

## Table des matières

<b>Introduction générale</b> .....	<b>4</b>
I. Relations entre les processus micro-évolutifs et les traits de vie des espèces marines à cycle benthopélagique .....	5
A. Durée de vie pélagique et dispersion larvaire : des flux de gènes à longue distance ? .....	5
B. Fécondité et taille des populations : des effets de dérive génétique limités ? .....	12
II. Les processus d'invasions biologiques en milieu marin vus sous l'angle de la biologie évolutive .....	14
A. Les introductions biologiques en milieu marin : nombres et vecteurs .....	17
B. Paradoxe de l'invasion et capacité d'adaptation : un lien possible ? .....	21
III. Présentation du modèle d'étude : <i>Crepidula fornicata</i> , une espèce invasive au cycle de vie typique des invertébrés marins .....	24
A. Une longue phase larvaire et un système de reproduction original .....	24
B. <i>Crepidula fornicata</i> , un gastéropode emblématique des processus d'invasions biologiques en milieu marin .....	28
IV. Présentation de la démarche et structure du document .....	30
<b>Première partie : Sélection et Invasion</b> .....	<b>32</b>
Introduction .....	33
<b>Chapitre 1</b> - Développement de nouveaux marqueurs génétiques chez l'espèce invasive <i>Crepidula fornicata</i> .....	38
I. Contexte de l'étude.....	38
II. Résultats principaux de l'étude .....	38
III. Article 1: « <i>In silico</i> mining and characterization of 12 EST-SSRs for the invasive slipper limpet <i>Crepidula fornicata</i> » .....	40
IV. Développement de marqueurs AFLPs.....	46
<b>Chapitre 2</b> - Le rôle de la sélection dans les processus d'invasion chez <i>Crepidula fornicata</i> .....	47
I. Contexte de l'étude.....	47
II. Résultats principaux de l'étude .....	48
III. Article 2: « Contrasting patterns of genome-wide polymorphism in the native and invaded distribution ranges of the marine invasive mollusk <i>Crepidula fornicata</i> » .....	50
<b>Deuxième partie : Dérive, Dispersion et Invasion</b> .....	<b>89</b>
Introduction .....	90
<b>Chapitre 1</b> - Dynamique démographique, sexuelle et reproductive d'une population de <i>Crepidula fornicata</i> en baie de Morlaix .....	93
I. <i>Crepidula fornicata</i> en baie de Morlaix : une installation certaine mais une prolifération modérée .....	93
II. Présentation des méthodes d'acquisition des données des suivis benthiques et pélagiques .....	96

A.	Etude de la population benthique .....	96
B.	Analyse de la population pélagique .....	98
III.	Synthèse des résultats .....	99
A.	Description de la structure d'âge et identification des périodes de recrutement .....	99
B.	Structure d'âge, structure sexuelle et dynamique de changement de sexe .....	103
C.	Cycle annuel de reproduction .....	106
IV.	Conclusions : des conditions idéales pour une loterie familiale ? .....	109
<b>Chapitre 2</b>	<b>- Dynamique de recrutement et diversité génétique des juvéniles de <i>Crepidula fornicata</i> en baie de Morlaix .....</b>	<b>111</b>
I.	Contexte de l'étude.....	111
II.	Résultats principaux de l'étude .....	112
III.	Article 3: « Inter-annual recruitment is accompanied by significant genetic drift in the invasive gastropod <i>Crepidula fornicata</i> » .....	114
<b>Chapitre 3</b>	<b>- L'hypothèse de loterie familiale chez <i>Crepidula fornicata</i> : comparaison de la diversité génétique au cours des différentes étapes du cycle de vie .....</b>	<b>143</b>
I.	Contexte de l'étude.....	144
II.	Résultats principaux de l'étude .....	144
III.	Article 4: « Genetic variance and kin-related individuals within the larval pool support sweepstakes reproductive success in the gastropod <i>Crepidula fornicata</i> ».....	146
<b>Chapitre 4</b>	<b>- Distribution de l'hétérozygotie multilocus, déséquilibre d'identité et taille des larves chez <i>Crepidula fornicata</i> .....</b>	<b>181</b>
I.	Contexte de l'étude.....	181
II.	Résultats principaux de l'étude .....	182
III.	Article 5: « Inter-annual recruitment is accompanied by significant genetic drift in the invasive gastropod <i>Crepidula fornicata</i> » .....	183
<b>Discussion et perspectives</b>	<b>.....</b>	<b>203</b>
I.	Processus de sélection dans l'aire de distribution naturelle de <i>C. fornicata</i> : un exemple de couplage entre barrières endogènes et exogènes ? .....	204
II.	<i>Crepidula fornicata</i> : une espèce introduite avec succès hors de son aire de distribution naturelle .....	208
III.	La phase larvaire : facteur de succès et moment critique du cycle de vie .....	214
<b>Bibliographie</b>	<b>.....</b>	<b>221</b>
<b>Annexes</b>	<b>.....</b>	<b>234</b>

---

# **Introduction générale**

---

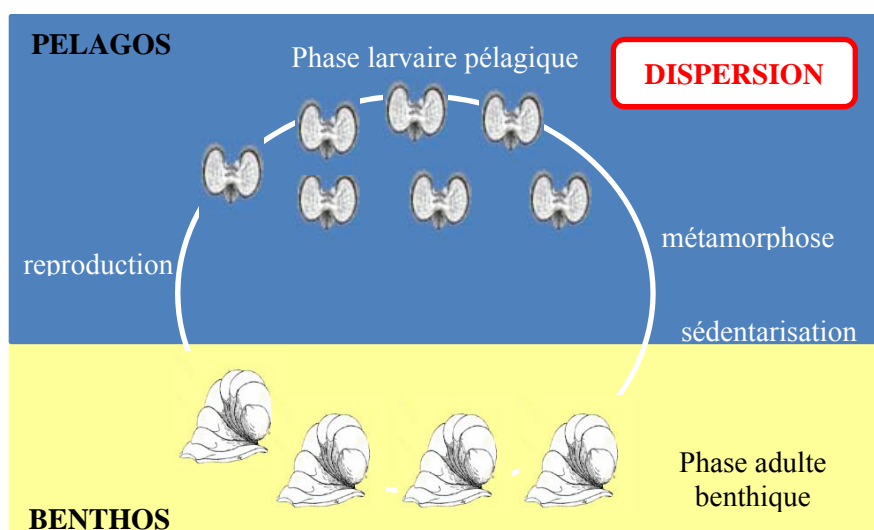


## I. Relations entre les processus micro-évolutifs et les traits de vie des espèces marines à cycle benthopélagique

Comprendre la façon dont les processus micro-évolutifs agissent dans les populations naturelles nécessite la prise en compte des particularités des espèces étudiées notamment en terme de cycle de vie et de reproduction. De très nombreuses espèces marines, en particulier chez les invertébrés, présentent des caractéristiques démographiques et reproductives originales notamment par rapport aux espèces animales terrestres : elles présentent souvent une très forte fécondité (par exemple, entre 20 et 100 millions d'ovules sont produits par une femelle du mollusque *Crassostrea gigas*), de grandes tailles de populations (par exemple, plus de mille individus par mètre carré chez *C. gigas*) et possèdent une phase larvaire souvent longévive assurant la dispersion (plusieurs semaines chez *C. gigas*). Schématiquement, ces caractéristiques devraient conduire à des diversités génétiques élevées et une dérive génétique faible ainsi qu'à des différenciations génétiques faibles (forts flux de gènes). Ces caractéristiques biologiques et leurs relations avec les propriétés génétiques des populations sont présentées et discutées ci-dessous.

### A. Durée de vie pélagique et dispersion larvaire : des flux de gènes à longue distance ?

Levin et Bridges (1995) soulignent qu'en milieu marin, sur 40 phylums dénombrés au moment de leur étude, seuls six ne présentent pas de phase dispersive. La grande majorité (plus de 70%) des invertébrés marins possède une larve planctonique et un cycle de vie benthopélagique (Mileikovski 1971). Ce type de cycle est défini par une alternance entre une phase adulte benthique peu ou pas mobile et une phase larvaire libre dans la colonne d'eau ; la larve est le principale, voire l'unique, vecteur de dispersion (Fig. 1).



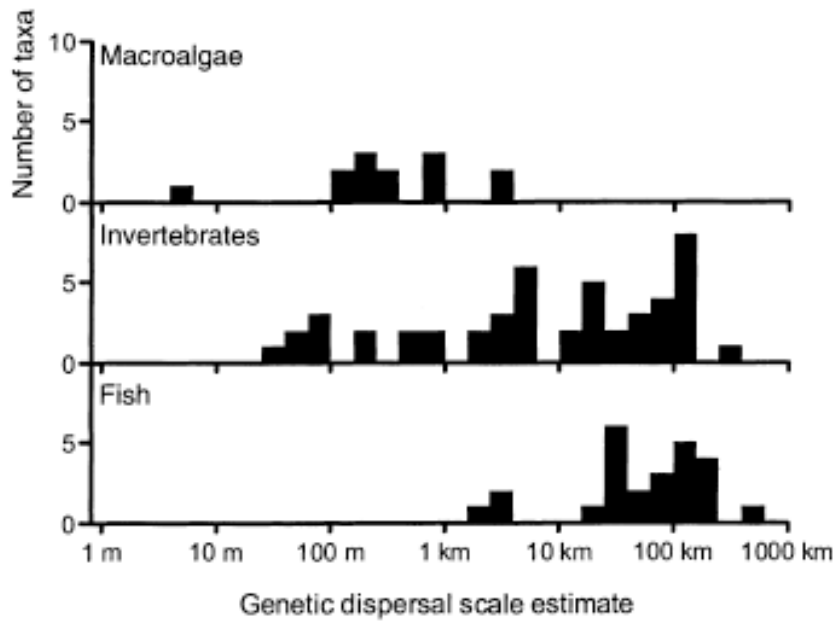
**Fig. 1** Exemple de cycle de vie benthopélagique : suite à la reproduction des adultes, les larves sont émises dans la colonne d'eau où elles peuvent être transportées loin de la population source, ou rester à proximité. Au cours de son développement, la larve acquiert les structures qui lui permettent de se métamorphoser, puis se métamorphose et se sédentarise dans les populations benthiques.

La dispersion est un processus écologique et évolutif majeur : elle joue un rôle fondamental dans la dynamique des populations et définit les flux de gènes et patrons de connectivité d'une espèce (Levin 2006, Pineda *et al.* 2007). La dispersion des larves va ainsi influencer de nombreux processus éco-évolutifs dont (1) la dynamique des populations en définissant à chaque génération les conditions initiales de la phase benthique (via la quantité et l'origine des larves), (2) les potentialités d'expansion ou de « fuite » des espèces en réponse à des changements des conditions environnementales ou (3) les capacités d'adaptations locales (Gaines et Roughgarden 1995, Shanks 2009, Levin 2006, Selkoe et Toonen 2011).

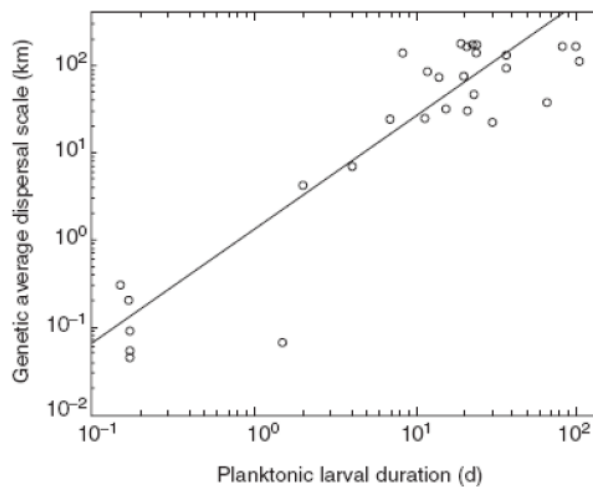
La durée de vie larvaire (ou « Pelagic Larval Duration », PLD) a largement été utilisée comme indicateur du potentiel de dispersion des espèces marines à cycle benthopélagique: l'hypothèse sous-jacente est que le temps passé par les larves dans la colonne d'eau avant leur sédentarisation est corrélée positivement avec la durée du transport larvaire et donc la distance de dispersion larvaire. A titre d'exemple, la distance réalisée par une larve du gastéropode *Haliotis rubra*, dont la durée de vie larvaire est de 6 jours, a été estimée inférieure à 15 m (Prince *et al.* 1987), celle d'une larve du gastéropode *Littorina littorea*, dont la PLD a été estimée à une trentaine de jours, peut aller jusqu'à 80km (Brenchley et Carlton 1983), ou encore celle d'une larve du gastéropode *Cymatium parthenopeum* (PLD estimée à 293 jours) jusqu'à 4400km (Scheltema 1971). De façon très schématique, la durée de la phase larvaire moyenne est estimée chez la plupart des invertébrés à cycle benthopélagique à quelques semaines (Shanks 2009). Dans un milieu *a priori* dispersif et sans véritable barrière (Palumbi 1994), on s'attend donc à une migration des larves sur de grandes distances. Cette relation entre PLD et distance de dispersion n'est toutefois pas triviale à établir. Les processus de dispersion restent mal compris pour la plupart des espèces à cycle benthopélagique (Cowen et Sponaugle 2009). En effet, si la PLD est très variable entre espèces (Levin et Bridges 1995, O'Connor *et al.* 2007), elle l'est également au sein d'une espèce (*e.g.* chez le mollusque *Crepidula fornicata*, elle varie de 2 à 7 semaines (Rigal 2009)). De plus, la dynamique du nuage larvaire, son origine et son devenir dans la population benthique restent difficiles à appréhender de façon directe, notamment à cause d'une observation et d'un suivi des larves *in situ* quasi-impossible en raison de leur petite taille (Levin 2006). Ce sont donc des méthodes indirectes qui sont le plus souvent employées pour analyser l'importance et/ou les trajectoires de la dispersion larvaire. Des bouées dérivantes ont été par exemple utilisées pour suivre les masses d'eau dans lesquelles sont transportées les larves (Natunewicz *et al.* 2001). Des modèles couplés biologie-physique de dispersion larvaire simulent la dispersion des larves en fonction de paramètres physiques, tels que les courants ou le vent, et biologiques, tels que la mortalité ou la durée de vie larvaire (Ayata *et al.* 2009). L'analyse des signatures bio-géochimiques peuvent permettre de déterminer les populations sources des larves en fonction de signatures chimiques contenues dans les structures construites au cours de leur développement (DiBacco et Levin 2000).

Les méthodes et outils issus de la génétique des populations sont également largement utilisés pour analyser les processus de dispersion, et ce dans tous les environnements (Broquet et Petit 2009). Ces approches permettent de quantifier les flux de gènes entre populations en analysant la distribution de la diversité génétique. Dans quelques cas, les larves ou les juvéniles ont été directement étudiés pour tenter, comme avec les méthodes bio-géochimiques, d'analyser l'origine des recrues dans une population, grâce à des tests d'assignation statistique ou de paternité (Manel *et al.* 2005). C'est ainsi que Saenz-Agudelo *et al.* (2011) ont étudié les patrons de connectivité entre 8 sites établis sur une trentaine de kilomètres et montré que 82% des recrues étaient des immigrants. Cependant, ces approches font généralement face à deux limites majeures pour l'étude des animaux marins à cycle benthopélagique : une structure génétique insuffisante pour caractériser les différentes populations sources et la difficulté de travailler directement sur la phase de dispersion, des larves souvent de petites tailles difficiles à échantillonner. Ainsi, la plupart du temps, les études génétiques sont basées sur des inférences indirectes de la dispersion à partir de l'analyse des populations benthiques adultes. Elles analysent les flux de gènes entre populations en se basant sur le calcul d'estimateurs d'indices de structure génétique tel que  $F_{ST}$  (Wright 1951, Weir et Cockerham 1984). Elles intègrent alors dans leur analyse non seulement le transport larvaire (tel qu'il peut être décrit avec des suivis de bouées dérivantes ou des modèles couplés biologie-physique), mais également les événements de sédentarisation, de recrutement et de mortalité dans la population benthique. Ces approches rendent compte de la dispersion dite « réalisée » ou « effective » sur plusieurs générations ou encore, pour reprendre la terminologie de Pineda *et al.* (2007), la « connectivité reproductive ».

En milieu marin, de nombreuses méta-analyses (*e.g.* Kinlan et Gaines 2003, Siegel *et al.* 2003, Shanks 2009) ont utilisé le modèle d'isolement par la distance dans sa version originale (Wright 1943, Slatkin 1993, Rousset 1997) ou modifiée (Palumbi 2003) pour estimer les distances de dispersion et leur relation avec la PLD (encadré 1). Ces analyses montrent que les distances de dispersions sont très variables chez les invertébrés marins (Fig. 2) et que la relation entre PLD et distance de dispersion est difficile à caractériser. En analysant 32 espèces, Siegel *et al.* (2003) ont montré l'existence d'une relation positive entre la PLD et la distance réalisée: plus la PLD est longue, plus des populations éloignées auront la possibilité d'échanger des gènes (Fig. 3). Mais ce résultat n'est pas validé par l'étude de Shanks (2009) réalisée sur 111 espèces. Ce dernier montre en effet que cette relation entre PLD et distance de dispersion est très faible quand la durée de vie larvaire devient supérieure à quelques jours (Shanks 2009). De plus, outre des problèmes dans le modèle utilisé par Siegel *et al.* (2003), Weersing & Toonen (2009) soulignent que l'utilisation d'une unique valeur de PLD pour une espèce donnée rend ces études peu fiables. En effet, la PLD est un caractère fortement variable y compris à l'échelle intra-spécifique. Elle varie notamment en fonction de la température de l'eau ou de l'abondance en nutriments (*e.g.* Houde 1989, Pepin 1991, McCormick et Molony 1995). Par exemple, chez le mollusque *Crepidula fornicata*, Rigal (2009) a montré d'importantes variations de la PLD



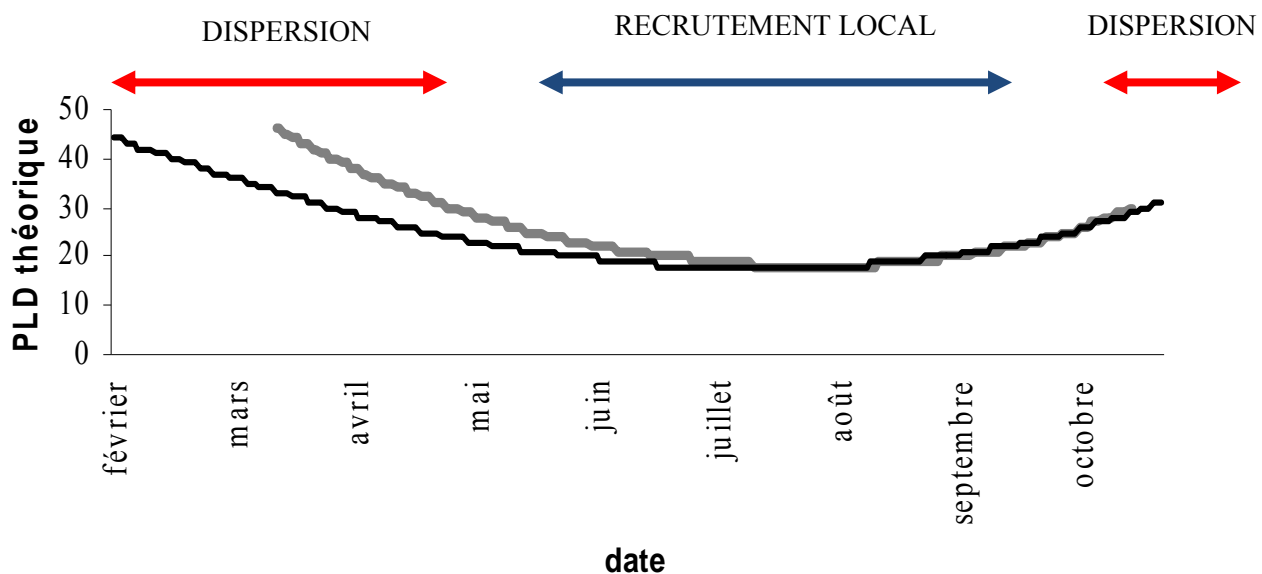
**Fig. 2** Distribution des dispersions moyennes estimées pour 90 organismes marins benthiques, basée sur la droite d'isolement par la distance de données issues de la littérature, pour les macroalgues (n=13), les invertébrés (n=48) et les poissons démersaux (n=25). D'après Kinlan et Gaines (2003)



**Fig. 3** Distance de dispersion (Dd) moyenne (en km) estimée grâce à l'utilisation de données génétiques basées sur un modèle d'isolement par la distance (axe des ordonnées) et durée de vie larvaire (axe des abscisses) pour 32 espèces (19 invertébrés, 12 poissons et une macroalgue) de divers groupes taxonomiques et origines géographiques. La durée de vie larvaire moyenne est estimée soit grâce aux otolithes de larves (poissons) soit par des élevages larvaires en laboratoire (invertébrés et algue). La droite de régression est définie par la relation :  $Dd=1,33 (PLD)^{1,30}$  et  $r^2=0.802$ . D'après Siegel *et al.* (2003)



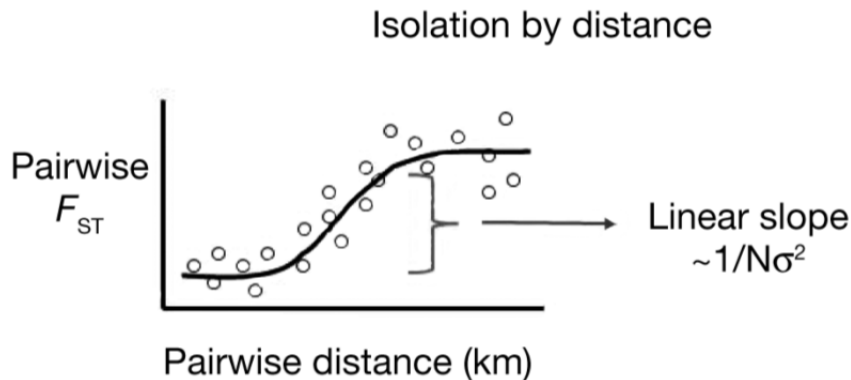
selon les saisons (en relation avec la température) : pour les larves émises dans la baie de Morlaix, la durée de vie larvaire est beaucoup plus courte en juillet, août et septembre qu'elle ne l'est en novembre, décembre ou janvier (Fig. 4). Rigal (2009) a proposé que cette variabilité de PLD pourrait permettre de disperser sur de longues distances ou au contraire de participer au recrutement local selon la date d'émission larvaire. Weersing & Toonen (2009) critiquent particulièrement l'utilisation d'une valeur « moyenne » comme indicateur de la durée de vie larvaire pélagique, cet estimateur étant selon eux particulièrement inapproprié par rapport à des indicateurs basés sur la PLD maximum. Les corrélations entre indices génétiques et PLD sont ainsi largement discutées et il semble peu probant d'utiliser des indices de structure génétique pour estimer des connectivités contemporaines (Lowe et Allendorf 2010, Selkoe et Toonen 2011). Ces études génétiques ont néanmoins permis (1) de montrer l'existence de flux de gènes entre populations très éloignées et (2) une relation entre durée de vie larvaire et distance de dispersion quand on compare des espèces caractérisées par des PLD très différentes (*e.g.* quelques jours *vs.* quelques semaines, Purcell *et al.* 2006) ou des espèces à développement direct par rapport à des espèces à développement indirect (Weersing et Toonen 2009, Selkoe et Toonen 2011).



**Fig. 4** Durée de vie larvaire estimée chez le gastéropode *Crepidula fornicata* en 2006 (trait gris) et 2007 (trait noir) en baie de Morlaix (Rigal 2009). De faibles températures ( $<15^{\circ}\text{C}$ ) seraient davantage liées une dispersion larvaire sur de longues distances, tandis que pour des températures plus élevées ( $>15^{\circ}\text{C}$ ), les larves seraient retenues au sein de la baie. Figure issue de Rigal (2009).

### Encadré 1- Etude de la dispersion par l'analyse de l'isolement génétique par la distance géographique

Si la dispersion est limitée dans l'espace, les populations les plus proches géographiquement devraient aussi être celles qui échangent le plus de migrants et de gènes. On devrait donc observer une augmentation de la différence génétique entre populations de plus en plus distantes. Ce principe a été formalisé sous le terme d'« isolement par la distance » par Wright (1943). Ce modèle peut être utilisé pour estimer les taux de dispersion en utilisant la pente de la relation linéaire entre l'indice  $F_{ST}$  (ou une fonction de cet indice) et les distances géographiques (Rousset 1997, 2001) comme schématisé sur le graphique ci-dessous issu de Selkoe et Toonen (2011).



Néanmoins, cet estimateur des taux de dispersion n'est calculable que pour des échelles d'espace appropriée fonction des distances de dispersion parents-descendants et si des données de densité des populations ( $N$ ) sont obtenues de façon indépendante, des conditions et données souvent difficiles à obtenir (mais voir Puebla *et al.* 2009).

