
       IV.2.1.2 Diversité génétique ......................................................................54  
       IV.2.1.3 Analyse démographique ..............................................................55
Table des matières

IV.2.2 ADN mitochondrial

IV.2.2.1 Diversité génétique et différenciation des populations

IV.2.2.2 Démographie

V. Acanthurus tractus

V. 1 Caractéristiques générales de l’espèce

V. 2 Résultats

V.2.1 Région non codante

V.2.1.1 Diversité génétique et différenciation des populations

V.2.1.2 Démographie

V.2.2 Cytochrome b

V.2.2.1 Diversité génétique et différenciation des populations

V.2.2.2 Démographie

VI. Discussion générale

VI.1 Structure génétique

VI.2 Variabilité génétique

VI.3 Démographie

VI.4 La problématique A. tractus/A. bahianus

VI.4.1 Les arguments de l’ADNmt

VI.4.2 Dispersion des larves ou des adultes

VI.4.3 Un ou plusieurs événements de dispersion

VI.4.4 Quelques remarques

VI.5 Perspectives

VII. Matériels et méthodes

VII.1 Echantillonnage

VII.2 Traitement des échantillons

VII.3 Extraction de l’ADN

VII.4 Amplification de l’ADN microsatellite

VII.5 Génotypage

VII.6 Amplification de l’ADN mitochondrial
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>VII.7</td>
<td>Séquençage de l’ADN mitochondrial</td>
<td>97</td>
</tr>
<tr>
<td>VII.8</td>
<td>Traitements préliminaires des données</td>
<td>98</td>
</tr>
<tr>
<td>VII.9</td>
<td>Analyses statistiques</td>
<td>98</td>
</tr>
<tr>
<td>VII.9.1</td>
<td>ADN microsatellite</td>
<td>98</td>
</tr>
<tr>
<td>VII.9.1.1</td>
<td>Structure populationnelle</td>
<td>98</td>
</tr>
<tr>
<td>VII.9.1.2</td>
<td>Diversité génétique</td>
<td>99</td>
</tr>
<tr>
<td>VII.9.1.3</td>
<td>Analyse démographique</td>
<td>100</td>
</tr>
<tr>
<td>VII.9.2</td>
<td>ADN Mitochondrial</td>
<td>101</td>
</tr>
<tr>
<td>VII.9.2.1</td>
<td>Analyse des séquences</td>
<td>101</td>
</tr>
<tr>
<td>VII.9.2.2</td>
<td>Diversité génétique et différenciation des populations</td>
<td>101</td>
</tr>
<tr>
<td>VII.9.2.3</td>
<td>Analyse démographique</td>
<td>103</td>
</tr>
<tr>
<td>References Bibliographiques</td>
<td></td>
<td>105</td>
</tr>
<tr>
<td>Annexes</td>
<td></td>
<td>121</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>137</td>
</tr>
<tr>
<td>Résumé</td>
<td></td>
<td>138</td>
</tr>
</tbody>
</table>
REMERCIEMENTS

Je suis reconnaissante à Pierre Capy, Marc Girondot, Serge Planes et Jorge Angulo d’avoir accepté de faire partie du jury de cette thèse et d’avoir consacré du temps à l’évaluation critique de ce mémoire.

Je souhaite évidemment exprimer ma reconnaissance à mes co-directeurs de thèse, Erik Garcia Machado et Didier Casane. À Erik, parce que c’est déjà la troisième (et la dernière) thèse qu’on fait ensemble. À Didier pour le travail supplémentaire qu’implique le fait d’avoir une étudiante étrangère. Merci à eux deux pour leur patience, pour m’avoir permis de travailler sur un sujet si intéressant qui commence à se développer à Cuba, et pour m’avoir guidée et laissée une indépendance dans le travail. Merci pour l’opportunité de faire une partie de ma formation dans un pays et un laboratoire où j’ai pu pratiquer une activité scientifique à son meilleur niveau.

Un remerciement spécial à Silvan à qui j’étais collée tout le temps jusqu’au moment où il a décidé de partir et où j’ai gagnée un peu d’indépendance en France. Merci pour tout ton temps, ta compagnie, pour être allé à la fac toutes les années avec moi, pour m’emmener à l’aéroport avec toutes mes valises, pour parler espagnol, et pour pleins d’autres choses. À Amandine pour ne pas faire moins.

À Isa pour son amitié, pour être toujours prête à m’aider et pour l’avoir fait dans plusieurs occasions, pour son hospitalité. Merci à Véro pour son affection et pour son aide.

À Hélène, et les filles de l’IRD, Claire, et Maud qui m’ont tendue la main lorsque j’ai commencé mes manipulations sur les microsats avec le séquenceur automatique, et qui ont fait une place pour faire passer mes plaques parmi les leurs.

Enfin à tous et toutes dans le labo avec qui j’ai partagé un peu de temps et qui m’ont fait oublier la distance : Gaëlle, François, les deux Céline, Florence, Magalie, Brigitte, David, Emily, Isabelle, Patrick… Merci à tous, même si j’oublie d’en nommer.
AGRADECIMIENTOS

En primer lugar doy gracias a Dios por haberme permitido llegar hasta aquí y quiero agradecer además a todos los que de una forma u otra, ya sea en lo profesional o lo personal me han brindado su apoyo para la feliz culminación de esta etapa de mi vida.

A Yoan por su amor y su paciencia y porque ya no conozco la vida sin él.

A mi mamá y Ricardo por estar siempre ahí, por sacarme de todos los apuros, porque llamo y siempre responden, por su apoyo incondicional.

A mi papá porque su ejemplo me impulsa siempre en la vida.

A mi prima Lety por su apoyo, por cocinar, por internet, por soportarme en mi casa del vedado día y noche y por hacer todo para que yo pudiera trabajar tranquila.

A mi familia en general: mi abuela y mi tía Ida, mi hermanita, mis primos, mis tíos, mi suegra por su amor y su apoyo siempre... me siento feliz de pertenecer a este clan.

A Eloisa, Anmari y Ariagna por su amistad. Gracias por la ayuda con tantas cosas que necesité y gracias por lo momentos de esparcimiento entre tantos días de escribir y escribir. A Ariagna además, porque aunque te hayas ido, tu trabajo en oceanología ha demostrado ser muy ventajoso.

A Volta por regalarme una portada tan linda y por permitirme opinar sobre otras que no lo fueron tanto.

En especial a Aymée y a mis compañeros de grupo Anna y Damir por su disposición y su ayuda en estos años.

A Adrian por el tiempo que me dedicó con los mapas y las distancias.

A Yuriem, Camila y Hanaina por el trabajo que hacen en la biblioteca, por su habilidad de buscar siempre la bibliografía que no tenemos.

A todos los muchachos del CIM, gracias por su ayuda cuando la necesité, ha sido un placer compartir con ustedes estos años: Orly, Leandro, Javier, Yuliet, Elaine, Jose Andrés, Alexei, Maickel, Patri, Maximo, Laura, Ivan, Lachi, el Coco, Pedrito.

También a Pedro por su guía y a los demás integrantes de esta familia por los momentos que hemos compartido: Teresa, Mirta, Betsy, Nelson y Yanisleidys, Harry, Henry y Lily, y Jessica.
À Yoa,

À mes parents
Introduction
I. INTRODUCTION

La dispersion des poissons de récifs dépend pour une large part, de l’existence de larves pélagiques qui peuvent être transportées au sein des courants marins loin de la population source (Cowen et coll., 2000). Ceci est mis clairement en évidence chez les espèces territoriales, celles ayant un mode de vie benthique ou celles ayant une phase adulte avec de faibles capacités de dispersion. De l’efficacité de ce transport dépendra donc l’existence de populations formant un continuum, plus ou moins connectées, tout le long des zones de récifs. Toutefois, ce phénomène de dispersion est beaucoup plus complexe et implique d’autres facteurs comme le comportement larvaire, la durée du stade larvaire pélagique, sa distribution vertical. D’autre part, la présence de barrières physiques, actuelles ou historiques peuvent limiter la dispersion de ces organismes. L’étude de cette question à Cuba nous aidera à déterminer les patrons de connectivité des récifs de l’archipel et aussi avec ceux de l’ensemble des Caraïbes. D’un point de vue pratique, cette information pourrait être utile à la mise en place de zones marines protégées.

L’étude de la dispersion par l’observation directe est très difficile pour deux raisons. Premièrement, au stade larvaire, les individus sont très petits, donc difficile à observer et par ailleurs ils sont très difficiles à identifier. Deuxièmement, de grandes distances isolent beaucoup de récifs parmi lesquels les individus adultes, ou les larves, se déplacent (Shulman et Bermingham, 1995; Rocha et coll., 2005b). Ces deux facteurs limitent donc le suivi direct des individus comme moyen d’étude de la dispersion de la plupart des poissons de récifs (Hellberg, 2007; mais voir Planes et coll., 2009). Il a donc été développé des méthodes qui, en utilisant des individus échantillonnés dans différentes zones, permettent à partir des informations obtenues par différents techniques d’inférer les patrons de migrations. Dans cette catégorie, on trouve les études de génétique des populations. Les populations peuvent être différentes du point de vue génétique à différents niveaux : 1) la présence de variants alleliques différents à un même locus, 2) des fréquences différentes pour un même variant à un locus, et 3) des déséquilibres de liaison différents entre variants à différents locus (Hellberg, 2007). Ces différences entre des populations s’accumulent ou se réduisent au cours du temps et avec des vitesses variables qui dépendent des effets simultanés de différentes forces évolutives dépendant elles-mêmes de facteurs environnementaux et de la biologie des espèces.

Il existe aujourd’hui une large gamme de marqueurs moléculaires qui peuvent être utilisés pour accomplir ce type de recherche. Cependant, pour qu’un marqueur
génétique soit potentiellement utile pour inférer si les populations sont ouvertes ou fermées à l'échange d'individus, il doit avoir un mode d'hérédité connu et être variable à l'intérieur de l'espèce (Hellberg et coll., 2002). Les allozymes ont été très utilisées (Lacson et Morizot, 1991; Planes et coll., 1996), mais elles ne révélaient qu'une partie limitée de la variabilité totale du génome (Nei, 1987) et c'est pourquoi elles ont cédé du terrain, en particulier à l'ADN mitochondrial et aux microsatellites. L'ADN mitochondrial a permis inférer des patrons de connectivité géographique. Il ne recombine généralement pas et sa taille de population effective est plus petite (1/4) que celle du génome nucléaire (Avise, 1994), de sorte que les changements qui ont lieu au niveau populationnel sont détectés par ce marqueur plus facilement qu'avec des gènes nucléaires. Les microsatellites, d'autre part, sont très variables au niveau du nombre des répétitions et par conséquent un polymorphisme de taille est généralement observé. Ils peuvent être analysés de façon rapide et peu chère (Hellberg et coll., 2002). Dans la pratique, la combinaison de divers marqueurs offre la plus grande quantité d'information pour l'étude d'une problématique particulière.

A Cuba, il a été encore fait trop peu d'études qui se servent des marqueurs moléculaires pour déterminer la structure génétique des populations de poissons des récifs et des patrons de connectivité entre les populations autour de l'île. Il existe certaines caractéristiques géographiques qui pourraient agir sur ce phénomène. Cuba comprend quatre groupes insulaires situés au nord et au sud et séparés entre eux par des zones beaucoup plus profondes où la plateforme est très étroite. D'autre part, Cuba est sous l'influence de divers systèmes de courants marins qui favorisent soit le mélange des différentes populations, soit l'autorecrutement. C'est pour ces raisons qu'on a souhaité étudier l'effet de la structure géographique de l'archipel cubain et la biologie reproductive des espèces sur la structure génétique et la connectivité des populations de poissons des récifs.

Dans cette perspective on a donc sélectionné trois espèces présentant des traits d'histoire de vies différents afin d'étudier leur influence sur la dispersion des poissons des récifs. Première espèce, S. partitus : c'est une espèce fortement territoriale qui fraie toute l'année mais préférentiellement pendant le printemps ; elle dépose ses œufs sur le substrat où ils sont protégés par le mâle. Les larves sont pélagiques pendant 28 jours dans la colonne d'eau. Deuxième espèce, H. flavolineatum : cette espèce fraie aussi toute l'année avec des pics pendant le printemps et l'automne ; les individus se regroupent pour la fraie, les œufs, puis les larves produits sont pélagiques, ces dernières ayant une durée d'existence très courte de 15 jours. Troisième espèce, A. tractus : cette espèce a un mode de vie nomade, les individus sont très mobiles même si des déplacements sur de grandes
distances n’ont pas été rapportés. Ils fraient aussi toute l’année avec des pics pendant les mois de janvier-février et aout-septembre, les individus fraient par paires ou en formant des groupes, ils produisent des œufs pélagiques qui donnent des larves elles aussi pélagiques pendant 69 jours.

D’autre part, on connait des facteurs environnementaux comme les fronts et courants marins, la topographie des fonds marins et de la côte qui affectent de la même façon la dispersion de toutes les espèces.

Si les facteurs environnementaux qui limitent ou favorisent la dispersion des organismes sont dominants sur les facteurs biologiques, on s’attend à trouver des patrons très semblables d’organisation de la diversité génétique des populations autour de Cuba, quelques soient les espèces considérées. Si les facteurs biologiques ont un effet dominant, on s’attend à des différences entre les patrons de structuration de la diversité génétique observés pour les différentes espèces. La combinaison des facteurs biologiques et environnementaux donnera des organisations de la diversité génétique plus ou moins semblables pour les différentes espèces en fonction du poids de ces différents facteurs.
Informations complémentaires
II. INFORMATIONS COMPLEMENTAIRES

II.1 Marqueurs moléculaires

II.1.1 ADN microsatellite

Les microsatellites sont des petites séquences d’environ quelques dizaines de nucléotides de longueur constituées par des éléments répétés dont la longueur peut varier entre 2 et 6 pb (Tautz, 1993). Tous les auteurs ne sont pas d’accord avec ce chiffre et la longueur des répétitions peut être fixée entre 2 et 8 pb (Armour et coll., 1999), 1 et 6pb (Goldstein et Pollock, 1997) ou même entre 1 et 5 pb (Schlötterer, 1998).


On les trouve de manière abondante chez les procaryotes autant que chez les eucaryotes. Chez l’homme, ils sont dispersés dans tout le génome et ils constituent environ 3% de celui-ci. La plupart du temps ils sont présents dans des régions non codantes, même s’il est connu qu’environ 8% sont localisés dans les régions codantes (Ellegren, 2000).

Les microsatellites ont été le marqueur par excellence pour beaucoup d’études de génétique des populations et d’écologie moléculaire (Schiffer et coll., 2007; Duvernell et coll., 2008; Edwards et coll., 2008; Godinho et coll., 2008; Schultz et coll., 2008; Väli et coll., 2008). L’élément clef de cette versatilité est la grande variabilité que l’on trouve de façon caractéristique à ces locus, et la rapidité et la fiabilité d’accès à ce type de données. Ce sont des marqueurs codominants qui abondent dans le génome ; ils sont relativement faciles à développer et peuvent être analysés à des coûts modérés.

L’absence de fonction de ces séquences répétées, leur neutralité sélective, a été remise en question. Kashi y Soller (1999) ont montré que des microsatellites peuvent être des éléments codants et de régulation de l’expression de gènes. De plus, il a été montré que la variation du nombre de répétitions peut produire des
Informations complémentaires


Il y a des nombreux facteurs qui agissent sur la dynamique évolutive des microsatellites. Parmi eux, on trouve : le nombre de répétitions, la séquence du motif répété, la longueur de l’unité répétée, les séquences flanquantes, la présence d’imperfections, le taux de recombinaison et de transcription, l’efficacité du système de réparation et l’âge des individus. (Schlötterer, 2000; Ellegren, 2000a).


Sous le modèle IAM chaque mutation produit un nouvel allèle et toutes les mutations ont la même probabilité d’apparition. Par contre, dans le cadre SMM, une mutation produit la pertes ou l’addition d’une unité répété et par conséquent on peut avoir comme résultat la création d’un allèle déjà présent dans la population (Estoup et Cornuet, 1999). Ainsi, les allèles avec des tailles très proches seraient plus proche phylogénétiquement que ceux avec des tailles plus différentes, et ce modèle possède donc une mémoire de la taille des allèles (Balloux et Lugon-Moulin, 2002). De plus, il est en accord avec un mécanisme moléculaire « mispairing slippage » à l’origine de l’addition ou de la délétion de répétitions et tient compte de l’homoplasie.

Il a été proposé deux variantes des modèles précédents : le Modèle à Deux Phases (TPM) c’est un cas particulier du SMM, développé par Di Rienzo et al. (1994) et le Modèle à K-Allèles (KAM) (Crow et Kimura, 1970) où il existe K allèles possibles avec une probabilité constante pour un allèle de muter vers n’importe quel des K-1 autres allèles.
En résumé, les microsatellites présentent un ensemble d’avantages qui ont déterminé son utilisation massive en génétique des populations : abondance, codominance, polymorphisme élevé et neutre en général. En plus, ils constituent de petits fragments d’ADN qui peuvent s’amplifier facilement pour PCR. Les travaux réalisés avec les microsatellites permettent de reconstruire l’histoire des migrations et de l’évolution des espèces, ainsi que évaluer la diversité biologique à différents niveaux d’organisation (Edwards et coll., 1992).

II.1.2 ADN mitochondrial

Le génome mitochondrial des animaux est formé par une petite molécule unique d’ADN circulaire dont la taille varie dans une gamme assez étroite chez les vertébrés (16-20kb) (Avise et coll., 1987; Boore, 1999). Sa séquence code pour 37 gènes dont 24 contiennent l’information pour une partie de la machinerie de traduction de la molécule elle-même (ARNt et ARNr), et les 13 autres pour les sous-unités de la chaine de transport des électrons. En plus, il est reconnu une région régulatrice (ou de contrôle) des processus de réplication et de transcription de l’ADNmt d’environ 0.8 Kb nommée D-loop chez les vertébrés (Avise et coll., 1987). Par ailleurs, il est décrit la présence, dans plusieurs ADNmt, de petits fragments non codants avec des fonction qui peuvent être aussi régulatrice (Boore, 1999).

L’utilité de l’ADNmt repose sur sa faible taille de population efficace, qui est généralement un quart de celle de l’ADN nucléaire et qui implique que la propagation d’un nouvel état d’un caractère dans une population va se produire, en moyenne, quatre fois plus vite que pour un marqueur autosomique. En plus, la taux de mutations synonymes est un ordre de grandeur plus grand que celui de l’ADN nucléaire (Birky et coll., 1989; Li et Graur, 1991).

L’ADN mitochondrial a été largement utilisé comme un outil pour déchiffrer l’histoire évolution et démographique des populations et des espèces (Ballard et Whitlock, 2004). Ce marqueur a démontré avoir une grande valeur en écologie moléculaire et en phyligéographie et 70% des études est basé sur l’analyse de la variation de la séquence de ce marqueur, de façon associée ou exclusive (Avise, 2000). Cette vaste application est due à une série d’avantages :

- Molécule à hérédité maternelle (les lignages maternels peuvent être suivis au cours du temps)
- Présent dans la cellule en copies multiples, ce qui est un avantage si l’on travaille avec du matériel dégradé.
- Évolution relativement rapide de la séquence, ce qui justifie son utilisation à l’échelle micro évolutive.
Informations complémentaires

- Homoplasmie (un seul haplotype par individu, bien que l’hétéroplasmie puisse exister)
- Molécule haploïde avec absence de recombinaison

Cette dernière affirmation a été parfois mise en question à cause des travaux qui ont décrit des cas de recombinaison dans la mitochondrie chez les champignons et les plantes (Birky, 2001), et aussi une espèce de moule (Ladoukakis et Zouros, 2001) ainsi que chez l’homme où la machinerie de recombinaison semble être présente (Kajander et coll., 2000).

Les limitations associées à ce marqueur sont :

- L’absence de recombinaison signifie que toute la molécule a la même histoire. Par conséquence, la généalogie construite a partir de l’ADNmt (comme n’importe quelle autre obtenu d’une seule molécule) ne reflétera forcément pas la totalité de l’histoire de l’espèce en question (Ballard et Whitlock, 2004).
- L’hérédité maternelle implique que l’information soit limitée lorsque la dispersion est liée au sexe (Ballard y Whitlock, 2004).
- L’emploi de la coalescence a montré que les donnés obtenus à partir de cette molécule (très utilisé de façon unique dans les décennies passées) sont souvent insuffisantes pour obtenir des estimations précises des paramètres contrairement aux données multi-locus qui ont amélioré de manière remarquable la performance des méthodes analytiques dérivées de la coalescence (Avise, 2010).

En général les taux de mutation de l’ADN mitochondrial sont plus grands que ceux de l’ADN nucléaire, mais le rapport entre la variation de l’un et de l’autre n’est pas constant et dépend des gènes en question (Ballard et Whitlock, 2004). Plusieurs facteurs ont été proposés comme cause de ce taux d’évolution élevé (Avise, 2009) :

- Mécanismes relativement inefficace de réparation de l’ADN ;
- Environnement riche en oxygène auquel les molécules d’ADNmt sont exposées dans la mitochondrie.
- Relâchement de contraintes fonctionnelles.
- L’ADNmt est nue (i.e. il n’est pas associé à des protéines histones qui sont elles-mêmes bien conservées and peuvent limiter le taux d’évolution de l’ADN nucléaire).
Par ailleurs, quelques positions dans la séquence de l’ADNmt sont plus susceptibles d’évoluer que d’autres, probablement du fait de pressions sélectives plus faible (Aquadro et coll., 1984). La plus grande vitesse avec laquelle le génome mitochondrial évolue par rapport à d’autres génomes (2% de divergence par million d’années chez les mammifères (Brown et coll., 1979), mais elle peut varier entre groupes (voir la revue de Avise, 2000) est attribuée premièrement aux changements dans ces sites-là au-delà desquels les différences s’accumulent plus lentement (Avise et coll., 1987).