Pour réaliser leur méta-analyse des distances de dispersion des organismes marins, Kinlan et Gaines (2003) ont utilisé une version modifiée de ce modèle en partant de résultats de simulations réalisées par Palumbi (2003). Ces simulations avaient pour but de déterminer des distances de dispersion à partir de données empiriques estimant l'accroissement de différence génétique par unité de distance. Elles utilisaient un modèle de migration de type stepping-stone (pas à pas) et exploraient des paramètres démographiques généralisables à un grand nombre d'espèces marines. Kinlan et Gaines (2003) ont réalisé un ajustement des données simulées et obtenu une relation de type exponentielle entre la distance de dispersion et la pente de la droite d'isolement par la distance :

$$\text{Distance de dispersion} = 0.0016 (\text{pente})^{-1.001}$$

Cette équation, dont les propriétés ont été (sont) débattues (cf. références citées dans Selkoe & Toonen 2011) a ensuite été utilisée pour d'autres études (*e.g.* Kinlan et Gaines 2005, Siegel *et al.* 2003, Shanks 2009).

Si les attendus de forte migration et fort flux de gènes ont été régulièrement mis en évidence, on peut toutefois noter certains cas pour lesquels des limites au flux de gènes ont été observées ainsi que de l'auto-recrutement (*i.e.* recrutement d'individus dont les parents sont présents localement ; *e.g.* Jones *et al.* 2005 pour un exemple chez un poisson clown). L'existence d'un couplage local (c'est-à-dire d'un fort taux de recrutement de larves à l'endroit où elles ont été émises) et de limites à la dispersion à petite échelle (par exemple au sein d'une baie ou d'un récif corallien ; Levin 2006) ont été démontrés par des approches de signature géochimiques (*e.g.* Jones *et al.* 1999), de modélisation (Ellien *et al.* 2004) ou encore de génétique des populations (Jones *et al.* 2005). Plusieurs raisons sont avancées pour expliquer ces dispersions à faibles distances du lieu de l'émission des larves, telles que la forte variabilité *in situ* de la durée de vie larvaire, des effets liés au comportement des larves (notamment leur réponse à des signaux chimiques dans l'environnement favorisant le retour des larves sur leur population d'émission) ou encore des effets liés à des structures hydrodynamiques de rétention locale des larves. Le fait que la dispersion et les flux de gènes qui en découlent sont plus limités que précédemment supposés peut avoir d'importantes conséquences sur le fonctionnement et l'évolution des populations. En particulier, de faibles flux de gènes entre populations pourraient favoriser l'apparition d'adaptation locale (Conover *et al.* 2006), *i.e.* pour un environnement donné et une espèce donnée, la valeur sélective des populations locales sera plus forte que celle des populations vivant dans des habitats différents. Ce processus intra-spécifique suppose néanmoins que les populations adaptées localement sont toujours, potentiellement ou dans les faits, capables d'échanger des gènes avec les autres populations de la même espèce (Kawecki et Elbert 2004). Chez les espèces marines, des patrons atypiques de structure génétique appelés « mosaïque génétique fluctuante » (« chaotic genetic patchiness ») sont également régulièrement mis en évidence (Johnson et Black 1982, Hedgecock 1994b). Ces mosaïques sont définies par l'existence de très fortes structures génétiques instables (variables d'une génération à une autre) à une échelle bien inférieure aux limites de dispersion des espèces (par exemple sur quelques centaines de mètre chez une espèce ayant une larve pouvant disperser pendant plusieurs semaines), qui plus est avec des structures génétique plus fortes que celles observées entre populations éloignées de cette même espèce. Une des explications avancées pour de telles observations est la forte influence de processus de dérive génétique, liée à une forte mortalité en phase larvaire et à une importante variance de succès reproducteurs individuels, à l'origine d'une « loterie familiale » appelée « sweepstakes reproductive events » par Hedgecock (1994a). Cette hypothèse est détaillée dans le chapitre 1 de la deuxième partie. Une des difficultés sous ce scénario est de pouvoir effectivement mettre en évidence des effets de dérive génétique chez les invertébrés marins à cycle benthopélagique.

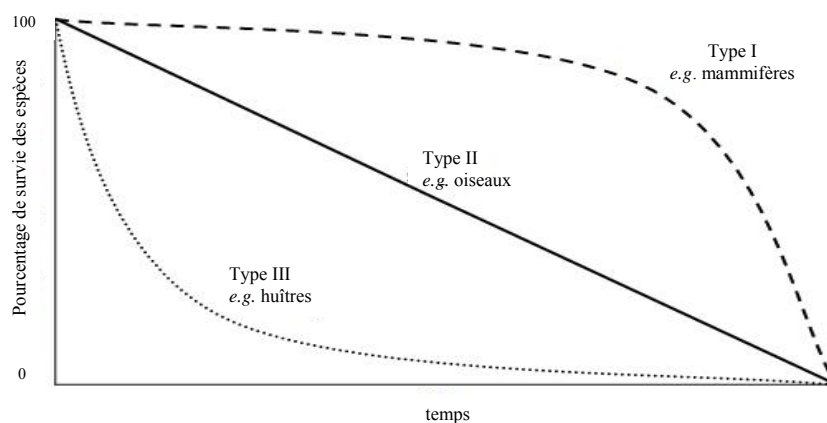
## B. Fécondité et taille des populations : des effets de dérive génétique limités ?

La plupart des espèces marines sont caractérisées par une grande taille de population (Waples 1998, Palsboll *et al.* 2007) et de fortes fécondités. Par exemple, l'étoile de mer épineuse *Acanthaster planci* pond en moyenne 12 à 24 millions d'œufs et jusqu'à 60 millions pour les femelles atteignant 40cm, l'ormeau *Haliotis rubra* pond entre 1.1 et 7.5 millions d'œufs, ou encore l'huître creuse *Crassostrea gigas* libère entre 50-200 millions d'œufs pendant une seule ponte. Pour de telles espèces, les effets de dérive génétique sont minimisés (*e.g.* David 1996, Launey 1998). Cependant, localement et transitoirement des populations de petites tailles peuvent exister et des effets de dérive génétique identifiés. Par exemple, Hansen *et al.* (2000) ont observé chez la truite brune *Salmo trutta* une perte de diversité génétique dans le cas d'un goulot d'étranglement, *i.e.* lorsque la taille de la population est transitoirement diminuée. De même, une perte de diversité génétique imputable aux effets de dérive génétique a été démontrée chez la crevette à pattes blanches *Litopenaeus vannamei* (de Francisco et Galetti 2005) dus à des effets de fondation, *i.e.* quand un petit nombre d'individus sont à l'origine d'une nouvelle population.

La taille de la population, c'est-à-dire le nombre d'individus dénombré dans la population (la « taille réelle » ou « census size »,  $N_c$ ), ne représente pas le nombre d'individus qui se reproduit effectivement, *i.e.* la taille efficace de la population ( $N_e$ ). En effet, la taille efficace de la population diminue par rapport à la taille réelle puisque seuls les adultes en âge de se reproduire (voire une partie de ceux-ci) contribuent à la génération suivante et cette diminution est accentuée lorsqu'il y a inégalité du rapport mâles/femelles, une variance du succès reproducteur, un chevauchement des générations etc. (Hedrick 2000). La taille efficace d'une population se définit comme la taille d'une population idéale (*i.e.* tous les individus ont la même chance de contribuer à la génération suivante) pour laquelle un paramètre génétique prend la même valeur que celle étudiée (Wright 1931). C'est la taille efficace d'une population qui influence les effets de dérive génétique. La dérive génétique influence à son tour le rythme de perte de la diversité génétique, le rythme de fixation des allèles, la consanguinité, mais également l'efficacité de la sélection en maintenant des allèles favorables (Berthier *et al.* 2002). Par exemple, on considère que pour que 99 % de la diversité génétique soit conservée par génération, il faudrait une taille efficace au minimum de 50 avec une égalité du sex-ratio. Les effets de la dérive génétique sont néanmoins tamponnés par la migration entre populations et un taux de migration même faible entre populations suffit à éliminer la différenciation génétique due à la dérive; par exemple, chez deux populations idéales de taille  $N$  qui échangent une proportion  $m$  par l'intermédiaire de migrants à chaque génération, aucune divergence significative ne se produira entre les populations si  $Nm > 1$  (Lewontin 1974, Slatkin 1985).

La taille efficace de la population est souvent estimée inférieure à la taille réelle de la population (Caballero 1994). Toutes espèces confondues (83 espèces étudiées), Palstra et Ruzzante (2008) ont

estimé une taille efficace moyenne de 260 et un ratio  $N_e/N_c$  de 0,14. Cependant ce ratio semble bien supérieur à celui estimé chez les espèces marines (Hauser et Carvalho 2008), bien que fortement variable d'une espèce à l'autre notamment du fait de la grande diversité des cycles de vie (Kapuscinski et Lannan 1986). Par exemple, une taille efficace de 176 (intervalle de confiance : 80-720) a été estimée chez le pagre commun *Pagrus auratus*, soit  $10^{-5}$  fois moins que la taille réelle de la population (Hauser *et al.* 2002), alors que pour la morue *Gadus morhua*, Poulsen *et al.* (2006) ont estimé une taille efficace supérieure à 500 individus et suggéré l'absence d'effet de dérive génétique dans les populations étudiées. Par ailleurs, chez la plupart des espèces marines, un chevauchement des générations (Nelson et Soule 1987) ainsi qu'un sex-ratio biaisé (Nunney 1993) contribuent à réduire la taille efficace de la population. Enfin, une forte variance des succès reproducteurs individuels a été suggérée pour les espèces marines présentant une forte fécondité et une forte mortalité des premiers stades (Hedgecock 1994a). En effet, de nombreux invertébrés marins et poissons présentent une courbe de survie de type 3, telle présentée dans la figure 5. Lors d'un évènement de reproduction, en raison de la forte mortalité des premiers stades de vie et des fortes fluctuations de l'environnement à micro-échelle en milieu marin (dues aux marées, courants, gradients de salinité etc.), un petit nombre d'individus vont survivre, créant ainsi une forte inégalité de contribution à la génération suivante entre individus reproducteur. Ainsi, la diversité des cycles de vie, des mécanismes d'accouplement particuliers, des générations chevauchantes et une variance du succès reproducteur sont autant de facteurs qui vont potentiellement diminuer la taille efficace des espèces marines malgré leur importante fécondité et grande taille de populations.



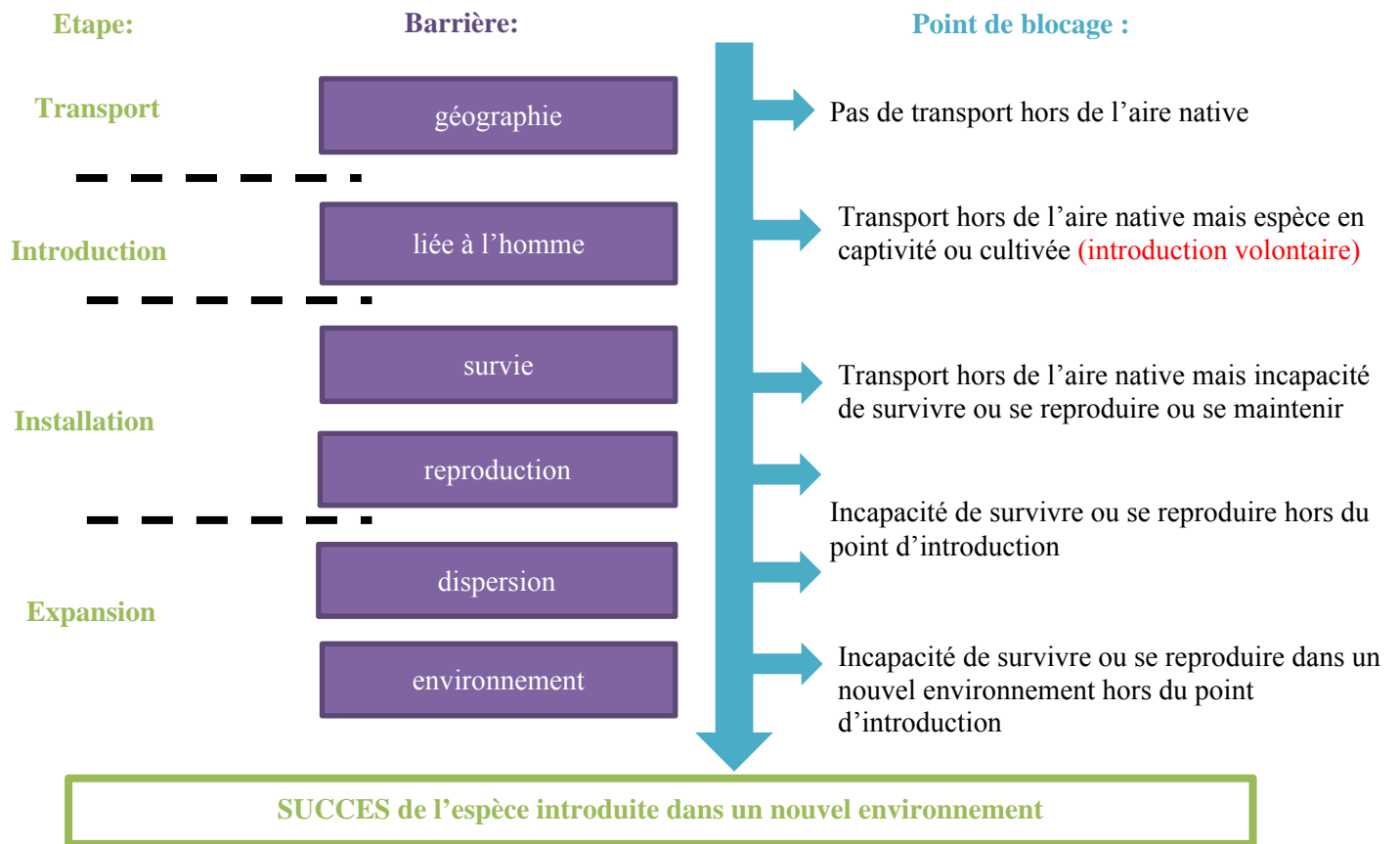
**Fig. 5** Courbes de survie théoriques des différentes espèces animales. Trois types de courbes ont été estimées, le type I, typique des mammifères, pour lequel la mortalité est plus importante en fin de vie, le type II, typique des oiseaux, pour lequel la mortalité est linéaire au cours du temps, et enfin le type III, typique des huîtres, pour lequel une forte mortalité pendant les premiers stades est observée. Source des données : MathWorld

## II. Les processus d'invasions biologiques en milieu marin vus sous l'angle de la biologie évolutive

La distribution géographique des espèces marines est modifiée par de nombreux processus naturels, tels que des changements physiques, géographiques et climatiques du milieu ou des modifications de l'environnement biotique (*e.g.* compétition, prédation, pathogènes). Diverses activités anthropiques affectent également cette distribution que ce soit par exemple la pêche, la pollution ou l'urbanisation qui peuvent notamment faire disparaître localement une espèce du fait de son exploitation intensive ou de la modification de son habitat. Les transports assistés par l'homme (*i.e.* introduction biologique) modifient également l'aire de distribution des espèces : une augmentation rapide et brutale de l'aire de distribution des espèces est observée lors de leur introduction en dehors de leur aire de répartition naturelle sur des échelles de temps très courtes.

Le processus d'introduction biologique est constitué de plusieurs étapes qui ont été historiquement schématisées de la façon suivante (Williamson 1996, Vermeij 1996, Colautti *et al.* 2006): (1) les individus sont transportés de leur aire d'origine dans un nouvel environnement et doivent rencontrer des conditions favorables à leur survie et reproduction pour fonder une population durable (étape d'introduction et installation), puis (2) la taille de la population introduite augmente et s'étend hors du point d'introduction (étape d'expansion) et enfin (3) l'espèce exotique se maintient sur le long terme dans sa nouvelle aire de distribution et établit de nouveaux liens écologiques avec les espèces natives ou introduites déjà installées (étape d'« intégration »). Ces trois étapes ne reflètent que partiellement la réalité du processus d'introduction, un processus complexe. Seule une petite proportion des espèces introduites vont « réussir » leur installation et à ce jour, il reste impossible de prévoir le devenir des espèces introduites (Simberloff 2009, Estoup et Guillemaud 2010), et ce malgré de nombreuses études et de nouveaux cadres théoriques régulièrement proposés pour rendre compte de la dynamique de ce processus, comme par exemple celui décrit dans la figure 6 (Blackburn *et al.* 2011, mais voir également Facon *et al.* 2006, Chapple *et al.* 2012, Occhipinti-Ambrogi 2007, Prentis *et al.* 2008). Une variété de terminologies et de définitions (*e.g.* Table 1) complique de plus la comparaison des patrons d'invasion et donc la compréhension des processus impliqués lors d'introductions biologiques (Blackburn *et al.* 2011).

De nombreuses espèces introduites sont à l'origine de modifications majeures et durables de la biodiversité (Sakai *et al.* 2001, Grosholz 2002). Certaines n'auront pas de conséquences visibles sur les autres espèces et/ou sur l'écosystème envahi, mais une fois établies, elles participent au fonctionnement et à l'évolution des peuplements et écosystèmes dans lesquelles elles ont été introduites (*e.g.* nouvelles proies, nouveaux prédateurs, hybridation avec des espèces autochtones).



Modifié d'après Blackburn *et al.* 2011

**Fig. 6** Les différentes étapes d'une introduction biologique proposées par Blackburn *et al.* (2011). Plusieurs étapes structurent le processus d'introduction biologique (en vert). A chaque étape, des barrières (en violet) doivent être surmontées pour passer à l'étape suivante ; si ces différents points de blocage ne sont pas surmontés (en bleu), l'introduction biologique de l'espèce échouera.

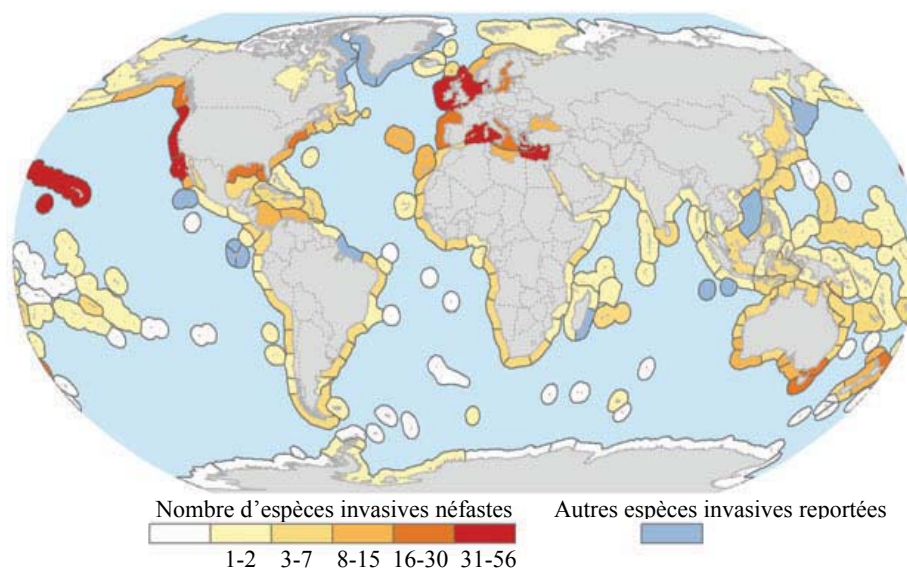
**Table 1** Illustration de la variété de terminologies et de définitions utilisées entre deux études portant sur les introductions biologiques, celle de Richardson *et al.* (2000) et Boudouresque (2008). D'après Boudouresque (2008).

Définition	Richardson <i>et al.</i> (2000)	Boudouresque (2008)
L'organisme, ou ses propagules, a franchi, grâce à l'homme, une barrière géographique	espèce introduite  synonymes: non native, non-indigène, alien, exotique	espèce non-indigène  synonymes: non native, alien, exotique
L'organisme se reproduit dans sa nouvelle région géographique, mais ne peut se maintenir à long terme	espèce occasionnelle  synonymes: espèce "errante", persistante après culture, échappée, "adventive"	adventice  synonyme: occasionnelle
L'organisme se reproduit régulièrement dans sa nouvelle aire géographique, sans l'aide de l'homme et se maintient sur le long terme	espèce naturalisée  synonyme: espèce établie	espèce introduite  synonyme: espèce naturalisée
Espèce naturalisée dont l'aire s'étend plus ou moins rapidement, y compris dans des habitats naturels (non perturbés)	espèce invasive	
Espèce naturalisée qui a un impact fort sur la diversité spécifique, sur le fonctionnement des écosystèmes indigènes, ou qui a un impact économique négatif	"weed" ( <i>i.e.</i> "mauvaise herbe")	espèce invasive
Espèce naturalisée qui bouleverse le fonctionnement des écosystèmes indigènes, éventuellement qui détermine la constitution d'un nouvel écosystème	transformateur	transformateur

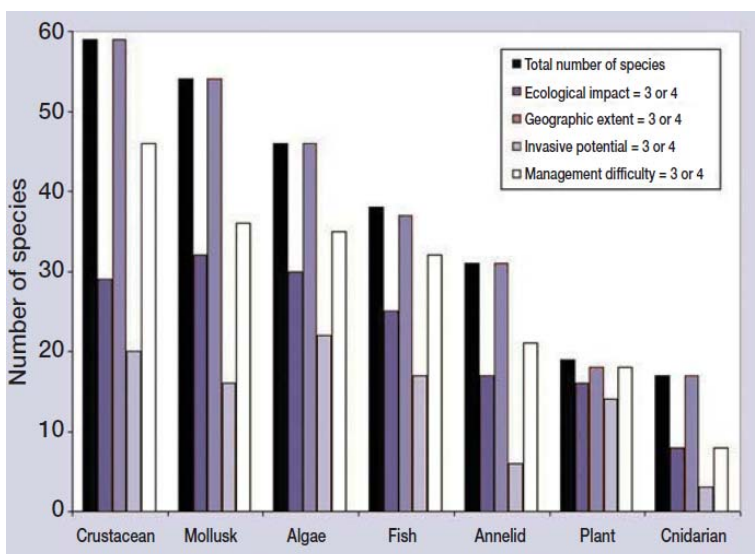


## A. Les introductions biologiques en milieu marin : nombres et vecteurs

En milieu marin, on dénombre la présence d'espèces introduites dans 84% des éco-régions marines (Fig. 7 ; Molnar *et al.* 2008), une éco-région étant une zone abritant un assemblage d'espèces différent de celui des zones voisines. Parmi 329 espèces étudiées par Molnar *et al.* (2008), 57%, appartenant à de très nombreux phylum, sont à l'origine d'impacts écologiques majeurs (Fig. 8). Parmi les 100 espèces invasives les plus menaçantes reportées par Lowe *et al.* (2004), 32 espèces proviennent du milieu marin. Notons que les critères utilisés dans leur sélection des espèces reposent sur l'impact connu de ces espèces sur la biodiversité et/ou les activités humaines. Par ailleurs, une seule espèce par genre a été sélectionnée.

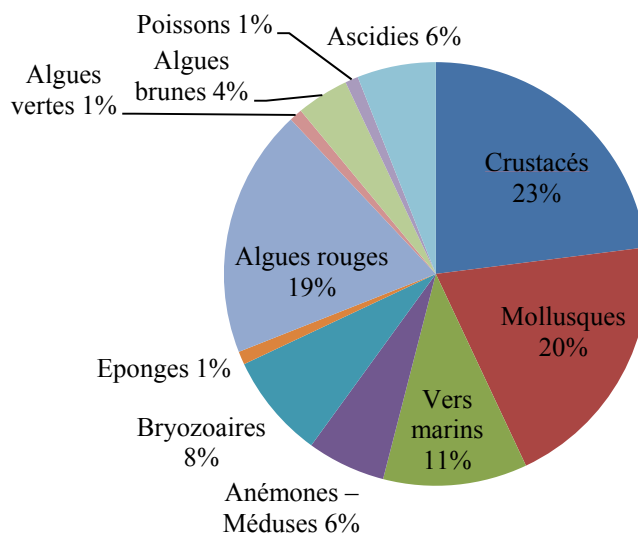


**Fig. 7** Distribution des espèces invasives à l'origine d'impacts écologiques dans 232 éco-régions marines analysées. Un impact sur la biodiversité a été montré pour 329 espèces invasives marines. Figure tirée de Molnar *et al.* (2008)



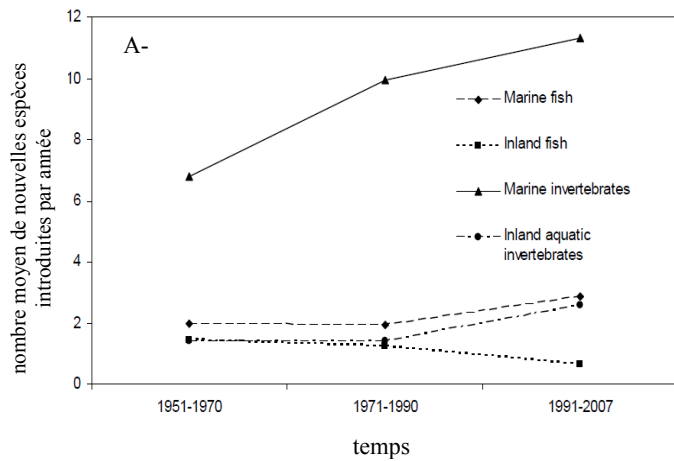
**Fig. 8** Répartition des 329 espèces invasives répertoriées par Molnar *et al.* (2008) au sein des sept phylums étudiés dans cette étude. Les histogrammes présentent le nombre total d'espèces invasives (noir) avec le nombre d'espèces à l'origine des plus forts impacts écologiques (violet foncé) ou dont l'étendue géographique (violet clair), le potentiel d'invasion (violet pâle) ou la difficulté de gestion (blanc) sont importantes. La notion de « fort impact » (valeur 3-4) est mesurée sur une échelle de 0 à 4, selon la classification détaillée dans Molnar *et al.* (2008). Figure issue de Molnar *et al.* (2008)

En Europe, on dénombre 987 espèces introduites, dont plus de 500 espèces en Méditerranée (Boudouresque 2008). A une échelle régionale, par exemple en Bretagne, 97 espèces introduites ont été répertoriées (source : Observatoire de la biodiversité et du patrimoine naturel en Bretagne), dont 74 animaux et 23 algues. Parmi ces 97 espèces, on note 31 espèces en phase d'expansion et 26 ayant des effets néfastes sur la biodiversité, les activités économiques et/ou la santé. Là encore, les espèces introduites regroupent de nombreux embranchements (Fig. 9). Il faut souligner que ces chiffres sous-estiment le nombre réel d'espèces introduites car il existe des biais dans ces analyses. Par exemple, pour certains groupes (*e.g.* bryozoaires, tuniciers), il y a des déficits en expertise taxonomique et pour d'autres (*e.g.* organismes microscopiques) des difficultés techniques pour réaliser les observations de terrain. Quel que soit le nombre réel, et bien que ces processus ne soient étudiés que depuis une soixantaine d'années (Elton 1958) pouvant légèrement sous-estimer le nombre d'introductions biologiques antérieures<sup>1</sup>, une importante accélération du nombre d'espèces introduites est notée au cours du 20<sup>ème</sup> siècle. En milieu marin, ces introductions se sont intensifiées au cours du vingtième siècle et en particulier depuis les années 1960 (Gouletquer *et al.* 2002, Hulme *et al.* 2008 ; Fig. 10). Cette évolution temporelle est fortement associée à l'augmentation des activités humaines à l'échelle mondiale, avec notamment le développement de l'aquaculture et du commerce maritime à une échelle mondiale (Fig. 11 ; Gouletquer et Le Moine 2002, Boudouresque 2008).

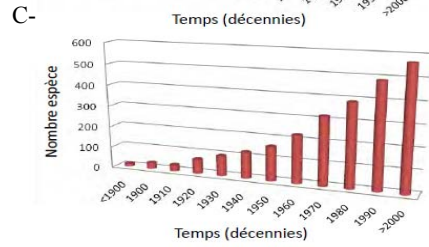
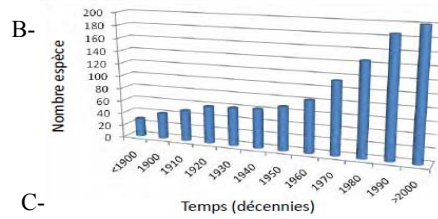


**Fig. 9** Proportion d'espèces introduites par phylum, dénombrée en Bretagne et basée sur 97 espèces introduites. Source des données : GIP Bretagne Environnement

<sup>1</sup> Les introductions biologiques existent depuis longtemps (Carlton 2011) et sont certainement plus nombreuses que ce qu'on ne le soupçonne (Carlton 1999). Par exemple, il a été proposé que le gastéropode *Littorina littorea*, originaire des côtes européennes de l'Atlantique ait pu traverser l'Atlantique avec les navigateurs danois il y a environ 1000 ans, ou encore plus récemment au début du siècle (Carlton *et al.* 1982, Zibrowius 1991).

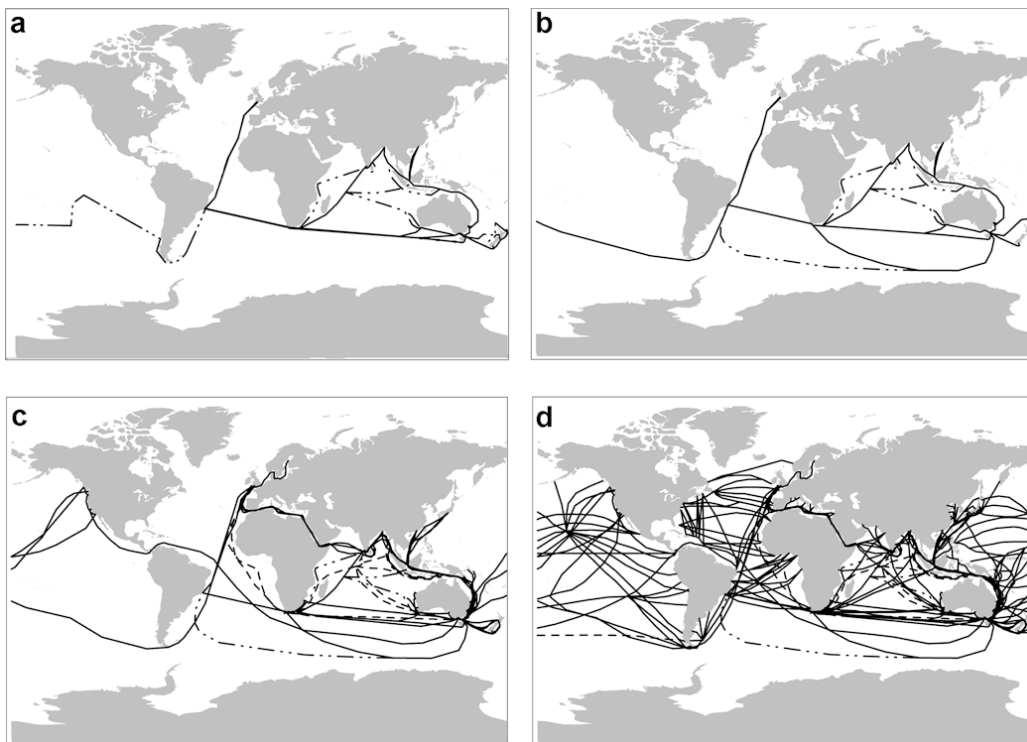


D'après Hulme *et al.* 2008



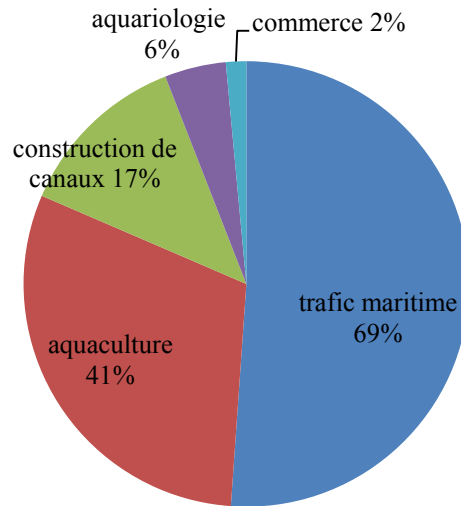
Source des données:  
DAISIE (Handbook of European Alien Species – Springer)

**Fig. 10** Nombre moyen d'espèces introduites nouvellement reportées chaque année depuis une soixantaine d'années (A-), du nombre d'espèces introduites le long des côtes Atlantiques (B-) et Méditerranéennes (C-) depuis un siècle.



**Fig. 11** Accroissement du trafic maritime au cours des derniers siècles : (a) avant 1839, le trafic maritime était essentiellement lié à l'exploration et de colonisation, (b) entre 1839 et 1851, le trafic maritime était essentiellement lié à l'immigration, puis entre 1852 – 1860, à la ruée vers l'Or, (c) entre 1861-1920 et (d) de 1921 jusqu'à aujourd'hui, les échanges commerciaux sont ensuite allées croissants. Les lignes continues représentent les itinéraires communément utilisés, celles en pointillés les itinéraires rares. Figure issue de Hewitt *et al.* (2004)

Plusieurs vecteurs ont été identifiés pour les introductions d'espèces dans de nouveaux environnements marins (Fig. 12) mais deux vecteurs sont majoritaires : le trafic maritime et l'aquaculture. Le trafic maritime est responsable d'introduction accidentelle par deux mécanismes (1) le transport dans les eaux de ballast utilisées pour lester les navires, et (2) le transport sur la coque du bateau (biofouling, *i.e.* les espèces sont fixées sur la coque). Les eaux de ballast sont des vecteurs majeurs pour les invertébrés présentant une phase larvaire (Carlton et Geller 1993), à l'instar par exemple du crabe bleu *Callinectes sapidus* introduit d'Amérique en Egypte (Galil 2000, 2001, Galil *et al.* 2002). En revanche, de nombreuses espèces benthiques ayant des larves à très courte durée de vie (ou aucune larve) peuvent également être transportées sur la coque de bateaux : ce mode de transport est proposé dans le cas du crustacé cirripède (balane) *Elminius modestus*, originaire de Nouvelle-Zélande et observé en 1945 en Angleterre (Harms 1999), de l'algue rouge *Anotrichium yagii*, originaire du Japon et de Corée, découverte au Brésil dans une région fréquentée par des bateaux de ces pays (Horta et Oliveira, 2000) ou encore de l'ascidie *Styela clava*, originaire de Corée et sans doute apportée à Plymouth (Angleterre) par des bateaux qui participaient à la Guerre de Corée (Lutzen 1999). L'aquaculture est à l'origine d'introduction délibérée d'espèces dans un nouvel environnement. Par exemple, entre 1971 et 1975, plus de 500 tonnes de l'huître creuse *Crassostrea gigas* ont été volontairement introduites du Canada en France, et plus de 10 000 tonnes de naissains du Japon au Canada afin de dynamiser les élevages conchylicoles français suite à des épisodes récurrents de mortalité des huîtres alors cultivées (Grizel et Héral 1991). Ces introductions délibérées sont également accompagnées d'introductions accidentelles. En effet, de nombreuses espèces ont également été introduites avec *C. gigas* (sur les coquilles de l'huître ou dans les emballages), à l'instar de la crépidule *Crepidula fornicata* (Blanchard 1995). Il existe naturellement de très nombreux autres vecteurs d'introduction d'espèces en milieu marin, mais qui participent proportionnellement moins à une échelle mondiale, même si localement ils peuvent avoir des effets majeurs. On peut citer, par exemple l'introduction de l'algue verte *Caulerpa taxifolia* en Méditerranée suite aux rejets de l'espèce de l'aquarium de Monaco ou le canal de Suez responsable de la majorité des introductions en Mer Méditerranée. Par ailleurs, l'introduction d'une espèce n'est pas nécessairement liée à un seul vecteur mais peut en combiner plusieurs, qui peuvent varier d'une région d'introduction à une autre. Par exemple, le wakamé *Undaria pinatifida*, algue originaire du Japon, a été introduit de trois façons différentes (aquaculture, espèce accompagnatrice de *C. gigas* et via le trafic maritime) hors de son aire native et Voisin *et al.* (2005) ont montré par des études génétiques l'importance jouée par le trafic maritime dans l'introduction en Nouvelle-Zélande et Australie tandis que l'aquaculture était le vecteur principal de son introduction en Europe.



**Fig. 12** Les principaux vecteurs d'introduction à l'échelle mondiale et la proportion d'espèces connues (ou potentiellement) introduites par ces vecteurs (nombre d'espèces introduites analysées = 329). Pour une même espèce, plusieurs vecteurs peuvent être identifiés. Modifié de Viard et Comtet (2009) utilisant les données de Molnar *et al.* (2008).

### B. Paradoxe de l'invasion et capacité d'adaptation : un lien possible ?

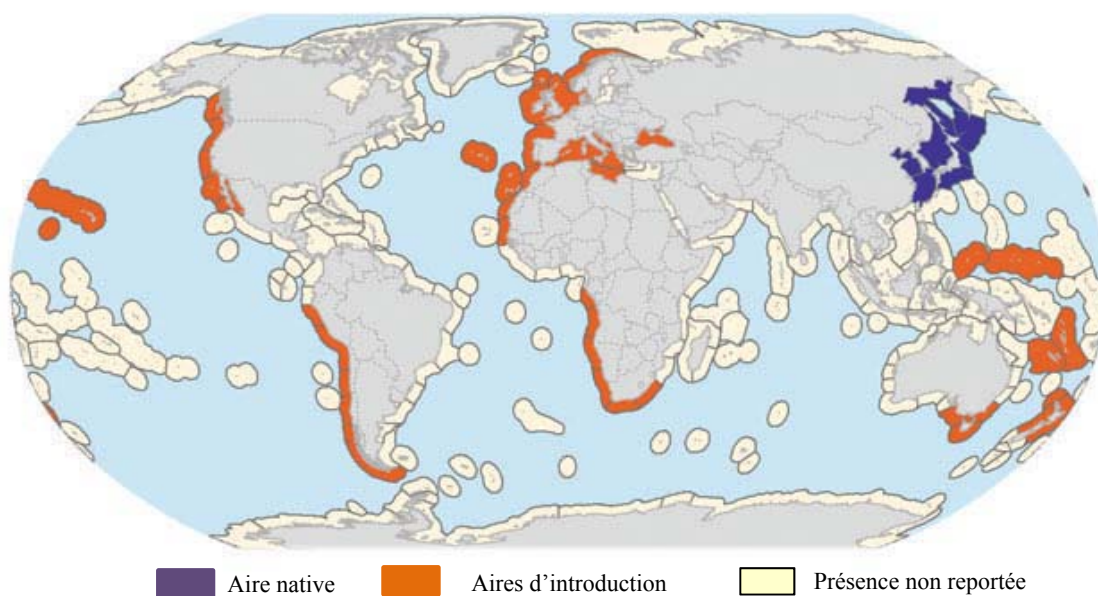
Les espèces introduites sont des modèles privilégiés pour étudier en temps réel des processus écologiques, comme les interactions entre espèces ou la dynamique des communautés, et/ou des processus évolutifs tels que les effets de fondation ou l'importance de la diversité génétique dans la réussite locale d'une espèce (Lee 2002, Sax *et al.* 2007). Par ailleurs, comprendre les raisons permettant l'installation d'une espèce dans un nouvel environnement est un enjeu de recherche en lui-même (*i.e.* résoudre le paradoxe de l'invasion ; Sax et Brown 2000). Par exemple, le crabe vert *Carcinus maenas* originaire des côtes Atlantiques d'Europe et d'Afrique du Nord, a été introduit notamment en Amérique du Nord où il est désormais présent à des températures plus froides que celles de son aire d'origine (Audet *et al.* 2003). Et il s'agit alors de comprendre comment il s'est adapté ou acclimaté à ces modifications environnementales.

Différentes explications peuvent être apportées à cet apparent « paradoxe de l'invasion ». La réussite des espèces non-indigènes dans de nouveaux environnements peut être liée à un contexte écologique nouveau (*e.g.* absence de prédateur et/ou de compétiteur, Willis *et al.* 1999, 2000) ou répondre à des processus évolutifs (Lee 2002, Sax *et al.* 2007, Lawson-Handley *et al.* 2011). Le succès d'une espèce introduite peut être lié à de la plasticité phénotypique, c'est-à-dire aux capacités des individus à moduler leurs phénotypes en fonction de l'environnement (Price *et al.* 2003). Par exemple, *C. maenas* présente une grande plasticité phénotypique qui pourrait avoir permis à l'espèce de s'établir dans de nouveaux milieux (Edgell et Hollander 2011). Quand les phénotypes introduits ne sont pas

bien adaptés au nouvel habitat et peu plastiques, d'autres types de mécanismes évolutifs, tels que des processus de sélection, ont été proposés (Stockwell *et al.* 2003, Sax *et al.* 2007). Sous l'effet de nouvelles conditions environnementales dans l'aire d'introduction, des traits conférant une plus grande valeur adaptative pourraient être sélectionnés. La diversité génétique introduite pourrait jouer un rôle déterminant dans ce processus car si les individus introduits possèdent déjà des allèles ou génotypes conférant un avantage dans le nouvel environnement, la sélection pourrait opérer efficacement à partir de ce polymorphisme préexistant (Barrett et Schlutter 2008). L'histoire d'introduction de l'espèce joue alors un rôle important sur sa réussite puisque l'accroissement du nombre d'individus introduits ainsi que des populations sources de l'introduction dans l'aire d'origine sont susceptibles d'augmenter la diversité génétique y compris la diversité génétique adaptative sur laquelle la sélection pourrait opérer. Plusieurs études portant sur des espèces invasives marines ont montré un niveau de diversité génétique entre aires d'origine et introduite similaire, voire supérieure dans le cas de multiples sources d'introduction (Simon-Bouhet *et al.* 2006 ; Roman et Darling 2007, Sax *et al.* 2007). De nouvelles mutations dans l'aire d'introduction peuvent également être à l'origine de phénotypes adaptatifs (Rainey et Travisano 1998). La sélection à partir de polymorphisme préexistant apparaît cependant un processus plus rapide que l'apparition de nouvelles mutations suivie de leur sélection. En effet, dans le cas de sélection sur de la variation génétique préexistante, les allèles favorables sont déjà présents dans le génome et en plus grandes fréquences que ne le seraient de nouveaux allèles (Innan et Kim 2004).

Les espèces invasives marines sont des espèces particulièrement intéressantes en regard de l'étude des processus de sélection et d'adaptation en milieu marin. En effet, comme nous l'avons vu précédemment, une grande majorité des espèces marines ont une forte fécondité et une phase larvaire assurant la dispersion sur de grandes distances. Dans ce contexte, des processus d'adaptation locale sont peu probables car l'adaptation locale est contrebalancée en permanence par les flux de gènes apportant des génotypes localement peu adaptés. Néanmoins, il existe des exemples dans la littérature ayant montré l'existence de processus de sélection et d'adaptation aux conditions environnementales locales. Par exemple, Koehn *et al.* (1980) ont montré de tels processus chez la moule *Mytilus edulis* le long du détroit de Long Island, aux Etats-Unis. Ce détroit est une baie présentant une salinité croissante du sud-ouest (embouchures de nombreuses rivières) au nord-est de la baie (océan Atlantique). En parallèle, la fréquence de l'allèle 94 de l'aminopeptidase augmente graduellement du sud-ouest (faible salinité) au nord-est (salinité océanique). Cette enzyme permettrait alors de réguler l'osmolarité intra-cellulaire en présence d'un milieu salin. Cet allèle est tout de même maintenu en faible fréquence à faible salinité, suggérant le maintien de flux de gènes le long du détroit. Wilding *et al.* (2001) ont également suggéré des mécanismes d'adaptation chez le gastéropode *Littorina saxatilis* le long de l'estran. Une distribution clinale de la forme et l'épaisseur de la coquille a été observée chez cette espèce le long de l'estran. La forme « H », caractérisée par une coquille fine, un large opercule et de petites stries, est majoritairement observée en haut de l'estran tandis que la forme « M »,

caractérisée par une coquille épaisse, un petite opercule et de larges stries, est majoritairement observée en bas de l'estran. La forme « H » est davantage résistante aux contraintes physiques (*e.g.* vent, vague), tandis que la forme « M » à la prédation. En analysant 290 marqueurs, 5% des loci ont montré une différenciation génétique plus forte qu'attendue, suggérant des effets de sélection le long de l'estran (directement ou indirectement). D'autres exemples d'adaptation locale sont reportés dans la littérature marine (*e.g.* Faure *et al.* 2008) et ce malgré d'importants flux de gènes. Une des hypothèses avancées pour expliquer cela est un flux de gènes moins important que celui supposé chez les espèces marines (cf. I.A). Dans ce contexte, les espèces invasives sont des modèles *a priori* intéressants d'étude car elles sont caractérisées par une rupture (ou en tout cas une très forte diminution) des flux de gènes entre deux groupes de populations, les populations de l'aire de distribution naturelle et les populations de l'aire d'introduction. Par ailleurs, de nombreuses espèces introduites montrent une large distribution au sein des aires d'introduction (souvent le long de fort gradient latitudinaux). Par exemple, l'huître creuse *Crassostrea gigas* introduite dans tous les océans, est présente sur d'importants gradients longitudinaux et latitudinaux (Fig. 13), aux environnements contrastés. Dans de nombreux cas, cette distribution s'explique au moins en partie par des introductions secondaires assistées par l'homme et donc, encore une fois, avec des flux de gènes naturels limités. Les espèces invasives semblent ainsi de bon modèle, d'une part pour étudier le paradoxe de l'invasion mais également pour tester la possibilité d'adaptation locale dans le nouvel environnement colonisé.



**Fig.13** Répartition mondiale de l'huître creuse *Crassostrea gigas* à l'échelle mondiale. Son aire d'origine (Nord-ouest Pacifique) est figurée en bleue, et les aires d'introduction en orange. Les zones où la présence de *C. gigas* n'est pas connue sont figurées en jaune. Figure modifiée de Molnar *et al.* (2008)

### III. Présentation du modèle d'étude : *Crepidula fornicata*, une espèce invasive au cycle de vie typique des invertébrés marins

#### A. Une longue phase larvaire et un système de reproduction original

La crépidule, *Crepidula fornicata*, est un mollusque de la classe des gastéropodes et de la famille des Calyptraeidae (Linné, 1758). Ce gastéropode marin est une espèce des milieux côtiers (baies et estuaires), occupant des habitats divers (fonds rocheux, graviers, sables ou vase) situés sur l'estran (*i.e.* zone de balancement des marées) jusqu'à 60m de profondeur. Cette espèce a un cycle de vie benthopélagique (Fig. 14), alternant une phase benthique de 6 à 8 ans (Le Gall 1980, Coum 1979) et une phase larvaire de 2 à 7 semaines (Rigal 2009) qui est le principal vecteur de dispersion (les adultes ne sont pas mobiles mais peuvent être déplacés sur de petites distances par les courants). La fécondation est interne et les femelles incubent les embryons au sein de la cavité palléale pendant 21-28 jours (Chipperfield 1951, Barranger 2011). Les individus mâles et femelles peuvent se reproduire plusieurs fois par saison de reproduction (*e.g.* 2-4 fois en Rade de Brest, Richard *et al.* 2006) et être ainsi à l'origine de plusieurs émissions larvaires. A chaque événement de reproduction, une femelle libère 5 000 à 20 000 larves au stade véligère dans la colonne d'eau, leur taille étant alors de l'ordre de 400µm (Pechenik et Lima 1984). Les larves se développent dans la colonne d'eau et acquièrent de nouvelles structures et fonctions physiologiques leur permettant d'atteindre un stade où elles sont capables de se métamorphoser en juvéniles benthiques. Les larves sont alors dites « compétentes » à la métamorphose et ont une taille de l'ordre de 900-1000µm. Suite à leur métamorphose, les post-larves (juvéniles) se fixent sur un substrat, généralement une coquille de crépidule. Certains traits larvaires, tels que la nutrition, le taux de croissance ou encore la métamorphose, ont fait l'objet de nombreuses études au sein du laboratoire (*e.g.* les travaux de thèse de F. Rigal (2009) et de F. Leroy (2011)) et dans l'équipe de Jan Pechenik (Tufts University, Boston)). Il a notamment été montré le rôle primordial joué par la température sur la dynamique du pool larvaire (Rigal 2009). La durée de vie larvaire est d'autant plus courte que la température du milieu est élevée. Par exemple, en 2006, la durée de vie larvaire en baie de Morlaix a été estimée à 18 jours pour des températures supérieures à 16°C et à 46 jours pour des températures inférieures à 10°C, suggérant que des températures froides pourraient favoriser une dispersion sur de longues distances, alors que des températures plus élevées pourraient favoriser un recrutement local. Par ailleurs, une grande variabilité des croissances larvaires a été observée chez cette espèce (Hilbish *et al.* 1999), y compris au sein d'une même ponte (Le Cam *et al.* 2009). Le Cam *et al.* (2009) ont montré que plus les larves étaient apparentées, plus leur taux de croissance était similaire. Le caractère « taux de croissance larvaire » semble héritable : Hilbish *et al.* (1999) estiment que 35% de la variance phénotypique observée pouvait être expliquée par une composante génétique.

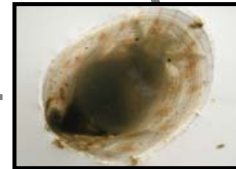


La **larve** est émise dans la colonne d'eau à une taille d'environ 400µm. Elle atteint la compétence à la métamorphose au bout de 2 à 7 semaines, à une taille de 800 - 1000µm.



Métamorphose

Le **juvénile** se fixe sur un substrat (généralement sur une coquille de crépidule) à une taille de 900-1000µm. Cette métamorphose peut être induite par la présence d'adultes conspécifiques mais elle peut être également spontanée



Pelagos  
Benthos

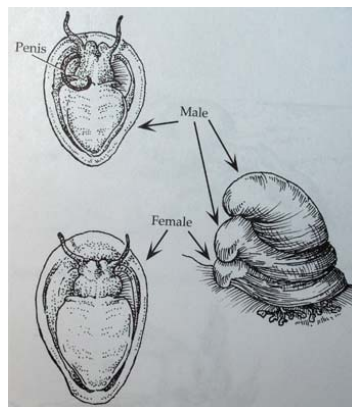


Après une fécondation interne, la **femelle** incube les embryons (flèche) et au bout de trois semaines, libèrent entre 5 000 à 20 000 larves, et ce plusieurs fois par saison de reproduction.