**II.1.2.1 Région Non Codante de l’ADN mitochondrial**

L’ADNmt contient une région non codante de grande taille située entre les gènes ARNt-Pro et ARNt-Phe. Cette région, habituellement nommée région de contrôle, contient les promoteurs de la transcription pour toute les deux chaînes, plus l’origine de réplication de la chaîne lourde (Chang et Clayton, 1986).


**II.2 Structure génétique des populations de poissons de récifs**

La connectivité des habitats marins et les conclusions faites à partir d’études de génétique des populations, sont des sujets largement présents dans la littérature scientifique (Shulman et Bermingham, 1995; Waters et coll., 2000; Riginos et Victor, 2001; Planes et Fauvelot, 2002; Rocha et coll., 2002; Chen et coll., 2004). Il est décrit que l’usage de marqueurs génétiques permet d’estimer d’une façon empirique, sous certaines conditions, les niveaux d’échanges entre les populations d’organismes marins (Hellberg, 2009). Cette information est de première importance pour la compréhension de l’évolution et l’écologie de ces organismes.
Du point de vue pratique, de tels travaux peuvent avoir des implications sur la conservation de la biodiversité soit à travers des programmes de gestion des ressources de pêche et/ou la définition de zones marines protégées (Lubchenco et coll., 2003).

Chez les espèces de poissons marins, il est souvent montré l’existence d’une faible structuration génétique due à d’importants flux génétiques entre les populations (Ward et coll., 1994; Waples, 1998; DeWoody et Avise, 2000). Les premiers travaux qui ont traité de ce sujet avec des populations de poissons de récifs reposent sur l’analyse de la variation des allozymes pour caractériser la structuration génétique. La plupart de ces travaux a trouvé une absence de différenciation génétique entre les localités étudiées, ce qui suggère l’existence de grandes populations panmictiques sur l’ensemble de grandes régions (Vawter et coll., 1980; Shaklee et Samollow, 1984; Lacson et Morizot, 1991; Lacson et Bassler, 1992).

Toutefois, Waples (1987) et Doherty et coll. (1995) ont trouvé que des populations de poissons de récifs peuvent être génétiquement différenciées sur l’ensemble de leur aire de répartition et ils ont suggéré que la cause de cette différenciation est en relation étroite avec la durée de la phase de larve pélagique. La plupart des espèces marines inclue dans son cycle de vie au moins une phase de ce type pendant laquelle les individus peuvent être transportés loin de la population source au sein des courants océaniques (Cowen et coll., 2000).

D’autres études qui ont comparé des espèces avec différents types de développement larvaire, loin de trouver une régularité, ont montré que cet aspect du cycle de vie ne prédit pas de façon simple les flux entre populations séparées (Shulman et Bermingham, 1995; Planes, 1998) et que des facteurs océanographiques comme les fronts, les courants et les gires pouvaient modeler de façon diverse la rétention larvaire et l’autorecrutement.

Actuellement, on connait divers facteurs qui agissent sur la structure génétique des populations de poissons de récifs, ce qui reflète la complexité de ce phénomène et son caractère multifactoriel. Du point de vue biologique, ces facteurs agissent plutôt sur le stade larvaire pélagique et on a défini (parmi d’autres) les suivants (Cowen et coll., 2000; Cowen, 2002; Sponaugle et coll., 2003):

- Comportement de fraient des adultes (lieu et moment par rapport les courants, gires et marées)
- Présence et durée des stades pélagiques des œufs et des larves
- Comportement (e.g. phylopatrie) et capacités sensorielles des larves
• Distribution horizontale et verticale des larves
• Potentiel de dispersion des juvéniles et adultes
• Environnement biotique (e.g. prédation et nourriture sur la zone inter établissement)
• Variation dans le succès reproductif (Hedgecock, 1994) qui peut produire un «chaotic genetic patchiness» (Johnson et Black, 1982)

La dispersion larvaire agit sur les niveaux d’échange génétique parmi les populations, pourtant, d’autres facteurs comme l’écologie des adultes peuvent avoir des effets du même ordre de grandeur (Hellberg, 2009).

Des facteurs écologiques et la présence d’habitats disponibles soit pour l’établissement des larves soit pour la colonisation par des adultes, peuvent limiter l’aire de distribution des espèces. Il a été constaté que l’adaptation aux habitats différents peut produire des lignées génétiquement différenciées dans la même espèce indépendamment des facteurs océanographiques et de la durée de la phase larvaire pélagique (Riginos et Nachman, 2001; Rocha et coll., 2005b). De la même façon, la capacité d’utiliser différents types d’habitats est en rapport direct avec l’importance de l’aire de distribution de nombreux poissons de récifs de la côte atlantique occidentale (Luiz et coll., 2011).

La structure génétique des populations de poissons de récifs a aussi une origine dans l’effet de facteurs physiques sur la dispersion de ces organismes. Même si les caractéristiques écologiques de l’habitat ne sont pas propices à leur développement, des facteurs océanographiques peuvent agir comme barrières en empêchant la dispersion larvaire et en produisant l’isolement (Planes et coll., 1996; Rocha et coll., 2002; Taylor et Hellberg, 2006; Galarza et coll., 2009) et ainsi agir comme des pièges et favoriser la rétention larvaire dans les récifs locaux (Swearer et coll., 2002). D’autre part, la circulation océanique transporte aussi les larves, ce qui permet la dispersion et colonisation d’habitats lointains (Muss et coll., 2001) ainsi que l’homogénéisation génétique à grandes distances (Horne et coll., 2008).

Un point fondamental dans l’interprétation des données génétiques est de pouvoir discerner les événements historiques du reste des facteurs qui agissent sur la structure génétique des populations (Fauvelot et Planes, 2002). L’échange de migrants qui a lieu dans le présent ne correspond pas nécessairement aux inférences sur la connectivité entre les récifs formulées sur la base des données génétiques, car la structuration génétique dépend des phénomènes génétiques, biotiques et abiotiques qui ont eu lieu dans le passé et qui peuvent être différents de ceux aujourd’hui à l’œuvre. Identifier et discriminer différents scénarios est
Informations complémentaires

essentiel, comme par exemple savoir si deux populations génétiquement similaires échangent quelques migrants dans le présent ou si elles se sont isolées dans un passé récent (Hellberg, 2007). La structure génétique des populations a été souvent mise en relation avec des événements historiques comme les changements de niveaux de la mer à cause des glaciations (Planes et coll., 1993; Bowen et coll., 2006a), des événements de colonisation (Rocha et coll., 2005a), l’émergence de terres (Muss et coll., 2001), etc.

Aujourd’hui, il existe une grande quantité d’information accumulée pour des espèces différentes, ce qui a permis de faire des études comparatives qui englobent dans une large mesure la diversité des poissons des récifs qui habitent les mers tropicales, et qui permet de faire des généralisations à grande échelle (Lester et Ruttenberg, 2005; Luiz et coll., 2011). Il est à noter qu’il y a une prédominance des travaux dont l’échantillonnage prend en compte l’ensemble de l’aire de distribution des espèces étudiées. Bien que ces travaux nous donnent une vision général des patrons de dispersion (Rocha et coll., 2005b; Bowen et coll., 2006b), ces résultats pourraient cacher des phénomènes d’échange que se manifeste à plus petite échelle (e.g. Planes et coll., 1996). Ainsi, d’après Doherty et coll. (1995), on doit faire plus attention à la stratégie d’échantillonnage car les erreurs de détection des variations à petite échelle pourraient surestimer la divergence apparente à grande échelle.

Dans le cas particulier de Cuba, on trouve très peu d’études antérieures à notre travail sur la structure génétique des populations d’organismes marins. Le seul travail de ce type portent sur les deux espèces les plus importantes de crevettes qui habitent la plateforme cubaine (*Litopenaeus schmitti* et *Farfantepenaeus notialis*). Les résultats montrent la présence de structuration génétique significative entre quelques populations proches géographiquement ainsi qu’une différenciation entre deux golfs séparés 350Km (Garcia Machado et coll., 2001; Espinosa et coll., 2003; Borrell et coll., 2004; Robainas, 2004; Robainas et coll., 2005). D’autre part, Paris et coll. (2005) ont simulé le transport larvaire de cinq espèces de vivaneaux autour de l’archipel cubain en utilisant un modèle de circulation océanique de haute résolution qui comprend des facteurs physiques et biologiques. Cette étude montre l’influence des courants marins sur la dispersion des larves, et indique un niveau d’autorecrutement entre 37 et 80%, ce qui implique une certaine identité des populations locales (figure 1.1).
Figure 1.1 Transport simulé des larves dans les jours a) 1, b) 7, c) 14 et d) 30 après le premier jour de l’ événement de frai de Lutjanus analis dans cinq sites autour de Cuba, correspondants aux couleurs. Le transporte larvaire dans cette étude a été considéré passive. Extrait de Paris et coll. (2005).

II.3 Modèles d’étude

Afin de réaliser notre travail, nous avons sélectionné trois des espèces parmi les plus communes qui vivent dans les récifs cubains : Stegastes partitus, Haemulon flavolineatum et Acanthurus tractus. Elles partagent quelques caractéristiques qui les rendent particulièrement appropriées pour cette étude : leur distribution est à peu près la même et couvre l’ensemble des récifs des Caraïbes, des îlots de la Floride et les Bahamas (Claro, 1994) ; ces espèces sont relativement faciles à échantillonner soit avec des filets, soit avec des fusils sous-marins ; il existe des informations scientifiques sur leurs caractéristiques écologiques et reproductives (e.g. Burke, 1995; Robertson et coll., 2005; Verweij et coll., 2006; Figueira et coll., 2008) ; des études de génétique des populations ont déjà été effectuées pour d’autres régions des Caraïbes et de l’Atlantique occidental (Lacson et Morizot, 1991; Lacson, 1992; Rocha et coll., 2002; Hepburn et coll., 2009). En revanche, ces espèces montrent des différences au niveau de quelques aspects de leurs traits d’histoire de vie que nous décrirons plus loin.

Informations complémentaires

**Phylum** Chordata

**Sous-phylum** Vertebrata

**Infra-phylum** Gnathostomatha

**Classe** Actinopterygii

**Sous-classe** Neoptrygii

**Infra-classe** Teleostei

**Super-ordre** Acanthopterygii

**Ordre** Perciformes

- **Sous-ordre** Labroidei
  - **Famille** Pomacentridae
    - **Genre** Stegastes
      - **Espèce** Stegastes partitus
        (Poey, 1868)

- **Sous-ordre** Percoidel
  - **Super-famille** Percoidae
    - **Famille** Haemulidae
      - **Genre** Haemulon
        - **Espèce** Haemulon flavolineatum
          (Desmarest, 1823)

- **Sous-ordre** Acanthuroidei
  - **Famille** Acanthuridae
    - **Genre** Acanthurus
      - **Espèce** Acanthurus tractus
        (Poey, 1860)
Stegastes partitus
III. Stegastes partitus

III. 1 Caractéristiques générales de l'espèce

L'espèce *Stegastes partitus*, (Poey, 1868), (figure 3.1), appelée « demoiselle bicolore » est une des neuf espèces de la famille Pomacentridae qui vivent dans la mer des Caraïbes (Allen, 1991). Sa distribution correspond à l’Atlantique Occidental, soit le sud de la Floride, les Bahamas et les Caraïbes, mais s’étend probablement jusqu’au Brésil (Cervigón, 1993) (figure 3.2).

![Figure 3.1. Stegastes partitus. A : larve ; B : juvénile ; C : juvénile à un stade avancé ; D : adulte. Extrait de Humann (2007).](image)

Les adultes atteignent une longueur maximale de 10,0 cm (LT)(Claro, 1994) ce qui est en accord avec nos observations qui montrent une valeur moyenne de 5,14cm (minimum 3,36 cm et maximum 9,9 cm) (n = 286). Ils présentent un patron de coloration dans lequel la partie antérieur du corps est habituellement noire et la partie postérieure blanchâtre. Il est courant de trouver un deuxième type de coloration avec une tache jaune située dans la région ventro-latérale. La limite entre la zone sombre et claire est variable; la zone antérieure sombre peut couvrir seulement la partie supérieure de la tête ou s’étendre plus postérieurement jusqu’à la moitié du corps ou encore plus en arrière jusqu’à la base de la nageoire caudale. Rarement, on trouve des individus présentant un patron de coloration inversé, avec la partie antérieure du corps claire et la partie postérieure sombre (Smith, 1997).
Les demoiselles bicolores sont des poissons diurnes et benthiques. Leur nourriture consiste principalement en débris organiques et organismes appartenant au phytobenthos, au zoobenthos et au zooplancton.

La demoiselle bicolore a été utilisée comme modèle dans différentes études de la biologie des poissons des récifs coralliens car elle partage quelques aspects de son cycle de vie avec d’autres espèces de poissons des récifs, comme l’existence de larves planctoniques et une très grande fidélité à son lieu de vie depuis son recrutement jusqu’au sa maturité (Chittaro et coll., 2006).

La reproduction se déroule de la façon suivante : les femelles déposent les œufs sur le substrat, puis les mâles fertilisent, protègent et entretiennent la ponte (Robertson et coll., 1988). Les mâles sont polygynes séquentiels et c’est pourquoi le nid contient souvent des groupes œufs de différentes femelles à différents stades de développement (Schmale, 1981).

La fraie se réalise préférentiellement à certains moments, ce qui est rapporté aussi pour la plupart des poissons tropicaux (García-Cagide et coll., 2001). Même si cette espèce se reproduit toute l’année, ils existent des maximums dans la taille relative des gonades et l’abondance des juvéniles qui sont conséquents avec un patron saisonnier de reproduction au niveau des populations et qui ont été registrés pendant le printemps (février-juillet) (Betancourt, 2005).

ont trouvé des valeurs différentes, minimum de 24 jours et maximum de 39 jours, pour les populations de *S. partitus* dans la localité de Banco Chinchorro, située à l’Est de la péninsule de Yucatán. D’après ces auteurs, la variabilité de la durée du stade larvaire chez *S. partitus* serait due aux différences de disponibilité d’habitat et de nourriture.

D’autre part, il est décrit que les larves sont transportés par les courants marins vers les eaux océaniques très proches du tombant (Valdes-Muñoz et Mochek, 1994) (à quelques centaines de mètres de la côte, < 5km) et qu’elles sont distribués dans les premiers 100 m de profondeur. Le recrutement et la métamorphose se produisent rapidement (Sponaugle et Cowen, 1996a). Une fois un individu installé sur un site, il reste sédentaire et occupe une zone relativement petite (<2 m²) (Robertson *et coll.*, 1988; Chittaro *et coll.*, 2006). Valdes-Muñoz et Mochek (1994) ont défini *S. partitus* comme une espèce territoriale, les individus sont isolés, ils occupent de façon permanente des espaces individuels, lesquels sont défendus, en particulier de l’entrée d’autres poissons. Le territoire défendu comprend le refuge, les zones d’alimentation et le site de fraie.


Ces événements de synchronisation avec le cycle lunaire sont cohérents dans une localité donnée pour différentes espèces du même genre. Le recrutement de *S. leucostictus* à la Barbade possède aussi un pic d’abondance dans le troisième quart lunaire (Sponaugle et Cowen, 1996b) et de la même façon Robertson (1992) a rapporté que le recrutement de cinq autres espèces du genre *Stegastes* qui vivent sur les côtes de Panamá est synchronisé avec la nouvelle.

### III. 2 Résultats

#### III.2.1 ADN microsatellite

Nous avons étudié un total de 248 individus de l’espèce *Stegastes partitus* de cinq localités autour de Cuba, en utilisant quatre locus microsatellites : SpGATA16,
Stegastes partitus

SpAAT40, SpAAC41 et SpGATA40T. La taille de l’échantillon par site est reportée dans le tableau 3.1.

Aucun des locus analysés n’a présenté d’allèles nuls d’après le logiciel Micro-Checker 2.2.3 (Van Oosterhout et coll., 2004). De la même façon, dans la comparaison par paires des locus analysés, aucune ne montre de déséquilibre de liaison après la correction de Bonferroni avec une valeur de probabilité corrigée de $\alpha=0.0083$ (annexe 1)

### Tableau 3.1. Localités et nombre d’individus échantillonnés pour les données microsatellites. Entre parenthèse les années d’échantillonnage.

<table>
<thead>
<tr>
<th>Localité</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Breton</td>
<td>45 (2008)</td>
</tr>
<tr>
<td>Ilot Avalos</td>
<td>52 (2005)</td>
</tr>
<tr>
<td>La Bajada</td>
<td>48 (2005)</td>
</tr>
<tr>
<td>Plage Baracoa</td>
<td>55 (2006)</td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>48 (2007)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>248</strong></td>
</tr>
</tbody>
</table>

Le calcul de la statistique $F_{IS}$ n’indique pour aucun des locus un écart significatif aux proportions de l’équilibre de Hardy-Weinberg une fois appliqué la correction de Bonferroni pour les tests multiples $\alpha=0,0025$ (400 randomisations) (tableau 3.2).

### Tableau 3.2. $F_{IS}$ par locus et localités d’échantillonnage de Stegastes partitus. Aucune valeur significative après correction de Bonferroni ($\alpha=0.0025$ et 400 randomisations).

<table>
<thead>
<tr>
<th></th>
<th>Plage Baracoa</th>
<th>Ilot Avalos</th>
<th>La Bajada</th>
<th>Ilot Coco</th>
<th>Ilot Breton</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpGATA16</td>
<td>0.119</td>
<td>0.202</td>
<td>0.194</td>
<td>0.187</td>
<td>0.156</td>
</tr>
<tr>
<td>SpAAT40</td>
<td>0.178</td>
<td>0.052</td>
<td>0.254</td>
<td>0.157</td>
<td>0.247</td>
</tr>
<tr>
<td>SpAAC41</td>
<td>0.266</td>
<td>0.155</td>
<td>0.131</td>
<td>0.280</td>
<td>0.129</td>
</tr>
<tr>
<td>SpGATA40T</td>
<td>0.163</td>
<td>0.136</td>
<td>0.266</td>
<td>0.277</td>
<td>0.127</td>
</tr>
</tbody>
</table>

**III.2.1.1 Structure populationnelle**

échantillonnées, clusters qui n’ont pas un rapport clair avec les localités où les individus ont été capturés (figures 3.2 et 3.3). Néanmoins la figure montre que la plupart des individus (68,75%) de la localité d’Ilot Coco ont une probabilité d’appartenance au cluster vert supérieur ou égal à 0,6, au contraire du reste des localités où moins de la moitié des individus possède cette caractéristique : Plage Baracoa (45,45%), La Bajada (33,33%), Ilot Breton (22,22%) et Ilot Avalos (21,15%). Cela pourrait indiquer une certaine différentiation génétique de cette localité par rapport aux autres.

Figure 3.2. Détermination du nombre de populations (K) de *S. partitus* a partir des données microsatellites. A : Moyennes et écart-types pour 20 runs des probabilités postérieures pour les valeurs de K. B : Méthode d’Evanno.
Figure 3.3. Probabilité d’appartenance de chaque individu aux 2 clusters, représentés par des couleurs différentes, et obtenue avec le logiciel STRUCTURE.

D’autre part, la matrice de valeurs de $F_{ST}$ pour les localités prises deux à deux nous montre une différentiation exprimée par des valeurs de $F_{ST}$ très faibles mais significatives d’un point de vue statistique. Cette différentiation, on la trouve pour toutes les localités sauf La Bajada (tableau 3.3).

Tableau 3.3. $F_{ST}$ pour les localités d’échantillonnage de *Stegastes partitus* prises deux à deux. Les valeurs en gras sont significatives après la correction de Bonferroni ($\alpha=0.005$ et 200 randomisations).

<table>
<thead>
<tr>
<th></th>
<th>Plage Baracoa</th>
<th>Iлот Avalos</th>
<th>La Bajada</th>
<th>Iлот Coco</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iлот Avalos</td>
<td>0.0092*</td>
<td>-</td>
<td>0.0059</td>
<td>0.0129*</td>
</tr>
<tr>
<td>La Bajada</td>
<td>0.0059</td>
<td>-0.0001</td>
<td>-</td>
<td>0.0130*</td>
</tr>
<tr>
<td>Iлот Coco</td>
<td>0.0129*</td>
<td>0.0130*</td>
<td>0.0076</td>
<td>-</td>
</tr>
<tr>
<td>Iлот Breton</td>
<td>0.0084*</td>
<td>0.0064*</td>
<td>0.0019</td>
<td>0.0095*</td>
</tr>
</tbody>
</table>

Une représentation de cette différentiation basée sur la méthode du cadrage multidimensionnel non-métrique et les distances génétiques de Nei (1978) (figure 3.4) montre que les localités de Plage Baracoa et Iлот Coco sont les plus éloigné du reste du point de vue génétique, cette dernière étant la plus isolée.
Figure 3.4. MDS basé sur la matrice des distances génétiques de Nei (1978).

Le test de Mantel ne nous a pas permis de rejeter l’hypothèse nulle d’absence de corrélation entre les matrices de $F_{ST}$ et les distances géographiques construites à partir du patron général des courants marins ($r=0.513$, $p=0.124$) ou en prenant en compte la distance minimum entre l’îlot Breton et l’îlot Coco ($r=0.539$, $p=0.109$).

III.2.1.2 Diversité génétique

La variabilité génétique a été estimée à partir des 4 locus microsatellites pour l’ensemble des données et en prenant chaque localité de façon indépendante.

En général, tous les locus possèdent un nombre élevé d’allèles qui oscillent entre 21 et 37 (locus SpAAT40 et SpGATA40T respectivement) (tableau 3.4).

La localité avec la plus grande variabilité lorsqu’on regarde le nombre moyen d’allèles par locus est l’îlot Avalos, l’îlot Breton présente le nombre le plus bas (tableau 3.4). Cela peut être attribué aux différences de la taille de l’échantillon analysé par site. On a donc calculé la richesse allélique qui mesure cette quantité indépendamment de la taille de l’échantillon. La localité avec la valeur la plus grande de richesse allélique reste l’îlot Avalos mais celle avec la valeur la plus basse est l’îlot Coco.