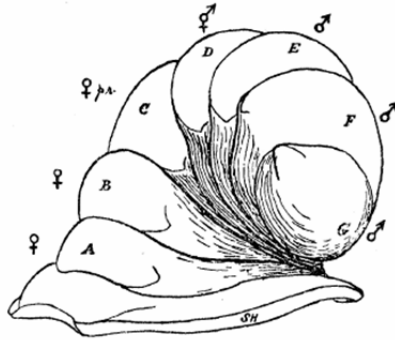
Sédentarisation

Les individus forment des **chaînes pérennes** d'individus ou plusieurs individus sont empilés. La reproduction se fait majoritairement au sein de la chaîne



Les individus sont d'abord **mâles** (présence d'un pénis) et sont situés au sommet de la chaîne. Puis, ils deviennent **femelles** (présence d'une papille vaginale) et sont situées à la base de la chaîne, en fonction de l'âge mais aussi du sex-ratio local. Schémas issus de Bertness 2007.

**Fig. 14** Cycle de vie de *Crepidula fornicata* (Linné 1758)



**Fig. 15** Dessin schématique d'une chaîne de crépidule d'après Orton (1909). La chaîne est constituée de sept individus, notés de A à G, et le sexe des individus est indiqué par ♂ pour les mâles, ♀ pour un individu en transition sexuelle, et ♀ pour les femelles. La base de la chaîne est une coquille d'huître vide, notée SH.

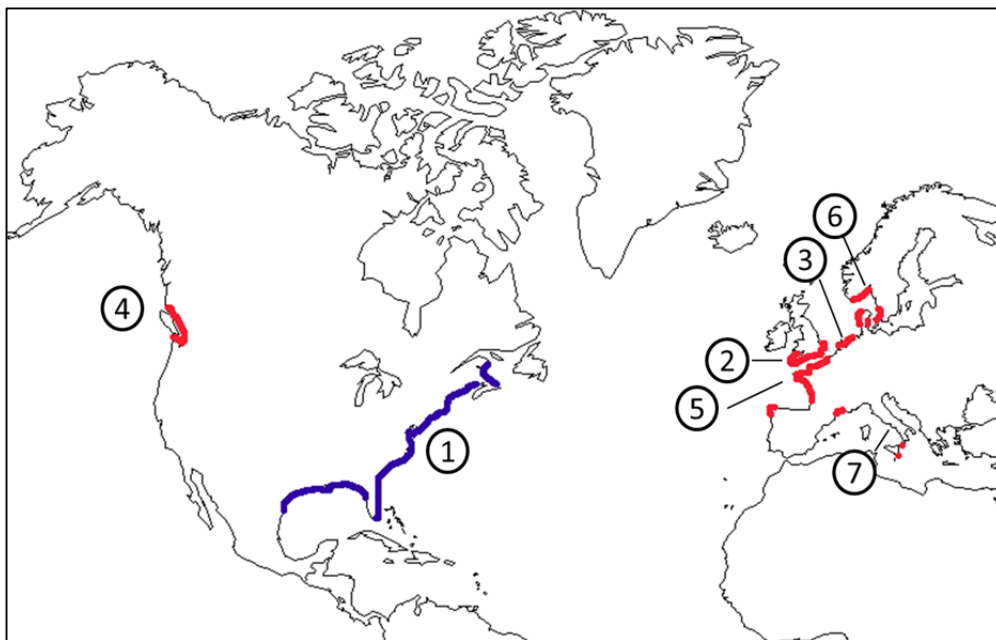


**Fig. 16** Observation directe de copulation au sein d'une chaîne de crépidules. La flèche rouge montre le pénis d'un mâle (en noir) fécondant une femelle située deux individus en dessous de lui.

Ce gastéropode présente également un système de reproduction original. Il présente un comportement fortement grégaire, formant des empilements pérennes d'individus, appelés « chaînes » (Fig. 15). Les chaînes sont constituées en général de deux à une quinzaine d'individus (Coe 1936) mais pouvant aller jusqu'à 51 individus (Sibeaux 2010). Les individus les plus jeunes sont situés au sommet de la chaîne, les plus âgés à la base. Des chaînes secondaires peuvent se greffer sur la chaîne principale. Au sein de ces chaînes sont trouvés des mâles et des femelles. Ce dimorphisme sexuel intra-chaîne est associé à de l'hermaphrodisme protandre (*i.e.* hermaphrodisme séquentiel caractérisé par un changement de sexe de mâle à femelle au cours de la vie de l'individu). Les individus les plus jeunes (au sommet de la chaîne) sont généralement des mâles, ceux situés à la base des femelles. Cette caractéristique commune au sein du genre *Crepidula* a fait de *Crepidula* sp. un modèle privilégié depuis le début du 20<sup>ème</sup> siècle pour l'étude des modalités de changement de sexe, notamment concernant les déterminismes pouvant influencer le moment du changement de sexe (Coe 1938, Hoagland 1978, Collin *et al.* 2005). Différentes études ont montré l'importance du sex-ratio local (*i.e.* intra-chaîne) comme paramètre déterminant le changement de sexe (Coe 1938, Hoagland 1978, Dupont *et al.* 2006). Lorsqu'une chaîne est constituée uniquement de mâles, rapidement, et quelle que soit sa taille, un ou plusieurs des mâles vont se transformer en femelles. Il faut cependant noter que les chaînes de crépidules présentent, comme classiquement observés chez les organismes protandres, un sex-ratio biaisé en faveur des mâles (Le Cam 2009). Différentes analyses de paternité (Le Cam *et al.* 2009, Dupont *et al.* 2006, Proestou *et al.* 2008, Barranger 2011) ont prouvé que la reproduction se fait majoritairement au sein de la chaîne : moins de 5-10% des larves sont assignées à un père n'appartenant pas à la chaîne où se trouve la mère. Ce résultat s'explique par la combinaison de l'agrégation des individus au sein des chaînes et de la protandrie qui permet la co-existence des deux sexes au sein de ces chaînes. Des observations directes de copulation ont d'ailleurs été notées sur des chaînes conditionnées au laboratoire (Fig. 16). Une chaîne de crépidules est ainsi un groupe autonome de reproduction (Le Cam *et al.* 2009). Par ailleurs, la paternité multiple semble être la règle chez cette espèce avec en moyenne 3 pères par ponte. Néanmoins, les paternités sont extrêmement biaisées. En effet, dans la majorité des cas (*e.g.* 4 cas sur 6 dans Le Cam *et al.* 2009, 26 cas sur 31 dans Barranger 2011), un (ou deux, Le Cam *et al.* 2009) père(s) contribue(nt) majoritairement à la ponte, avec entre 39 et 100% des larves qui lui (leur) sont assignées. Ce père se situe souvent au-dessus de la femelle, ou est séparé d'elle au maximum par deux individus.

## B. *Crepidula fornicata*, un gastéropode emblématique des processus d'invasions biologiques en milieu marin

Outre le fait d'être un modèle privilégié pour l'étude du développement embryonnaire et larvaire (Henry *et al.* 2010) et l'étude des modalités de la protandrie (Coe 1938, Hoagland 1978), *C. fornicata* est également largement étudiée du fait de son statut d'espèce introduite et invasive en Europe. *C. fornicata* est une des 104 espèces invasives répertoriées par Gouletquer (2002) le long des côtes atlantiques de la Manche. Elle est à l'origine d'une large gamme de modifications de la diversité biologique et du fonctionnement des écosystèmes côtiers dans lesquels elle a été introduite (Thieltges *et al.* 2006). Emblématique des processus d'introduction biologique en milieu marin, elle a réussi avec succès son introduction et installation et ce, sur une période relativement courte (*e.g.* 40 ans en France). L'existence d'une longue phase larvaire assurant la dispersion de l'espèce et une stratégie de reproduction efficace (hermaphrodisme protandre, paternité multiple) sont des traits de vie qui ont été proposés comme déterminants pour la réussite de l'installation et de l'expansion de *C. fornicata* en Europe et sur la côte ouest des Etats-Unis (Richard *et al.* 2006, Dupont *et al.* 2006, Viard *et al.* 2006).



**Fig. 17** Aires de distribution naturelle (en bleu) et introduite (en rouge) de *Crepidula fornicata* et étapes de l'invasion, d'après Blanchard (1997) : 1 aire d'origine, 2 : (1880) Grande Bretagne, 3 : (1910 -1950) : Belgique, Allemagne et Pays-Bas, 4 : (1930) côtes ouest des Etats-Unis, 5 : (1940) sud de la Grande-Bretagne et France, 6 : (1950) Danemark, Norvège, Suède, 7 : (1970) Méditerranée et Espagne

*C. fornicata* est originaire de la côte atlantique de l'Amérique du Nord, où elle est présente du Canada jusqu'aux Caraïbes (Fig. 17). L'histoire de son introduction et de sa dissémination en Europe et sur la côte ouest des Etats-Unis est intimement liée à celle de l'ostréiculture. En effet, les premières

observations de crépidules en Europe se font en Grande-Bretagne suite à l'importation de l'huître américaine *Crassostrea virginica* à la fin du 19<sup>ème</sup> siècle (Crouch 1894, 1895 cités dans Adam et Leloup 1934, Orton 1912, Hoagland 1978). *C. fornicata* est ensuite observée au début du 20<sup>ème</sup> siècle le long des côtes européennes (Belgique, Allemagne, Pays-Bas et France; Blanchard 1997). La longue phase larvaire a certainement favorisé la dispersion de l'espèce le long des côtes européennes (Adam et Leloup 1934), mais cette dissémination apparait également associée à la culture de l'huître plate *Ostrea edulis* donnant lieu à des échanges de naissains entre ostréiculteurs qui se sont involontairement accompagnés de l'introduction d'autres espèces, dont la crépidule (Blanchard 1997). Les densités observées restent cependant modérées et il ne semble pas que des populations importantes se soient jusqu'alors installées, à l'exception des Pays-Bas (Brouwer 1934). L'introduction massive de *C. fornicata* le long des côtes atlantiques et méditerranéennes françaises date des années 70. L'apparition simultanée de la crépidule sur divers sites ostréicoles coïncide avec le début de la culture intensive de l'huître creuse *Crassostrea gigas* importée notamment de la côte ouest des Etats-Unis après son introduction depuis l'Asie. A cette époque, *C. fornicata* est également déjà installée sur la côte ouest des Etats-Unis ; son introduction y date des années 1930 suite à l'importation de l'huître américaine *Crassostrea virginica* (Kozloff 1974). L'introduction volontaire de *C. gigas* de la côte ouest des Etats-Unis en Europe s'est alors accompagnée de l'introduction accidentelle de *C. fornicata*. La culture de *C. gigas* a donné lieu à d'importants échanges entre ostréiculteurs (Smaal *et al.* 2009), facilitant l'installation de *C. fornicata* sur de nombreux sites ostréicoles (Blanchard 1995, 1997). Par la suite, la phase larvaire dispersive a certainement favorisé l'expansion secondaire de cette espèce le long du littoral (Dupont *et al.* 2003, Dupont *et al.* 2007b, Viard *et al.* 2006).

Aujourd'hui, *C. fornicata* est présente de l'Espagne à la Norvège (Fig. 17) avec des surfaces colonisées et des densités variables. Par exemple, en France, elle occupe environ 800km<sup>2</sup> en Baie de Saint-Brieuc (soit 25% de la zone), 150km<sup>2</sup> en Baie du Mont-Saint-Michel (soit 61% de la zone) ou encore 115km<sup>2</sup> en Rade de Brest (soit 90% de la zone subtidale, Pagny *et al.* 2010), atteignant des biomasses de 250 000 tonnes en baie de Saint-Brieuc, 160 000 tonnes en baie du Mont-Saint Michel et 210 000 tonnes en Rade de Brest (Blanchard 2009, Pagny *et al.* 2010). En revanche, dans la baie d'Arcachon, sa biomasse est estimée à 55 tonnes et n'est présente que sur 5% de la zone (de Montaudouin *et al.* 2001).

#### IV. Présentation de la démarche et structure du document

La plupart des espèces marines à cycle benthopélagique présentent des caractéristiques démographiques originales, telles qu'une forte fécondité, une grande taille de population ainsi qu'une phase larvaire dispersive. Notre modèle d'étude, l'espèce invasive *Crepidula fornicata*, présente de telles caractéristiques. Nous avons vu dans les parties précédentes que ces propriétés biologiques devraient conduire à des effets de dérive génétique limités et de forts flux de gènes, contrant l'émergence d'adaptations locales. Cependant, dans le cas d'espèces introduites, le brusque changement d'environnement et la rupture de flux de gènes associée lors du transport de l'espèce en dehors de son aire d'origine pourrait faciliter la mise en évidence d'effets de sélection (postérieurs à l'introduction). Par ailleurs, des effets de dérive génétique ont été régulièrement mis en évidence empiriquement chez les invertébrés marins (e.g. ratio  $N_e/N_c$  faible, expliqué en partie par d'importantes mortalités larvaires). Chez *C. fornicata*, ces effets de dérive génétique pourraient être accentués du fait des modalités de la reproduction et de dispersion des larves (e.g. sex-ratio et paternité biaisés, émission de pool de larves apparentées). Cette thèse a ainsi eu pour objectif d'analyser l'influence de deux forces évolutives majeures, la sélection et la dérive génétique chez l'espèce *Crepidula fornicata*, au sein de son aire d'introduction en Europe.

Mes travaux de thèse se sont articulés en deux parties :

- la première partie de cette thèse s'intéresse à la dispersion assistée par l'homme (i.e. introduction *stricto sensu*). Elle vise plus spécifiquement à rechercher des preuves d'évènements de sélection ayant agi postérieurement à l'introduction de *C. fornicata* en dehors de son aire d'origine. De nouveaux outils moléculaires, des locus microsatellites basés sur l'exploitation d'une banque d'EST (article 1) et des AFLPs ont été développés pour répondre à cet objectif. Une approche de scan génomique (basée sur 344 marqueurs AFLPs et microsatellites) utilisant 22 populations échantillonnées dans l'aire d'origine et les aires d'introduction a ensuite été menée afin de tester l'hypothèse de processus de sélection post-introduction (article 2).
- dans une deuxième partie, nous nous sommes intéressées au cycle de vie de la crépidule et aux effets de dérive génétique qui pourraient en résulter. Contrairement à la première partie, les analyses réalisées ici se sont focalisées sur une petite échelle spatiale mais intégrant une dimension temporelle (intra- et inter-génération). Nous avons étudié une population introduite de *C. fornicata* localisée en baie de Morlaix. Les questionnements et les échantillons étudiés sont issus des résultats d'un suivi des populations benthique et pélagique initié en 2004 (benthos) et 2005 (pelagos) en un site (Château du Taureau) de cette baie. Les principaux résultats obtenus au cours de ces suivis auxquels j'ai contribué au cours de ma thèse sont présentés dans le chapitre 1. Nous avons ensuite conduit une série

d'analyses pour 1) estimer la taille efficace de la population (article 3), 2) tester l'hypothèse de « loterie familiale » (scénarios de type « sweepstake » proposé par Hedgecock 1994a) dans laquelle un nombre limité de familles assurerait le renouvellement complet de la population (article 4) et 3) analyser la relation entre la composition génotypique multi-locus des individus et leur taille pour tester l'hypothèse de corrélation hétérozygotie-valeur sélective chez les larves de ce mollusque (article 5).

Note : A l'exception du chapitre 1 de la seconde partie, les études menées sont présentées sous forme d'articles scientifiques publiés, soumis ou à soumettre dans des revues internationales. En préambule de ces articles, un résumé en français est proposé en début d'article pour rappeler le contexte de l'étude et présenter les principaux résultats obtenus. A la fin de ce résumé, un petit encart en italique indique un référencement d'annexes. Ces annexes sont des protocoles ou des résultats d'analyse complémentaires qui ne sont pas présentés dans les articles.





---

**Première partie :**

**Sélection et Invasion**

---



Les différences morphologiques, physiologiques ou comportementales observées entre individus d'une même espèce relèvent pour une large part de leur réponse au milieu. En favorisant les allèles/phénotypes les plus avantageux dans un environnement donné, la sélection naturelle favorise une meilleure adéquation entre l'individu et son environnement. Cette sélection s'exerce sur des caractères variables et héréditaires. Comprendre les mécanismes moléculaires et les bases génétiques à l'origine des processus adaptatifs est un des enjeux majeurs en biologie évolutive (Luikart *et al.* 2003, Vasemagi et Primmer 2005, Nielsen *et al.* 2009). Pour cela, plusieurs approches impliquant des études moléculaires peuvent être utilisées, dont :

- l'approche « gène candidat » : cette approche prend comme point de départ un gène donné, dont le polymorphisme ou l'expression est relié à un phénotype qui confère un avantage adaptatif (Philips 2005). La fonction du gène étudié est supposée connue, soit directement chez l'espèce cible, soit chez une autre espèce chez laquelle le gène est susceptible d'avoir la même fonction. Les gènes choisis pour cette approche sont souvent des gènes structuraux ou impliqués dans des processus physiologiques (exemples de gènes dans Ford 2002, Nielsen *et al.* 2009). Cette approche nécessite donc de connaître la fonction du gène étudié, d'identifier les polymorphismes existant sur ce gène et/ou de ses environs, et de relier la fonction putative du gène à un trait adaptatif, rendant difficilement envisageable cette approche chez certaines espèces pour lesquelles des données génomiques sont peu disponibles.

- l'approche « QTL » (Quantitative Trait Locus) : la plupart des traits phénotypiques ont des caractères quantitatifs à déterminisme génétique complexe ; cette approche consiste à mesurer l'association entre le génotype à différents locus et la valeur du trait phénotypique associé. On recherche des marqueurs génétiques significativement associés à un caractère quantitatif (*i.e.* caractère pour lesquels les individus présentent une valeur du caractère se répartissant de manière continue dans une certaine gamme). L'identification de QTL se fait généralement par l'étude de croisements entre lignées qui se différencient phénotypiquement pour les traits d'intérêt, permettant d'établir des associations entre les génotypes à différents locus en regard du phénotype obtenu. Les traits phénotypiques quantitatifs étudiés sont souvent liés à la croissance, à la morphologie ou encore à certains traits d'histoire de vie (*e.g.* date de floraison chez les plantes), mais peuvent également être l'expression d'un gène (eQTL, Gilad *et al.* 2008). L'approche « QTL » repose donc essentiellement sur un grand nombre de marqueurs, l'identification des traits adaptatifs et l'établissement de lignées, limitant souvent cette approche à des espèces facilement cultivables/élevables pour lesquelles des croisements entre lignées peuvent être réalisés.

- l'approche dite de balayage génomique ou « scan génomique » (terme dérivé de l'expression anglaise « genome scan » que j'utiliserai dans ce manuscrit) : cette approche ne nécessite ni l'identification de

trait phénotypique adaptatif précis ni de connaissance *a priori* sur le génome. Elle repose sur l'analyse simultanée d'un grand nombre de marqueurs distribués largement sur le génome afin de discriminer des locus sous sélection ou liés à des locus sous sélection des autres marqueurs. Cette approche est détaillée dans l'encadré 2.

Les connaissances acquises sur *C. fornicata* ne permettaient pas de conduire dans le cadre de ce travail de thèse d'approches gènes candidats ou QTLs pour explorer d'éventuels processus de sélection ayant agi lors du processus d'introduction. En effet, *C. fornicata* est une espèce pour laquelle très peu d'informations sur le génome et le transcriptome sont disponibles (Henry *et al.* 2010). Par ailleurs, il existe peu d'études documentant des traits phénotypiques adaptatifs sur lesquels de telles approches pourraient être menées. Quelques études suggèrent que la croissance larvaire puisse être un de ces traits. En effet, c'est un caractère héritable (Hilbish *et al.* 1999) et variable (Pechenik et Levine 2007, Le Cam *et al.* 2009). Les gènes impliqués dans le développement larvaire pourraient être des candidats potentiels. Des travaux ont été initiés à ce sujet dans l'équipe (post-doctorat de Nicolas Taris) qui ont permis d'identifier des gènes impliqués dans l'acquisition de la compétence et/ou la métamorphose (cf. Synthèse et Perspectives) mais les avancées réalisées étaient insuffisantes pour être mises à profit dans le cadre de cette thèse. Par ailleurs, il n'existe aucune donnée documentant des changements de développement entre aire d'origine et aire d'introduction de la crépidule. La plus importante variation observée chez la crépidule entre populations natives et introduites concerne le nombre d'individus moyens par chaîne. Au sein de l'aire native, ce nombre varie entre 2,1 et 3,2 (Hoch et Cahill, 2012) tandis qu'il est de 9,8 au sein de l'aire introduite (données acquises en baie de Morlaix, Le Cam 2009). Un grand nombre d'individus au sein d'une même chaîne pourrait favoriser une reproduction plus efficace et être un signe d'une augmentation de la compétition interspécifique dans l'aire d'introduction. Néanmoins, il n'existe aucune étude suggérant que ce caractère puisse être en partie au moins génétiquement déterminé. Par ailleurs, la difficulté d'élever et de croiser en laboratoire *C. fornicata* empêche de mener à bien des approches de type QTL. Cette série de difficultés nous a conduits à développer une approche plus globale basée sur des balayages génomiques, l'approche par scan génomique. De plus, cette approche nous paraissait la plus appropriée compte-tenu de nos objectifs, qui étaient en premier lieu de rechercher des traces d'évènements de sélection en lien avec le processus d'introduction. Nécessitant l'utilisation de nombreux marqueurs et en l'absence de données génomiques pour l'espèce cible, nous avons choisi d'utiliser des marqueurs AFLP (Amplified Fragment Length Polymorphism). Cette technique est basée sur la détection de mutations ponctuelles sur le génome situées au niveau des sites de coupure des enzymes de restriction. Elle ne nécessite pas de connaissance préalable sur le génome et permet d'obtenir des profils génétiques sur de nombreux individus pour de nombreux marqueurs, y compris chez des espèces non modèles (Bonin *et al.* 2007).

Au cours de ce travail de thèse, nous avons également développé de nouveaux locus microsatellites pour l'étude de *C. fornicata*. Ces développements ont été menés en exploitant une banque d'EST (Expressed-Sequence Tags) préalablement construite chez *Crepidula fornicata*. Cela nous a permis d'une part d'accroître la diversité des marqueurs utilisés dans notre étude de scan génomique (cf. article 2) et, d'autre part, de disposer de nouveaux outils pour l'étude de la diversité génétique des populations de *C. fornicata* (études présentées dans la deuxième partie de ce manuscrit de thèse). Plusieurs marqueurs microsatellites avaient été déjà développés chez cette espèce, (Dupont et Viard 2003, Proestou 2006, Dupont *et al.* 2007a) mais leur étude dans de nombreuses populations naturelles et dans le cadre d'analyses de paternité avait mis en évidence la présence d'allèles nuls pour certains d'entre eux. Nous souhaitons ainsi à la fois augmenter le nombre de marqueurs disponibles et minimiser le nombre de locus avec des allèles nuls.

Suite au développement de ces différents marqueurs (AFLP et microsatellites), détaillé dans le chapitre 1 (article 1 et annexe 2), l'hypothèse de processus de sélection au cours du processus d'introduction de *C. fornicata* en dehors de son aire d'origine a été testée par une approche de scan génomique. Cette étude fait l'objet du chapitre 2 de cette première partie.

## Encadré 2 - « Scan génomique » : les méthodes statistiques de détection de signatures de sélection

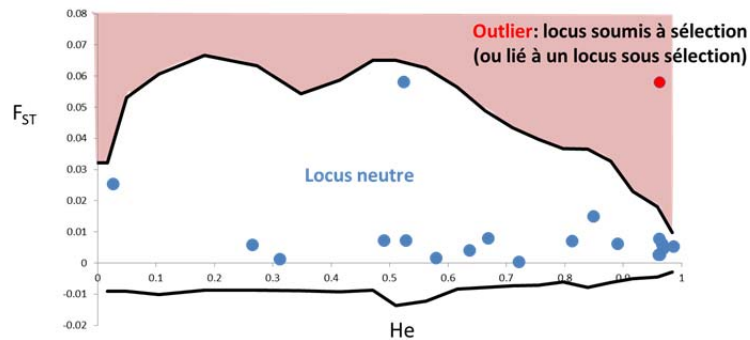
Les effets de sélection sont propres à certain locus tandis que les effets démographiques (dérive génétique, migration) affectent de manière identique l'ensemble du génome (Cavalli-Sforza 1966). Une méthode pour identifier les bases génétiques de la sélection consiste à détecter ce nombre restreint de locus se comportant différemment de la plupart des autres locus du fait d'effets sélectifs. Les étapes conduisant à la détection de ces locus « atypiques » ou « outliers » (terme anglais que j'utiliserai dans ce manuscrit), soumis à sélection ou en déséquilibre avec des locus soumis à sélection, peuvent être schématisés de la façon suivante : (1) la distribution du polymorphisme génétique d'un grand nombre de locus est estimée afin d'inférer la manière dont la démographie affecte le polymorphisme sous l'hypothèse de neutralité et d'un modèle connu de structure des populations ; (2) cette distribution est comparée entre locus afin d'identifier des marqueurs présentant un polymorphisme singulier (*i.e.* se distinguant de l'histoire commune partagée par tous les locus).

C'est en 1973 que ce raisonnement est formalisé pour la première fois (test de Lewontin et Krakauer) : tous les individus partagent une histoire démographique commune et les variations des fréquences alléliques entre populations sont expliquées par des effets de dérive génétique (hypothèse de neutralité); par conséquent tous les locus ont un  $F_{ST}$  comparable ( $F_{ST}$  de Wright, 1951, utilisé ici pour décrire les différences de fréquences alléliques entre populations) sauf ceux affectés par de la sélection. Lewontin et Krakauer (1973) comparent la distribution des valeurs de  $F_{ST}$  avec celle d'un  $\chi^2$  avec (n-1) degré de liberté (hypothèse de distribution normale des fréquences alléliques, et n le nombre de populations), et la variance inter-locus des valeurs de  $F_{ST}$  observées est comparée aux valeurs attendues sous l'hypothèse d'un modèle neutre ( $\sigma^2=2 F_{ST}^2/(n-1)$ ), permettant ainsi d'accepter ou rejeter le modèle neutre.

Mais, ce modèle s'avère très restrictif et est critiqué (*e.g.* Nei et Maruyama 1975, Robertson 1975) : il ne tient pas compte de différents paramètres ou histoires évolutives affectant la démographie (*e.g.* migration entre deux populations ; histoire commune récente des populations) et les approximations faites supposent entre autres de grands échantillons, un nombre identique d'allèles à chaque locus, des  $F_{ST}$  égaux entre populations. Enfin, ce modèle ne permet pas d'identifier de locus en particulier. Point de départ des méthodes statistiques pour détecter des signes de sélection, ce modèle subit des modifications successives afin de prendre en compte des scénarios démographiques plus réalistes, augmentant ainsi la fiabilité du modèle.

Une première modification est apportée par Bowcock *et al.* (1991) : son modèle estime la distribution des  $F_{ST}$  conditionnellement aux fréquences alléliques générées sous l'hypothèse d'un modèle neutre et d'un scénario démographique fixe ; les marqueurs atypiques montrent des excès significatifs des valeurs de  $F_{ST}$  en comparaison aux valeurs attendues sous le modèle neutre. Mais ce modèle s'avère extrêmement dépendant du scénario évolutif inféré initialement et peu robuste aux hypothèses alternatives.

Le modèle développé par Beaumont et Nichols (1996), repose sur des simulations réalisées à partir d'un modèle en îles, permettant d'estimer pour chaque locus la distribution théorique des  $F_{ST}$  conditionnellement à l'hétérozygotie attendue ( $H_e$ ) sous l'hypothèse neutre. Cette distribution conjointe des valeurs de  $F_{ST}$  et  $H_e$  se révèle robuste sous différents scénarios démographiques (différents taux de mutation, différentes tailles de populations, différents modèles de migration, présence de consanguinité...). Les locus présentant une valeur conjointe  $F_{ST}/H_e$  plus forte ou plus faible que celle attendue sous l'hypothèse neutre sont interprétés comme étant sous l'action de la sélection. L'identification de ces locus atypiques se fait de la manière suivante : (1) pour chaque locus étudié, on estime une valeur de  $F_{ST}$  et de  $H_e$  ; (2) le  $F_{ST}$  multi-locus moyen est estimé (moyenne des  $F_{ST}$  calculés sur l'ensemble des locus) ; cette valeur correspond alors à la valeur du  $F_{ST}$  attendue sous l'hypothèse de neutralité ; (3) des simulations sont réalisées, ciblant le  $F_{ST}$  moyen précédemment calculé, et permettant d'inférer l'évolution des fréquences sous le modèle neutre (modèle en îles). La distribution conjointe théorique neutre des valeurs de  $F_{ST}$  et  $H_e$  est ainsi estimée; (4) on définit l'enveloppe « neutre » (quantiles de la distribution conjointe  $F_{ST}$  et  $H_e$  attendue sous l'hypothèse de neutralité), dans laquelle 99% des couples  $F_{ST}/H_e$  sont théoriquement observés ; (5) les locus potentiellement soumis à sélection apparaissent au-dessus/en-dessous de l'enveloppe de confiance (enveloppe neutre), montrant une différenciation plus forte ou plus faible qu'attendue.



**Fig. 1** Identification des locus sous sélection selon l'approche développée par Beaumont et Nichols (1996). L'enveloppe neutre est symbolisée par les lignes noires, les locus se comportant comme des locus neutres sont symbolisés en bleu et les locus soumis à sélection en rouge.

Bien que ce modèle apparaisse plus réaliste que les modèles précédemment développés, les locus situés hors de l'enveloppe neutre ne sont pas pour autant réellement soumis à l'action de la sélection. En effet, de nombreux faux-positifs sont détectés, liés notamment à la multiplicité des tests réalisés. Par ailleurs, des biais dans les résultats sont identifiés, liés à des écarts importants aux hypothèses initiales (taux de mutation faible par rapport à la migration, corrélation des fréquences alléliques entre dèmes) (Storz 2005). Cette approche est implémentée dans le logiciel Dfdist (pour des marqueurs dominants).

Beaumont et Balding (2004) modifient la méthode développée par Beaumont et Nichols (1996), en apportant au modèle une approche hiérarchique bayésienne. Cette approche bayésienne est utilisée pour générer la distribution théorique des  $F_{ST}$  conditionnellement à l'hétérozygotie, et permet des variations de paramètres entre populations (mutation, taille des populations). Enfin, la valeur moyenne du  $F_{ST}$  ciblée attendue sous l'hypothèse neutre est la valeur moyenne du  $F_{ST}$  auquel on enlève 30% de valeurs les plus hautes et les basses, ces locus pouvant être influencés par la sélection. Par le scénario démographique inféré et l'approche bayésienne, cette méthode semble plus robuste que les précédentes (Foll et Gaggiotti 2006), réduisant les biais dus au modèle démographique utilisé et ciblant un  $F_{ST}$  supposé neutre.

Foll et Gaggiotti (2008) ont également développé une approche bayésienne, permettant notamment de faire varier les paramètres démographiques entre populations (mutation et taille de la population). Le modèle est une régression logistique estimant directement la probabilité postérieure de chaque locus d'être sous l'action de la sélection ou non. Le coefficient  $F_{ST}$  se décompose en une composante locus-spécifique ( $\alpha$ ) partagé par toutes les populations, et en une composante population-spécifique ( $\beta$ ) partagé par tous les locus (combinaison linéaire). Deux modèles alternatifs (avec et sans  $\alpha$ ) sont testés et lorsque la composante  $\alpha$  est nécessaire pour expliquer le patron de diversité observé d'un locus, ce locus est influencé par la sélection. Ce modèle a l'avantage de prendre en compte la multiplicité des tests réalisés, minimisant ainsi les faux-positifs et a été montré comme l'un des modèles les plus fiables pour détecter des signatures de sélection (Pérez-Figueroa *et al.* 2010). Cette approche est implémentée dans le logiciel BayeScan.

Contrairement aux approches développées jusque-là supposant différents scénarios démographiques, le modèle de Vitalis *et al.* (2001) s'affranchit de l'histoire démographique, souvent inconnue, pour détecter la sélection « population-spécifique ». Deux populations sont comparées, ces populations partageant une histoire commune (équilibre démographique) avant de diverger. Cette approche permet de réduire le nombre de paramètres du modèle, d'identifier les locus qui n'ont pas divergé de la même façon entre les deux populations et de détecter la sélection « locale ». Mais ce modèle n'intègre ni la dispersion, ni d'hypothèse alternative (*i.e.* sans sélection) et la multiplicité des tests (tests par paire de populations) augmentent la détection de faux-positifs.





# Chapitre 1 - Développement de nouveaux marqueurs génétiques chez l'espèce invasive *Crepidula fornicata*

---

## I. Contexte de l'étude

L'identification de marqueurs soumis à sélection, de manière directe ou indirecte, par une approche de scan génomique nécessite l'utilisation d'un grand nombre de marqueurs moléculaires répartis le plus uniformément possible dans le génome. Plusieurs types de marqueurs ont récemment été utilisés pour identifier des marqueurs soumis à sélection (Luikart *et al.* 2003, Nosil *et al.* 2009), dont notamment les microsatellites (*e.g.* Negrini *et al.* 2012), AFLPs (Amplified Fragment Length Polymorphism, *e.g.* Mattersdorferou *et al.* 2012) ou encore des SNPs (Single Nucleotide Polymorphism, *e.g.* Limborg *et al.* 2012).

Par ailleurs, l'utilisation de marqueurs microsatellites (ou single sequence repeats -SSRs) présente de nombreux avantages en génétique des populations : ce sont des marqueurs génétiques riches en information du fait notamment de leur codominance et multi allélisme. Ils sont abondants dans le génome et répartis le long du génome. Les marqueurs microsatellites apparaissent ainsi comme des marqueurs de choix pour estimer la diversité génétique au sein et entre populations.

Chez *Crepidula fornicata*, un petit nombre de marqueurs microsatellites avait été mis au point (Dupont et Viard 2003, Proestou *et al.* 2008, Dupont *et al.* 2006), dont sept utilisés en routine au laboratoire. Des analyses de paternité ont cependant révélé la présence d'allèles nuls en grande fréquence pour trois de ces marqueurs (Dupont *et al.* 2006, 2007a), pouvant biaiser l'estimation de la diversité génétique et de la structure génétique (Chapuis et Estoup 2007). Dans cette étude, nous avons donc cherché à développer de nouveaux marqueurs microsatellites, et ainsi, augmenter le nombre de marqueurs utilisables pour une meilleure estimation de la diversité génétique (pour inférer avec plus de précision la structure génétique).

## II. Résultats principaux de l'étude (détaillés dans l'article 1 et l'annexe 1)

Par un criblage *in silico* d'une banque d'ADNc construite sur l'espèce *Crepidula fornicata* et comportant 2669 unigènes, 361 EST contenant un motif microsatellite (EST-SSRs) ont été identifiés, représentant donc 13.5% du nombre initial d'unigènes. Parmi ces 361 EST-SSRs, 85 ont été sélectionnés, présentant un motif microsatellites répétés au minimum 15 fois et une longueur attendue variant entre 100 et 400pb. Parmi ces 85 marqueurs, 42 ont été correctement amplifiés sur gel d'agarose et 12 locus, dont 3 di-, 2 tri-, 3 tétra- et 4 penta-nucléotides, ont été sélectionnés en vue de l'optimisation des conditions d'amplification pour du génotypage en routine à partir de larves, juvéniles et adultes de *C. fornicata*. Ces 12 nouveaux marqueurs peuvent être amplifiés grâce à trois multiplexes (*i.e.* co-amplification de plusieurs locus lors d'une seule PCR). Au final, le développement

de microsatellites à partir d'un criblage *in silico* d'une banque EST a montré un rendement similaire (3.3% de motifs utilisables parmi les 2669 unigènes initiaux) au développement de marqueurs microsatellites à partir de banque génomique enrichie en motifs microsatellites à partir d'ADN total, pour *C. fornicata* (2.3%, Dupont et Viard 2003) et pour une autre espèce proche de *C. fornicata* (*C. convexa* : 1.4%, Daguin-Thiébaud *et al.* 2009). La phase critique du développement de marqueurs microsatellites est dans ces deux cas l'étape d'optimisation pour du génotypage en routine (et l'obtention de profils non ambigus, faciles à lire et répétables).

Les 12 microsatellites identifiés sont hautement spécifiques de *Crepidula fornicata*, comme en témoignent les résultats négatifs d'amplification sur 8 autres espèces du genre, et peuvent donc être utilisés comme marqueurs diagnostics pour notamment identifier *C. convexa* de *C. fornicata*, difficilement distinguables sur des critères morphologiques (McGlashan *et al.* 2008).

A partir d'une descendance larvaire, nous avons pu confirmer l'hérédité mendélienne de 11 de ces locus (le douzième locus étant monomorphe dans la descendance testée). Ces résultats sont présentés en Annexe 1. Cette analyse a permis de montrer que les locus étaient facilement utilisables pour des analyses d'ADN extrait de larves (donc en faible quantité et en qualité moins bonne que chez des adultes). L'analyse de cette descendance n'a pas mis en évidence d'allèles nuls (mais une seule descendance a été testée ici).

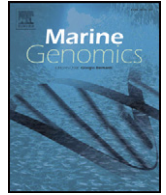
Le polymorphisme de ces 12 locus a été testé en analysant 88 adultes échantillonnés dans une population (baie de Morlaix ; Bretagne ; France). Ces marqueurs montrent un polymorphisme variable avec deux et quarante allèles et une hétérozygotie attendue de 0.011 à 0.953 selon les locus. La qualité des génotypes obtenus est très bonne, suggérant que les erreurs de génotypages devraient être limitées. La présence d'allèles nuls, communément reportés chez les invertébrés marins (Yu et Li 2008, David *et al.* 1997, Lemer *et al.* 2011), ne peut cependant pas être complètement exclue dans ce nouveau jeu de marqueurs surtout pour les locus les plus polymorphes.

Ces nouveaux marqueurs ont été développés à partir de banques EST (Expressed-Sequence-Tag) et donc dans les régions exprimées du génome, potentiellement davantage soumises à sélection que les marqueurs microsatellites développés à partir de banques génomiques « anonymes ». Nous avons comparé les informations données par ces marqueurs lors de l'analyse des 88 individus adultes à celles obtenues avec les 7 locus microsatellites anonymes déjà développés chez cette espèce. Leurs propriétés (diversité, écart aux attendus d'Hardy-Weinberg, etc.) sont similaires à celles obtenues avec les 7 locus microsatellites « anonymes ». L'utilisation combinée de ces 19 locus résulte en une très forte valeur de probabilité d'exclusion ( $p=0.99$ ), soulignant l'utilité de ces locus pour des analyses à très fine échelle telles que des analyses de parenté et paternité.

**III. Article 1: « *In silico* mining and characterization of 12 EST-SSRs for the invasive slipper limpet *Crepidula fornicata* »**

« Identification *in silico* et optimisation de 12 locus EST-SSRs chez l'espèce invasive *Crepidula fornicata* »

Florentine Riquet, Marion Ballenghien, Arnaud Tanguy et Frédérique Viard – publié dans *Marine Genomics*



Short communication

## *In silico* mining and characterization of 12 EST-SSRs for the invasive slipper limpet *Crepidula fornicata*

Florentine Riquet<sup>a,b,\*</sup>, Marion Ballenghien<sup>a,b,1</sup>, Arnaud Tanguy<sup>b,c</sup>, Frédérique Viard<sup>a,b,\*\*</sup>

<sup>a</sup> UPMC Univ Paris 06, UMR 7144, Team Diversity and Connectivity in Coastal Marine Landscapes, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>b</sup> CNRS, UMR 7144, Lab. Adaptation and Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>c</sup> UPMC Univ Paris 06, UMR 7144, Team GAME, Station Biologique de Roscoff, 29680 Roscoff, France

### ARTICLE INFO

#### Article history:

Received 15 June 2011

Received in revised form 29 July 2011

Accepted 3 August 2011

#### Keywords:

*Crepidula fornicata*

EST-SSRs

Non-indigenous species

Relatedness analysis

Mollusc

### ABSTRACT

*In silico* mining of an expressed sequence tags (ESTs) library was found to be efficient at isolating simple sequence repeats (SSRs) loci in the non-indigenous marine mollusc *Crepidula fornicata*. Twelve SSR loci were developed for routine genotyping. Cross-species amplification to 8 other *Crepidula* species showed that the 12 loci are highly specific for *C. fornicata*. Mendelian inheritance was shown for 11 of them (1 being monomorphic in the analyzed offspring array). The genetic diversity for 88 adults was found to be variable across the 12 loci (2–40 alleles, expected heterozygosity between 0.023 and 0.898) with a high overall exclusion probability of 0.99. The degree of genetic polymorphism found here is similar to that shown for 7 anonymous SSRs previously developed and here used on the same samples. This set of 12 specific loci is relevant to perform reliable population and relatedness analyses in *Crepidula fornicata*.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

The increased amount of genomic resources in marine species, including in non-model organisms, represent a novel opportunity for finding markers to be used for intra-specific studies in ecological and evolutionary applications, such as population genomics, conservation genetics and molecular ecology (Allendorf et al., 2010). In particular, *in silico* mining of genomic data can be a reliable way to obtain rapidly microsatellite loci (i.e. Simple Sequence Repeats—SSRs) in marine species (e.g. Vogiatzi et al. (2011) for a recent example in sea bream).

*Crepidula fornicata* is a marine mollusc native to the American North East coast that has been targeted in recent developmental, ecological and genomic research, increasing its potential as a model

system for evolutionary studies (see Henry et al., 2010 for a review). For instance, Taris et al. (2009, 2010) identified cues for metamorphosis and their effects on gene expression during larval development. *C. fornicata* is also widely studied as an emblematic invasive species in Europe, where it has been introduced repeatedly during the 20th century (Blanchard, 1997).

In marine non-indigenous species (NIS) with a free-swimming larval stage, like *C. fornicata*, larval supply has been suggested to be a key component of invasion success (Rigal et al., 2010; Roman and Darling, 2007). Unfortunately, larvae cannot usually be tracked directly; thus understanding recruitment, dispersion and colonization dynamics requires the development of indirect approaches, in particular the molecular estimation of dispersal parameters (Broquet and Petit, 2009; Selkoe et al., 2008). Microsatellite-based studies were previously performed in *C. fornicata* to assess paternity (Dupont et al., 2007b, 2006; Le Cam et al., 2009; Viard et al., 2006). However, the number of reliable markers was limited, largely because of the presence of null alleles (Dupont et al., 2006, 2007a) decreasing dramatically the number of reliable loci that could be used in this species (Le Cam et al., 2009).

We here focus on the development of novel microsatellites relevant for parentage and dispersion analyses in the marine mollusc *Crepidula fornicata*, by screening an Expressed Sequence Tags (ESTs) library previously developed for this species, a rapid and low-cost development of microsatellite markers compared with the development of a dedicated enriched genomic library.

\* Correspondence to: F. Riquet, Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France. Tel.: +33 298292544; fax: +33 298292324.

\*\* Correspondence to: F. Viard, Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France. Tel.: +33 298292312; fax: +33 298292324.

E-mail addresses: [riquet@sb-roscoff.fr](mailto:riquet@sb-roscoff.fr) (F. Riquet), [viard@sb-roscoff.fr](mailto:viard@sb-roscoff.fr) (F. Viard).

<sup>1</sup> Present address: CNRS, UMR 5554, Institut des Sciences de l'Evolution, Lab. Phylogénie et Evolution moléculaire, Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex05, France.

## 2. Materials and methods

### 2.1. EST-SSRs detection, primer design and development of multiplex PCRs

As part of a previous transcriptomic study, an ESTs library had been developed for *Crepidula fornicata* (N. Taris, A. Tanguy, C. Dreanno and F. Viard, unpublished data). Briefly, several cDNA and subtractive (SSH) libraries were obtained from total RNA extracted from cephalic tissues from 6 males and 6 females, pool of > 1000 larvae and juveniles (all tissues) using RNable (Eurobio, les Ulis, France). Poly(A+) mRNA was isolated from total RNA using the PolyATtract® mRNA Isolation System (Promega, Madison, WI, USA) according to the manufacturer's instructions. Subtractive libraries were obtained using the PCR-select cDNA subtraction kit (Clontech, Palo Alto, CA, USA). A non-normalized full-length cDNA library was constructed using the SMART PCR cDNA synthesis kit (Clontech). From 5662 clones sequenced, a local database of 2669 ESTs (unigenes) was obtained after redundancy elimination. For the present study, EST-SSRs were identified from these 2669 unigenes using the command line version of the SSR primer program which is integrating the SSR repeat finder SPUTNIK and the PCR primer design software Primer3 (Robinson et al., 2004). The SPUTNIK procedure was restricted to SSRs motifs with a minimum number of 15 repeats and an expected length between 100 bp and 400 bp. From the output of the program, the sequences containing EST-SSRs were retrieved and blasted in a local database created with these sequences and with previously developed microsatellites (Dupont and Viard, 2003; Dupont et al., 2006; Proestou, 2006) to avoid any redundancy. PCR primers were designed using Primer3 and MPrimer3 (Rozen and Skaletsky, 2000; Koressaar and Remm, 2007) with an optimal size of the core repeated region around 20 bp and a range from 20 to 25.

DNA of nine adults of *C. fornicata* obtained according to the protocol of Dupont et al. (2006) were used for preliminary tests with touch-down PCR programs carried out with a PTC200™ thermocycler (MJ Research), started with 5 min at 95 °C, 5 cycles with 30 s at 95 °C, 30 s at initial temperature (annealing temperature ( $T_A$ ) + 5 °C = 58, 55, 52 or 45 °C) then decreasing by 1 °C/cycle, 45 s at 72 °C followed by 32 cycles at 95 °C for 30 s,  $T_A$  for 30 s, 72 °C for 45 s and finished at 72 °C for 10 min. Based on banding patterns checked on agarose and polyacrylamide sequencing gels (LI-COR 4200™ sequencer, LI-COR® Biosciences), EST-SSRs with clear banding patterns were selected for routine genotyping on a capillary sequencer. Before being processed on an ABI 3130xl DNA sequencer (Applied Biosystems™), amplification products were obtained using touch-down PCRs (program as described above). Briefly, 1 µL of DNA (stock DNA diluted 1:100) was mixed with 0.9× buffer (Thermoprime, ABgene®), 0.13 mM of each dNTP, 1 mM MgCl<sub>2</sub>, 0.05 µg/µL of bovine serum albumin, 0.4U of Taq polymerase (Thermoprime, ABgene®) and 0.17 µM of each primer. Following amplification, 2 µL of water diluted 1:10 amplicon were pooled with 9.7 µL of deionized formamide and 0.3 µL GeneScan-600LIZ Size Standard (Applied Biosystems™). Scoring of alleles was performed using GeneMapper® version 4.0 (Applied Biosystems™). Following PCR optimization and polymorphism analysis of each primer pair, a multiplex procedure was performed: 0.17 µM of each primer was initially used in all multiplex reactions and primer concentrations were modified subsequently to score peak heights between 600 and 3000 relative fluorescent units for each marker.

### 2.2. Cross-species amplification

These newly developed EST-SSRs were investigated for cross-species amplification in 9 related calyptraeid species (Collin, 2003): *Crepidula fornicata*, *Crepidula astrasolea*, *Crepidula onyx*, *Crepidula depressa*, *Crepidula fecunda*, *Crepidula dilatata*, *Crepidula convexa*, *Crepidula plana* and *Crepidula coquimbensis*. DNA was extracted from two adults of each species using a Nucleospin® Multi-96 Tissue Kit

(MACHEREY-NAGEL) following the manufacturer's protocol, and amplified over the 12 EST-SSR loci as previously described.

### 2.3. Mendelian inheritance and null alleles

*C. fornicata* adults are long-lived and sedentary, with a stacking behavior and internal fertilization; experimental progeny-arrays are thus difficult to obtain in the laboratory. So, to get insights about Mendelian inheritance of the optimized EST-SSRs, we used results of a published paternity analysis (Le Cam et al., 2009) by re-analyzing 25 larvae previously assigned to a single pair of parents with 4 anonymous SSRs (family D in Le Cam et al., 2009). Using the software Colony 2.0.0.1 (Wang and Santure, 2009; Wang, 2004), we first confirmed the paternal assignment with our new set of loci. Then, a  $\chi^2$  analysis was performed using the software R 2.10.1. (R Development Core Team 2007) to test inheritance of the others EST-SSRs (null hypothesis: segregating alleles do not deviate from Mendelian inheritance ratio at a significance level  $\alpha=0.05$ ). For an expected ratio of 1:1, 10,000 simulations were performed. Presence of null alleles was checked first using this paternity analysis and then using Micro-Checker (Van Oosterhout et al., 2004).

### 2.4. Polymorphism evaluation

The genetic polymorphism of the EST-SSRs was analyzed by genotyping 88 adults sampled by scuba diving in the Bay of Morlaix (Brittany, France) in April 2007. Genotypes of every individual were checked twice and for some of them up to 15 times during the different steps during the optimization procedure. The software GENEPOP 4.0 (Rousset, 2008) was used to estimate expected heterozygosity ( $H_e$ ), fixation index ( $F_{IS}$ ) and to test departure from Hardy–Weinberg equilibrium and linkage disequilibrium between loci. Results obtained with the EST-SSRs were compared to those obtained with 7 anonymous SSRs previously developed (Dupont and Viard, 2003; Dupont et al., 2006; Proestou, 2006). The same set of 88 individuals was genotyped following protocols described in the papers cited above. Exclusion probability of the 12 EST-SSRs and the 7 anonymous SSRs was estimated using Gerud 2.0 (Jones, 2005).