Les estimations de hétérozygotie observées ($Ho$) et attendues ($He$) sont élevées, que ce soit par locus ou par localité d’échantillonnage, les fréquences attendues des hétérozygotes étant plus grandes que celles observées (tableau 3.4).
Le rang allélique le plus grand est observé avec le locus SpGATA40T qui présente une différence de 134 entre le nombre maximum et minimum de répétitions (tableau 3.5).

Tableau 3.4. Statistiques descriptives sur les microsatellites étudiés pour *S. partitus* par localité d'échantillonnage. Ba : Plage Baracoa, Av : Ilot Avalos, Bj : La Bajada, Co : Ilot Coco, Br : Ilot Breton

<table>
<thead>
<tr>
<th>Locus</th>
<th>Ba</th>
<th>Av</th>
<th>Bj</th>
<th>Co</th>
<th>Br</th>
<th>Na</th>
<th>Ho</th>
<th>He</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpGATA16</td>
<td>17</td>
<td>26</td>
<td>25</td>
<td>22</td>
<td>20</td>
<td>33</td>
<td>0.795</td>
<td>0.940</td>
</tr>
<tr>
<td>SpAAT40</td>
<td>16</td>
<td>17</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>21</td>
<td>0.697</td>
<td>0.922</td>
</tr>
<tr>
<td>SpAAC41</td>
<td>28</td>
<td>29</td>
<td>25</td>
<td>21</td>
<td>22</td>
<td>36</td>
<td>0.829</td>
<td>0.949</td>
</tr>
<tr>
<td>SpGATA40T</td>
<td>28</td>
<td>24</td>
<td>26</td>
<td>28</td>
<td>26</td>
<td>37</td>
<td>0.833</td>
<td>0.953</td>
</tr>
<tr>
<td>M</td>
<td>22.25</td>
<td>24</td>
<td>23</td>
<td>21.75</td>
<td>21</td>
<td>21.75</td>
<td>0.756</td>
<td>0.807</td>
</tr>
<tr>
<td>Ra</td>
<td>19.61</td>
<td>21.21</td>
<td>21.43</td>
<td>19.88</td>
<td>20.22</td>
<td>19.88</td>
<td>0.721</td>
<td>0.788</td>
</tr>
<tr>
<td>Ho</td>
<td>0.756</td>
<td>0.807</td>
<td>0.747</td>
<td>0.721</td>
<td>0.788</td>
<td>0.721</td>
<td>0.92 ±0.03</td>
<td></td>
</tr>
<tr>
<td>He ±</td>
<td>0.92 ±0.93 ±0.94 ±0.93 ±0.94 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.t.</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

Na : Nombre total d’allèles, Ho : Hétérozygotie observée, He : Hétérozygotie attendue, Ra : Richesse allélique, M : nombre moyen d’allèles par locus

Tableau 3.5. Rang allélique pour les quatre locus microsatellites de *S. partitus*

<table>
<thead>
<tr>
<th>Locus</th>
<th>Rang allélique</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpGATA16</td>
<td>100</td>
</tr>
<tr>
<td>SpAAT40</td>
<td>57</td>
</tr>
<tr>
<td>SpAAC41</td>
<td>99</td>
</tr>
<tr>
<td>SpGATA40T</td>
<td>134</td>
</tr>
<tr>
<td>Moyenne</td>
<td>97.5</td>
</tr>
<tr>
<td>E.T.</td>
<td>31.52</td>
</tr>
</tbody>
</table>

**III.2.1.3 Analyse démographique**

L’écart de la valeur observée de diversité génique dans la population par rapport à sa distribution théorique a été testé en utilisant le test de Wilcoxon (Wilcoxon sign rank test) disponible dans le logiciel Bottleneck. D’une part on a analysé toutes les localités indépendamment et d’autre part on les a regroupées en prenant en compte des résultats des analyses de structure génétique. On a trouvé des
arguments en faveur d’un goulot d’àtranglement, dans tous les cas pour le modèle IAM, et dans certaines localités pour les modèles TPM et SMM (tableau 3.6).

**Tableau 3.6.** Test de Wilcoxon pour la détection de goulots d’àtranglement. p-excès d’hétérozygotes. En gras les valeurs significatives pour \( \alpha = 0.05 \).

<table>
<thead>
<tr>
<th>Localité</th>
<th>IAM</th>
<th>TPM</th>
<th>SMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plage Baracoa</td>
<td>0.03125</td>
<td>0.15625</td>
<td>0.90625</td>
</tr>
<tr>
<td>Ilot Avalos</td>
<td>0.03125</td>
<td>0.09375</td>
<td>0.90625</td>
</tr>
<tr>
<td>La Bajada</td>
<td>0.03125</td>
<td>0.03125</td>
<td>0.03125</td>
</tr>
<tr>
<td>Ilot Breton</td>
<td>0.03125</td>
<td>0.03125</td>
<td>0.03125</td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>0.03125</td>
<td>0.09375</td>
<td>0.90625</td>
</tr>
<tr>
<td>Toutes sans Ilot Coco</td>
<td>0.03125</td>
<td>0.06250</td>
<td>0.90625</td>
</tr>
</tbody>
</table>

L’analyse démographique avec le logiciel Lamarc (Kuhner, 2006) a été réalisé en prenant en compte de manière indépendante les groupes : Ilot Coco et le reste des localités regroupées. Ce logiciel recommande d’utiliser le programme Structure (Pritchard et coll. 2000) pour déterminer la structure des populations à analyser, étant donné que si le taux migration est élevé on n’obtiendra pas de résultats précis. L’analyse a montré que les deux populations se trouvent dans des conditions démographique différentes. La valeur de l’estimateur de croissance pour Ilot Coco est \( g = -492,68 \) (IC95% -499,55 à -447,30) ce qui indique une réduction de la taille de la population. Par contre, pour le reste des localités cette estimation indique une stabilité démographique \( g = -0,002 \) (IC95% -0,040 à 0,149). La valeur de thêta pour ilot Coco est de \( \theta = 0,044 \) (IC95% 0,035 à 0,054), et pour le reste \( \theta = 9,961 \) (IC95% 9,788 à 9,996). En ce qui concerne les taux de migration (M), la valeur de M vers la population de l’ilot Coco est \( M = 0,010 \) (IC95% 0,010 à 0,011) et celle dans le sens inverse \( M = 0,091 \) (IC95% 0,050 à 0,140). Le taux de migration calculé en terme de \( 4Nm \) a été 0,004 individus vers l’ilot Coco est 0,89 vers le reste des localités.
III.2.2 ADN mitochondrial

III.2.2.1 Diversité génétique et différenciation des populations

On a séquencé un fragment de 379 pb de la région non codante de l’ADN mitochondrial chez 231 individus de *S. partitus* échantillonnés autour de Cuba (tableau 3.6). L’analyse de l’ensemble des individus a permis d’identifier 82 haplotypes différents et 74 sites polymorphes (tableau 3.7). Les valeurs relatives de composition nucléotidique montrent une prédominance des A et T (68,92%) par rapport à C et G (31,08%).

Le test d’homogénéité des fréquences n’a montré aucune différence entre les localités échantillonnées au cours du temps (annexe 2). C’est pourquoi on a mélangé les haplotypes des mêmes localités obtenues pour différentes années.

<table>
<thead>
<tr>
<th>Localité</th>
<th>n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Breton</td>
<td>35</td>
<td>(2008)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>231</strong></td>
<td></td>
</tr>
</tbody>
</table>

On a construit un réseau d’haplotypes en utilisant la méthode de « median joining », et une optimisation basé sur le maximum de parcimonie (figure 3.5). On a obtenu une représentation en forme d’étoile dans laquelle il apparaît un haplotype en fréquence majoritaire (111 individus) dont les autres dérivent, séparés par un ou deux pas mutationnels. Beaucoup moins représentés, on trouve deux haplotypes chez 13 et 14 individus qui représentent eux aussi de petits centres de diversification. Enfin on distingue quelques variants haplotypiques partagés par deux, trois ou quatre individus appartenant ou pas à la même localité. En règle générale, pour ce qui est de la relation entre haplotypes et localités...
d’échantillonnage, on trouve à peu près la même proportion d’haplotypes partagés et uniques par localité.

La distance moyenne entre tous les individus séquencés est $d=0,0043 \pm 0,001$, obtenue avec le modèle TPM1uf+$\Gamma$ ($\alpha=0,437$). Cette distance est à peu près la même lorsqu’on estime la distance entre les individus de l’Ilot Coco et les autres ($d=0,004 \pm 0,001$).

Figure 3.5. Réseau d’haplotypes obtenu à partir d’un fragment de la région non codante de l’ADNmt de *S. partitus*. Les connexions entre haplotypes indiquent un pas mutationnel à moins qu’on les signale avec des barres transversales. Dans ce cas, celles-ci correspondent au nombre de mutations. Les chiffres indiquent le nombre d’individus présentant l’haplotype, sauf pour les haplotypes présent chez un seul individu.

On a évalué la différentiation entre les localités prises deux à deux en utilisant la statistique Snn (Hudson, 2000). Cette analyse, comme pour les données microsatellites a trouvé des différences significatives entre l’Ilot Coco et les autres localités échantillonnés, sauf Plage Baracoa. Le résultat de la comparaison entre les îlots Breton et Avalos est aussi significatif (tableau 3.7).

L’analyse du polymorphisme a révélé une valeur élevée de diversité haplotypique et la plus petite valeur pour l’estimation de la diversité nucléotidique parmi les trois espèces analysées (tableau 5.8), ce qui correspondrait aussi à une événement d’expansion récente et une accumulation de mutations à partir une population
d’effectif efficace faible (Grant et Bowen, 1998; Avise, 2000). Ces résultats sont supportés par les valeurs négatives et significatives des statistiques D de Tajima et \( F_S \) de Fu (tableau 3.8).

| Tableau 3.7. Valeurs de la statistique \( S_{ne} \) obtenues pour les localités d’échantillonnage prises deux à deux. * : Valeurs significatives, p<0.05. |
|-------------------------------|-----------------|-----------------|-----------------|
| Ilot Breton                   | Ilot Avalos     | La Bajada       | Plage Baracoa   |
| Ilot Avalos                   | 0.56*           |                 |                 |
| La Bajada                     | 0.51            | 0.52            |                 |
| Baracoa                       | 0.53            | 0.51            | 0.51            |
| Ilot Coco                     | 0.57*           | 0.55*           | 0.55*           | 0.50            |

### III.2.2.2 Démographie

Le nombre observé de différences entre haplotypes pris deux à deux (mismatch distribution) a produit une distribution unimodale (Somme des carrés des écarts SSD=0,00001 ; p=0,99) (figure 3.6) ou la plupart des haplotypes sont séparés par une mutation, avec une valeur de \( R^2 = 0,01 \); \( p = 0,001 \) (tableau 3.9). Cette statistique, comme \( F_S \) de Fu, est un bon outil pour détecter la croissance des populations (Ramos-Onsins et Rozas, 2002). Ce résultat est en accord avec l’hypothèse d’une population qui a connu une expansion démographique récente (Rogers et Harpending, 1992). Les résultats de l’estimation bayésienne des paramètres populationnels obtenus avec le logiciel Lamarc 2.1.6 pour la population d’ilot Coco et le reste des localités regroupées montre qu’actuellement elles sont en croissance : \( g = 831,34 \); IC95% = 504,34-999,99 pour Ilo t Coco et pour le reste \( g = 945,27 \); IC95% = 837,12 - 1000,00 avec des valeurs de thêta pour Ilo t Coco \( \theta = 0,15 \) (IC95% = 0,03-0,31) et pour le reste \( \theta = 0,32 \) (IC95% = 0,15-0,50). Les taux de migration mesurés, vers et depuis les deux entités analysés, sont très élevés : \( M \) vers Ilo t Coco =705,98 (IC95% = 349,36-999,96) et \( M \) vers le reste =273,14 (IC95% = 70,99-488,66).
**Tableau 3.8.** Indices de diversité moléculaire pour un fragment de la région non codante de l’ADNmt de l’espèce *S. partitus*.

<table>
<thead>
<tr>
<th></th>
<th>Ilot Coco</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nombre de séquences</td>
<td>52</td>
<td>231</td>
</tr>
<tr>
<td>Nombre d’haplotypes</td>
<td>20</td>
<td>82</td>
</tr>
<tr>
<td>Transitions</td>
<td>16</td>
<td>53</td>
</tr>
<tr>
<td>Transversions</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>Indels</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Sites polymorphiques</td>
<td>23</td>
<td>74</td>
</tr>
<tr>
<td>Diversité haplotypique ($h$) ± e.t.</td>
<td>0.723 ± 0.067</td>
<td>0.763 ± 0.031</td>
</tr>
<tr>
<td>Diversité nucléotidique ($\pi$) ± e.t.</td>
<td>0.004 ± 0.003</td>
<td>0.004 ± 0.003</td>
</tr>
<tr>
<td>$K$ ± e.t.</td>
<td>1.358 ± 0.853</td>
<td>1.573 ± 0.941</td>
</tr>
<tr>
<td>$D$ de Tajima</td>
<td>-2.34 p=0.0004</td>
<td>-2.612 p=0.00</td>
</tr>
<tr>
<td>$F_S$ de Fu</td>
<td>-19.57 p=0.00</td>
<td>-27.76 p=0.00</td>
</tr>
</tbody>
</table>

*K* : Moyenne des différences nucléotidiques entre les séquences prises deux à deux

**Figure 3.6.** Distribution du nombre de différences entre haplotypes pris deux à deux (mismatch distribution) pour *S. partitus*. 

---

*Stegastes partitus*
Tableau 3.9. Paramètres démographiques obtenues pour *S. partitus* basés sur un fragment de la région non codante de l’ADNmt

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>IC2.5%-IC97.5%</th>
<th>$\theta_0$</th>
<th>IC2.5%-IC97.5%</th>
<th>SSD</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,66</td>
<td>0,18-3,77</td>
<td>0,26</td>
<td>0,00-0,68</td>
<td>0,00001</td>
<td>0,010</td>
</tr>
</tbody>
</table>

$\tau$: (tau) et intervalle de confiance de 95% ; $\theta_0$: thêta avant l’expansion ; SSD : somme des carrés des écarts ; R2 : statistique de Ramos-Onsins et Rozas (2002).
Haemulon flavolineatum
IV. HAEMULON FLAVOLINEATUM

IV. 1 Caractéristiques générales de l’espèce


![Figure 4.1. Haemulon flavolineatum. A : juvénile initial ; B : juvénile dans la nuit ; C : juvénile tardif ; D : adulte. Extrait de Humann (2007).](image)

Durant la phase adulte, elles atteignent une taille maximale de 30cm (Guitart, 1977). Nos observations pour les individus échantillonnées montrent une valeur moyenne de 15,54 cm (minimum 10,80 cm et maximum 19,20 cm).

Le corps présente des bandes jaunes obliques, sur un fond gris argenté, qui passent par le centre des écailles. Au dessus de la ligne latérale, ces bandes se disposent de manière horizontale. Les nageoires sont jaunes. Pendant la phase
juvénile, on observe une tache noire de chaque côté de la nageoire caudale (Poey, 1868).

Ce sont des poissons suprabenthiques avec des habitudes nocturnes. On les trouve au niveau des récifs coralliens et des mangroves éloignées des estuaires (Burke, 1995). Lorsque la nuit tombe, les gorettes migrent vers les sites où elles se nourrissent d’une grande variété d’invertébrés benthiques qui vivent dans les herbiers marins et dans les plaines sablonneuses qui les entourent (Claro, 1994).

Figure 4.2. Distribution de *Haemulon flavolineatum*. L’échelle de couleurs indique la probabilité d’occurrence. Extrait de Fishbase (Froese et Pauly, 2011)

Pendant la journée, elles passent la plupart du temps sur le fond du récif et sont peu mobiles. Elles se déplacent en groupes qui peuvent être composés par peu d’individus ou en rassembler des milliers (Humann, 1994). A Cuba, nos observations directes nous suggèrent que, la plupart du temps, elles forment des groupes relativement petits et elles forment des rassemblements de milliers d’individus pendant la période de reproduction.

Les individus sont matures normalement quand ils atteignent la longueur totale d’environ 12 cm. La proportion d’individus matures atteint un maximum dans l’intervalle entre 17 et 19 cm, et probablement la taille moyenne des individus matures est 15,5 cm (Gaut et Munro, 1983). Nos observations montrent que les individus avec des gonades bien développés possèdent des tailles supérieures à 13,1 cm.

Haemulon flavolineatum

Le recrutement suit un cycle de 15 jours en accord avec les quarts lunaires, et c’est plus clairement évident pendant la période de mai à décembre. C’est une différence avec beaucoup d’autres espèces de poissons des récifs pour lesquelles le recrutement, bien que cyclique, se concentre essentiellement dans une période particulière de l’année (Williams et Sale, 1981).

IV. 2 Résultats

IV.2.1 ADN microsatellite

Le test d’homogénéité des fréquences entre les localités échantillonnées à différentes années n’a pas détecté des différences au cours du temps, sauf pour le locus HfAAC3 dans la localité « Ilot Avalos » (annexe 3). Ceci nous a permis de regrouper tous les données microsatellites pour chaque localité. Les chiffres qui correspondent aux individus échantillonnées sont décrits dans le tableau 4.1.

On a détecté la présence d’allèles nuls pour le locus HfAAC46 avec le logiciel Micro-Checker 2.2.3 (Van Oosterhout et coll., 2004). Dans ce cas, on a estimé les fréquences corrigées des allèles et des génotypes en tenant compte de la présence des allèles nuls.

Parmi les comparaisons par paire des locus analysés, aucune ne montre de déséquilibre de liaison après la correction de Bonferroni avec une valeur de probabilité corrigée de α=0,0033 (annexe 4).
L’écart aux fréquences de Hardy-Weinberg a été testé à partir du calcul de la statistique $F_{IS}$ qui peut adopter des valeurs négatives ou positives, avec la valeur 0 indiquant un ajustement parfait. Ce test n’a identifié que le locus HfAAC41 comme présentant des fréquences statistiquement différentes de celles attendues (tableau 4.2).

### Tableau 4.1. Localités et nombre d’échantillons pour les données microsatellites. Entre parenthèse l’année d’échantillonnage.

<table>
<thead>
<tr>
<th>Localité</th>
<th>n</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Breton</td>
<td>44 (2008)</td>
<td>-----</td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>24 (2008)</td>
<td>-----</td>
</tr>
<tr>
<td>Total</td>
<td>204</td>
<td></td>
</tr>
</tbody>
</table>

### Tableau 4.2. Valeur de $F_{IS}$ et de la probabilité de l’écart aux fréquences de HW pour chaque locus étudié.

<table>
<thead>
<tr>
<th>Locus</th>
<th>$F_{IS}$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>HfAAC3</td>
<td>-0,052</td>
<td>0,828</td>
</tr>
<tr>
<td>HfAAC10</td>
<td>0,076</td>
<td>0,317</td>
</tr>
<tr>
<td>HfAAC37</td>
<td>-0,010</td>
<td>0,601</td>
</tr>
<tr>
<td>HfAAC41</td>
<td>0,097</td>
<td>0,012</td>
</tr>
<tr>
<td>HfAAC43</td>
<td>0,088</td>
<td>0,450</td>
</tr>
<tr>
<td>HfAAC46</td>
<td>0,112</td>
<td>0,110</td>
</tr>
</tbody>
</table>

### IV.2.1.1 Structure populationnelle

valeur est autour de 0,5 (figure 4.3). Ceci nous donne à peu près la même probabilité d’avoir une origine dans l’un ou l’autre des clusters. On voit la difficulté de ce test d’identifier K=1 comme la valeur la plus vraisemblable de K, et cela nous permet de conclure que la valeur réelle de ce paramètre pour l’ensemble des données est 1 au lieu de 2. Ce résultat est confirmé par la matrice de valeurs de la statistique $F_{ST}$ entre les localités d’échantillonnage pour lesquels on trouve des valeurs très faibles et non significatives (tableau 4.3).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.2.png}
\caption{Détermination du nombre de populations (K) de \textit{H. flavolineatum} à partir des données microsatellites. A : Moyennes et écart-types pour 20 runs des probabilités postérieures pour les valeurs de K ; B : Méthode d’Evanno \textit{et coll.} (2005).}
\end{figure}
Figure 4.3. Probabilités d’appartenance de chaque individu à chacun des clusters définis pour le logiciel STRUCTURE (Pritchard et coll., 2000) pour \( K = 2 \).

### Tableau 4.3. \( F_{ST} \) pour les localités d’échantillonnage prises deux à deux. Toutes les valeurs sont non significativement différentes de 0 après correction de Bonferroni (\( p<0,05 \)).

<table>
<thead>
<tr>
<th></th>
<th>Ilot Breton</th>
<th>Ilot Avalos</th>
<th>La Bajada</th>
<th>Plage Baracoa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>0,0105</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Bajada</td>
<td>0,0012</td>
<td>0,0006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plage Baracoa</td>
<td>0,0101</td>
<td>-0,0002</td>
<td>0,0002</td>
<td></td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>0,0041</td>
<td>0,0014</td>
<td>-0,0029</td>
<td>-0,0008</td>
</tr>
</tbody>
</table>

En appliquant le test de Mantel nous n’avons trouvé aucune corrélation entre les matrices de différentiation génétique, \( F_{ST} \), et les distances géographiques, calculées sur la seule base du patron général de courants marins (\( r=0,065, p=0,820 \)) ou en prenant en compte la distance minimum entre l’ilot Breton et l’ilot Coco (\( r=0,057, p=0,797 \)).

**IV.2.1.2 Diversité génétique**

La variabilité génétique a été estimée à partir des 6 locus microsatellites : HfAAC3, HfAAC10, HfAAC37, HfAAC41, HfAAC43 et HfAAC46.

La population dans son ensemble présente un nombre modéré d’allèles par locus avec des valeurs qui oscillent entre 12 et 22 (locus AAC3 et AAC46 respectivement) (tableau 4.4). D’autre part, les valeurs d’hétérozygotie sont élevées, sauf pour le locus HfAAC3 qui est fortement en dessous de la moyenne car il y a un allèle très fréquent par rapport à tous les autres. Le rang allélique est similaire pour l’ensemble des locus sauf pour le locus AAC46 pour lequel la différence entre le nombre maximum et minimum de répétitions est beaucoup plus élevée.

### Tableau 4.4. Statistiques descriptives des microsatellites étudiés chez H.
IV.2.1.3 Analyse démographique

L’écart de la valeur observée de diversité génique dans la population par rapport à sa distribution théorique a été testé en utilisant le test de Wilcoxon (Wilcoxon sign rank test) disponible dans le logiciel Bottleneck (Cornuet et Luikart, 1996). On n’a trouvé aucun argument en faveur d’un goulot d’étranglement quel que soit le modèle de mutation utilisé (tableau 4.5). Toutefois, on observe un déficit d’hétérozygoties pour les modèles TPM et SMM.

### Tableau 4.5. Test de Wilcoxon pour la détection de goulots d’étranglement.

<table>
<thead>
<tr>
<th></th>
<th>p-déficit d’hétérozygotes</th>
<th>p-excès d’hétérozygotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAM</td>
<td>0,92188</td>
<td>0,21875</td>
</tr>
<tr>
<td>TPM</td>
<td>0,03906</td>
<td>0,97656</td>
</tr>
<tr>
<td>SMM</td>
<td>0,00781</td>
<td>1,00000</td>
</tr>
</tbody>
</table>

L’analyse démographique en utilisant le logiciel Lamarc 2.1.6 (Kuhner, 2006) a montré que pour la taille on est en présence d’une population stable avec une valeur de $\theta = 9,052$ (IC95% 9,040-9,064) et presque nulle pour la croissance, $g = -0,103$ (IC95% -0,095 à -0,111).
IV.2.2 ADN mitochondrial

IV.2.2.1 Diversité génétique et différenciation des populations

On a séquencé 367 pb du premier fragment hypervariable de la région non codante de l’ADN mitochondrial et on a obtenu au total les séquences de 207 individus qui représentent toutes les localités échantillonnées pour l’analyse spatiale et au cours du temps (tableau 4.6). En analysant tous les individus confondus nous avons trouvé 80 sites polymorphes qui définissent 91 haplotypes différents (tableau 4.7). Les valeurs relatives de composition nucléotidique montrent une prédominance des A et T (61,64%) par rapport à C et G (38,36%).

Les résultats du test d’homogénéité des fréquences montrent que seule la localité de « Plage Baracoa » montre une différence significative au cours du temps, $\chi^2=47,40$, $p=0,04$ (annexe 5) est pourquoi, comme pour les données microsatellites, on a mélangé les haplotypes des mêmes localités obtenues pour différentes années.

Pour montrer de façon graphique les relations entre les haplotypes analysés, on a construit un réseau en utilisant la méthode de « median joining ». Cette représentation a dévoilé la présence de nombreux variants haplotypiques uniques, de même que des haplotypes communs et présents dans toutes, ou presque toutes, les localités échantillonnées (figure 4.4). La distribution trouvée ne montre pas un regroupement évident des individus par rapport à leur localité d’échantillonnage.
La distance moyenne entre les individus et calculée selon la méthode TPM2uf+I+I (\(\alpha = 0.403\)) est \(d = 0,012 \pm 0,003\).

Le calcul de la statistique \(S_{mn}\) (Hudson, 2000) pour les localités prises deux à deux, n’a pas détecté de différentiation génétique entre les localités (tableau 4.7), la valeur globale de \(S_{mn} = 0,18154, \ p = 0,9180\). Considérant l’absence de différentiation parmi les localités échantillonnées, on a décidé de les traiter comme une seule population.

Figure 4.4. Réseau d’haplotypes obtenu à partir d’un fragment de la région non codante de l’ADNmt de *H. flavolineatum*. Les connexions entre haplotypes indiquent un pas mutationnel à moins qu’on les signale avec des barres transversales. Dans ce cas, celles-ci correspondent au nombre de mutations. Les chiffres indiquent le nombre d’individus portant l’haplotype, sauf pour les haplotypes uniques.
Tableau 4.7. Indices de diversité moléculaire pour un fragment de la RNC de l’ADNmt de l’espèce *H. flavolineatum*

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nombre de séquences</td>
<td>207</td>
</tr>
<tr>
<td>Nombre d’haplotypes</td>
<td>91</td>
</tr>
<tr>
<td>Transitions</td>
<td>52</td>
</tr>
<tr>
<td>Transversions</td>
<td>18</td>
</tr>
<tr>
<td>Indels</td>
<td>17</td>
</tr>
<tr>
<td>Sites polymorphes</td>
<td>80</td>
</tr>
<tr>
<td>Diversité haplotypique</td>
<td>0,960 ± 0,00005</td>
</tr>
<tr>
<td>Diversité nucléotidique</td>
<td>0,01123 ± 0,00041</td>
</tr>
<tr>
<td>K</td>
<td>7,926 ± 3,70</td>
</tr>
<tr>
<td>$D$ de Tajima</td>
<td>-1,579 (p=0,024)</td>
</tr>
<tr>
<td>$F_s$ de Fu</td>
<td>-24,464 (p=0,0003)</td>
</tr>
</tbody>
</table>

K : Moyenne des différences nucléotidiques entre les séquences prises deux à deux.

Tableau 4.8. Valeurs de la statistique $S_{mn}$ obtenues pour les localités d’échantillonnage prises deux à deux. Toutes les valeurs sont non significatives, p>0,05.

<table>
<thead>
<tr>
<th></th>
<th>Ilot Breton</th>
<th>Ilot Avalos</th>
<th>La Bajada</th>
<th>Plage Baracoa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>0,57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Bajada</td>
<td>0,52</td>
<td>0,54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baracoa</td>
<td>0,48</td>
<td>0,43</td>
<td>0,47</td>
<td></td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>0,50</td>
<td>0,55</td>
<td>0,51</td>
<td>0,46</td>
</tr>
</tbody>
</table>

L’analyse du polymorphisme a révélé une valeur élevée de diversité haplotypique et une très faible diversité nucléotidique (tableau 4.8), ce qui peut correspondre à une croissance rapide de la population et une accumulation de mutations à partir une population d’effectif efficace faible (Grant et Bowen, 1998; Avise, 2000).
**IV.2.2.2 Démographie**

Les valeurs des statistiques D de Tajima et $F_S$ de Fu, sont négatives et significatives pour l’ensemble des données (tableau 4.8) ce qui suggère un événement d’expansion de la population.

La distribution du nombre observé de différences entre haplotypes pris deux à deux (mismatch distribution) a produit une distribution unimodale (somme des carrés des écarts SSD=0,0007, $p=0,9$) (figure 4.5) avec une valeur de $R^2=0,033$, $p=0.027$ (tableau 4.9). Ce résultat est en accord avec l’hypothèse d’une population qui a connu une expansion démographique récente (Rogers et Harpending, 1992). Dans cette même perspective, les paramètres populationnels estimés avec le logiciel LamArc 2.1.6 (Kuhner, 2006) suggèrent aussi que la population est en croissance compte tenu de la valeur fortement positive du taux de croissance ($g=840,65$, IC95%=631,94-991,14) avec une valeur de thêta $\theta=1.31$ (IC95%=0.74-1.90).

**Figure 4.5.** Distribution du nombre de différences entre haplotypes pris deux à deux (mismatch distribution) pour *H. flavolineatum.*

<table>
<thead>
<tr>
<th>τ</th>
<th>IC2,5%-IC97,5%</th>
<th>θ₀</th>
<th>IC2.5%-IC97,5%</th>
<th>SSD</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0</td>
<td>2,268-6,809</td>
<td>0,490</td>
<td>0,0018-0,1354</td>
<td>0,0007</td>
<td>0,033</td>
</tr>
</tbody>
</table>

τ : (tau) et intervalle de confiance de 95% ; θ₀ : thêta avant l’expansion ; SSD : somme des carrés des écarts ; R² : statistique de Ramos-Onsins et Rozas (2002).
Acanthurus tractus
V. ACANTHURUS TRACTUS

V. 1 Caractéristiques générales de l’espèce


![Figure 5.1 Acanthurus tractus](image)

**Figure 5.1** *Acanthus tractus*. A : larve ; B : post-larve ; C : juvenile ; D : adulte. Extrait de Humann (2007).

Plus précisément, on trouve cette espèce dans l’Atlantique Nord tropical, du Massachusetts et Bermudes jusqu’à Trinidad et Tobago, dans le Gulf de Mexique et les îles des Caraïbes (Bernal et Rocha, 2011) (figure 5.2).

Récemment, sur la base de différences de coloration du bord des nageoires caudale et dorsale et sur celle de la distance génétique, Bernal et Rocha (2011) ont élevé au statut d’espèces différentes les lignées Caribéenne et Atlantique d’*Acanthus bahianus*, avec une distribution restreinte au nord et au sud de l’embouchure des rivières Orénoque-Amazone respectivement. Ces auteurs ont repris le nom *Acanthus tractus* proposé par Poey (1860) pour la lignée des Caraïbes. Les résultats présentés ci-dessous nous montrent la présence de la lignée sud dans les Caraïbes et ils mettent donc en doute l’existence de deux entités spécifiques étant donné qu’on ne sait pas s’il existe un isolement reproducteur entre elles. Comme tous les aspects du cycle de vie chez *A. tractus* correspondent dans la littérature au
nom *A. bahianus* dans ce mémoire on parlera d’*Acanthus bahianus/A. tractus* pour évoquer cette espèce.

*A. bahianus/A. tractus* est observé entre 2 et 40 mètres de profondeur (Desoutter, 1990), mais il est plus fréquent entre 2 et 25 mètres (Baensch et Debelius, 1997). Après la phase larvaire, on le trouve associé aux récifs peu profonds, maximum 25m, à la différence de *A. chirurgus* et *A. coerulus*, les deux autres espèces de poissons chirurgiens les plus communs dans l’atlantique et qui ont été observés dans des zones beaucoup plus profondes (Nagelkerken *et coll.*, 2000).

![Figure 5.2](image.png)

**Figure 5.2.** Distribution d’*Acanthus tractus*. L’échelle de couleurs indique la probabilité d’occurrence. Modifié d’après Fishbase (Froese et Pauly, 2011)

Les deux espèces présentent une coloration gris bleuté ou marron homogène, à la différence de son congénère *A. chirurgus* qui montre de traits verticaux de chaque côté du corps. Les nageoires caudale, dorsale et anale sont bordées de blanc ou de bleue pâle chez *A. tractus* et une bande claire peut apparaitre à la base du pédoncule caudal (Humann, 1994). La taille moyenne estimée pour cette espèce sur la base de notre échantillonnage (468 individus) est de 14,71 cm, minimum 8,54 cm et maximum 20,0 cm.

Les chirurgiens sont des poissons qui vivent dans de grands territoires dont ils utilisent les ressources alimentaires de façon extensive. Cette manière caractéristique d’occuper leur territoire implique un mode de vie nomade (Valdes-Muñoz et Mochek, 1994). Ils préfèrent les zones benthique et suprabenthique, ils sont très mobiles et forment fréquemment des groupes. Leur alimentation est essentiellement composée d’algues épiphytes qui croissent sur les coraux ou dans
Acanthurus tractus

les herbiers proches. D’après Robertson (1985) on peut les trouver seuls, car ils sont territoriaux dans une certaine mesure, ou en bandes qui se déplacent pour se nourrir. Ainsi, même si quelques adultes marqués ont été recapturés jusqu’à un kilomètre du site original de capture (Randall, 1962), Bardach (1958) a rapporté que des bancs surveillés pendant 30 minutes nagent d’habitude dans une zone de moins de 500m².

Reeson (1983) a suggéré que cette espèce se reproduit pendant toute l’année avec des pics de l’activité de fraie pendant janvier-février et aout-septembre.


Reeson (1983) a montré que la taille des individus arrivant à maturité est autour de 11 cm, mais que la plupart des poissons sont matures pour une taille de 15-16cm de longueur totale.
V. 2 Résultats

V.2.1 Région non codante

V.2.1.1 Diversité génétique et différenciation des populations

On a analysé un fragment de 274 pb de la région non codante de l’ADN mitochondrial d’individus de l’espèce *A. tractus* échantillonnés autour de Cuba (tableau 5.1). On a séquencé ce fragment pour 73 individus qui correspondent à un même nombre d’haplotypes avec 159 sites polymorphes. Les valeurs relatives de composition nucléotidique montrent une prédominance des A et T (67,29%) par rapport à C et G (32,71%). Les transitions sont plus nombreuses que les transversions et le ratio transition/transversion a été estimé égal à 4,3:1.

<table>
<thead>
<tr>
<th>Localité</th>
<th>RNC</th>
<th>Plage Baracoa</th>
<th>Ilot Breton</th>
<th>Ilot Coco</th>
<th>La Bajada</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytb</td>
<td>22(20)</td>
<td>23</td>
<td>25</td>
<td>33</td>
<td>26</td>
<td>149</td>
</tr>
</tbody>
</table>

Le résultat de l’analyse d’assignation utilisant l’approche bayésienne implémentée dans le logiciel BAPS 5.3 est présenté figure 5.3. La méthode a regroupé tous les individus dans deux clusters avec une représentation différente des haplotypes dans chacun d’eux. Néanmoins, les échantillons appartenant aux différentes localités sont distribués dans les deux clusters ($\chi^2=9,49$ p=0,388) et par conséquent aucune partition géographique n’a été observé.
Figure 5.3. BAPS pour un fragment de la RNC d’A. tractus. Les individus ont été assignés à deux clusters représentés par des couleurs différentes.

La moyenne des différences entre les haplotypes est $k=24,41$ substitutions (tableau 5.2). Le réseau d’haplotypes, construit avec la méthode de « median joining » montre les relations entre les haplotypes de la région non codante. On a mis en évidence l’existence de deux lignées séparées par plus de 20 mutations (figure 5.4) qui correspondent aux deux clusters définis par la méthode bayésienne (BAPS).

La divergence moyenne des séquences entre les deux lignées est $d=0,199\pm 0,037$, estimée avec le modèle $TrN + \Gamma (\alpha=0,758)$. La lignée (I) contient le plus grand nombre d’échantillons et montre une distance moyenne entre individus égale à $d=0,057 \pm 0,009$. La lignée (II), composé par six haplotypes, présente une distance moyenne entre individus relativement plus élevé, $d=0,073\pm 0,014$. 
Figure 5.4. Réseau d’haplotypes obtenu à partir d’un fragment de la région non codante de l’ADNmt d’*A. tractus*. Les connexions entre haplotypes indiquent un pas mutationnel à moins qu’on les signale avec des barres transversales. Dans ce cas, celles-ci correspondent au nombre de mutations décrit dans la légende. Les nombres romains correspondent aux lignées I et II. Chaque haplotype est présent chez un seul individu.

Compte tenu de la grande différenciation existant entre les deux lignées observées chez l’espèce *A. tractus*, et l’absence de différenciation entre les localités d’échantillonnage, les calculs des indices de diversité génétique (tableau 5.2) et les analyses démographiques ont été faites pour la lignée I d’une part et pour l’ensemble des données d’autre part.

Les estimations de diversité haplotypique et diversité nucléotidique sont très élevés pour la lignée I et pour l’espèce en son entier tableau 5.2.

**V.2.1.2 Démographie**

Les tests de neutralité, soit par la statistique *D* de Tajima ou *F*_5 de Fu, donnent des valeurs négatives, mais elles ne sont pas toutes significatives au seuil 5% (tableau 5.2). Lorsqu’on enlève les haplotypes correspondant à la lignée II, les deux statistiques sont significatives. Les deux valeurs négatives et la valeur absolue de *F*_5 particulièrement élevé suggèrent un événement d’expansion démographique de la population.
### Tableau 5.2. Indices de diversité génétique pour un fragment de la RNC de l’ADNmt de l’espèce A. tractus.

<table>
<thead>
<tr>
<th></th>
<th>Lignée I</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nombre de séquences</td>
<td>67</td>
<td>73</td>
</tr>
<tr>
<td>Nombre d’haplotypes</td>
<td>67</td>
<td>73</td>
</tr>
<tr>
<td>Transitions</td>
<td>127</td>
<td>146</td>
</tr>
<tr>
<td>Transversions</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>Indels</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Nombre de sites polymorphiques</td>
<td>144</td>
<td>159</td>
</tr>
<tr>
<td>Diversité haplotypique ($h$) ± e.t.</td>
<td>1,00±0,0026</td>
<td>1,00±0,0023</td>
</tr>
<tr>
<td>Diversité nucléotidique ($n$) ± e.t.</td>
<td>0,067±0,034</td>
<td>0,089 ± 0,044</td>
</tr>
<tr>
<td>$K$ ± e.t.</td>
<td>18,49±8,30</td>
<td>24,41±10,84</td>
</tr>
<tr>
<td>$D$ de Tajima</td>
<td>-1,62 (p=0,024)</td>
<td>-1,37 (p=0,056)</td>
</tr>
<tr>
<td>$F_S$ de Fu</td>
<td>-24,20 (p=0,0002)</td>
<td>-24,09 (p=0,0001)</td>
</tr>
</tbody>
</table>

K : Moyenne des différences nucléotidiques entre les séquences prises deux à deux.

L’analyse démographique par la méthode de « mismatch distribution », a montré que la distribution du nombre observé de différences entre haplotypes pris deux à deux pour l’espèce en son entier est bimodale. Les valeurs observées ne s’ajustent pas à celles attendues sous le modèle d’expansion soudaine : somme des carrés des écarts = 0.029, p=0.0006 (figure 5.5A). Le deuxième pic correspond aux différences entre les deux lignées. Lorsqu’on analyse les haplotypes appartenant à la première lignée, on trouve que les données sont en accord avec l’hypothèse d’un événement d’expansion de la population (figure 5.5B ; tableau 5.3). Les résultats de l’estimation bayésienne des paramètres démographiques de la population réalisée avec le logiciel Lamarc 2.1.6 (Kuhner, 2006) montrent une valeur positive du taux de croissance ($g$=97,84 ; IC95%=[69,58-129,66]) indiquant qu’actuellement la population est en croissance. La valeur estimée de $\theta = 4.05$ (IC95% = [2,30-7,88]).
**Acanthurus tractus**

Figure 5.5. Distribution du nombre de différences entre haplotypes pris deux à deux (mismatch distribution) pour un fragment de la RNC d’*A. tractus*. **A** : Ensemble des données. **B** : Lignée I

<table>
<thead>
<tr>
<th>V.2.2 Cytochrome b</th>
</tr>
</thead>
</table>

V.2.2.1 Diversité génétique et différenciation des populations

L’analyse d’un fragment du gène mitochondrial cytochrome b, a permis de définir 61 haplotypes pour les 149 individus séquencés (tableau 5.1). Ce fragment présente 74 sites polymorphes, tous situés en troisième position des codons. Les valeurs relatives de composition nucléotidique montrent une prédominance modérée des A et T (54,45%).

Le test d’homogénéité de fréquences n’a pas mis en évidence des différences significatives entre les échantillonnages (années 2006 et 2008) effectués pour la localité d’Ilot Avalos ($\chi^2_{0.05}=32,67$ ; $p=0,46$). Les individus de cette localité ont donc été regroupés pour les analyses.
L’approche bayésienne implémentée dans le logiciel BAPS 5.3 (Corander et Tang, 2007; Corander et coll., 2008) a mis en évidence l’existence de trois clusters pour l’ensemble des données (figure 5.6). Les clusters 1 et 2 sont les plus proches : la différence du log de la vraisemblance marginale est la plus proche de zéro pour ces deux groupes (annexe 6). De la même façon que pour la région non codante aucune partition géographique n’a été trouvée en rapport avec ses clusters ($\chi^2=2,12 ; p=0,9$).

**Figure 5.6** Résultats de l’analyse BAPS pour le fragment du gène **cyt b** d’*A. tractus*. Les individus ont été assignés à trois clusters représentés par des couleurs différentes.

Le réseau d’haplotypes construit avec la méthode de « median joining » est présenté dans la figure 5.7. La valeur moyenne des différences entre tous les individus est égale à $k = 4,00 \pm 2,01$. Les relations généalogiques révèlent l’existence de deux lignées divergentes. La première (I) est composée par les deux haplotypes les plus fréquents (qui correspondent aux haplotypes ABA1 et ABA2 décrits par Rocha et coll. (2002) pour les Caraïbes et plus récemment assignés à *Acanthurus tractus* et d’un ensemble d’haplotypes rares qui dérivent de ces haplotypes majeurs. Ils correspondent aux clusters vert et rouge obtenus par le logiciel BAPS (figure 5.6). Le nombre moyen de différences nucléotidiques dans cette lignée est égale à $k = 2,93 \pm 0,034$ (tableau 5.4). La deuxième lignée (II) est connectée à la première par neuf mutations (huit transitions et une transversion). Elle n’est composée que par sept individus qui représentent toutes les localités d’échantillonnage et cinq d’entre eux partagent le même haplotype. Le nombre moyen de différences nucléotidiques dans cette lignée est égale à $k = 1,143 \pm 0,182$. Dans ce groupe d’individus on trouve aussi les haplotypes ABA3 et ABA4 décrits par Rocha et coll. (2002) qui représentent l’espèce *A. bahianus* (sensu Bernal et Rocha (2011)) laquelle d’après ces auteurs est restreinte à l’Atlantique.
Sud. Ces haplotypes sont connectés aux haplotypes cubains par une à quatre transitions. La distance génétique moyenne estimée entre les individus des deux lignées est égale à $d = 0,020 \pm 0,006$. Nous l’avons obtenue en utilisant le modèle TrN + $\Gamma$ ($\alpha=0,259$). D’autre part les valeurs de diversité haplotypique et nucléotidique sont relativement élevées (tableau 5.4).