## 3. Results and discussion

### 3.1. EST-SSRs type and distribution in the expressed sequences

Among 2669 unigenes, 361 EST-SSRs (13.4%) were identified, most of them (50.7%) being di-nucleotide SSRs (Table 1); 130 (36%) of these were selected with the following criteria: a minimum of 15 repeats and a predicted size of 100–400 bp. From these, primers for 85 EST-SSRs (65.4%) were designed: 29.4% of the primer pairs were synthesized for di-nucleotides (25) and tetra-nucleotides (25), and less for tri-nucleotides (20) and penta-nucleotides (15). Using agarose gel, reliable amplification was obtained for 42 EST-SSRs. After checking the amplification patterns on acrylamide gel and then on capillary sequencer, 12 EST-SSRs (GenBank IDs: JK312475 to JK312486) were selected for routine genotyping: 3 di-nucleotides, 2 tri-nucleotides, 3 tetra-nucleotides and 4 penta-nucleotides. Three multiplex PCRs were developed to amplify these 12 EST-SSRs. The final primer sets, loci combinations, their annealing temperatures, and MgCl<sub>2</sub> concentrations are detailed in Table 2.

Using a conventional microsatellite-enriched genomic library built for *C. fornicata*, Dupont and Viard (2003) isolated 7 loci from 300 sequenced clones. In another species of the same genus, namely *C. convexa*, Daguin-Thiebaut et al. (2009) isolated 11 loci from 768 clones for routine genotyping. The output of SSRs characterization from *in silico* mining in *C. fornicata* is thus similar to the yield of microsatellite development using a conventional genomic library (EST-SSRs: 3.3%; anonymous SSRs: 2.3% for *C. fornicata* and 1.4% for

**Table 1**

Number of EST-SSRs sorted according to the core repeat type, at each step of the characterization process.

	Dinucleotide	Trinucleotide	Tetranucleotide	Pentanucleotide	Total
SSRs found in the ESTs library	183	85	60	33	361
EST-SSRs with >15 repeats and predicted size of 100–400 bp	42	32	36	20	130
EST-SSRs for which primers were designed	25	20	25	15	85
EST-SSRs with a correct banding pattern on agarose gel <sup>(1)</sup>	13	11	10	8	42
EST-SSRs selected for multiplex and routine genotyping	3	2	3	4	12

(1) A “correct” banding pattern refers to the occurrence of a single (or two) clear band(s) with a size fitting with the expected size based on the cDNA sequence.

*C. convexa*). *In silico* mining of SSRs is well worth doing, as it is less costly and much less time-consuming.

**3.2. Cross-species amplification of 12 EST-SSRs within the genus *Crepidula***

Because EST-SSRs are located in transcribed genomic regions, they are expected to be more conserved and thus more transferable across taxonomic boundaries than anonymous SSRs (Ellis and Burke, 2007). The 12 EST-SSRs were tested within the genus *Crepidula* across 8 species widely distributed in the phylogenetic tree of the Calyptraeidae (Collin, 2003). The 12 EST-SSRs amplified exclusively DNAs of *C. fornicata*. The absence of cross-species amplification may be explained by a rapid divergence of the flanking sequences of the microsatellite motif, due to the combination of the high mutation rates of microsatellites (range: 10<sup>-2</sup> to 10<sup>-6</sup>; Dallas, 1992; Goldstein et al., 1995) and the large population size typical of molluscs (Waples, 1998; Palsboll et al., 2007). Besides unstable flanking sequences, failure of cross-species amplification may also be due to deletion in the targeted species of the whole sequence where the primers were designed in the source species. For example, by studying three gastropods, McInerney et al. (2011) suggested a role of genomic rearrangement processes mediated by transposable elements containing microsatellites. The absence of cross-species amplification nevertheless points out the usefulness of the 12 EST-SSRs in distinguishing among species, an important point considering that *C. fornicata* and *C. convexa* can be mis-identified when using morphological criteria alone (McGlashan et al., 2008).

**3.3. Mendelian inheritance and presence of null alleles**

Using paternity analysis, the EST-SSRs tested codominantly segregated according to Mendelian expectations (Table S1 provided as Electronic Supplementary Material). Null alleles in SSRs are commonly reported in marine invertebrates (e.g. Yu and Li, 2008; David et al., 1997; Lemer et al., 2011). The use of EST-SSRs is expected to decrease the problem of null alleles as compared to anonymous SSRs, mainly due to functional constraints in transcribed regions of the genome (Li et al., 2002). In agreement with this expectation, we did not observe null alleles when analyzing the progeny-array. This result was however only partly confirmed when using Micro-Checker on the population dataset of 88 individuals. Four loci out of twelve indeed showed evidence for null alleles (BI13YE17, DA8YN14, DA1YN22 and DA1YL23). The low frequency of null alleles (<5%) in these loci may have prevented their detection with the analysis of a single progeny-array. We only studied the segregation of 32 parental alleles (across 11 loci in one family) and we cannot rule out that some of the PCR amplifications failed at these four loci with routine genotyping of individuals at population level. Eight loci did not show any null alleles in our dataset and should be given priority for future population genetic analyses.

**3.4. Polymorphism evaluation of the 12 EST-SSRs**

The analysis of 88 adults of *C. fornicata* showed the absence of any significant linkage disequilibrium between loci; moreover all 12 loci were found to be polymorphic (2–40 alleles, Table 2). Estimates for He

**Table 2**

Characteristics of 12 EST-SSRs newly developed for *Crepidula fornicata* with loci sorted according to the composition and length of the repeated region.

Locus name	ABI Dyes (multiplex subset)	Motif repeats	Primer sequences (5'-3')	T <sub>A</sub> (°C)	[primer] (μM)	Size range (base pairs)	A	He	F <sub>IS</sub> (P-value)
DA1YL23	PET® (3)	(GT) <sup>19</sup> AT(GT) <sup>7</sup>	F: TCTCACCTAACTCTCT TAAAAACG R: GCACACAGACGCTCATTGTCTG	53	0.2	185–278	38	0.953	0.263 (P<10 <sup>-4</sup> )
DA1YN22	6FAM™ (2)	(TA) <sup>9</sup>	F: ATAAGCATGGCTTGACGCTC R: GAAAGAACAAGATGACCAAAGCAG	50	0.27	232–291	13	0.635	0.426 (P<10 <sup>-4</sup> )
CL83	VIC® (3)	(ATC) <sup>6</sup>	F: TGGCCATGATGGTGGCGGTT R: CAGCAAGAGCACACAAGGGGTT	53	0.06	225–263	8	0.595	-0.490 (P<10 <sup>-4</sup> )
BI13YE17	PET® (1)	(CAA) <sup>5</sup> TAA(CAA) <sup>2</sup> TAA(CAA) <sup>3</sup>	F: ACGTGACTGCCTGTGTGGCG R: TCTAGCCCTCGCCCTAAAGAGGGAG	53	0.17	206–254	31	0.940	0.238 (P<10 <sup>-4</sup> )
DA8YN14	VIC® (1)	(CCA) <sup>10</sup>	F: CGATGTGTCATACTGCTGA R: GCTAATAAACAGGGATTGTAAG	53	0.27	117–252	40	0.951	0.134 (P<10 <sup>-4</sup> )
CL322	6FAM™ (1)	(ATTG) <sup>4</sup>	F: GTGGTGAGATACTGCTTGAAC R: TGCTGTGCTACTTCTGGAC	53	0.2	235–247	9	0.443	0.054 (P=0.117)
DA1YA23	6FAM™ (3)	(CAAA) <sup>4</sup>	F: AAAGTAACAGCAAATGTCAA R: TTCGTCGGCACAATCTGG	53	0.33	137–247	6	0.398	-0.152 (P=0.396)
DA5YM24	6FAM™ (1)	(CTTC) <sup>4</sup>	F: ATCAGAGGAACAACAACACC R: GAAAAGACGGACGGAGGAAGA	53	0.13	103–129	10	0.537	0.089 (P=0.140)
CL191	NED™ (1)	(GTTGG) <sup>3</sup>	F: AAGAAAGGGCCGATTGAGT R: ATGTGTCATCCCCATCTGCT	53	0.13	176–183	2	0.503	-0.040 (P=0.831)
CL270	PET® (2)	(TTGAA) <sup>10</sup> TTA(TTGAA) <sup>2</sup> TTCAA(TTGAA) <sup>3</sup>	F: GCCACTACAACAGTAGTCCACGC R: ACATGGCCCCATCTCCAGGCA	50	0.05	197–297	21	0.812	0.006 (P<10 <sup>-4</sup> )
CL291	NED™ (3)	(TAAAA) <sup>3</sup>	F: ATGTGTTGATGTTGCCATAA R: TAGGTCCAAACTGCTATTCA	53	0.17	214–225	2	0.011	-0.003 (P=1)
DA4Y003	NED™ (2)	(ATTAC) <sup>3</sup>	F: ATTTCAACTCTTTCCCAAC R: GAACTAACCAATGCTCTTTC	50	0.23	263–273	6	0.676	0.191 (P=0.007)

T<sub>A</sub>: locus-specific annealing temperature, A: number of alleles, H<sub>e</sub>: expected heterozygosity, F<sub>IS</sub>: fixation index and P-value from exact test for Hardy–Weinberg equilibrium.

varied between 0.011 (CL291) and 0.953 (DA1YL23). Significant deviations from HWE were shown for 7 loci with a large variation in  $F_{IS}$  values from  $-0.490$  to  $0.426$ . The multilocus  $F_{IS}$  value was lower for EST-SSRs than for anonymous SSRs (0.093 and 0.182 respectively). Departure from Hardy–Weinberg equilibrium (HWE) is common for marine invertebrates (Addison and Hart, 2005; David et al., 1997; Hare et al., 1996; Hedgecock et al., 2004; Raymond et al., 1997) and part of departure from HWE is likely due to null alleles at four loci as highlighted above. However, other biological reasons may explain this HWE departure in *Crepidula fornicata*. In marine invertebrates, a large genetic variance in reproductive success, neglected for a long time and now considered as a major driver in marine species evolution, can lead to spatial and temporal genetic patchiness and explain some of the HWE departure through Wahlund effects (Cassista and Hart, 2007; Pudovkin et al., 1996). For instance, the 88 individuals used in our study were not from a single generation. The species we used lives up to 10 years; there is no relevant growth model that can be used to delineate individual generations for individuals more than 3 years-old. This “long-lived” property can amplify genetic variance and temporal Wahlund effects. The 12 loci did not display the same pattern (excess or deficit of heterozygotes or being at HWE), which agrees with expectations for a Wahlund effect as rather than for inbreeding, for which all loci are affected in the same way (Hedrick, 2000). However, the large deviations observed cannot be explained by Wahlund effect only and support the occurrence of PCR problems at some loci. Sources of heterozygote reduction deserve further study, i.e. investigation of additional families and populations, to check for alternative hypotheses explaining heterozygote deficiencies and/or the design of new primers (Lemer et al., 2011).

Estimations of population parameters with EST-SSRs are often congruent with those based on anonymous SSRs (Woodhead et al., 2005; Kim et al., 2008). However, because EST-SSRs are obtained from the expressed portion of the genome, they may be prone to functional constraints. As such, they are often considered as good candidate loci for marker–trait association or genome scan studies (Coyer et al., 2011). When using EST-SSRs in population genetics, neutrality tests should be used before studying neutral processes (e.g. migration). Balancing selection could, for example, explain the excess of heterozygotes observed at CL83. In the absence of sequence data, the most powerful tests are  $F_{ST}$ -based (Zhai et al., 2009; Narum and Hess, 2011) but such tests cannot be used on the present dataset, as all individuals came only from one population. When the 12 EST-SSRs were tested for selection using Ewens–Watterson homozygosity test (Ewens, 1972; Watterson, 1978), only one locus (CL191) showed some evidence of selection ( $p = 0.02$ ). In the study progeny-array, the segregation pattern observed at this locus, however, fitted with expectations (Table S1). Local reduction in genetic variation can be expected in EST-SSRs under positive selection (Scott, 2001; Oleksyk et al., 2010). When comparing results obtained with the 7 anonymous SSRs previously developed (Table 3) and the EST-SSRs (Table 2),

slightly lower mean values of  $A$  ( $<40$  for 12 EST-SSRs and  $<49$  for 7 anonymous SSRs) and  $H_e$  (0.620 and 0.760 for 12 EST-SSRs and 7 anonymous SSRs, respectively) were observed. However, the high exclusion probability computed across the 12 EST-SSRs ( $p = 0.992$ ) was similar to the value obtained across the 7 anonymous SSRs ( $p = 0.995$ ). This comparison suggests that there are no major discrepancies between the genetic variations observed with the two categories of markers. In addition, the genetic polymorphism of EST-SSRs is large enough to perform paternity analyses, which require high allelic diversity for statistical power.

#### 4. Conclusions

The 12 newly isolated EST-SSRs are specific to *Crepidula fornicata*, are Mendelian-inherited, and display an appropriate level of polymorphism for a large range of population studies. Combined or not with anonymous SSRs, they offer a set of reliable microsatellites that should enable further molecular ecology studies in *Crepidula fornicata*, for instance for genome scan studies or to better assess the degree of relatedness between newly metamorphosed recruits in the wild.

Supplementary materials related to this article can be found online at doi:10.1016/j.margen.2011.08.001.

#### Acknowledgements

The authors are grateful to the divers and mariners of the “Service Mer et Observation” [Marine Operations Department] (FR2424) at the Station Biologique de Roscoff (SBR) and to Sabrina Le Cam, Jan Pechenik, Rachel Collin and Antonio Brante for providing specimens of *Crepidula* species. We thank the Bioinformatic Department of SBR for help in bioinformatics analyses and Jan Pechenik for help in editing the manuscript. The genotyping was carried out on the BiogenOuest platform at SBR. Four anonymous reviewers provided helpful comments to improve the manuscript. This work was funded by the Agence Nationale de la Recherche (*in silico* mining of ESTs: Project Mirage; ANR-05-BLAN-0001; SSRs optimization: Project Hi-Flo ANR-08-BLAN-0334). FR acknowledges a PhD grant from the Centre National pour la Recherche Scientifique and the Conseil Régional de Bretagne (SEA-FLO ARED project).

F.R. is a PhD student in F.V.'s lab working on dispersal and adaptation in *Crepidula fornicata*.

M.B. is an engineer specialized in development of molecular and genomic tools.

A.T. is a UPMC lecturer at the Station Biologique de Roscoff. His interests include genomic and transcriptomic studies to study adaptation in marine species from extreme environments.

F.V. is a permanent CNRS researcher at the Station Biologique de Roscoff. Her interests include molecular ecology and evolutionary biology to examine historical and contemporary dispersal and

**Table 3**

Characteristics and polymorphism of 7 anonymous SSRs previously developed for *Crepidula fornicata* (see Reference below). In previous studies, the 7 anonymous SSRs were genotyped by electrophoresis using polyacrylamide sequencing gels (LI-COR®Biosciences). In the present study, the amplification products were run in an ABI prism 3130XL DNA sequencer (Applied Biosystems™). Each locus forward primer was labelled with the ABI dyes indicated. A: number of alleles,  $H_e$ : expected heterozygosity,  $F_{IS}$ : fixation index and  $p$ -value from exact test for Hardy–Weinberg equilibrium.

Locus	Reference	ABI dyes	Size range (pb)	$A$	$H_e$	$F_{IS}$	$P_{HWE}$
CfCA4	Dupont and Viard (2003)	6FAM™	147–174	18	0.806	0.008	$P = 0.306$
CfCA2	Dupont and Viard (2003)	PET™	217–222	4	0.526	0.006	$P = 0.873$
CfCATGT	Dupont and Viard (2003)	PET™	209–291	27	0.823	0.241	$P < 10^{-4}$
CfGT9	Dupont and Viard (2003)	VIC®	219–301	49	0.975	0.440	$P < 10^{-4}$
Cf8	Proestou (2006)	VIC®	115–264	37	0.943	0.151	$P < 10^{-4}$
H7	Dupont et al. (2006)	NED™	102–153	14	0.885	0.029	$P = 0.041$
E4	Viard, Kruse and Lacut (unpublished) <sup>a</sup>	NED™	312–353	8	0.335	0.483	$P < 10^{-4}$

<sup>a</sup> E4 is a locus which has been obtained from a second screening of the genomic library developed by Dupont and Viard (2003). Briefly, the motif repeat is (GT) 10 (GAGT) 3, the forward and reverse primers are 5'-AAAGAGGAAGGAAAGGTTGTATG-3' and 5'-TCACCAGCAAATCCAGAACG-3' respectively (primer concentration for PCR of 0.75  $\mu$ M) and the annealing temperature is 62 °C.

adaptation processes of marine coastal species, in particular invasive species.

## References

- Addison, J.A., Hart, M.W., 2005. Spawning, copulation and inbreeding coefficients in marine invertebrates. *Biol. Lett.* 1 (4), 450–453.
- Allendorf, F.W., Hohenlohe, P.A., Luikart, G., 2010. Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11 (10), 697–709.
- Blanchard, M., 1997. Spread of the slipper limpet *Crepidula fornicata* (L. 1758) in Europe. Current state and consequences. *Sci. Mar.* 61 (sup. 2), 109–118.
- Broquet, T., Petit, E.J., 2009. Molecular estimation of dispersal for ecology and population genetics. *Annu. Rev. Ecol. Evol.* 40, 193–216.
- Cassista, M.C., Hart, M.W., 2007. Spatial and temporal genetic homogeneity in the Arctic surfclam (*Mactromeris polynyma*). *Mar. Biol.* 152 (3), 569–579.
- Collin, R., 2003. Phylogenetic relationships among calyptraeid gastropods and their implications for the biogeography of marine speciation. *Syst. Biol.* 52 (5), 618–640.
- Coyer, J.A., Hoarau, G., Pearson, G., Mota, C., Jüterbock, A., Alpermann, T., John, U., Olsen, J.L., 2011. Genomic scans detect signatures of selection along a salinity gradient in populations of the intertidal seaweed *Fucus serratus* on a 12 km scale. *Mar. Genom.* 4 (1), 41–49.
- Daguin-Thiebaud, C., Le Cam, S., Viard, F., 2009. Isolation of 11 microsatellite markers in *Crepidula convexa* (Gastropoda, Calyptraeidae) for parentage analyses. *Mol. Ecol. Res.* 9 (3), 917–920.
- Dallas, J.F., 1992. Estimation of microsatellite mutation-rates in recombinant inbred strains of mouse. *Mamm. Genome* 3 (8), 452–456.
- David, P., Perdieu, M.A., Permot, A.F., Jarne, P., 1997. Fine-grained spatial and temporal population genetic structure in the marine bivalve *Spisula ovalis*. *Evolution* 51 (4), 1318–1322.
- Dupont, L., Bernas, D., Viard, F., 2007a. Sex and genetic structure across age groups in populations of the European marine invasive mollusc, *Crepidula fornicata* L. (Gastropoda). *Biol. J. Linn. Soc.* 90 (2), 365–374.
- Dupont, L., Ellien, C., Viard, F., 2007b. Limits to gene flow in the slipper limpet *Crepidula fornicata* as revealed by microsatellite data and a larval dispersal model. *Mar. Ecol. Prog. Ser.* 349, 125–138.
- Dupont, L., Richard, J., Paulet, Y.M., Thouzeau, G., Viard, F., 2006. Gregariousness and protandry promote reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of larval paternity. *Mol. Ecol.* 15 (10), 3009–3021.
- Dupont, L., Viard, F., 2003. Isolation and characterization of highly polymorphic microsatellite markers from the marine invasive species *Crepidula fornicata* (Gastropoda :Calyptraeidae). *Mol. Ecol. Notes* 3 (4), 498–500.
- Ellis, J.R., Burke, J.M., 2007. EST-SSRs as a resource for population genetic analyses. *Heredity* 99 (2), 125–132.
- Ewens, W.J., 1972. Sampling theory of selectively neutral alleles. *Theor. Pop. Biol.* 3 (1), 87–112.
- Goldstein, D.B., Linares, A.R., Cavalli Sforza, L.L., Feldman, M.W., 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. U. S. A.* 92 (15), 6723–6727.
- Hare, M.P., Karl, S.A., Avise, J.C., 1996. Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. *Mol. Biol. Evol.* 13 (2), 334–345.
- Hedgecock, D., Li, G., Hubert, S., Bucklin, K., Ribes, V., 2004. Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. *J. Shellfish. Res.* 23 (2), 379–385.
- Hedrick, P.W., 2000. *Genetics of Populations*, 2nd ed. Jones and Bartlett Publishers, Inc, Boston.
- Henry, J.J., Collin, R., Perry, K.J., 2010. The slipper snail, *Crepidula*: an emerging lophotrochozoan model system. *Biol. Bull.* 218 (3), 211–229.
- Jones, A.G., 2005. GERUD 2.0: a computer program for the reconstruction of parental genotypes from half-sib progeny arrays with known or unknown parents. *Mol. Ecol. Notes* 5 (3), 708–711.
- Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23 (10), 1289–1291.
- Kim, K.S., Ratcliffe, S.T., French, B.W., Liu, L., Sappington, T.W., 2008. Utility of EST-Derived SSRs as population genetics markers in a beetle. *J. Hered.* 99 (2), 112–124.
- Lemer, S., Rochel, E., Planes, S., 2011. Correction method for null alleles in species with variable microsatellite flanking regions, a case study of the black-lipped pearl oyster *Pinctada margaritifera*. *J. Hered.* 102 (2), 243–246.
- Li, Y.C., Korol, A.B., Fahima, T., Beiles, A., Nevo, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.* 11 (12), 2453–2465.
- Le Cam, S., Viard, F., Cagnon, M., Pechenik, J.A., 2009. Role of multiple paternity in causing variation in larval growth rates in the gastropod *Crepidula fornicata*. *Integr. Comp. Biol.* 49, E98–E.
- McGlashan, D.J., Ponniah, M., Cassey, P., Viard, F., 2008. Clarifying marine invasions with molecular markers: an illustration based on mtDNA from mistaken calyptraeid gastropod identifications. *Biol. Inv.* 10 (1), 51–57.
- McInerney, C.E., Allcock, A.L., Johnson, M.P., Bailie, D.A., Prodohl, P.A., 2011. Comparative genomic analysis reveals species-dependent complexities that explain difficulties with microsatellite marker development in molluscs. *Heredity* 106 (1), 78–87.
- Narum, S.R., Hess, J.E., 2011. Comparison of FST outlier tests for SNP loci under selection. *Mol. Ecol. Res.* 11, 184–194.
- Oleksyk, T.K., Smith, M.W., O'Brien, S.J., 2010. Genome-wide scans for footprints of natural selection. *Philos. T. R. Soc. B* 365 (1537), 185–205.
- Palsboll, P.J., Berube, M., Allendorf, F.W., 2007. Identification of management units using population genetic data. *Trends Ecol. Evol.* 22 (1), 11–16.
- Proestou, D.A., 2006. Isolation and characterization of microsatellite markers in the Atlantic slipper shell *Crepidula fornicata* for use in paternity analysis. *Mol. Ecol. Notes* 6 (2), 437–439.
- Pudovkin, A.I., Zaykin, D.V., Hedgecock, D., 1996. On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics* 144 (1), 383–387.
- Rigal, F., Viard, F., Ayata, S.D., Comtet, T., 2010. Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? *Biol. Inv.* 12 (9), 3171–3186.
- Raymond, M., Vaanto, R.L., Thomas, F., Rousset, F., deMeeus, T., Renaud, F., 1997. Heterozygote deficiency in the mussel *Mytilus edulis* species complex revisited. *Mar. Ecol. Prog. Ser.* 156, 225–237.
- Robinson, A.J., Love, C.G., Batley, J., Barker, G., Edwards, D., 2004. Simple sequence repeat marker loci discovery using SSR primer. *Bioinformatics* 20 (9), 1475–1476.
- Roman, J., Darling, J.A., 2007. Paradox lost: genetic diversity and the success of aquatic invasions. *Trends Ecol. Evol.* 22 (9), 454–464.
- Rousset, F., 2008. GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Resour.* 8 (1), 103–106.
- Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386.
- Selkoe, K.A., Henzler, C.M., Gaines, S.D., 2008. Seascape genetics and the spatial ecology of marine populations. *Fish. Fish.* 9 (4), 363–377.
- Scott, K.D., 2001. Microsatellites derived from ESTs, and their comparison with those derived from other methods. In: Henry, R.J. (Ed.), *Plant Genotyping—The DNA Fingerprint of Plants*. CABI Publishing, CABI International, Wallingford, Oxon, UK, pp. 225–237.
- Taris, N., Comtet, T., Stolba, R., Lasbleiz, R., Pechenik, J.A., Viard, F., 2010. Experimental induction of larval metamorphosis by a naturally-produced halogenated compound (dibromomethane) in the invasive mollusk *Crepidula fornicata* (L.). *J. Exp. Mar. Biol. Ecol.* 393 (1–2), 71–77.
- Taris, N., Comtet, T., Viard, F., 2009. Inhibitory function of nitric oxide on the onset of metamorphosis in competent larvae of *Crepidula fornicata*: a transcriptional perspective. *Mar. Genom.* 2 (3–4), 161–167.
- Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4 (3), 535–538.
- Viard, F., Ellien, C., Dupont, L., 2006. Dispersal ability and invasion success of *Crepidula fornicata* in a single gulf: insights from genetic markers and larval-dispersal model. *Helgol. Mar. Res.* 60 (2), 144–152.
- Vogiatzi, E., Lagnel, J., Pakaki, V., Louro, B.V.M., Canario, A., Reinhardt, R., Kotoulas, G., Magoulas, A., Tsigonopoulos, C.S., 2011. In silico mining and characterization of simple sequence repeats from gilthead sea bream (*Sparus aurata*) expressed sequence tags (EST-SSRs): PCR amplification, polymorphism evaluation and multiplexing and cross-species assays. *Mar. Genom.* 4 (2), 83–91.
- Wang, J., Santure, A.W., 2009. Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics* 181 (4), 1579–1594.
- Wang, J.L., 2004. Sibship reconstruction from genetic data with typing errors. *Genetics* 166 (4), 1963–1979.
- Waples, R.S., 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* 89 (5), 438–450.
- Watterson, G.A., 1978. Homozygosity test of neutrality. *Genetics* 88 (2), 405–417.
- Woodhead, M., Russell, J., Squirrell, J., Hollingsworth, P.M., Mackenzie, K., Gibby, M., Powell, W., 2005. Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Mol. Ecol.* 14, 1681–1695.
- Yu, H., Li, Q., 2008. Exploiting EST Databases for the Development and Characterization of EST-SSRs in the Pacific Oyster (*Crassostrea gigas*). *J. Hered.* 99 (2), 208–214.
- Zhai, W., Nielsen, R., Slatkin, M., 2009. An investigation of the statistical power of neutrality tests based on comparative and population genetic data. *Mol. Biol. Evol.* 26 (2), 273–283.