Figure 5.7. Réseau d’haplotypes obtenu à partir d’un fragment du gène cyt b de l’ADN mt d’A. tractus. Les connexions entre haplotypes indiquent un pas mutationnel à moins qu’on les signale avec des barres transversales. Dans ce cas, celles-ci correspondent à chacun des types de mutations décrits dans la légende. Les chiffres indiquent le nombre d’individus, sauf pour les haplotypes uniques. Les nombres romains correspondent aux lignées I et II ; * : lignée sud décrite par Rocha et coll. (2002).
Tableau 5.4. Indices de diversité génétique pour un fragment du gène cytb de l'ADNmt de l'espèce A. tractus.

<table>
<thead>
<tr>
<th></th>
<th>Lignée I</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nombre de séquences</td>
<td>142</td>
<td>149</td>
</tr>
<tr>
<td>Nombre d'haplotypes</td>
<td>58</td>
<td>61</td>
</tr>
<tr>
<td>Transitions</td>
<td>57</td>
<td>64</td>
</tr>
<tr>
<td>Transversions</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Indel</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sites polymorphiques</td>
<td>66</td>
<td>74</td>
</tr>
<tr>
<td>Diversité haplotypique (h) ± e.t.</td>
<td>0,856±0,026</td>
<td>0,868±0,024</td>
</tr>
<tr>
<td>Diversité nucléotidique (n) ± e.t.</td>
<td>0,005±0,003</td>
<td>0,007±0,004</td>
</tr>
<tr>
<td>K ± e.t.</td>
<td>3,04±1,593</td>
<td>4,002±2,012</td>
</tr>
<tr>
<td>D de Tajima</td>
<td>-2,357 (p=0,000)</td>
<td>-2,237 (p=0,0003)</td>
</tr>
<tr>
<td>Fₜ de Fu</td>
<td>-26,431 (p=0,000)</td>
<td>-25,860 (p=0,000)</td>
</tr>
</tbody>
</table>

Dans cette même perspective, le calcul de la statistique Sₘ entre les localités prises deux à deux (tableau 5.5) est cohérent car il indique une homogénéité génétique en prenant seulement les individus appartenant à la lignée I. Lorsqu'on a fait ce calcul en incluant toutes les individus les résultats ont été les mêmes (annexe 7).

Tableau 5.5. Valeurs de la statistique Sₘ obtenues pour les localités d'échantillonnage prises deux à deux. Toutes les valeurs sont non significatives, p>0,05.

<table>
<thead>
<tr>
<th>Ilot Breton</th>
<th>Ilot Avalos</th>
<th>La Bajada</th>
<th>Plage Baracoa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ibot Avalos</td>
<td>0,46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Bajada</td>
<td>0,46</td>
<td>0,51</td>
<td></td>
</tr>
<tr>
<td>Baracoa</td>
<td>0,46</td>
<td>0,52</td>
<td>0,46</td>
</tr>
<tr>
<td>Ibot Coco</td>
<td>0,43</td>
<td>0,46</td>
<td>0,49</td>
</tr>
</tbody>
</table>

K : Moyenne des différences nucléotidiques entre les séquences prises deux à deux.
V.2.2.2 Démographie

Comme attendu, l'information démographique obtenue par les tests de neutralité (D de Tajima et $F_S$ de Fu) confirme celle obtenue pour la région non codante (tableau 5.4). La distribution du nombre observé de différences entre paires d’haplotypes « mismatch distribution » pour l’ensemble des données (SSD=0,0057, p=0,574) et pour la lignée I s’ajuste aux fréquences attendues sous le modèle d’expansion soulande, comme observée pour la région non codante (figure 5.8, tableau 5.6). Les paramètres populationnels obtenus avec le logiciel Lamarc 2.1.6 (Kuhner, 2006) montrent la même tendance que les résultats pour la région non codante : g=976,55 ; IC95%=[595,71-998,65 et thêta $\theta=0,16$ (IC95%=0,10-0,23).

![Graphique de fréquences mismatch distribution]

**Figure 5.8.** Distribution du nombre de différences entre haplotypes pris deux à deux (mismatch distribution) pour le gène cytb d’A. tractus. A : Ensemble des données ; B : Lignée I.

<table>
<thead>
<tr>
<th>Tableau 5.6.</th>
<th>Paramètres démographiques obtenus pour A. tractus (lignée I) basés sur un fragment du gène cytb de l’ADNmt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>IC2,5%-IC97,5%</td>
</tr>
<tr>
<td>3,79</td>
<td>1,13-6,48</td>
</tr>
<tr>
<td></td>
<td>p=0,61</td>
</tr>
</tbody>
</table>

$\tau$ (tau) et intervalle de confiance de 95% ; $\theta_0$ : thêta avant l’expansion ; SSD : somme des carrés des écarts ; R2 : statistique de Ramos-Onsins et Rozas (2002).
Discussion générale
VI. DISCUSSION GENERALE

VI.1 Structure génétique

Dans ce travail, nous avons fait une étude comparative de la structuration génétique des populations de trois espèces de poissons des récifs avec des histoires naturelles contrastées, mais exposées aux mêmes contraintes environnementales. Nous avons découvert que ce sont probablement les aspects biologiques (i.e. le comportement reproducteur) que déterminent l'existence de populations génétiquement différenciées chez les espèces de poissons étudiées.

L'analyse de la structure génétique des populations nous a montré la façon dont la diversité génétique est distribuée parmi les localités étudiées de *S. partitus*, *H. flavolineatum* et *A. tractus* autour de l'archipel cubain. Cette structure a été décrite sur la base des marqueurs d'ADN mitochondrial et microsatellite.

Au cours des dernières années les preuves se sont accumulées sur l'existence de discontinuités génétiques dans les systèmes marins. Néanmoins, il est admis que la plupart des populations de poissons marins présentent une structuration génétique faible due à des taux de flux génétique élevés entre elles (Ward *et coll.*, 1994; Waples, 1998; DeWoody et Avise, 2000; Palumbi, 2003).

Pour les trois espèces étudiées, et selon le marqueur analysé, nous avons observé des résultats différents mais qui sont en accord avec cette tendance.

Les deux espèces pour lesquelles on a pu faire l’analyse de locus microsatellites, *H. flavolineatum* et *S. partitus*, sont caractérisées par des valeurs faibles de *F*<sub>ST</sub>, ce qui est attendu pour des espèces marines qui ont un grand potentiel de dispersion (Hoarau *et coll.*, 2002; Knutsen *et coll.*, 2003; Ovenden et Street, 2003; Knutsen *et coll.*, 2009; McCusker et Bentzen, 2010). Pour *H. flavolineatum* il a été mis en évidence l’absence de structure génétique significative entre toutes les localités analysées. Ceci a été aussi montré par Schlüeter (1998) qui a étudié des populations d’ilots de la Floride et qui a détecté des quantités significatives de flux génique qui homogénéise la diversité génétique de ces populations. En revanche, une étude à plus grande échelle géographique, le long des Caraïbes, des Bahamas et des ilots de la Floride (Purcell *et coll.*, 2006) a décrit une structuration génétique faible suivant un patron d’isolement par la distance, ce qui suggère des niveaux plus bas d’échanges à l’échelle régionale qui seraient dus aux limites de capacité de dispersion de cette espèce.

Par ailleurs, chez *S. partitus* nous avons obtenu des indices de différentiation génétique ($F_{ST}$) faible mais statistiquement significatifs entre toutes les localités comparées deux à deux sauf pour les comparaisons avec La Bajada. L’analyse de la structuration de la diversité réalisées avec le logiciel STRUCTURE (Pritchard et coll., 2000) et la visualisation de la divergence des populations par une approche MDS basées sur les distances génétiques de Nei (1978) s’accordent pour indiquer que l’îlot Coco est le plus différencié par rapport aux autres localités analysées. Le marqueur mitochondrial n’a pas permis de détecter des différences entre toutes les localités, mais à permis de distinguer l’îlot Coco des autres populations, ce qui pourrait être dû au taux de mutation plus élevé qui caractérise l’évolution moléculaire des microsatellites (Hedrick, 1999). En règle générale, chez *S. partitus*, la différentiation génétique se manifeste à une échelle relativement petite entre localités séparées de quelques centaines de kilomètres (entre 300 km et 700 km). D’autres études chez cette espèce ont montré que des populations séparées par des distances inférieures entretiennent des échanges génétiques suffisants pour permettre l’homogénéisation des fréquences alléliques (Christie et coll., 2010), mais que la différentiation peut suivre un patron d’isolement par la distance lorsqu’on travaille à des échelles régionales (Purcell et coll., 2009).

La structuration observée chez *S. partitus* indique des flux génétiques limités, comme décrit chez d’autres espèces de poissons et en raison de différents facteurs (Ruzzante et coll., 2000; Planes et coll., 2001; Riginos et Nachman, 2001; Planes et Fauvelot, 2002; Rocha et coll., 2005b; Wilson, 2006). Lorsqu’on regarde les comparaisons des localités prises deux à deux, on trouve que la différenciation est largement indépendante de la distance entre sites d’échantillonnage. Ceci est en accord avec l’absence de structuration trouvée pour les deux autres espèces échantillonnées aux mêmes endroits, et nous permet d’écart er l’influence de ce facteur « distance » sur la structure génétique existant. White et coll. (2010) ont montré, au moyen de simulation de la dispersion larvaire basée sur les observations des courantes océaniques, que les distances géographiques ne prédissent pas la structuration du polymorphisme génétique. Ceci parait vrai aux échelles suffisamment faibles où la dispersion est également probable entre toutes
les populations et il n’y a donc pas de relation entre la distance et le flux génétique (Hellberg, 2009).

L’impact des distances géographiques sur la structuration génétique des populations dépend étroitement de différents aspects de la biologie de l’espèce comme le type d’œufs produits, benthiques ou pélagiques, la présence de stades larvaires pélagiques et la durée de ces phases de développement. De ces facteurs dépendent le recrutement local ou la dispersion sur de longues distances (Bowen et coll., 2006a). Shulman et Bermingham (1995) ont trouvé que les espèces avec des œufs non-pélagiques ont, en moyenne, des capacités de dispersion plus restreintes et une différentiation génétique supérieure à celles qui ont des œufs pélagiques. Par ailleurs, le lien entre la durée du stade larvaire et la structuration génétique des populations a été objet de nombreuses études, sans aboutir dans la plupart des cas à une relation clairement établie (Shulman et Bermingham, 1995; Taylor et Hellberg, 2005; Bay et coll., 2006; Bowen et coll., 2006a).

Les données suggèrent que les courants marins peuvent aussi modeler la structuration génétique des populations par le transport des larves d’une population à une autre population ou en agissant comme des barrières avec différents degrés de perméabilité (Gaylord et Gaines, 2000; Muss et coll., 2001; Taylor et Hellberg, 2006). Au niveau de la Caraïbe, Shulman et Bermingham (1995) en analysant un groupe d’espèces avec des traits de vie divers n’ont pas pu mettre en évidence des différences significatives entre des localités séparées par les principaux courants, mais ils suggèrent l’importance possible des courants à petite et moyenne échelle qui pourraient avoir une influence local sur l’isolement de populations d’une même espèce. Dans cette perspective sont développés des modèles qui permettent de prédire les patrons de dispersion larvaire et de structuration génétique basées sur des données océanographiques plus précises (Paris et coll., 2005; Galindo et coll., 2006; White et coll., 2010).

La circulation des courants peut être impliquée dans la dispersion des larves des poissons qui fraient au niveau de la plateforme insulaire et favoriser de cette façon le flux de gènes entre populations. Toutefois la différenciation génétique trouvée n’est pas homogène pour les trois espèces étudiées, ce qui pourrait indiquer l’importance des aspects spécifiques du cycle de vie sur les limites de dispersion.

Figure 6.1. Patron des courants marin autour de Cuba. Extrait de Claro et coll. (2001).

*Haemulon flavolineatum* produit œufs pélagiques (Mc Farland et coll., 1985) qui produisent des larves planctoniques d’une durée moyenne de 15 jours, durée parmi les plus courtes mesurées pour des poissons des récifs coralliens. De plus, Garcia-Cagides et coll. (1994) ont suggéré que les individus réalisent pour frayer des migrations des eaux intérieures de la plate-forme vers les régions proches du tombant. Même si cette espèce ne réalise pas les grandes migrations pour frayer décrites pour d’autres espèces comme les lutjanidés (Claro et Lindeman, 2003), la taille des populations augmentent pendant la reproduction en raison du regroupement d’individus venant de différentes endroits (Claro, communication personnelle).

*Acanthurus tractus* fraient en couples ou en grands groupes et produit des œufs pélagiques (Colin et Clavijo, 1988; Weber et Brown, 2009). L’œuf éclos et produit une larve connu sous le nom d’acronurus qui est visiblement adapté a une vie pélagique (Randall, 1961a; Leis et Rennis, 1983) d’une durée de 69 jours. De plus quelques auteurs considère que les espèces qui composent la famille des Acanthuridés appliquent une stratégie reproductive qui permet aux œufs, puis aux larves, d’être transportés sur de longue distance dans l’océan en choisissant pour
frayer des localités avec une forte influence des courants marins dirigés vers l’océan (Johannes, 1978; García-Cagides et coll., 1994).

*Stegastes partitus*, à l’opposé des autres espèces étudiées, pond des œufs dans un nid sur le fond marin et qui sont protégés par le mâle jusqu’à l’éclosion. La période larvaire pélagique a une durée moyenne de 28,8 jours (maximum de 39 jours) (Villegas-Hernández et coll., 2008). Cette espèce est sédentaire et les individus pondent leurs œufs à quelques dizaines de mètres des zones où ils se nourrissent (Robertson, 1990). Les larves des pomacentridés comme *S. partitus* sont généralement collectés près de la côte (Leis, 1991; Cowen et Castro, 1994) et probablement elles restent dans l’environnement du récif pendant toute la durée de leur stade pélagique (Leis, 1991; Sponaugle et Cowen, 1996a), affectés par un régime de courants associés à la côte (Cowen et Castro, 1994). C’est pourquoi les larves de cette espèce ne sont pas exposées dans la même mesure aux courants prédominants qui peuvent modeler la distribution des larves et la structure génétique de *H. flavolineatum* et *A. tractus*.

Par ailleurs, il a été proposé que les caractéristiques suivantes du comportement peuvent induire que les larves se développent près de la côte plutôt qu’elles se dispersent (Riginos et Victor, 2001). En effet, les larves provenant des œufs benthiques sont habituellement plus avancées dans leur développement que celle qui proviennent d’œufs pélagiques (Cowen et Sponaugle, 1997), et ainsi elles pourraient avoir une meilleure capacité pour nager de façon active et de maintenir leur position dans la colonne d’eau (Stobutski et Bellwood, 1994; Fisher et coll., 2000) et éviter la dispersion (Doherty et coll., 1996). Leis et coll. (2007) ont observé, dans une expérience *in situ*, que 80% des larves nagent de façon orientée vers le récif. Quelques auteurs suggèrent que, probablement, elles utilisent leurs capacités sensorielles pour accomplir ces déplacements en direction des récifs (voir la revue de Sale, 2004). Le comportement larvaire pourrait être un paramètre important pour comprendre la connectivité entre les populations des poissons des récifs (Bowen et coll., 2006a).

En résumé, selon le modèle proposé par Paris et coll. (2005) on peut prévoir que les espèces tels que *A. tractus* et *H. flavolineatum*, avec des capacités de dispersion semblables à celles des lutjanidés, montrent un certain isolement génétique dans la région de l’îlot Coco par rapport le reste des localités étudiées. Les résultats ne soutiennent pas cette idée pour ces espèces et indiquent que l’effet sur l’isolement de la direction des courants marins dans certaines régions est plus marqué pour des espèces telles que *S. partitus* qui présente un comportement reproducteur qui peut favoriser une rétention local des larves de manière plus marquée. D’ailleurs,
cette espèce présente un niveau de fragmentation génétique plus important, même dans des localités affectées par des patrons de courants marins qui favorisent le mélange des larves provenant de récifs différents, comme les localités du sud et sud-ouest de l’archipel (figure 6.1, 7.1, tableau 3.3).

Malgré l’homogénéité génétique observée pour les sites échantillonnés chez la gorette jaune et le chirurgien marron, il a été décrit que des populations peuvent être démographiquement indépendantes du point de vue écologique, même si la migration est suffisamment forte pour empêcher la différenciation génétique. L’absence de cette dernière n’indique donc pas nécessairement des échanges fréquents entre les populations (Waples, 1998; Waples et Gaggiotti, 2006). La littérature scientifique récente sur ce sujet indique qu’un pourcentage important des larves retournent au récif natal (Swearer et coll., 1999; Jones et coll., 2005; Almany et coll., 2007; Planes et coll., 2009; Christie et coll., 2010). Ainsi, Swearer et coll. (2002) ont conclu que l’autorecrutement a lieu pour différentes espèces présentant différents traits d’histoire de vie et vivant dans différents habitats. La capacité des larves à rejoindre leur population d’origine peut être un phénomène commun à différentes espèces marines. À Cuba, ce comportement a été observé dans des populations de quelques espèces de gorettes et de vivaneaux vivants au sud ouest de l’île (Lindeman et coll., 2001). D’autre part, Paris et coll. (2005) ont estimé des taux « virtuels » d’autorecrutement entre 37 et 80% pour les larves de mérous et de vivaneaux ; espèces qui font des migrations pour frayer et dont les larves sont fréquemment observées hors de la plateforme (Claro et Lindeman, 2003), ce qui favorise par ailleurs l’homogénéisation.

VI.2 Variabilité génétique

L’abondante littérature scientifique qui traite de la diversité génétique des poissons marins indique, en général, une grande variabilité de celle-ci (Buonaccorsi et coll., 2001; Waldick et coll., 2002; Adams et coll., 2006). Pour les indicateurs de la diversité au niveau des locus microsatellites, ils oscillent généralement entre des valeurs modérées et grandes (Buonaccorsi et coll., 2001; Adams et coll., 2006; Purcell et coll., 2006; Wilson, 2006), en raison du taux de mutation élevé qui caractérise ce type de marqueur (entre $10^{-2}$ chez l’homme (Weber et Wong, 1993) et $10^{-6}$ chez la drosophile (Schug et coll., 1997)). Toutefois, dans le cas des espèces menacées (DeSalle et Amato, 2004) avec des tailles des populations faibles, cette diversité peut être réduite.

Les estimations de diversité génétique obtenues au cours de ce travail sont un peu plus élevées que celles obtenues pour des locus microsatellites dans d’autres travaux.

La variabilité génétique de la région non codante de l’ADNmt présente des caractéristiques différentes chez les trois espèces analysées. Nous avons classé ces espèces dans des catégories différentes en fonction des estimations de la diversité haplotypique (*h*) et de la nucléotidique (*π*), comme proposé par Grant et Bowen (Grant et Bowen, 1998) et Avise (2000).

Chez *Stegastes partitus* la diversité génétique est caractérisée par une valeur élevée de *h* et une valeur faible de *π* (tableau 6.1). Ces valeurs sont à l’intérieur de l’intervalle de variation observé par Bay et coll. (2006) pour huit espèces de pomacentridés, et par Shulman et Bermingham (1995) pour deux espèces de cette même famille. En règle générale, une diversité génétique caractérisée par une diversité nucléotidique faible et une diversité haplotypique élevée est observée de façon récurrente chez les poissons marins (Grant et Bowen, 1998). Cette organisation de la diversité génétique suggère une croissance rapide à partir d’une population ancestrale d’effectif efficace (*Nₑ*) faible, et un temps écouté suffisamment long pour générer une variabilité haplotype par mutations, mais insuffisamment long pour permettre l’accumulation d’un grand nombre de différences entre les séquences analysées. Ce scenario est en accord avec l’existence d’un événement fondateur ou avec des fluctuations de la taille de la population avec des goulots d’étranglement.
Chez *Acanthurus tractus*, la diversité génétique est caractérisée par des valeurs élevées de diversité haplotypique et nucléotidique (tableau 6.1). Cette dernière valeur, qui dépasse 8%, est remarquable compte tenu que π varie en règle générale entre 0 et 10% (Grant et Bowen, 1998). Les acanthuridés pourraient avoir un taux de mutation particulièrement élevé dans la région non codante de l’ADNmt. Ainsi, chez trois espèces du genre *Naso*, on observe 7,3% < π < 13,6%, cette dernière valeur étant la plus grande trouvée pour ce marqueur chez les poissons (Klanten et coll., 2007; Horne et coll., 2008). D’autres espèces de poissons marins ont aussi montré une évolution particulièrement rapide de cette région de l’ADN mitochondrial (McMillan et Palumbi, 1997; Bowen et coll., 2006b). Une diversité forte, tant au niveau haplotypique que nucléotidique, peut correspondre à l’existence de lignées bien différenciées qui coexistent après un contact secondaire entre populations séparées pendant un temps long, ou à des populations de taille efficace grande et stable pendant une longue période de leur histoire évolutive et jusqu’à aujourd’hui (Grant et Bowen, 1998). Néanmoins, s’il existe vraiment chez cette espèce un taux de mutation de l’ADNmt particulièrement élevé, les scénarios démographiques précédemment décrits pourraient être remis en question.

Enfin, chez *Haemulon flavolineatum*, la diversité génétique présente un état intermédiaire à ceux observés chez les deux autres espèces, une diversité haplotypique grande mais entre les valeurs observées pour les deux autres espèces et une diversité nucléotidique modérée qui est aussi incluse entre les valeurs observées pour les deux autres espèces (tableau 6.1).

Les relations entre les haplotypes des individus échantillonnés, présentées sous la forme d’un réseau d’haplotypes, nous permettent de faire des hypothèses sur l’histoire évolutive des populations analysées. Chez *S. partitus*, nous avons obtenu une représentation sous forme d’étoile, avec un haplotype très abondant et plusieurs haplotypes dérivés qui s’en distinguent par un ou deux pas mutationnels. Cet haplotype est présent dans toutes les localités, ce qui peut indiquer l’existence
de flux génétiques récents entre elles (Horne et coll., 2008). Il est probablement le plus ancien, étant donné qu’il existe une corrélation positive entre la fréquence et l’âge des haplotypes, et il a probablement été présent dans les populations pendant un temps long (Castelloe et Templeton, 1994; voir la revue de Posada et Crandall, 2001). En conséquence, la plupart des nouveaux haplotypes dérivent des haplotypes les plus communs, ce qui implique que les variants les plus rares correspondent aux mutations les plus récentes (voir aussi la revue de Posada et Crandall, 2001; Templeton, 2006). Ces hypothèses sont généralement vérifiées pour les séquences sélectivement neutres et en l’absence de subdivision des populations.

Le réseau d’haplotypes obtenu chez H. flavolineatum diffère du précédent. Au lieu d’un haplotype prédominant, on observe quelques variants plus fréquents, et d’autres peu fréquentes et qui dérivent des premiers, avec de nombreux haplotypes uniques interconnectés de façon complexe. Cette absence de résolution peut s’expliquer par l’homoplasie (des mutations reverses ou parallèles) (Posada et Crandall, 2001). Même si la plupart des haplotypes diffèrent de leur voisin le plus proches par un seul pas mutationnel, comme chez S. partitus, la moyenne des différences nucléotidiques entre les séquences prises deux à deux (k) est près de 8 et la diversité nucléotidique est plus grande que 1%, supérieure à ce qu’on trouve chez S. partitus. A. tractus présente la particularité de fortes divergences entre les haplotypes des individus qui constituent les populations autour de Cuba.

VI.3 Démographie

Nous avons utilisé différentes méthodes pour tester l’hypothèse de neutralité des séquences d’ADNmt. Ceux-ci donnent aussi des arguments en faveur de l’existence de changements démographiques dans les populations étudiées. Les tests de neutralité D de Tajima et Fs de Fu ont produit des estimations avec des valeurs négatives et significatives pour les trois espèces étudiées, ce qui suggère soit l’effet d’un balayage sélectif, d’une expansion de la taille des populations ou un goulot d’étranglement ancien. Parmi ces trois hypothèses, des cas d’expansion des populations ont été corroborés pour les trois espèces par l’analyse de « mismatch distribution » où la distribution du nombre des différences entre les haplotypes pris deux à deux est unimodale, ce qui est en accord avec le modèle d’expansion soudaine. Toutefois, lorsqu’on regarde les paramètres de croissance obtenus pour les trois espèces on se rend compte que les temps correspondant au processus d’expansion sont différents, ce qui indique soit, que ces événements ne se sont pas produit au même moment, et par conséquent ne sont pas le produit d’un même
facteur causal (e.g. changement écologique) ou qu’un même agent n’a pas affecté les espèces avec la même ampleur. D’autre part les résultats des analyses bayésiennes réalisées avec la région de contrôle de l’ADNmt en utilisant le logiciel Lamarc montrent que les populations continuent à croître. Malgré les différences génétiques trouvées chez *S. partitus* entre l’ilot Coco et le reste des individus regroupés, il existe des taux élevé et différent de flux génétique entre ces deux populations identifiées, plus précisément le flux est plus faible vers l’ilot Coco. Ces résultats confirment que la population de l’ilot Coco, bien que de façon faible, est plus isolée du reste.

L’espèce *S. partitus* est, parmi les trois espèces, la plus proche de l’événement d’expansion soudaine, avec une valeur de $\tau = 1,66$ ce qui correspond au temps mesuré sur la base du taux de mutation de la région non codante. Dans cette perspective les évènements d’expansion de *H. flavolineatum* et *A. tractus* sont plus éloignées dans le temps avec des valeurs de $\tau = 6,0$ et $\tau = 15,91$ respectivement. Les estimations de $\theta$ montrent aussi des différences pour les trois espèces analysées, la plus petite valeur étant obtenue pour *S. partitus*. Ces trois espèces sont parmi des plus communes dans les récifs cubains (Claro, 1994), mais *S. partitus* est la deuxième parmi les poissons les plus abondants dans les récifs des petites Antilles et Cuba (Claro et coll., 1998), avec des populations composées de milliers d’individus dans des localités situées au nord de La Havane (González-Sansón et Aguilar, 2003). Toutefois, une grande variance du succès reproducteur peut provoquer la diminution de la taille efficace des populations et la coalescence tôt des lignées existantes (Hedgecock, 1994).

Lorsqu’on regarde le signal produit par les marqueurs microsatellites chez *S. partitus* et *H. flavolineatum*, on ne détecte aucun changement dans la taille des populations analysées. Elles semblent plutôt stables avec des taux de croissance exponentielle ($g$) très proches du zéro. Cette contradiction apparente peut refléter les caractéristiques intrinsèques des marqueurs, c’est-à-dire que dans certaines cas il n’y a pas de coïncidence entre les données nucléaires et les données mitochondriales (Zink et Barrowclough, 2008). Comme le temps de coalescence est une fonction positive de la taille efficace de la population ($N_e$), qui est quatre fois plus grand pour l’ADN nucléaire que pour l’ADNmt, ces marqueurs donnent une information sur la démographie et la dispersion ayant eu lieu dans des temps différents, le passé plus récent pour l’ADNmt et qui correspond aux générations entre $4N_e$ et $N_e$ et le passé lointain pour l’ADN nucléaire, aux générations avant $4N_e$ (Templeton, 2006). C’est pourquoi si les événements d’expansion ont eu lieu
entre 4Ne et Ne générations on n’est pas capable de les détecter avec les marqueurs microsatellites.

D’autre part, et avec ce même marqueur on a obtenu des arguments en faveur de l’occurrence d’un goulot d’étranglement chez les populations de S. partitus, pour le modèle de mutation IAM. Strictement parlant, l’excès de diversité génique par rapport à la diversité génique attendue à l’équilibre mutation dérive (principe du test utilisé par le logiciel Bottleneck) a été seulement démontré par ce modèle par Maruyama et Fuerst (1985). Même si on ne connait pas le modèle selon lequel les microsatellites étudiés mutent, il a été décrit que lorsqu’ils deviennent légèrement du modèle SMM, fait qui est très commun (Balloux et Lugon-Moulin, 2002), ils montreront un excès de diversité génique en conséquence d’un goulot d’étranglement (Cornuet et Luikart, 1996).

VI.4 La problématique A. tractus/A. bahianus

VI.4.1 Les arguments de l’ADNmt

Lorsqu’on compare A. tractus et A. bahianus avec les deux autres espèces du genre qui habitent dans la même région (A. chirurgus et A. coerulus) on se rend compte qu’elles sont très sensibles aux barrières écologiques imposées par le flux d’eau douce produit par l’embouchure des rivières Orénoque et Amazone (Rocha et coll., 2002). C’est pourquoi ce cas a été considéré comme un exemple classique de forte structuration de populations produit par une barrière géographique qui s’est maintenue pendant une période évolutive longue (Fig 17.1 chez Helfman et coll., 2009). Notre analyse des relations entre les haplotypes d’A. tractus trouvés autour de Cuba a regroupé sept individus avec les haplotypes ABA3 et ABA4 de la lignée sud décrite par Rocha et coll. (2002) comme A. bahianus. Ces haplotypes sont séparés par neuf mutations du clade qui défini l’espèce A. tractus des Caraïbes (Bernal et Rocha, 2011).

Nos données ne réfutent pas les résultats préalables qui montrent une divergence génétique importante entre des populations des deux provinces biogéographiques. Toutefois, la présence d’haplotypes du sud dans les Caraïbes (i.e. Cuba) suggère que des individus du sud ont été capables de traverser la barrière d’eau douce mentionnée ci-dessus. Etant donné que la taille de notre échantillon est quatre fois plus grande que celle de Rocha et coll. (2002) (149 vs 38), et que les haplotypes des individus du sud représentent seulement 5% de notre échantillon, il est compréhensible que ces haplotypes n’aient pas été détectés dans la première étude sur un petit échantillon.
VI.4.2 Dispersion des larves ou des adultes


VI.4.3 Un ou plusieurs événements de dispersion

Les données disponibles sont insuffisantes pour estimer précisément la date de l’événement de dispersion vers les Caraïbes. Les haplotypes identifiés dans notre étude ont des positions terminales et internes dans le réseau (figure 5.7), mais ils représentent seulement une petite fraction de la diversité révélée par le travail précédent. Comme par ailleurs le nombre de substitutions nucléotidiques entre les haplotypes du sud trouvés à Cuba et les séquences pour cette région qui sont disponibles, les haplotypes ABA3 et ABA4, est faible (un à quatre mutations), il est probable que la migration d’*A. bahianus* vers la région des Caraïbes ait eu lieu relativement récemment. La distribution actuelle des haplotypes suggère aussi
qu’au moins un événement ancien de colonisation a eu lieu beaucoup plus tôt dans l’histoire évolutive d’*A. bahianus/A. tractus*. Si on fait l’hypothèse de 2% de divergence par million d’année, taux qui a été largement utilisé pour estimer des temps de divergence à partir du gène cytb (*Brown et coll.*, 1979), la divergence de 2.0% entre les séquences de ces deux espèces indique qu’elles ont divergé il y a environ un million d’années, soit bien après la mise en place de l’organisation actuelle du bassin des rivières Orénoque-Amazone (c. 10 million d’années *Lovejoy et coll.*, 1998). Ceci fait penser qu’un événement de colonisation a eu lieu à l’occasion d’une dispersion restreinte à travers la barrière déjà établie. Cette colonisation par une population de petite taille pourrait expliquer la monophylie réciproque des populations (ou espèces) présentent des deux cotés de la barrière.

**VI.4.4 Quelques remarques**

L’existence des haplotypes d’*A. bahianus* dans la province biogéographique des Caraïbes soutient l’hypothèse d’enrichissement de la faune dans cette province (*Robertson et coll.*, 2006; *Rocha et coll.*, 2008). Le patron de couleurs qui caractérise *A. bahianus* n’a pas été observé dans l’échantillon analysé. Il est toutefois possible qu’on n’est pas fait assez attention à ce point lors de la collecte des individus car il y a une assez forte variabilité de la coloration chez cette espèce. A l’occasion de futurs échantillonnages, il sera nécessaire de rechercher les différences morphologiques que séparent *A. tractus* et *A. bahianus* pour voir s’il y a une correspondance avec la divergence génétique. Plus de données, en particulier sur l’ADN nucléaire, sont maintenant nécessaires pour élucider la question de l’isolement génétique entre ces deux espèces d’acanthuridés. On peut se poser la question suivante : est-ce que *A. bahianus* est réellement présent dans les Caraïbes ou ces haplotypes sont-ils le produit événements anciens d’hybridation qui ont amené à de l’introgression mitochondriale ? Déterminer si ces deux espèces se croisent ou si seulement elles cohabitent là où se superposent leurs aires de répartition est un aspect crucial de futurs travaux sur ces poissons.

**VI.5 Perspectives**

Ce travail de thèse représente la première étude de la structure génétique des populations de poissons marins autour de Cuba. Il nous a permis de révéler des caractéristiques des populations de trois espèces et montrer qu’au moins a l’échelle étudiée et dans les conditions écologiques de l’archipel cubain les caractéristiques biologiques des espèces peuvent déterminer le niveau d’isolement géographique des populations. Les estimations de paramètres démographiques ont aussi montré...
Discussion générale

des disparités entre les espèces, ce qui peut être un indice d’histoires évolutives différentes, ainsi que des sensibilités différentes aux événements écologiques qui se sont produits dans la région pendant les derniers milliers d’années. Les résultats nous apportent de nouvelles interrogations pour des travaux futurs.

Il serait utile d’obtenir plus de données sur les locus microsatellites chez *S. partitus* afin d’augmenter la puissance pour mesurer la structure génétique des populations.

L’analyse de l’ADN nucléaire chez *A. tractus* et *A. bahianus* permettra de répondre à la question de l’isolement génétique entre ces deux espèces d’acanthuridés.

Il faudrait échantillonner et ajouter dans les analyses des individus provenant de la zone sud orientale de l’île, sachant que des travaux préalables ont montré des pourcentages élevés d’autorecrutement dans cette zone.


Enfin, nous étendrons ce type d’étude à d’autres organismes marins avec des capacités de dispersion et des besoins écologiques différents pour pouvoir faire des généralisations et identifier les spécificités des patrons phylogéographiques qui caractérisent les populations autour de Cuba.
Matériels et Méthodes
VII. MATERIELS ET METHODES

VII.1 Echantillonnage

Nous avons échantillonné des individus adultes ( mâles et femelles) des trois espèces étudiées (*Stegastes partitus, Haemulon flavolineatum* et *Acanthurus tractus*) dans les cinq localités suivantes situées autour de l’archipel cubain : Ilot Avalos (Archipel des Canarreos), La Bajada (Péninsule de Guanahacabibes), Plage Baracoa (au nord de La Havane), Ilot Coco (Archipel des Jardins du Roi) et Ilot Breton (Archipel des Jardins de la Reine) pendant la période de 2005-2010 (figure 7.1, tableau 7.1). L’échantillonnage a été répété dans le temps dans quelques localités et pour les espèces *A. tractus* et *H. flavolineatum*. L’échantillonnage a été réalisé avec des filets pour l’espèce la plus petite et moins rapide (*S. partitus*) et avec des fusils sous-marins pour les autres deux.

Figure 7.1 Localités d’échantillonnage autour de l’île de Cuba pour les espèces *S. partitus, H. flavolineatum* et *A. tractus*. 
Tableau 7.1. Dates et localités d’échantillonnage

<table>
<thead>
<tr>
<th>Localité</th>
<th>Espèces</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S. partitus</td>
<td>H. flavolineatum</td>
<td>A. tractus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fév. 2010</td>
<td>Fév. 2010</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Plage Baracoa</td>
<td>Jui, 2006</td>
<td>Mai, 2007</td>
<td>Fév. 2006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mar, 2009</td>
<td>Nov. 2009</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fév. 2008</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Ilot Breton</td>
<td>Déc. 2008</td>
<td>Déc. 2008</td>
<td>Déc. 2008</td>
<td></td>
</tr>
</tbody>
</table>

VII.2 Traitement des échantillons

La plupart des échantillons a été traitée aussitôt après capture. Néanmoins, pour quelques cas, cela n’a pas été possible et les animaux ont été congelés à -20°C jusqu’à leur traitement ultérieur. Par la suite des fragments de tissus correspondants aux nageoires caudales, pectorales et dorsales ont été coupés et stockés dans l´éthanol 95% pour conserver le matériel génétique.

VII.3 Extraction de l’ADN

L’ADN a été extrait soit en utilisant du phénol-chloroforme et la méthode de Sambrook et coll. (1989), en prélevant un fragment de tissu d´environ 8,0 mm², soit en utilisant la méthode d´extraction avec résine Chelex™ (Walsh et coll., 1991) selon Estoup et coll. (1996), à partir de 2,0 mm² de tissu.

La qualité des extractions obtenues par une ou l’autre de ces deux méthodes a été vérifiée au moyen de l´électrophorèse en gel d´agarose 1% en utilisant le tampon TBE 0.5% (Tris-Acide Borique 40 mM, EDTA 1 mM). Pour la révélation, on a immersé le gel dans une solution de Bromure d´éthidium 0.5 μg/ml et ensuite regardé le résultat avec un transilluminateur de lumière UV.
VII.4 Amplification de l’ADN microsatellite

Les marqueurs microsatellites pour *S. partitus* et *H. flavolineatum* ont été choisis parmi ceux décrits par Williams *et coll.* (2003; 2004) (tableaux 7.2 et 7.3). Aucun marqueur microsatellite n’a été pas utilisé pour *A. tractus*.

La plupart des génotypes des individus ont été obtenus de façon automatique (avec un séquenceur capillaire), sauf quelques uns qui ont été obtenus en utilisant l’électrophorèse en gel dénaturant de polyacrylamide. Les conditions d’amplification par PCR sont les suivantes (volume final de 10 µL pour l’automatique et de 20 µL pour le manuel) : 5-20 ng d’ADN dilué dans le tampon TE 1X, du tampon d’amplification 1X, 200 µM de dNTPs, 1,5 mM de MgCl₂, 0,5 µM (0,2 µM pour le procédé manuel) de chacune des amorces et 0,5 u (1u pour le procédé manuel) de Taq DNA Polymérase (GoTaq Promega). Pour le locus AAC43 de *H. flavolineatum* 1u de polymérase a été utilisé dans tous les cas. Le volume final a été complété avec de l’eau désionisée stérile.

Pour la méthode automatique, les amorces ont été marquées avec les fluorochromes VIC, 6-FAM, NED et PET (tableaux 7.2 et 7.3). Le cycle d’amplification par PCR a été le suivant : 3 min à 95°C de dénaturation initiale et 34 cycles de 45s à 92°C pour la dénaturation de l’ADN, 45s à 53°C pour l’hybridation des amorces à l’ADN cible, et 1min à 72°C pour l’extension des chaînes nucléotidiques. Au bout du programme on a ajouté 10min à 72°C pour compléter l’extension des produits d’amplification.

**Tableau 7.2. Amorces pour l’amplification des locus microsatellites utilisés dans l’étude des populations de *S. partitus* de l’archipel cubain. Chaque fluorochrome est marqué en couleur.**

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>AMORCES</th>
</tr>
</thead>
</table>
| *SpGATA16* | Direct 5´-CCTTTTTCAACTAAATTAGATAGA-3´  
Reverse 5´-VIC-CTTGTGTCTTGTGTAGTGTGT-3´ |
| *SpAT40* | Direct 5´-NED-CTGGGCTTCTCTTTTTGTT-3´  
Reverse 5´-AGTGGCTCAGAGTTTTCTC-3´ |
| *SpAAC41* | Direct 5´-6-FAM-CGAAGGCTGTGTTTAGTTACA-3´  
Reverse 5´-GCCCCGTCACTGACAGTCTGCTG-3´ |
| *SpGATA40T* | Direct 5´-ATGCCTCACAATGATGTATATTT-3´  
Reverse 5´-PET-TTGCCCTGCTGATAATTACG-3´ |
Matériels et Méthodes


<table>
<thead>
<tr>
<th>LOCUS</th>
<th>AMORCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>HfAAC3</td>
<td>Direct  5´-<strong>VIC</strong>-GCTCCATGCTGAATGTGAT-3´</td>
</tr>
<tr>
<td></td>
<td>Reverse 5´-ATGGTCTGTGCTTTGAGGCTGTGAT-3´</td>
</tr>
<tr>
<td>HfAAC10</td>
<td>Direct  5´-CACGGAGAAAGGAATAGCAT-3´</td>
</tr>
<tr>
<td></td>
<td>Reverse 5´-<strong>PET</strong>-TGCGAGAGAGAGGTTTCA-3´</td>
</tr>
<tr>
<td>HfAAC37</td>
<td>Direct  5´-CCACAGAGAGGTACAGAC-3´</td>
</tr>
<tr>
<td></td>
<td>Reverse 5´-<strong>NED</strong>-CCGCTGTTTTCAATTTATCT-3´</td>
</tr>
<tr>
<td>HfAAC41</td>
<td>Direct  5´-<strong>6-FAM</strong>-CTCTGCTGGTTTTCTCTCTTGTGAT-3´</td>
</tr>
<tr>
<td></td>
<td>Reverse 5´-CAGGACGGTTTCAAGTGTCAG-3´</td>
</tr>
<tr>
<td>HfAAC43</td>
<td>Directo 5´- <strong>NED</strong>-CTGACCTTGTGACAAAAA-3´</td>
</tr>
<tr>
<td></td>
<td>Reverse 5´-CGTGAGTCAAACATACACTGA-3´</td>
</tr>
<tr>
<td>HfAAC46</td>
<td>Direct  5´-CCAGGGAGGAGAGAGT-3´</td>
</tr>
<tr>
<td></td>
<td>Reverse 5´-<strong>VIC</strong>-AGGGGTCAAAAGGAAAGTCCATG-3´</td>
</tr>
</tbody>
</table>

* L’unité de base répétée de tous les microsatellites est AC sauf pour le locus HfAAC43 où elle est GT.

VII.5 Génotypage

Les individus ont été analysés dans un séquenceur automatique ABI3130 XL Genetic Analyser en utilisant GS500 (-250) LIZ comme marqueur de taille et avec le logiciel Genemapper 3.0 (Applied Biosystems).

Pour les échantillons analysés sur gel d’acrylamide, la taille des allèles a été déterminée en utilisant comme marqueurs des génotypes préalablement connus. Dans le cas des allèles présentant des difficultés pour la lecture de la taille on a refait migrer l’échantillon sur gel pour vérifier les résultats.
VII.6 Amplification de l’ADN mitochondrial


Les PCR ont été effectuées avec les mêmes concentrations de réactifs que pour l’analyse manuelle des microsatellites, mais dans un volume final de 50 µL. Nous utilisons le programme décrit précédemment pour les microsatellites, mais en faisant 40 cycles et des températures d’hybridation entre 52°C et 54°C en fonction des paires d’amorces utilisées. Pour le cytb, on a modifié le programme avec les cinq premiers cycles à 47°C et le reste à 50°C comme température d’hybridation.

Avant le séquençage, les produits d’amplification ont été purifiés pour éliminer l’enzyme et les amorces résiduelles. A cette fin, on a utilisé le kit Nucleospin Gel and PCR clean-up Gel extraction Nucleospin®Extract II (Macherey-Nagel) (www.mn-net.com).