#### **IV. Développement de marqueurs AFLPs**

L'objectif de l'étude présentée dans l'article 1 était de développer de nouveaux marqueurs microsatellites, d'une part pour une meilleure estimation de la variabilité génétique entre et au sein des populations de *C. fornicata* et d'autre part pour augmenter (et diversifier) le nombre (et le type) de marqueurs utilisables dans le cadre d'un scan génomique. Le nombre de marqueurs microsatellites disponibles (19 au total dont quelques-uns suspectés d'allèles nuls) est trop limité pour identifier des marqueurs sous sélection par une approche de scan génomique. Aussi en parallèle, des protocoles de typage AFLP ont été développés. Ils sont détaillés en annexe 2.



## Chapitre 2 - Le rôle de la sélection dans les processus d'invasion chez *Crepidula fornicata*

---

### I. Contexte de l'étude

Des changements évolutifs sous l'influence de la sélection naturelle sont attendus suite à l'introduction d'une espèce en dehors de son aire d'origine. L'hypothèse de changements adaptatifs rapides dans l'aire d'introduction est en effet proposée pour expliquer le succès des espèces introduites dans un environnement dans lequel elles n'avaient jusqu'alors pas évolué. Bien que cette hypothèse soit largement reprise dans la littérature, pour le moment, peu d'études *in situ* se sont attachées à étudier le rôle de la sélection lors d'introduction biologique ainsi qu'à identifier les variations adaptatives sous-jacentes (Prentis *et al.* 2010). Les espèces invasives en milieu marin pourraient être dans ce contexte de bons modèles d'étude car elles présentent souvent une forte diversité génétique dans les populations introduites (Roman et Darling 2007) sur laquelle la sélection pourrait agir.

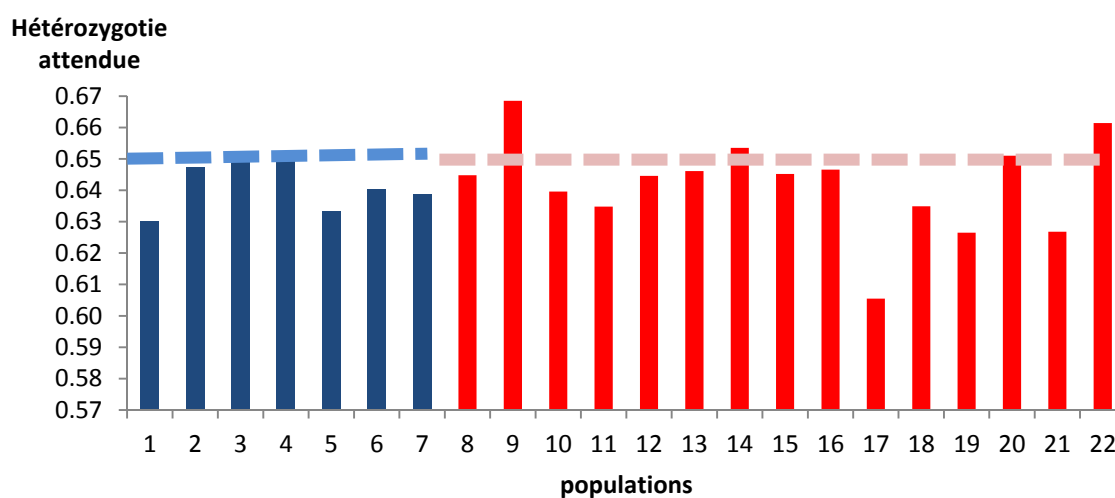
Originaire de la côte est de l'Amérique du Nord et introduit sur la côte Ouest de Etats-Unis et en Europe à plusieurs reprises à la fin du 19<sup>ème</sup> siècle, le gastéropode *Crepidula fornicata* en est un exemple typique. Des études antérieures ont montré une forte diversité génétique chez cette espèce, que ce soit dans l'aire d'origine ou dans l'aire d'introduction (Hoagland 1985, Dupont *et al.* 2003). Ces patrons avaient été interprétés comme résultant d'une forte pression de propagules (multiples introductions et/ou multiples sources d'introduction). Cependant, l'échantillonnage réalisé dans ces études ne couvrait pas dans une même étude l'ensemble de l'aire d'origine ni celles d'introduction. Par ailleurs, ces études avaient été réalisées à l'aide d'un petit nombre de marqueurs allozymiques ne permettant pas des analyses fines des patrons de diversité génétique.

Dans cette étude, nous avons recherché si des processus de sélection pouvaient être mis en évidence en relation avec le processus d'introduction de *C. fornicata* en Europe et sur la côte ouest des Etats-Unis. Nous avons choisi d'aborder cette question par une approche de scan génomique ne nécessitant pas de connaissances préalables du génome de l'espèce étudiée ou des traits pouvant être soumis à sélection.

A l'aide de 17 marqueurs microsatellites et 327 marqueurs AFLPs, 683 individus ont été génotypés. Ces individus ont été échantillonnés dans sept populations de l'aire d'origine (du Massachussets à la Floride) et quinze populations des aires d'introduction (14 populations européennes, et une population de la côte ouest des Etats-Unis) afin de 1) confirmer le scénario d'introduction suggéré par les études antérieures (forte pression de propagules), 2) rechercher des preuves de mécanismes de sélection post-introduction et 3) comparer les résultats des scans génomiques réalisés au sein de l'aire d'origine et entre aire d'origine et aires d'introduction.

## II. Résultats principaux de l'étude (détaillés dans l'article 2)

Dans un premier temps, nos résultats ont montré une forte diversité génétique, quelque soit les marqueurs utilisés et les populations analysées (natives ou introduites) (Fig. 1). Ces résultats sont en accord avec ceux précédemment documentés, sur la base d'étude de polymorphisme enzymatique. De plus, une faible structure génétique a été observée entre aires d'origine et d'introduction, confirmant un schéma d'introduction caractérisé par une forte pression en propagules et de multiples introductions de différentes sources, un scénario souvent observé chez les espèces marines invasives.

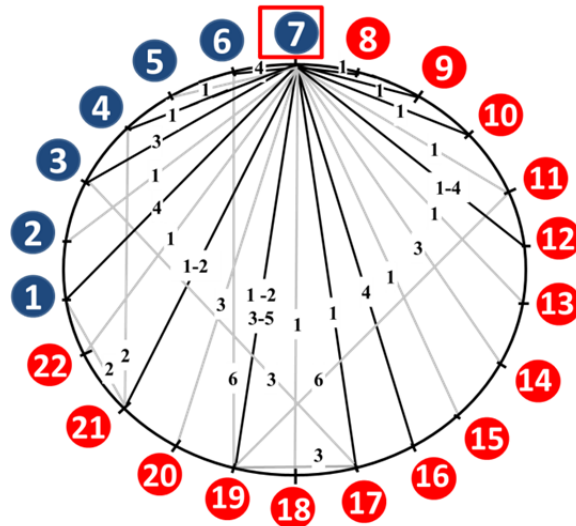


**Fig. 1** Hétérozygotie attendue estimée à partir des 17 marqueurs microsatellites, pour chaque population échantillonnée au sein de l'aire d'origine (en bleu) et au sein des aires introduites (en rouge). Les populations échantillonnées au sein de l'aire d'origine sont numérotées de 1 à 7, et du Nord (Massachusetts) au Sud (Floride). La population introduite échantillonnée sur la côte ouest des Etats-Unis est numérotée 8, et les populations européennes de 9 à 22, du Nord (Suède) au Sud (France, Méditerranée). L'hétérozygotie attendue moyenne des populations natives est indiquée par la ligne en pointillé bleu, celle des populations introduites par la ligne en pointillé rouge.

Par ailleurs, la population la plus méridionale de l'échantillonnage dans l'aire d'origine, *i.e.* celle située sur la côte ouest de la Péninsule de Floride, présente une structure génétique très différente de toutes les autres populations, quels que soient les marqueurs utilisés. Les analyses de structure génétique permettent de l'exclure comme source potentielle d'introduction. En outre, les marqueurs microsatellites ont mis en évidence un patron de structure génétique en Europe qui pourrait être expliqué par deux événements d'introduction indépendants, l'un étant lié à une (plusieurs) source(s) située(s) dans l'aire native, l'autre, plus récent, étant lié à une (plusieurs) source(s) située(s) sur la côte ouest des Etats-Unis, où la crépidule était déjà introduite (« population relai »).

Malgré le nombre important de marqueurs utilisés (344 marqueurs) et contrairement à nos attendus, aucun mécanisme de sélection n'a été mis en évidence suite à l'introduction de *C. fornicata* en Europe et sur la côte ouest des Etats-Unis.

En revanche, huit outliers (2.3%) ont été identifiés au sein de l'aire d'origine, entre la population située dans le Golfe de la Floride et les autres populations (Fig. 2), ces locus caractérisant une barrière génétique située le long des côtes de la Péninsule de Floride.



**Fig. 2** Identification des outliers identifiés dans notre étude (numérotés de 1 à 6) entre chaque paire de population (22 populations étudiées, 231 comparaisons menées). Les outliers identifiés par deux logiciels sont figurés par une ligne noire, ceux uniquement par un des deux logiciels en gris. Les populations sont numérotées de 1 à 22 avec en bleu, les populations de l'aire d'origine (du Nord au Sud), et en rouge les populations introduites (8 : population échantillonnée sur la côte ouest des Etats-Unis, puis du Nord au Sud).

Cette singularité génétique localisée le long des côtes de la Floride, coïncide avec la barrière biogéographique bien connue le long de cette péninsule (Avisé 2000). Ce résultat suggère une évolution sur le long terme d'adaptations locales dans l'aire d'origine de la crépidule.

Aucun mécanisme de sélection ayant agi suite à l'introduction de *C. fornicata* en Europe et sur la côte ouest des Etats-Unis n'a été mis en évidence, suggérant d'une part, que les différentiels de sélection, s'ils existent, sont trop faibles pour laisser une trace détectable par des scans génomiques, et d'autre part, que la réussite de *C. fornicata* a également pu dépendre d'autres processus, tels que la plasticité phénotypique. La méthode par scan génomique apparaît robuste quand il s'agit d'analyser des processus de sélection sur le long terme puisque huit locus « outliers » ont été identifiés au sein de l'aire native, mais ne semble pas appropriée pour observer des événements de sélection récents, comme dans le cas d'introductions biologiques. L'approche par scan génomique est efficace pour révéler une évolution sur le long terme d'adaptations locales, reflétant l'accumulation d'incompatibilités génétiques, mais semble moins efficace pour identifier de récentes adaptations, suite à de la sélection sur de la variation préexistante. D'autres études d'espèces introduites examinant le rôle de la sélection sont nécessaires pour mettre en perspective nos résultats.

Cette étude a été présentée sous forme de poster à la conférence Jacques Monod « Développements théoriques et empiriques en génomique évolutive » (Avril 2012) ; le poster est présenté en annexe 3.

Des résultats complémentaires, non inclus dans l'article 2, sont présentés dans l'annexe 4 (analyses d'assignation menées avec le logiciel Structure pour 3 groupes identifiés), l'annexe 5 (détails de l'identification des outliers pour les 22 populations) et l'annexe 6 (détails de l'identification des outliers pour les 231 comparaisons des populations deux à deux).

### **III. Article 2: « Contrasting patterns of genome-wide polymorphism in the native and invaded distribution ranges of the marine invasive mollusk *Crepidula fornicata* »**

« Comparaison des patrons de polymorphismes génomiques entre aire d'origine et aires d'introduction chez le mollusque marin invasif *Crepidula fornicata* »

Florentine Riquet, Claire Daguin-Thiébaud, Marion Ballenghien, Nicolas Bierne et Frédérique Viard - soumis à *Molecular Ecology*

1 **Contrasting patterns of genome-wide polymorphism in the native and invaded**  
2 **distribution ranges of the marine invasive mollusk *Crepidula fornicata***

3 Florentine Riquet<sup>1, 2</sup> – Claire Daguin-Thiébaud<sup>1, 2</sup> – Marion Ballenghien<sup>1, 2, 3, 4</sup> – Nicolas Bierne<sup>3, 5</sup>  
4 Frédérique Viard<sup>1, 2</sup>

5 <sup>1</sup> UPMC Univ Paris 06, UMR 7144, Team Diversity and Connectivity in Coastal Marine Landscapes,  
6 Station Biologique de Roscoff, 29680 Roscoff, France

7 <sup>2</sup> CNRS, UMR 7144, Lab. Adaptation and Diversity in the Marine Environment, Station Biologique de  
8 Roscoff, 29680 Roscoff, France

9 <sup>3</sup> Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier, France

10 <sup>4</sup> CNRS, UMR 5554, Institut des Sciences de l'Evolution, Lab. Phylogénie et Evolution moléculaire,  
11 Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France

12 <sup>5</sup> CNRS, Institut des Sciences de l'Evolution, UMR 5554, Station Méditerranéenne de  
13 l'Environnement Littoral, 2 rue des Chantiers, 34200 Sète, France

14

15 Correspondence: F. Riquet: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de  
16 Roscoff, 29680 Roscoff, France; Tel: +33298292544, Fax: +33298292324; e-mail: [riquet@sb-roscoff.fr](mailto:riquet@sb-roscoff.fr)

17 F. Viard: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680  
18 Roscoff, France; Tel: +33298292312, Fax: +33298292324; e-mail: [viard@sb-roscoff.fr](mailto:viard@sb-roscoff.fr)

19 Running title: Selection and invasion in *C. fornicata*

20 Keywords: *Crepidula fornicata*, gastropod, biological invasion, adaptation, AFLP, microsatellites

21

22 Abstract: Selection processes are believed to be an important evolutionary driver behind the successful  
23 establishment of non-indigenous species. However, evidence supporting this assumption is still scarce.  
24 Using microsatellite- and AFLP-based genome scans, we looked for evidence of selection following  
25 the introduction of the mollusk *Crepidula fornicata*. Native to the northwestern Atlantic, this  
26 gastropod has become an emblematic invader since its introduction during the 19<sup>th</sup> and 20<sup>th</sup> centuries  
27 in the northeastern Atlantic and northeastern Pacific. We examined 683 individuals from 7 native and  
28 15 introduced populations spanning latitudinal gradients over which genome scans have often  
29 identified loci with atypical patterns of genetic differentiation (i.e. outliers) indicative of selection  
30 processes in other species. Our results confirmed the previously documented high genetic diversity in  
31 native and introduced populations with little genetic structure between the two ranges, a pattern typical  
32 of marine invaders. No outliers were detected in the introduced range or between the introduced and  
33 native populations, but eight outliers were identified within the native range. This genetic singularity  
34 was congruent with a well-known biogeographic break in the Florida Peninsula. Our results call into  
35 question the effectiveness of AFLP genome scans in detecting adaptation on the time-scale of  
36 biological invasions. Genome scans often reveal long-term adaptation but are less effective in  
37 detecting recent local adaptation from pre-existing variation on polygenic traits. This study advocates  
38 other methods to detect selection effects –e.g. at the phenotypic level — in the study of biological  
39 invasions, although genome scans may remain useful for elucidating introduction histories.

40



## 41 **Introduction**

42 Natural selection is involved in adaptive responses to a variety of environmental changes driven by  
43 natural or anthropogenic modifications (e.g. climate change, habitat alterations, fisheries, etc.).  
44 However, studying selection is difficult. Distinguishing the effects of selection from those of other  
45 evolutionary forces can be challenging because selection does not frequently leave distinguishable  
46 footprints in the genome (Li *et al.* 2012). For instance, gene surfing (i.e. alleles that reach high  
47 frequencies just by chance during population expansion) has often been interpreted as evidence of  
48 selection events, although it is just a sign of genetic drift (Excoffier & Ray 2008). In addition, even  
49 when local selection maintains phenotypic differences, it does not necessarily produce genetic  
50 differentiation at all of the genes underlying the trait (Le Corre & Kremer 2012), and those that do  
51 show genetic structure are expected to affect neutral variation on a very small chromosomal scale  
52 (Charlesworth *et al.* 1997, Bierne 2010). Finally, adaptation can be locally impeded or even offset by  
53 gene flow (i.e. “gene swamping”, Lenormand 2002), a pattern exemplified in many marine species.  
54 Large population sizes and the long-lived dispersive phases of many marine species mean that  
55 populations are only weakly genetically structured (i.e. connected by high gene flow), opposing local  
56 adaptation (Nielsen *et al.* 2009). For example, selection has not been demonstrated in contrasted  
57 environments in the highly dispersive periwinkle *Echinolittorina hawaiiensis* (Tice & Carlon 2011).  
58 Nevertheless, selection is known to play a key role in maintaining adaptive polymorphism in  
59 dispersive marine species (e.g. Koehn *et al.* 1980).

60 In this context, invasive marine invertebrates make interesting case studies. Non-indigenous species  
61 successfully survive, reproduce and disperse into new environments characterized by different  
62 ecological conditions from those prevailing in the native range. However, introduced phenotypes may  
63 not be the best-adapted to the new habitat, and adaptive changes following introduction are thought to  
64 be an important process behind invasion success (Stockwell *et al.* 2003; Sax *et al.* 2007). Selection of  
65 advantageous traits may be rapid, conferring a competitive advantage to the introduced species over  
66 the native species within a few generations (Reznick & Ghalambor 2001; Prentis *et al.* 2008).

67 Identifying such adaptive changes can help elucidate the reasons for the successful establishment of  
68 non-indigenous species, a central issue in the study of biological invasions (Sax *et al.* 2007). Selection  
69 processes may be particularly effective in marine invasive species, which generally display large  
70 populations and a high level of genetic diversity in introduced populations (e.g. Simon-Bouhet *et al.*  
71 2006; Roman & Darling 2007 for a review). This high genetic diversity may increase the likelihood of  
72 encountering adapted genotypes, creating new genetic combinations and limit the effect of genetic  
73 drift as well as favoring selection on pre-existing genetic variation (Stockwell *et al.* 2003; Facon *et al.*  
74 2006; Barrett & Schuller 2008).

75 We here tested for the occurrence of post-introduction selection processes by looking for genomic  
76 regions shaped by selection in one of the most emblematic invasive mollusks, the slipper limpet  
77 *Crepidula fornicata*. This long-lived species (living up to 10 years) displays a typical marine benthopelagic  
78 life cycle alternating a 2 to 7 week free-floating larval stage and a sessile adult stage. It  
79 inhabits shallow bays and estuaries. Native to the northwestern Atlantic, this marine gastropod was  
80 first introduced in Great Britain at the end of the 19<sup>th</sup> century and beginning of the 20<sup>th</sup> century,  
81 probably with spats of the American oyster *Crassostrea virginica* (Hoagland 1984). The species  
82 expanded during the 20<sup>th</sup> century along the coasts of Europe (France, Belgium, Netherlands and  
83 Germany). In addition, massive accidental introductions occurred in the 1970s along the French  
84 Atlantic coasts and in the Mediterranean Sea with the intentional introduction of the Pacific oyster  
85 *Crassostrea gigas* from the northeastern Pacific (Puget Sound), where the Pacific oyster and the  
86 slipper limpet had previously been introduced (in the 1930s; Blanchard 1997; Hoagland 1984). Today,  
87 *C. fornicata* is well established in the northeastern Pacific and in Europe (mainly from Norway to  
88 Spain). Previous studies of *C. fornicata* have documented high genetic polymorphism in both native  
89 and introduced populations (Hoagland 1984; Dupont *et al.* 2003), concurring with its complex  
90 introduction history. However, the previous studies failed to cover both the full native and introduced  
91 ranges of *C. fornicata* and only a limited number of allozyme markers were used, preventing efficient  
92 detection of selection footprints. One increasingly used approach to detect selection effects is the use

93 of a large number of markers spread throughout the genome (i.e. the genome scan approach; Storz  
94 2005) to identify loci that display levels of population differentiation higher than those under neutral  
95 expectations. These so-called outlier loci are directly under selection or, more likely, linked with a  
96 locus under selection (Bierne *et al.* 2011). Outlier loci have often been identified along latitudinal or  
97 altitudinal gradients characterized by noticeable environmental and climatic variations (e.g. Bonin *et*  
98 *al.* 2006). The large latitudinal distribution of *C. fornicata* in its native and invaded ranges makes this  
99 model species appropriate for the genome scan approach. In non-model species for which genome data  
100 are not available, the AFLP technique is a fast and cost-effective method to provide the required  
101 number of markers over the whole genome (Bonin *et al.* 2006). This strategy has been successfully  
102 applied in native species to analyze adaptive response to an environmental gradient (e.g. Coyer *et al.*  
103 2011; Nunes *et al.* 2011). In this study, we developed a genome-scan approach using 344 polymorphic  
104 loci (17 microsatellite and 327 AFLP markers) to compare the genetic diversity of native populations  
105 with populations sampled in two distinct areas of introduction of *C. fornicata*. We had three  
106 objectives: (1) to confirm the previously suggested scenario of introduction (i.e. large propagule pool  
107 associated with a large genetic diversity); (2) to test for the presence of post-introduction selection  
108 events in European and American introduced populations; (3) to compare the results of a genome scan  
109 conducted in the native range — where long-term evolution interferes with local adaptation — with a  
110 genome scan conducted in the invaded range — where adaptation, if any, can only have evolved on a  
111 short term.

## 112 **Materials and Methods**

### 113 *Sample and DNA collection*

114 Twenty-two populations of *Crepidula fornicata* were sampled in 2001-2002, except for the one from  
115 Fort Pierce (FL, USA), which was sampled in April 2010. Seven populations were sampled in the  
116 natural range of the species (along the eastern coast of North America), and 15 in introduced ranges:  
117 one from the Pacific coast of the USA and 14 from Europe (North Sea, Channel, Atlantic coast and

118 Mediterranean Sea) (Fig. 1, Table 1). An average of 32 individuals per population was examined.  
119 Individuals were either frozen at -80°C or preserved in 96% ethanol for genetic analyses. Genomic  
120 DNA was extracted from adults using either the phenol-chloroform method or using the DNeasy  
121 Tissue Kit (QIAGEN) following the manufacturer's protocol.

## 122 *Microsatellite and AFLP genotyping*

123 **Microsatellite genotyping.** Individuals were genotyped at 17 microsatellite loci (12 EST-SSRs:  
124 BI13YE17, CL191, CL322, DA5YM24, DA8YN14, CL270, DA1YN22, DA4YO03, CL291, CL83,  
125 DA1YL23, DA1YA23 as defined by Riquet *et al.* (2011), and 5 anonymous markers: CfH7 (Dupont *et*  
126 *al.* 2006), Cf8 (Proestou 2006), CfCA4, CfCA2, (Dupont & Viard 2003) and E4 (details in Riquet *et*  
127 *al.* 2011). Loci were amplified by polymerase chain reactions (PCR) following protocols detailed in  
128 Riquet *et al.* (2011). Amplification products were separated by electrophoresis on an ABI 3130xl  
129 DNA sequencer (Applied Biosystems™) and alleles were scored using GeneMapper® v. 4.0 (Applied  
130 Biosystems™). A part of the whole dataset (11.2%) was genotyped twice with the same results.