VII.7 Séquençage de l’ADN mitochondrial


Le mélange de réaction a été préparé à partir de 20 à 40 ng de produit de PCR, 3 µL de tampon 5X, 1 µL d’amorces (3,3 µM) et 2 µL de Mix BigDye (Promega) pour être utilisé avec le programme suivant: 1 min à 96°C, 24 cycles de: 10s à 96°C, 5 s à 50°C et 4 min à 60°C. Ensuite, l’ADN a été précipité en rajoutant 2 µL de EDTA 125 mM, AcNa 3M et 50 µL d’éthanol absolu et en laissant 15 min à l’abri de la lumière. Puis on a centrifugé pendant 20 min à 13 500 rpm et le surnageant a été soigneusement éliminé. L’étape suivant est un lavage avec de l’éthanol 70%, suivi par une centrifugation de 7 min à la même vitesse. Finalement, l’alcool a été enlevé et le culot séché à 37°C une heure minimum, voire une nuit. Le culot est resuspendu dans de la diformamide 5%.
VII.8 Traitements préliminaires des données

Chaque allèle microsatellite a été inclus dans une classe d’allèle nommée en accord avec son rang de taille (procédé connu sous le nom de “binning”) étant donné que la longueur théorique du motif répété n’est pas toujours en accord avec celle déterminé par le séquenceur automatique (Amos et coll., 2007). Ce procédé a été implémenté dans le logiciel FLEXIBIN, qui offre le plus de précision et de vitesse (Amos et coll., 2007).

Le programme Micro-Checker 2.2.3 (Van Oosterhout et coll., 2004) a été utilisé pour faire des corrections de lecture et détecter la présence d’allèles nuls.

VII.9 Analyses statistiques

VII.9.1 ADN microsatellite

Pour tester l’homogénéité de fréquences au cours du temps pour chacune des localités d’échantillonnage on a utilisé la statistique $\chi^2$ avec le logiciel CHIRXC (Zaykin et Pudovkin, 1993).

Préalablement à l’analyse des données microsatellites, on a testé le déséquilibre de liaison entre chaque paire de locus en utilisant le test basé sur le rapport de valeurs de vraisemblance (G-test) disponible dans le logiciel FSTAT 2.9.3 (Goudet, 2001). Ce test permet de tester l’association entre génotypes pour toutes les paires de locus.

VII.9.1.1 Structure populationnelle

Afin de déterminer l’existence d’une structure génétique pour l’ensemble de données indépendamment des localités d’échantillonnage, on a utilisé le logiciel STRUCTURE (Pritchard et coll., 2000). Le modèle de groupement implémenté dans le logiciel permet, en utilisant plusieurs locus, d’assigner les individus aux (K) populations avec une certaine valeur de probabilité. Il est requis que les locus soient en équilibre de liaison et il est fait l’hypothèse d’équilibre de Hardy-Weinberg dans les populations. Cette méthode a l’avantage de donner des résultats fiables d’assignation avec un nombre modeste de locus. Les paramètres ont été fixés pour utiliser l’« admixture model », ce que permet aux individus d’avoir plusieurs origines et le « correlated alleles frequencies », qui tient compte, d’après Falush et coll. (2003), que les fréquences alléliques de populations différentes sont probablement similaires à cause de la migration ou du polymorphisme ancestral partagé. Vingt itérations ont été réalisées pour des valeurs de K comprises entre 1
et 8 avec une période de burn-in de $10^4$ et un nombre de répétitions de la MCMC de $10^6$. La meilleure partition est déterminée à partir des logarithmes des valeurs de probabilité postérieure pour chaque valeur de $K$. On a aussi utilisé la méthode décrit par Evanno et coll. (2005), qui permet estimer le paramètre $K$. Celui-ci correspond à la quantité de variation entre les vraisemblances et les variances de ces probabilités pour des valeurs successives de $K$.

On a estimé les valeurs de $F_{ST}$ et sa signification statistique pour les localités d’échantillonnage prises deux à deux, en utilisant le logiciel FSTAT 2.9.3 (Goudet, 2001).

Pour montrer graphiquement les relations entre les différentes localités d’échantillonnage, on a utilisé le cadrage multidimensionnel non-métrique (Lessa, 1990) appliqué à la matrice des distances génétiques (Nei, 1978) avec le logiciel XLSTAT Version 2012.2.01 (disponible à l’adresse http://www.xlstat.com). Les distances sont transformées en conservant au mieux les différences originales entre les données et représentées dans un espace bidimensionnel. L’ajustement (goodness of fit) entre les distances est évalué en utilisant un test de stress.

Pour tester l’existence de corrélation entre les valeurs de différentiation génétique $F_{ST}$ et les distances géographiques (modèle « stepping stone ») on a utilisé le test de Mantel (Mantel, 1967) avec le logiciel XLSTAT Version 2012.2.01) Ce test repose sur la détermination de la corrélation entre deux, ou plus, matrices et permet de produire une distribution empirique de la statistique standardisée de Mantel, le coefficient de corrélation de Pearson, $r$. S’il n’existe pas de relation entre les matrices, la valeur de $r$ ne s’écartera pas de la distribution des valeurs obtenues par permutations et on acceptera l’hypothèse nulle d’absence de corrélation. Les distances géographiques plus courtes entre les localités d’échantillonnage ont été déterminées par le logiciel SIG : ArcGis 9.3. Deux matrices de distances géographiques ont été construites : une première sur la base du patron général des courants marins, et la deuxième en prenant en compte la distance minimale entre les plate-formes sud-est (Ilot Breton) et nord-est (Ilot Coco).

**VII.9.1.2 Diversité génétique**

La variabilité génétique a été déterminée à partir du calcul du nombre d’allèles par locus (Na) et de leur fréquence. On a aussi estimé les valeurs de richesse allélique (Ra) avec le logiciel FSTAT 2.9.3 (Goudet, 2001) et l’hétérozygotie observée (Ho) et attendue (He) d’après Nei (1987) avec le logiciel ARLEQUIN 3.5 (Excoffier et Lischer, 2010). Les déviations aux fréquences de l’équilibre de Hardy-Weinberg (HWE) ont été évaluées en estimant la valeur exacte de la probabilité de la
statistique $F_{IS}$ (Wright, 1965) modifié par Weir y Cockerham (1984) a partir de la méthode de chaînes de Markov implémenté dans le logiciel GENEPOP 4.0 (Rousset, 2008).

En général, en cas de comparaisons multiples, la valeur du seuil de rejet a été ajustée en réalisant la correction de Bonferroni (Rice 1989) pour un niveau de rejet de 5%. Ce test, disponible dans le logiciel FSTAT 2.9.3 (Goudet, 2001) nous permet de diminuer la probabilité d’obtenir des valeurs significatives (rejeter l’hypothèse nulle alors qu’elle est vrai).

**VII.9.1.3 Analyse démographique**

Pour détecter si les population ont traversé des réductions démographiques drastiques (goulots d´étranglement) relativement récentes on a utilisé le logiciel Bottleneck (Cornuet et Luikart, 1996). Dans une population qui a souffert récemment une réduction drastique de sont effectif efficace ($N_e$), la diversité génique sera supérieure à la diversité génique attendue à l´équilibre mutation dérive à partir du nombre d´allèles observés dans une population de taille constante (Luikart et coll., 1998). Ce logiciel estime pour chaque locus et population la distribution de la diversité génique attendue qui est comparé à la valeur observée pour déterminer s’il y a un ajustement entre les distributions ou si la distribution observée indique un excès ou déficit d’allèles. La valeur théorique est obtenue par simulation du processus de coalescence des gènes en considérant trois modèles possibles de mutation des microsatellites : IAM, TPM et SMM. L’occurrence d’autres événements démographiques tels que le flux génétique ainsi que le changement de la taille de la population au cours du temps ont été estimés en utilisant l´analyse bayésienne avec le logiciel Lamarc 2.1.6 (Kuhner, 2006). Thêta ($\Theta$) qui représente $4N_e\mu$ est calculé ($N_e$ : taille efficace de la population, $\mu$ : le taux de mutation par site par génération). Il est spécifique des organismes diploïdes. Lamarc mesure le taux de migration comme $M = m/\mu$, ou $m$ est la probabilité d’une lignée d’immigrer par génération. Une valeur de $M$ égale à 1 indique qu’il est également probable pour une séquence de migrer que pour un site de muter. Pour obtenir les taux de migration en termes de $4Nm$ l’auteur recommande de multiplier $M$ par l’estimation de thêta de la population réceptrice. La croissance de la population est estimée par le taux de croissance exponentiel $g$. Une valeur négative de $g$ indique que la population décroit. Par contre une valeur positive montre qu’elle est en croissance. Lorsque $g$ est égal zéro la population possède une taille constante dans le temps. Les estimations de $g$ ne sont pas symétriques en ordre de grandeur, étant donné qu’une valeur de $g = 10$ indique une croissance faible, mais
si $g = -10$ on est en présence d’une réduction significative de la taille de la population pour la plupart des valeurs de thêta.

**VII.9.2 ADN Mitochondrial**

**VII.9.2.1 Analyse des séquences**

Les séquences ont été corrigées avec le logiciel Bioedit 7.0.8.0 (Hall, 1999). La taille moyenne des fragments de la RNC obtenus était de 380 pb pour *S. partitus*, 367 pb pour *H. flavolineatum* et 273 pb pour *A. tractus*, pour lequel la taille obtenue du fragment du gène cytb a été de 608 pb. Les séquences ont été alignées en utilisant l’algorithme Clustal-W implémenté dans le logiciel MEGA 4.0 (Tamura et coll., 2007).

Le modèle de substitutions nucléotidiques le plus probable pour décrire la variation des séquences analysées a été sélectionné en utilisant le logiciel jModelTest 0.1.1 (Guindon et Gascuel, 2003; Posada, 2008) et suivant le « Bayesian Information Criterium » (BIC; Schwarz, 1978).

La neutralité des mutations dans les séquences comparées a été estimée en utilisant les statistiques $D$ de Tajima (Tajima, 1989) et $F_S$ de Fu (Fu, 1997) avec le logiciel ARLEQUIN 3.5 (Excoffier et Lischer, 2010). Sous l’hypothèse de neutralité on attend des valeurs proches de zéro, et par conséquent un écart significatif à cette valeur nous permet de considérer des hypothèses de sélection naturelle ou de changements démographiques. Le statistique $D$ de Tajima produit des valeurs négatives sous l’effet de balayage sélectif, d’expansion de population et de goulots d’étranglement. Par contre, si la sélection naturelle favorise le maintien du polymorphisme, il y aura plus d’haplotypes aux fréquences moyennes qu’on attendait sous la neutralité et $D$ serait positif (Templeton, 2006).

**VII.9.2.2 Diversité génétique et différenciation des populations**

Pour tester l’hypothèse nulle d’homogénéité des fréquences haplotypiques au cours du temps, les échantillons d’une même localité obtenus à des années différentes ont été comparés en utilisant le test $\chi^2$. Pour cela on a utilisé le programme CHIRXC (Zaykin et Pudovkin, 1993) qui applique la méthode de simulation de Monte Carlo suivant Roff et Bentzen (1989).

La diversité génétique a été décrite en considérant la quantité d’haplotypes, le nombre de sites polymorphiques, le nombre total de transitions, de transversions et
d’insertions-délétions obtenus avec le logiciel ARLEQUIN 3.5 (Excoffier et Lischer, 2010).

Pour l’analyse du polymorphisme des séquences, la diversité nucléotidique (\( \pi \)) et haplotypique (\( h \)) (Nei, 1987) ainsi que les écarts types ont été calculées. Le premier paramètre donne le nombre moyen de différences nucléotidiques par site entre deux séquences, indépendamment du nombre d’haplotypes et le deuxième donne une information sur le nombre et la fréquence des différents allèles d’un locus, indépendamment du rapport entre les séquences, et elle est équivalente à l’hétérozygotie de l’échantillon. Ces calculs ont été faits en utilisant le logiciel DnaSP 5.10 (Librado et Rozas, 2009).

Les réseaux d’haplotypes ont été construits afin d’identifier leurs relations ainsi qu’explorer graphiquement l’existence de possibles associations entre haplotypes et distribution géographique. Cette représentation permet de rendre compte des processus évolutifs qui ont lieu au niveau de l’espèce (Posada et Crandall, 2001). On a utilisé le méthode de « Median Joining » (MJ) (Bandelt et coll., 1999) suivi par un traitement de Maximum de Parcimonie implémentés dans le logiciel Network 4.5.1.6 software (Fluxus Technology).

Pour évaluer l’existence de populations différenciées on a utilisé la méthode bayésienne implémentée dans le logiciel BAPS 5.3 (Corander et Tang, 2007; Corander et coll., 2008). Cette approche permet de trouver le nombre de clusters le plus probables pour l’ensemble des données, l’appartenance de chaque individu à un cluster correspondant, et une information sur l’origine des ancêtres pour chaque individu en tenant compte du nombre de clusters. Les analyses ont été réalisées soit sans donner d’informations sur l’origine des individus soit en les associant à leurs localités d’échantillonnage. L’affinité aux clusters a été déterminée en calculant le changement de valeur du log de la vraisemblance marginale du modèle de mélange (« admixture ») qui a lieu lorsqu’un individu est déplacé d’un cluster source vers un cluster cible. Les fréquences des haplotypes dans les clusters identifiés ont été comparées avec le logiciel CHIRXC (Zaykin et Pudovkin, 1993).

La différentiation des localités prises deux à deux a été estimée en utilisant la statistique Snn (Hudson, 2000) avec le logiciel DnaSP 5.10 (Librado et Rozas, 2009). Cette statistique est plus puissante que toutes les statistiques basées sur les séquences, décrit par Hudson et coll. (1992), lorsque le niveau de variation est faible à modéré.
VII.9.2.3 Analyse démographique

On a examiné la distribution du nombre observé de différences entre paires d’haplotypes « mismatch distribution » (Rogers et Harpending, 1992). Cette distribution observée est comparée avec une distribution attendue des différences par paires en utilisant les paramètres d’expansion populationnel. Elle est habituellement multimodale lorsqu’on analyse des populations stables. Au contraire elle est unimodale pour les populations qui sont passées par une expansion démographique récente (Rogers et Harpending, 1992). Cette analyse nous a permis d’estimer le paramètre populationnel thêta avant (θ₀) et après (θ₁) l’événement d’expansion, ainsi que tau (τ), le moment de la croissance mesurée en unités de temps mutationnelles. Thêta dépend de la taille de la population des femelles par la formule θ = 2Neu (Ne : taille efficace de la population ; u : taux de mutation). Cette analyse a été réalisé avec les logiciels ARLEQUIN 3.5 (Excoffier et Lischer, 2010) en utilisant l’approche non linéaire des moindres carrés et DnaSP 5.10 (Librado et Rozas, 2009) pour obtenir la statistique R² (Ramos-Onsins et Rozas, 2002).

Le taux de migration et le flux génétique ainsi que le changement de la taille de la population au cours du temps ont été estimés en utilisant l’approche bayésienne avec le logiciel Lamarc 2.1.6 (Kuhner, 2006), comme pour les microsatellites.

Le calcul de Fs de Fu (1997) nous permet aussi faire des estimations démographiques et il a été donc utilisé comme une estimation indépendant des changement de la taille des populations. Il est négatif lorsqu’on a un excès d’allèles rares (allèles aux fréquences faibles) et Fu (1997) a montré qu’il est un test approprié pour détecter la croissance des populations et le « genetic hitchhiking ». De façon complémentaire on a calculé la statistique R² (Ramos-Onsin et Rozas 2002) qui est basée sur les différences entre le nombre de mutations de type singletons et la moyenne des différences nucléotidiques. Ramos-Onsins et Rozas (2002) ont estimé à partir d’une grande variété de cas que Fs de Fu et R² sont les tests les plus puissants pour détecter la croissance des populations. Le comportement de R² est supérieur pour les populations de tailles réduites, Fs est meilleur pour les populations de tailles.
REFERENCES BIBLIOGRAPHIQUES


Références bibliographiques


References bibliographiques


Références bibliographiques


Références bibliographiques


Références bibliographiques


Références bibliographiques


Références bibliographiques


**ANNEXES**

**Annexe 1.** Valeurs des probabilités associées au test de déséquilibre de liaison par paire des locus microsatellites pour *Stegastes partitus* ($\alpha = 0.0083$ et 120 randomisations).

<table>
<thead>
<tr>
<th></th>
<th>GATA16</th>
<th>AAT40</th>
<th>AAC41</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAT40</td>
<td>0.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC41</td>
<td>0.44</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>GATA40T</td>
<td>0.68</td>
<td>1.00</td>
<td>0.34</td>
</tr>
</tbody>
</table>

**Annexe 2.** Valeurs de $\chi^2$ pour le test d’homogénéité des fréquences au cours du temps pour un fragment de la RNC de *S. partitus* (*: p<0,05)*.

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>31.41</td>
<td>0.16</td>
</tr>
<tr>
<td>La Bajada</td>
<td>30.14</td>
<td>0.56</td>
</tr>
<tr>
<td>Plage Baracoa</td>
<td>33.92</td>
<td>0.56</td>
</tr>
</tbody>
</table>

**Annexe 3.** Valeurs de $\chi^2$ pour le test d’homogénéité des fréquences au cours du temps pour 6 locus microsatellite de *H. flavolineatum* (***: p<0,01).

<table>
<thead>
<tr>
<th></th>
<th>AAC3</th>
<th>AAC10</th>
<th>AAC37</th>
<th>AAC41</th>
<th>AAC43</th>
<th>AAC46</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>15.51**</td>
<td>21.03</td>
<td>16.92</td>
<td>18.31</td>
<td>19.68</td>
<td>22.36</td>
</tr>
<tr>
<td>La Bajada</td>
<td>14.07</td>
<td>19.68</td>
<td>14.07</td>
<td>19.68</td>
<td>21.03</td>
<td>27.59</td>
</tr>
<tr>
<td>Plage Baracoa</td>
<td>12.59</td>
<td>25.00</td>
<td>18.31</td>
<td>15.51</td>
<td>21.03</td>
<td>25.00</td>
</tr>
</tbody>
</table>
Annexe 4. Valeurs de probabilités associées au test de déséquilibre de liaison par paire de locus microsatellites pour *Haemulon flavolineatum* (*α* =0,0033 et 300 randomisations).

<table>
<thead>
<tr>
<th></th>
<th>AAC3</th>
<th>AAC10</th>
<th>AAC37</th>
<th>AAC41</th>
<th>AAC43</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC10</td>
<td>0.360</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC37</td>
<td>0.483</td>
<td>0.783</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC41</td>
<td>0.047</td>
<td>0.677</td>
<td>0.067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC43</td>
<td>0.363</td>
<td>0.397</td>
<td>0.180</td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td>AAC46</td>
<td>0.140</td>
<td>0.823</td>
<td>0.583</td>
<td>0.543</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Annexe 5. Valeurs de χ² pour le test d'homogénéité des fréquences au cours du temps pour un fragment de la RNC de *H. flavolineatum* (*:* p<0,05).

<table>
<thead>
<tr>
<th></th>
<th>χ²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>43,77</td>
<td>0,91</td>
</tr>
<tr>
<td>La Bajada</td>
<td>43,77</td>
<td>0,19</td>
</tr>
<tr>
<td>Plage Baracoa</td>
<td>47,40</td>
<td>0,04*</td>
</tr>
</tbody>
</table>

122
**Annexe 6.** Changement de la vraisemblance marginal du « mixture clustering model » qui a lieu lorsque les individus sont resitués du cluster source (cluster 2) aux clusters cibles (clusters 1 et 3).

Changement de la vraisemblance

**Annexe 7.** *A. tractus.* Valeurs de la statistique $S_{nn}$ obtenues pour le locus cytb et les localités d’échantillonnage prises deux à deux en utilisant tous les individus. Toutes les valeurs sont non significatives, $p>0,05$.

<table>
<thead>
<tr>
<th></th>
<th>Ilot Breton</th>
<th>Ilot Avalos</th>
<th>La Bajada</th>
<th>Plage Baracoa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>0,47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Bajada</td>
<td>0,44</td>
<td>0,52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baracoa</td>
<td>0,44</td>
<td>0,51</td>
<td>0,44</td>
<td></td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>0,44</td>
<td>0,45</td>
<td>0,50</td>
<td>0,49</td>
</tr>
</tbody>
</table>

**Annexe 8.** Valeurs de la statistique $\chi^2$ obtenues chez *H. flavolineatum* pour la comparaison des fréquences haplotypiques entre les localités d’échantillonnage prises deux à deux. Toutes les valeurs sont non significatives, $p>0,05$.

<table>
<thead>
<tr>
<th></th>
<th>Ilot Breton</th>
<th>Ilot Avalos</th>
<th>La Bajada</th>
<th>Plage Baracoa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilot Avalos</td>
<td>58,19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Bajada</td>
<td>47,07</td>
<td>55,86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baracoa</td>
<td>56,77</td>
<td>59,81</td>
<td>52,48</td>
<td></td>
</tr>
<tr>
<td>Ilot Coco</td>
<td>40,27</td>
<td>43,32</td>
<td>39,62</td>
<td>35,69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>ESPECES</th>
<th>AMORCES</th>
</tr>
</thead>
</table>
| **RNC** | Commons | CR-A 5´- AATTCTACCCCTAGCTCCAAAG -3´  
BR E 5´- CCTGAAGTAGGAACCAGATG -3´ |
| **Stegastes partitus** | CRAint 5´- CTATTCTCTCGAGATCTGC -3´  
CREint 5´- TGCCAGGAATGACTAAGTG -3´  
RNC-Bf 5´- TTTAAAGTAGGCATTCAAGA -3´  
RNC-Br 5´- AACTGATGGTCGTTACCTACTACAT -3´ |
| **Haemulon flavolineatum** | Hf Aintf 5´- TCACCCCTAGCTCCAAAGCC -3´  
Hf Eintr 5´- CCAGATGCCAGGAATAGTGC -3´  
Hf 189f 5´- TGAGTGCCAGGAATCGAATA -3´  
Hf 226r 5´- TACTCGTGAGGGGTATTACT -3´ |
| **Acanthurus tractus** | Abint1F 5´- CACCCCTAGCTCCAAAGC -3´  
Abint1R 5´- AGGAACCAGATGCCAGGA3´  
Abint2F 5´- GTATTCTATTAATCTATT -3´  
Abint2R 5´- AGGAACCAGATGCCAGGA3´ |
| **Cyto b** | Glufish 5´- CCAAGACTTTGAAAGACCAGGTGG -3´  
CB3 5´- GGCAATAGGAARTATCATTC -3´ |
The surgeonfish, Acanthurus bahianus, has crossed the Amazon – Orinoco outflow barrier.

Jessy Castellanos-Gell\textsuperscript{a,b}, Aymée Robainas-Barcia\textsuperscript{a*}, Didier Casane\textsuperscript{b}, Pedro Chevalier-Monteagudo\textsuperscript{c}, Fabián Pina-Amargós\textsuperscript{d} and Erik García-Machado\textsuperscript{a}.

\textsuperscript{a} Centro de Investigaciones Marinas, Universidad de La Habana, Calle 16, No. 114 entre 1ra y 3ra, Miramar, Playa, Ciudad Habana 11300, Cuba
\textsuperscript{b} Laboratoire Evolution Génomes et Spéciation, UPR9034 CNRS, Université Paris-Diderot and Université Paris-Sud, 91198 Gif-sur-Yvette, France.
\textsuperscript{c} Acuario Nacional de Cuba, Calle 60, Miramar, Playa, Ciudad Habana 11300, Cuba.
\textsuperscript{d} Centro de Investigaciones de Ecosistemas Costeros, Cayo Coco, Morón, Ciego de Ávila, CP 69 400, Cuba.

*: Present address: Universidad de Barcelona, Diagonal 643, 08028 Barcelona, España.

**Corresponding author:** Erik García-Machado
Centro de Investigaciones Marinas, Universidad de La Habana, Calle 16, No. 114 entre 1ra y 3ra, Miramar, Playa, Ciudad Habana 11300, Cuba.
Email: egarcia@cim.uh.cu

**Running title:** Acanthurus bahianus crossed the Amazon outflow barrier.

**Keywords**
Acanthurus bahianus, Acanthurus tractus, Amazon-Orinoco barrier, Caribbean, cytochrome \textit{b}, dispersal, mitochondrial DNA, reef fish

Aymée Robainas-Barcia and Jessy Castellanos-Gell contributed equally to this work.
Abstract

Dispersal varies among species according to different biological and environmental factors. It is known that there is strong genetic subdivision within the Ocean Surgeonfish, (*Acanthurus tractus*) and the Barber Surgeonfish (*Acanthurus bahianus*), in the Caribbean and southern Atlantic biogeographic provinces with relation to the Amazon-Orinoco outflows. We analysed *cytb* gene sequence diversity from 149 individuals collected at five localities around Cuba between October, 2006 and February, 2010. As expected, most individuals had haplotypes identical or closely related to those previously reported for the Caribbean. However, South Atlantic lineage haplotypes were also found in all surveyed localities with frequencies around 5%. This finding suggests that *A. bahianus* has dispersed in recent times across the Amazon-Orinoco barrier, probably because environmental perturbations have aided dispersal.
Introduction

The ability of marine species to disperse during different life stages, and how this influences the geographic structure of natural populations, has been one of the most important issues in marine ecology and biodiversity studies (Muss et al. 2001; Lubchenco et al. 2003; Rocha et al. 2008). However, direct and accurate estimates of dispersal are hard to obtain because many physical and biological factors impede observations of this process (Hellberg 2007). In this sense, population genetic analyses have been of the utmost importance in inferring the dynamics, dispersal patterns and connectivity of populations (Palumbi 2003; Purcell et al. 2006; Rocha et al. 2008). For example, a study of three Atlantic surgeonfishes (*Acanthurus bahianus*, *A. coeruleus* and *A. chirurgus*), distributed throughout the tropical western Atlantic from North America to southern Brazil, showed that different habitat preferences and sensitivities to ecological conditions could explain the contrasting levels of population genetic subdivision in these three species across the Amazon-Orinoco outflow barrier (Rocha et al. 2002). *A. chirurgus* shows unrestricted gene flow between both regions, while *A. coeruleus* has an intermediate condition (i.e. statistically significant genetic differentiation but no reciprocal monophyly). In contrast, there is strong genetic subdivision within *A. bahianus* with endemic haplotypes in the Caribbean and South Atlantic biogeographic province. Based on caudal and dorsal fin colour differences and genetic distances, Bernal and Rocha (2011) proposed that the Caribbean and the South Atlantic lineages of *A. bahianus* are different species, and they suggested the use of the name *A. tractus* proposed by Poey (Poey 1860) for the Caribbean lineage.

We have analysed the genetic diversity of *A. tractus* around Cuba. In this paper we provide evidence for the occurrence of the South Atlantic lineage (i.e. *A. bahianus* Bernal and Rocha, 2011) on Cuban reefs.

Materials and Methods

*Sample collection and DNA processing*

A total of 149 adult fish (males and females) previously recognized as *Acanthurus bahianus* Castelnau, 1855 (Castelnau 1855) (but recently *Acanthurus tractus* Poey, 1860) were collected from five localities around Cuba between October, 2006 and February, 2010. The sample sites were: Avalos Key, Los Canarreos Archipelago (n =
La Bajada, Guanahacabibes peninsula (n = 26); Baracoa beach, northwest Havana (n = 23); Breton Key, Jardines de la Reina Archipelago (n = 25); and Coco Key, Jardines del Rey Archipelago (n = 33) (Fig. 1). Samples were collected using a polespear while snorkelling or scuba diving. Immediately after capture fin clips were stored in 95% ethanol.

Figure 1. Location of sampling sites of *Acanthurus tractus/bahianus* around Cuba.

Total DNA was extracted by proteinase K digestion and phenol/chloroform extraction using the Phase Lock Gel kit (Eppendorf). The primers Glufish 5´-CCAATGACTTGAARAACCAYCGTTG-3´, modified from GluDGL (Meyer et al. 1990), and CB3 5´-GCCAAATAGGAARTATCATTC-3´ (Palumbi 1996) were used to amplify 830bp of the cytochrome b gene (*cytb*). Polymerase chain reactions (PCRs) were carried out in 50 µl reaction volume with 5 ng to 100 ng of genomic DNA containing 1 unit of GoTaq DNA polymerase (Promega), 0.2 µM of each primer, 0.2 nM of dNTPs, and 1.5 mM MgCl2. The reaction profile was 94°C for 3 min for initial denaturing, followed by 40 cycles at 94°C for 45 s, 50°C for 1 min, 72°C for 1:30 min and a final extension at 72°C for 5 min. PCR products were purified using the NucleoSpin Extract II kit (Macherey-Nagel), and cycle-sequenced in both directions, using an ABI Prism Big Dye terminator sequencing kit V.3 (Applied Biosystems). The fragments were resolved with an ABI 3100 automated sequencer (Applied Biosystems). The sequences of selected haplotypes were deposited in GenBank: accession numbers FJ905179 and FJ905180. The complete data set is available from the authors on request.
**DNA data analysis**

Sequences were edited using Bioedit 7.0.8.0 (Hall 1999) and aligned with the ClustalX alignment method using MEGA 5.04 (Tamura et al. 2011). The length of the analysed sequences was restricted to 608 bp in order to identify the haplotypes with respect to those previously defined by Rocha et al. (2002). The substitution model used in the phylogenetic analysis was the Tamura and Nei (1993) substitution matrix with a gamma distribution of substitution rate variation across sites (TrN + Γ with $\alpha = 0.259$) as suggested by the Bayesian Information Criterion (BIC) (Schwarz 1978) implemented in jModelTest version 0.1.1 (Guindon and Gascuel 2003; Posada 2008). The neighbor joining (NJ) method was used to estimate the phylogenetic tree. The robustness of the nodes was assessed by bootstrap with 1,000 replicates. The available sequences of the species *A. bahianus* (accession numbers AY029306 and AY029307) and *A. tractus* (accession numbers AY029308 and AY029309) were included in the analysis. Additionally, the sequences of *A. chirurgus* (AY029304 and AY029305) and *A. coeruleus* (AY029311) were included as outgroups.

Haplotype relationships were represented by a network constructed using the program Network 4.5.1.6 software (Fluxus Technology). The network was obtained by applying the Median Joining (MJ) network algorithm (Bandelt et al. 1999), and was post processed using maximum parsimony calculations to reduce the number of superfluous network links.

**Results**

The 608 bp of the mtDNA *cytb* gene fragment analysed from Cuban samples had a total of 74 polymorphic sites, all at third codon positions, and a total of 61 haplotypes for the 149 sequenced individuals.

The NJ tree inferred with the *cytb* sequences is shown in Fig. 2A. The Cuban sequences belong to two clades. The first is comprised of most of the individuals ($n = 142$) as well as the Caribbean haplotypes ABA1 and ABA2 representing *A. tractus*. The second one is comprised of seven sequences and the haplotypes ABA3 and ABA4 representing the species *A. bahianus*, which previous studies have only detected in the South Atlantic (Rocha et al. 2002; Bernal and Rocha 2011).
The relationships among the cytB haplotypes from Cuban localities further support the existence of two major haplotype groups separated by 9 mutational changes (eight transitions and one transversion) (Fig. 2B).

Figure 2. A) Phylogenetic tree showing the relationships among Caribbean and Brazilian haplotypes of *A. tractus/bahianus* with other Atlantic acanthurids. Av: Avalos key; Ba: Baracoa beach; Bj: La Bajada; Br: Breton key; Co: Coco key. B) *Acanthurus tractus* and *A. bahianus* Maximum Parsimony network of cytB haplotypes found around Cuba. Numbers inside circles indicate number of individuals excepted for haplotypes found in only one individual. Numbers outside circles correspond to haplotype identification. ABA1, ABA2, ABA3 and ABA4 haplotypes were available from Genebank (Rocha et al. 2002). Network branches correspond to one mutation unit except when indicated by slashes. Southern haplotypes were distributed as follow: 59 from Av; 60 from Bj; 61 from Co, Br and Ba.

The first group contains two common haplotypes which correspond to haplotypes ABA1 and ABA2. These haplotypes were found in 50 and 19 individuals respectively (Supplementary material 1), and appear to represent the diversification cores of *A. tractus* mtDNA haplotypes. Among the other haplotypes, eight were found in two to nine individuals and 48 were found in only one individual. The average number of nucleotide differences within this cluster was $k = 2.93 \pm 0.034$. The haplotypes from this group were identified as At-C1 to At-C58.
The second Cuban group comprised three haplotypes connected by one to four transitions substitutions to haplotypes ABA3 and ABA4 (Fig. 2B). Of the 3 haplotypes, one was present in 5 of the 7 individuals (Supplementary material 1), and was distributed among the sampled localities. The average number of nucleotide differences within this cluster (excluding haplotypes ABA3 and ABA4) was $k = 1.143 \pm 0.182$. The mean genetic distance estimated between groups ($d = 0.020 \pm 0.006$) was roughly the same as estimated previously by Rocha et al. (2002). Haplotype and nucleotide diversity estimates for *A. tractus* were relatively high $h = 0.856 \pm 0.026$ and $\pi = 0.0048 \pm 0.0003$. On the other hand the *A. bahianus* like clade was represented by only 3 haplotypes ($h = 0.524 \pm 0.209$ and $\pi = 0.0019 \pm 0.0009$) in Cuba.

**Discussion**

*The mtDNA evidence*

Compared to the other two species of the genus within the same area (*A. chirurgus* and *A. coeruleus*), *A. bahianus* and *A. tractus* appears highly sensitive to the ecological constraint imposed by the Orinoco-Amazon freshwater outflows (Rocha et al. 2002) and has been considered a textbook example of deep population structuring due to long evolutionary separation by a geographic barrier (Fig 17.1 in Helfman et al. 2009). Our analysis of cyt* b* haplotype relationships of *A. tractus* from Cuba grouped seven Cuban individuals within the southern lineage with the ABA3 and ABA4 haplotypes described as the South Atlantic *A. bahianus* (Rocha et al. 2002). These haplotypes were separated by nine mutations from the second cluster that defines the resurrected *A. tractus* for the Caribbean (Bernal and Rocha 2011).

While our results do not negate previous results showing deep population structuring, the presence of mtDNA haplotypes from the southern region in the Caribbean (*i.e.* Cuba) suggests that *A. bahianus* may be able to overcome the Orinoco-Amazon freshwater outflow barrier. Our survey sample size is about four times larger (149 vs. 38) than that of Rocha et al. (2002). The southern haplotype individuals represent only about 5% of our sample, therefore these haplotypes may not be detected in a much smaller sample despite being widespread in the Caribbean.

*Larval or adult dispersal*

The juvenile and adult stages of both *A. bahianus* and *A. tractus* are highly vagile (Robertson et al. 2005). However, its paucity in deep and muddy sea beds and its
occurrence almost exclusively in shallow reefs have been considered major factors in explaining the isolation between the Caribbean and South Atlantic lineages of these species (Rocha et al. 2002). On the other hand, it lays pelagic eggs and the larvae have a planktonic phase of up to 69 days (Sponaugle and Cowen 1996). Both these characteristics are known to enhance opportunities for dispersal driven by marine currents (Muss et al. 2001). The widespread occurrence of *A. bahianus* from Brazil to the mid-Atlantic ridge, with no genetic differentiation, suggests that low salinity conditions around the Amazon-Orinoco outflow barrier are more effective at preventing larval dispersal than current patterns (Rocha et al. 2002).

Lessios and Robertson (2006) and Floeter et al. (2008) shed considerable light on this issue when they compared population structure and gene flow in several fish species from the Pacific and Atlantic Ocean. Both studies showed that there is no single shared pattern of genetic divergence and separation times between the populations of different species. This suggests that fish larvae may sporadically disperse through a barrier as a result of environmental fluctuations that eventually modify the effectiveness of the barrier. This appears also to apply for the Amazon-Orinoco barrier in the case of *A. bahianus*. The dispersal of *A. tractus* to the South Atlantic cannot be ruled out but it is less likely as the main current runs from south to the Caribbean. Nonetheless, this needs to be verified, as it is known that other species migrate across both sides of the barrier (Floeter et al. 2008; Rocha et al. 2008).

**One or more dispersal events**

The available data is insufficient to estimate the approximate time of the dispersal event to the Caribbean. The haplotypes identified in the present study have terminal and internal positions in the network (Fig. 2B) but they do represent just a small fraction of the diversity revealed by the previous study. However, as the number of nucleotide substitutions between the southern haplotypes found in Cuba and the available sequences of ABA3 and ABA4 haplotypes are small (one to four changes), it suggests that the migration of *A. bahianus* into the Caribbean region has occurred quite recently. The current haplotype distribution also suggests that at least one ancient colonization event occurred much earlier in the *A. bahianus* / *A. tractus* evolutionary history. If we assume the widely used divergence rate of 2% million years\(^{-1}\) for estimating evolutionary time (Brown et al. 1979), the sequence divergence of *d* = 2.0% for the *cytb* gene between these two species indicates that they diverged about 1 million years ago,
well after the origin of the present configuration of the Amazon–Orinoco basin (c. 10 million years ago, Lovejoy et al. 1998). This suggests that a colonization event mediated by restricted dispersal across the established freshwater barrier, and probably associated with or followed by demographic reduction, explains the reciprocal monophyly and population divergence on each side of the barrier.

Final considerations

The occurrence of *A. bahianus* haplotypes in the Greater Caribbean biogeographic province gives further support to the faunal enrichment hypothesis (Robertson et al. 2006; Rocha et al. 2008). The color pattern identifying *A. bahianus* was not observed in the analyzed sample. However, its low frequency at the surveyed localities and the natural variation that have coloration tonalities of this species may lead them eventually undetected to a rapid inspection. During future sampling the collectors need to carefully examine the morphological difference that separates *A. tractus* and *A. bahianus* of the fish sampled to assess the match between them and genetic patterns. We suggest that further data (including nuclear DNA and color) are necessary to clarify *Acanthurus* speciation. In particular, we can ask the following question: is *Acanthurus bahianus* really present in the Caribbean or are these mtDNA haplotypes a result of ancient hybridization that led to introgression? Assessing whether these two species are interbreeding or simply coexisting in overlap areas is a crucial aspect of future work. Since the two lineages split around 1 million years ago, resolution of this situation would provide important information about how long it takes sister species to become fully reproductively isolated.
Acknowledgements

We thank Oscar Valmaña and Luis Sanchez for essential support in field work, and Gaspar González Sansón and Consuelo Aguilar Betancourt for the valuable discussions about the ecology of marine fish species. We thank Cushla Metcalfe for improving of the English of the manuscript. This study was partially financed by the Research Grant A4139-1 from the International Foundation for Science assigned to ARB and the Embassy of France in Cuba.
References


ABSTRACT

Dispersal of marine fishes mainly depends on passive transport of planktonic larvae driven by ocean currents. It allows populations to be connected through thousands of kilometers but also make them sensitive to factors limiting larval dispersion. In this regard it is known that there are several physical and biological factors that determine the distribution of marine species larvae and therefore population connectivity.

The present work aims to study the effects of geographic traits of Cuban archipelago and biological characteristics of reef fishes inhabiting it on the genetic structure of their populations. Three different species with distinct biological characteristics and broadly distributed within the Caribbean were selected: Stegastes partitus, Haemulon flavolineatum and Acanthus tractus.

Individuals were sampled from five localities distributed around Cuba during the time period from 2005 to 2010. We analyzed mitochondrial DNA (mtDNA) polymorphism: a fragment of the non coding region (NCR) for the three species and cyt b for A. tractus. The polymorphism of nuclear microsatellite loci was studied for S. partitus and H. flavolineatum.

Genetic diversity, assessed by means of Ho and He have high values when analyzing microsatellite loci. These molecular markers revealed the presence of contrasting geographic structure patterns for the two species. Low but significant pairwise \( F_{ST} \) estimates were obtained for S. partitus while H. flavolineatum was genetically homogenous. These results were supported by mtDNA analyses. Likewise, A. tractus showed no evidence of significant genetic differentiation at the level of the NCR and cyt b gene. Differences in population genetic structure of surveyed species could be the result of extant marine currents acting distinctly on species dispersion according to their reproductive behaviour (i.e. spawning sites, benthonic or pelagic eggs and capacity of larvae to staying close to the natal reef or being driven away from it).

Mitochondrial markers showed high levels of haplotypic diversity for the three species and contrasted nucleotide diversity, low for S. partitus, intermediate for H. flavolineatum and high for A. tractus. According to these observations and results of neutrality tests and mismatch distribution analyses, it is suggested that recent population expansion occurred in these species. However differences in population parameter estimations suggest that the time and the rate of population expansion were different for the different species.

The analysis of A. tractus cyt b haplotype relationships grouped some Cuban individuals with two haplotypes previously described as A. bahianus. This latter species has been recently proposed as restrict to South Atlantic and separated from A. tractus by the Amazon-Orinoco outflow barrier. The presence of mtDNA haplotypes from the southern region in Cuba suggests that A. bahianus may be able to overcome this barrier.
La dispersion des poissons de récifs coralliens dépend, pour une large part, de l’existence de larves pélagiques qui peuvent être transportées au sein des courants marins loin de la population source. De l’efficacité de ce transport dépendra l’existence de populations formant un continuum, plus ou moins connectées, tout le long des zones de récifs. Toutefois, ce phénomène de dispersion est complexe et il implique aussi d’autres facteurs, physiques et biologiques, qui influent sur la dispersion de ces organismes.

L’objectif de cette thèse a été d’étudier les effets des caractéristiques géographiques de l’archipel cubain et de la biologie des espèces sur la structure génétique et la connectivité des populations de poissons de récifs en utilisant comme modèle trois espèces présentant des traits d’histoire de vie différents : *Stegastes partitus*, *Haemulon flavolineatum* et *Acanthurus tractus*.

Nous avons échantillonné des individus adultes dans cinq localités situées autour de l’archipel cubain pendant la période 2005-2010. Le polymorphisme de marqueurs microsatellites a été étudié chez *S. partitus* et *H. flavolineatum*, celui de la région non codante de l’ADNmt chez les trois espèces et celui du gène *cyt b* chez *A. tractus*.

L’analyse de la diversité génétique des locus microsatellites révèle des taux d’hétérozygotie observés (Ho) et attendus (He) élevés à chaque locus et pour les deux espèces, mais elles présentent différents patrons d’organisation de la diversité génétique. Chez *S. partitus*, la comparaison des localités prises deux à deux nous montre une différenciation exprimée par des valeurs de $F_{ST}$ très faibles mais significatives, alors que chez *H. flavolineatum* aucune structuration géographique n’est observée. Ces résultats sont confirmés au niveau de l’ADN mitochondrial, lequel indique aussi une absence de différenciation d’*A. tractus* pour l’ensemble des localités étudiées. Les différences de structuration génétique des espèces étudiées pourraient être le résultat des effets des courants marins autour de l’île qui agissent de façon différente sur la dispersion des espèces en fonction de leur comportement reproducteur (e.g., sites de fraie, œufs pélagiques ou benthiques et capacité des larves de rester prêt des récifs natals ou d’être transportés au loin).

Pour les marqueurs mitochondrial, nous avons obtenu des valeurs élevées de diversité haplotypique pour les trois espèces, mais la diversité nucléotidique est très différente selon les espèces, faible chez *S. partitus*, grande chez *H. flavolineatum*, et très grande chez *A. tractus*. La distribution du nombre de différences entre les séquences suggère que ces populations sont passées par des périodes d’expansion démographique. Ceci est confirmé par les estimations des paramètres de tests de neutralité. Les paramètres caractérisant la croissance des populations qui ont été obtenus pour les trois espèces ne sont pas semblables, ce qui suggère que le temps et l’ampleur des événements d’expansion ont été différents pour les trois espèces.

L’analyse du gène *cyt b* nous a permis d’identifier des haplotypes chez *A. tractus* qui n’avaient été observés préalablement que chez *A. bahianus*. L’aire de répartition de cette dernière espèce a été récemment restreinte à l’Atlantique Sud sur la base de la divergence génétique avec *A. tractus* et de la séparation géographique qui est maintenu par la barrière que constitue l’embouchure des rivières Orénoque et Amazone. La présence d’haplotypes du sud à Cuba suggère que des individus du sud ont été capables de traverser cette barrière.