131 **AFLP genotyping.** The AFLP protocol from Vos *et al.* (1995) was adapted to *Crepidula fornicata*  
132 following Bonin *et al.* (2005) with minor modifications. Briefly, genomic DNA (between 50 and  
133 150 ng) was digested 3 h at 37°C with restriction endonucleases *MseI* and *EcoRI* (5 U, New England  
134 Biolabs, NEB). *EcoRI* and *MseI* double-stranded adapters were ligated to restriction fragments with  
135 T4 DNA ligase (5 U, NEB) overnight at 16°C. Pre-selective amplification reactions were carried out  
136 with two primer pairs (*EcoRI*+A/*MseI*+A and *EcoRI*+C/*MseI*+C). For the selective amplification, four  
137 primer pairs (Table 2) were selected out of 62 tested, ranging from 50 to 600 bp. They yielded  
138 polymorphic peaks with high signal above noise, thus maximizing scoring peaks with reliability. The  
139 four selective PCR products were pooled and 2 µL were electrophoresed on an ABI 3130xl DNA  
140 sequencer (Applied Biosystems™) with 0.3 µL of GeneScan 600LIZ Size Standard (Applied  
141 Biosystems™) and 9.7 µL of deionized formamide. Negative controls at each step of the AFLP  
142 protocol were included to detect any contamination or artefactual peaks. More than 12% of the

143 samples were randomly chosen and genotyped twice for each primer combination to test for the  
144 repeatability of the AFLP profiles, as recommended by Bonin *et al.* (2005) and as required for error  
145 rate estimation in AFLPscore (Whitlock *et al.* 2008; see below). AFLP profiles were carefully  
146 examined with GeneMapper v.4.0 (Applied Biosystems™). Peak heights (in relative fluorescence  
147 units, RFU) for all individuals for each primer combination were exported into the R script AFLPscore  
148 v. 1.4 (Whitlock *et al.* 2008). Replicate samples were used to optimize the AFLPscore threshold (i.e.  
149 threshold that minimizes genotyping error and maximizes the number of valuable markers) and to  
150 determine repeatability rates for the AFLP markers (% of markers that are consistently scored).  
151 Markers likely to contribute high error rates to the data were discarded and the threshold that gave the  
152 lowest error rates for each primer combination was used to determine the final genotype for all  
153 individuals.

#### 154 *Microsatellite and AFLP data analyses*

155 **Genetic diversity.** For microsatellites, allele frequencies, the average number of alleles ( $N_{all}$ ) and  
156 allelic richness ( $A_r$ , i.e. the expected number of alleles corrected for sampling size, based on a  
157 rarefaction method) were estimated for each sample using FSTAT 2.9.3 (Goudet 1995). We used  
158 GENEPOP 4.1 software (Rousset 2008) to estimate expected heterozygosity ( $H_e$ ), fixation index ( $F_{IS}$ )  
159 and to test for departure from Hardy-Weinberg equilibrium in each population (10,000  
160 dememorization steps, 500 batches and 5,000 iterations per batch). To adjust the p-values for multiple  
161 tests, we computed the q-values using the QVALUE package in the R software (Storey 2002). For  
162 AFLPs, estimates of allelic frequencies were computed according to a Bayesian method with non-  
163 uniform prior distribution of allele frequencies (Zhivotovsky 1999) using the software AFLPsurv  
164 (Vekemans 2002). For each sample, allele frequencies were used to calculate the proportion of  
165 polymorphic loci (PLP, i.e. loci with allele frequencies from 5% to 95%), as well as unbiased  
166 estimates of expected heterozygosity ( $H_j$ , or Nei's gene diversity). Band richness ( $B_r$ ), an analogue of  
167 allelic richness, was computed on the AFLP data using the software AFLPDIV (Coart *et al.* 2005).

168 **Genetic population structure.** The genetic structure between sampled populations was computed by  
169 calculating an estimate of  $F_{ST}$  (Weir & Cockerham 1984) using GENEPOP 4.1 for microsatellites.  
170 Exact tests for population differentiation (10,000 dememorization steps, 500 batches and 5,000  
171 iterations per batch) were carried out to test for differences in allele distributions among native and  
172 introduced populations and between pairwise populations. For AFLPs,  $F_{ST}$  estimates were calculated  
173 and tested based on a permutation procedure (10,000 permutations) using AFLPsurv. The overall  
174 genetic structure among population samples was depicted using a Correspondence Analysis (CA)  
175 computed on allelic frequencies with GENETIX software (Belkhir *et al.* 2004) for microsatellites and  
176 the R package adegenet 1.3-4 (Jombart & Ahmed 2011) for AFLPs.

177 Genetic structure was also investigated by analyzing both microsatellite and AFLP data with an  
178 individual-based Bayesian clustering implemented in the software STRUCTURE 2.3.3 (Pritchard *et*  
179 *al.* 2000). All analyses were performed using the University of Oslo Biportal  
180 (<https://www.bioportal.uio.no/>). For each value of K (ranging from 1 to 25), we ran 30 replicate chains  
181 of 150,000 MCMC iterations and discarded the 50,000 burn-in iterations. An admixture model with  
182 correlated allele frequencies was applied without using *a priori* information on population origin. The  
183 most likely number of clusters was ascertained using the method proposed by Evanno *et al.* (2005)  
184 that determines the maximum value of  $\Delta K$ , a measure for the second-order rate of change in the  
185 likelihood of K, using the online software Structure Harvester (Earl & von Holdt 2011). Infiles  
186 produced by Structure Harvester were used in CLUMPP (Jakobsson & Rosenberg 2007) to determine  
187 individual assignment probabilities that best matched with all replicate runs. We used Distruct  
188 (Rosenberg 2004) to visualize the individuals' assignment to the different clusters.

189 **Outlier detection.** To identify outliers among all population samples and between each pair of samples  
190 (22 populations, 231 comparisons), we used two approaches. Both are methods that are based on the  
191 identification of loci showing genetic differentiation ( $F_{ST}$ -based estimates) significantly different from  
192 the estimated genome average. For the first approach, we used Beaumont and Nichols's (1996)  
193 method, using the software Lositan (Antao *et al.* 2008) (for codominant microsatellite markers) and

194 Dfdist (for dominant AFLP markers). Briefly, for each marker, this method uses simulations to  
195 estimate the distribution of  $F_{ST}$  according to the degree of heterozygosity under neutral expectations.  
196 Loci that do not fit neutral expectations (i.e. unusually high or low  $F_{ST}$  values) are identified as  
197 outliers. In this study, the trimmed  $F_{ST}$  (the highest and lowest 30% of empirical  $F_{ST}$  values are  
198 removed) was used as the targeted  $F_{ST}$  value in simulations, as recommended by Caballero *et al.*  
199 (2008). A total of 50,000 loci were simulated to generate  $F_{ST}$  values, and different values for the  
200 mutation rate  $\mu$  were initially tested for the simulations ( $10^{-3}$  to  $10^{-6}$ ) with identical results. The false  
201 discovery rate was controlled and q-values were computed using the QVALUE R-package (Storey  
202 2002). For the second approach, we used the software BayeScan, which implements the method  
203 developed by Foll & Gaggiotti (2008). It uses a Bayesian method and a logistic regression model to  
204 estimate directly the posterior probability (i.e. posterior odds) that a given locus is subject to selection.  
205 Briefly, locus-population  $F_{ST}$  values are decomposed as a linear combination of a locus-specific  
206 component ( $\alpha$ ) and a population-specific component ( $\beta$ ). A locus is assumed to be under selection  
207 when  $\alpha$  is necessary to explain the observed pattern of diversity. For each locus, a reversible-jump  
208 MCM explores two alternative models (with or without  $\alpha$ ) and estimates their relative posterior  
209 probabilities. The Jeffrey Interpretation scale (Jeffrey 1961) was used to identify the outliers (Foll and  
210 Gaggiotti 2008). Outliers found in both methods were examined in detail.

## 211 **Results**

### 212 *Level of genetic diversity in the native and introduced ranges*

213 Table 1 provides the estimates of genetic diversity for each population, based on microsatellite and  
214 AFLP markers. All 17 microsatellite markers were polymorphic with 2 (CL291) to 84 (DA8YN14)  
215 alleles per locus. The number of alleles per population sample ranged from 8.06 (Chesapeake) to 14.47  
216 (Fort Pierce). Allelic richness ( $A_r$ ) was quite stable, ranging from 6.59 (Yerseke) to 7.56 (Long  
217 Island). The same held for  $H_e$ , ranging from 0.606 (Port-en-Bessin) to 0.669 (Tjarno). Departures from  
218 Hardy-Weinberg equilibrium were found in each population mainly due to a deficiency in

219 heterozygotes at eight loci (BI13YE17, DA8YN14, DA1YN22, DA1YL23, CL83, Cf8, E4 and Ca4).  
220 Because these heterozygote deficiencies may be partly due to null alleles, all analyses (diversity,  
221 structure and outlier analyses) were performed with and without these eight loci, and results were  
222 identical. For AFLPs, a total of 327 polymorphic markers were selected, ranging from 50 to 500 bp.  
223 The number of markers varied among primer pairs, yielding from 37 to 110 polymorphic AFLP  
224 markers (Table 2). Of the 683 individuals genotyped, individuals that were not scored for the four  
225 primer combinations were excluded from the dataset (no missing values). The average mismatch error  
226 rate for the four primer combinations was  $2.11 \pm 0.76$ , the average  $\epsilon_{1,0}$  error rate was  $8.87 \pm 1.30$  and  $\epsilon_{0,1}$   
227 was  $0.09 \pm 0.05$  (Table 2). This dataset is the best trade-off between the number of markers and their  
228 quality (Holland *et al.* 2008; Whitlock *et al.* 2008). The values of PLP,  $H_j$  and  $B_r$  were similar across  
229 samples: they ranged from 27.5 (Port-en-Bessin) to 34.3 (Portsmouth, Arcachon, Sète), from 0.092  
230 (Port-en-Bessin) to 0.124 (Somers Point) and from 1.30 (Port-en-Bessin) to 1.44 (Bourgneuf),  
231 respectively. The gene diversity of native and introduced populations was not significantly different  
232 for either type of marker (group comparison with FSTAT software:  $p=0.99$  and  $p=0.85$  for AFLPs and  
233 microsatellites, respectively).

#### 234 *Genetic structure among and between native and introduced populations*

235 Genetic structure was estimated using either 17 microsatellite markers (dataset 1), 327 AFLP markers  
236 (dataset 2) or 319 AFLP markers, excluding the identified outliers (see below; dataset 3). The exact  
237 test revealed a significant difference in allele frequencies among the 22 sampled populations using any  
238 of the three datasets (“All population samples” in Table 3). In all datasets, the genetic structure among  
239 samples from the native range was higher than the genetic structure between the native and introduced  
240 ranges. In the native range, overall significant genetic structure was mainly due to the genetic  
241 differentiation of the Longboat Key (LBK) sample, located at the southern edge of the native range:  
242 lower  $F_{ST}$  estimates were observed when LBK was discarded (Table 3). Many pairwise comparisons  
243 revealed significant differences, but the only consistent results were observed with LBK, showing  
244 significant genetic differences with all of the 21 other populations (Supporting Information - Table



245 S1). As shown in Table 3 (and Table S1), the genetic structure estimated with 327 AFLP markers was  
246 higher than with microsatellites and in the AFLP set excluding outliers, especially when LBK was  
247 involved. In addition, low genetic structure was observed among introduced populations.

248 The correspondence analysis (CA) output is shown in Figure 2 for microsatellites. The CA showed  
249 clear differentiation between LBK and all other populations for microsatellites, a pattern also observed  
250 using AFLPs. However, CA was not very informative with AFLPs because less than 2% of the total  
251 variance was explained on the two first axes (Supporting Information - Fig. S1), even when rare alleles  
252 were removed from the data prior to analyses (Jombart *et al.* 2009). Using microsatellites only, three  
253 clusters were distinguished along the second axis: (1) the Chesapeake population, (2) a second group  
254 with five native populations plus five introduced populations (Mud Bay in US, the three UK  
255 populations and Port-en-Bessin in France), and (3) a third group clustering only introduced  
256 populations from northern Europe (Tjarno, Limfjord, Sylt Island, Yerseke) plus French populations  
257 except Port-en-Bessin.

258 Two clusters were identified from all samples in the STRUCTURE analyses (maximum  $\Delta K$  at  $K=2$ )  
259 whatever the marker set. Individual Bayesian assignments revealed that the first cluster (pictured in  
260 white in Fig. 3) was mainly defined by LBK individuals whatever the marker set: 100% of the LBK  
261 individuals showed individual assignment with probabilities higher than 50% to this cluster. Some  
262 introduced populations also showed a singular pattern for the microsatellites only set: 45 to 66% of the  
263 individuals sampled in Mud Bay (US), Lawrenny (UK), Canvey (UK) and Port-en-Bessin (France)  
264 were assigned with a probability higher than 50% to the first cluster (Fig. 3), a pattern also observed in  
265 native populations, except Nahant. This was not the case for the other introduced populations. These  
266 results were globally congruent to the results of the CA.

#### 267 *Signature of positive selection*

268 No outliers were detected using the microsatellite set, whatever the comparison (i.e. of 22 samples, 21  
269 samples excluding LBK or 231 pairwise population comparisons).

270 For the AFLP marker set, detection of outliers was first performed using all populations (i.e. 22  
271 populations). Six loci out of 327 (1.8%) were consistently identified as outliers with both methods  
272 (thresholds: 99% confidence levels in Dfdist; with a posterior probability higher than 91% in  
273 BayeScan), numbered from 1 to 6 in Fig. 4. With Dfdist, two additional outliers were observed (locus  
274 nos. 9 and 10, Fig 4A), and with BayeScan, three other loci were identified (locus nos. 7, 11 and 12,  
275 Fig. 4B). We then performed 231 pairwise comparisons using both methods and the same confidence  
276 level as described above. Using Dfdist, from the 231 comparisons, 104 outliers were detected whereas  
277 using BayeScan, only 5 outliers, also detected by Dfdist, were identified (Fig. 4C). These five outliers  
278 were observed in 12 out of 231 comparisons, and, interestingly, all involved LBK (Fig. 4D, black  
279 lines). In addition, these 12 comparisons involved at least one outlier previously identified in the  
280 global analysis (locus nos. 1 to 4 and 7; Fig. 4). Locus nos. 1 to 4 and 7 were also identified in 13  
281 other pairwise comparisons, of which 9 involved LBK, still at the 99% confidence level for Dfdist, but  
282 with a posterior probability lower than 90% for BayeScan (gray lines in Fig. 4D). An additional locus,  
283 named locus no. 8 (Figs. 4C and D), was revealed in several comparisons without LBK by Dfdist and  
284 by BayeScan (ranked as “substantial” with the Jeffrey Interpretation scale). To summarize, eight loci  
285 were consistently identified as outliers: 1) locus nos. 1 to 4 in global and pairwise analyses, 2) locus  
286 nos. 5 and 6 only in global analyses, and 3) locus nos. 7 and 8 in pairwise analyses. Allelic frequencies  
287 of these outliers are given in Figure S2. These eight loci, likely to be under the influence of selection  
288 processes, were those excluded from dataset 3 (see “Genetic structure analyses”).

## 289 **Discussion**

290 In this study, we used 344 loci to analyze the genetic diversity and structure in 7 native and 15  
291 introduced (Europe and USA) populations of the marine gastropod *Crepidula fornicata*. Apart from  
292 additional insights on the introduction process, our genome scan study revealed adaptive  
293 differentiation in the native range. However, our results do not support the hypothesis of selection  
294 following introduction to explain the success of this invasive species.

295 Several eco-evolutionary hypotheses involving selection processes have been proposed to explain the  
296 success of invasive species (Lee 2002; Facon *et al.* 2006; Sax *et al.* 2007), and studies on invasive  
297 plants (e.g. Barrett *et al.* 2008) or insects (e.g. Zayed & Whitfield 2008), for instance, support these  
298 hypotheses. Evolutionary changes may result from selection from pre-existing genetic variation  
299 [hypothesis 1], a process exemplified by coat color changes in oldfield mice according to the color of  
300 land colonized (Steiner *et al.* 2007), or new beneficial alleles supplied by mutation [hypothesis 2] (e.g.  
301 Rainey & Travisano 1998). Hypothesis 1 is the most likely to occur on a short-time scale, but  
302 introduction processes need a large enough pool of genetic diversity for selection to take place (Barrett  
303 & Schlutter 2008). Although populations of some non-indigenous marine species have reduced genetic  
304 diversity due to founder events (e.g. the invasive Chinese mitten crab *Eriocheir sinensis*; Herborg *et*  
305 *al.* 2007), many marine invasive species have similar, or even higher genetic diversity in their  
306 introduced as in their native populations, especially when they derive from multiple source  
307 populations and involve population admixture (e.g. Goldstien *et al.* 2011; reviewed in Roman &  
308 Darling 2007). *C. fornicata* falls in the latter category: genetic diversity was similar in populations  
309 from the introduced and native ranges. We also sampled large latitudinal gradients on both sides of the  
310 Atlantic and expected to find adaptive polymorphisms as in other marine species (e.g. Bradbury *et al.*  
311 2010; Gagnaire *et al.* 2012). Despite high genetic diversity and ample sampling, no outlier loci were  
312 revealed in an overall comparison of the introduced and native ranges. Only comparisons involving  
313 the LBK population revealed outliers, which can be attributed to selection processes occurring within  
314 the native range (see below). Further, our analysis did not provide any evidence to support hypothesis  
315 2 either, even though genome scans are expected to identify “hard local sweeps” better than selection  
316 processes acting on previously neutral pre-existing variation (Przeworski *et al.* 2005; Teshima *et al.*  
317 2006).

318 Weak selection, particularly if it involves polygenic traits, is the most likely explanation for our  
319 results. *C. fornicata* is found in similar environments in the native and introduced ranges (shallow  
320 bays and estuaries) and weak divergent selection between native and introduced populations of *C.*

321 *fornicata* may have impeded the identification of outliers. The genome-scan approach can successfully  
322 identify genomic regions under selection, but usually when pre- and post-reproductive isolation  
323 mechanisms have evolved to build up a strong genetic barrier, e.g. after colonization of sticklebacks in  
324 freshwater environments (Makinen *et al.* 2008). Conversely, this approach cannot effectively  
325 demonstrate weak selection (see also Przeworski 2002; Pennings & Hermisson 2006; Narum & Hess  
326 2011), even with genome-wide analysis (Burke *et al.* 2010). Our results thus cast some doubt as to the  
327 effectiveness of genome scans in demonstrating selection processes following introduction, at least  
328 when there are no major environmental changes between the native and the introduction ranges.  
329 Ideally, QTL mapping to screen for genetic markers closely linked to genes that underlie a phenotypic  
330 trait of interest (i.e. target of selection in the introduction range) would be needed to examine selection  
331 processes and get more direct insights into the additive genetic variance associated with these  
332 phenotypic traits (Lee 2002; Stinchcombe & Hoekstra 2008). There are good candidate phenotypic  
333 traits, known to be variable and heritable in *C. fornicata*, such as larval growth rate (Hilbish *et al.*  
334 1999; Le Cam *et al.* 2009). However, QTL analyses require controlled crosses, a difficult task in *C.*  
335 *fornicata* whose individuals are long-lived, form stacks that are impossible to manipulate non-  
336 destructively, and practice internal fertilization. Alternative approaches include quantitative genetic  
337 differentiation analyses among populations (expressed as  $Q_{ST}$ ; Spitze 1993) on other candidate traits,  
338 such as stacking behavior (Le Cam *et al.* 2009). Finally, if selection is not a prominent driver of the  
339 introduction success in *C. fornicata*, non-adaptive phenotypic plasticity may play a role. High  
340 ecological tolerance (euryhaline and eurythermal) has indeed been reported in *C. fornicata* (Blanchard  
341 1995), suggesting that *C. fornicata* can cope with a range of environments.

342 Although no outliers were identified in the introduction range, we nevertheless detected outlier loci in  
343 the native range. Significant genetic structure was observed between the native populations from the  
344 Atlantic (from Nahant to Fort Pierce) and the one located in the Gulf of Mexico, Longboat Key  
345 (LBK). A large part of this genetic structure was explained by a subset of loci (2.2%) that behaved like  
346 non-neutral loci. This percentage of outliers (2.2%) is similar to those reported in the literature: 1-5%

347 outliers have been detected in non-model species (animal or vegetal, vertebrate or invertebrate, marine  
348 or terrestrial) based on AFLP genome scans (Wood *et al.* 2008). Reduced gene flow for this subset of  
349 the genome occurred along the southern coast of Florida, coinciding with the well-known transition  
350 zone between temperate and subtropical marine ecosystems (Hare & Avise 1996; Pelc *et al.* 2009).  
351 Studying mtDNA markers of *C. fornicata* and two other direct-developers from the same  
352 Calyptraeidae family, Collin (2001) also suggested a similar genetic break in the same area. It is  
353 tempting to impute the observed genetic differentiation to local adaptation alone, with outlier loci  
354 being signs of selection. However, the observed  $F_{ST}$ -outliers may be due to the coupling between  
355 endogenous and exogenous barriers trapped at environmental boundaries rather than local adaptation  
356 alone (Bierne *et al.* 2011). In this interpretation, although the location of the genetic break would be  
357 due to exogenous factors (e.g. temperature, salinity, oceanic currents), the barrier to gene flow would  
358 be mainly driven by many endogenous incompatibilities that restrict gene flow in part of the genome.  
359 The observed genetic break would thus be trapped by the natural hydrographic barrier occurring at the  
360 tip of Florida (Pelc *et al.* 2009). The endogenous factors are unfortunately difficult to identify (Barton  
361 & Bengtsson 1986; Bierne *et al.* 2011).

362  $F_{ST}$ -based comparisons have successfully proven their use in identifying genomic regions under  
363 selection in contrasted environments (e.g. Wilding *et al.* 2001; Nunes *et al.* 2011), or in hybrid zones  
364 (Gagnaire *et al.* 2011). Here, they were also useful in shedding light on the history of introduction.  
365 Although no evidence of selection events was identified between native and introduced ranges or  
366 within the introduced range, the outliers identified within the native range proved helpful for  
367 unambiguously excluding the LBK population as a source population. Source populations of the  
368 European and north Pacific introductions are unlikely to be identified more precisely. From Nahant to  
369 Fort Pierce, native populations are weakly genetically structured whatever the markers used (enzymes  
370 (Hoagland 1984; Dupont *et al.* 2003), mtDNA markers (Collin 2001), microsatellites and AFLPs (this  
371 study)). The exact origin of introduction cannot be identified, a common problem in marine invaders  
372 with long-lived pelagic larvae and large population sizes (Geller *et al.* 2010; Fitzpatrick *et al.* 2012).

373 The high level of genetic diversity in the introduction range supports a scenario of introduction with  
374 large propagule pressure and several sources (Roman & Darling 2007). Although the increase in the  
375 number of markers provided the same overall results in terms of genetic diversity as previous studies  
376 carried out with isozymes (Hoagland 1984; Dupont *et al.* 2003), new inferences could be made. In  
377 particular, genetic structure was observed among introduced populations and two groups could be  
378 defined based on CA and STRUCTURE analyses using microsatellites. Two populations from the UK,  
379 one French population (Port-en-Bessin) and the American introduced population showed similar  
380 genetic composition, contrasting with the other European populations. However, AFLPs did not  
381 strongly support this genetic structure at the European scale, suggesting that, in parallel to genotyping  
382 hundreds of bi-allelic loci (e.g. AFLPs), highly variable codominant markers, (e.g. microsatellites), are  
383 still useful for reconstructing the history of invasion. In the absence of selection events, as shown here,  
384 two distinct groups of European populations appear to have contrasting introduction histories (date  
385 and sources). Invasive species have often been shown to have contrasting genetic footprints depending  
386 on the introduction processes involved (e.g. Voisin *et al.* 2005; Goldstien *et al.* 2011). The  
387 introduction of *C. fornicata* in England and in the Puget Sound (Mud Bay) was reported in the early  
388 20<sup>th</sup> century as the consequence of accidental introductions during the importation of American oysters  
389 from the North American Atlantic coast. This report is congruent with the genetic similarity observed  
390 between two UK populations and Mud Bay. The second group of European populations may be better  
391 explained by the repeated releases of oyster spats in the 1970s: Japanese oysters previously introduced  
392 in the North Pacific (Puget Sound) — where *C. fornicata* had already been imported (in the early 20<sup>th</sup>  
393 century) — were introduced along the Mediterranean and Atlantic coasts of France (Blanchard 1995;  
394 1997). The Pacific oyster subsequently spread to northern Europe (Smaal *et al.* 2009) possibly with  
395 the slipper limpet. This introduction pathway may explain the genetic similarity of all the other  
396 European populations, with the exception of Port-en-Bessin. It is unlikely that this secondary  
397 introduction pathway was accompanied by important gene surfing events. Genetic drift and founder  
398 events may play an important role in invasions, but are less efficient in large populations such as those  
399 established by *C. fornicata* (Excoffier & Ray 2008; Lehe *et al.* 2012). Two historical pathways and

400 sources may thus explain the pattern observed here: a first wave of introduction in Mud Bay and the  
401 UK from a primary source in the native range, and a second wave in France and North Europe from a  
402 secondary source located in an already invaded area (i.e. Puget Sound).

### 403 **Conclusion**

404 No evidence for selection processes following introduction was revealed using a genome-scan  
405 approach. No genome scans in other invasive marine species have been published to date with which  
406 to compare our results. Such a comparison would be relevant to either generalize the pattern observed  
407 here to marine invertebrates with large population size or to underline the singularity of the study  
408 species *C. fornicata*. Selection events were nonetheless demonstrated within the native range, and  
409 likely to result from long-term adaptation processes. We suggest that selection between native and  
410 introduced ranges and within the introduced range may be particularly difficult to identify with  
411 genome scans because it results from recent adaptation. In invasive species with high gene flow and  
412 little genetic structure, which are notoriously challenging to examine using a molecular approach,  
413 genome scans may nevertheless be useful in recovering the history of introduction (e.g. exclusion of  
414 source populations), as demonstrated here. In light of the increasing number of genome scan studies,  
415 this study also advocates pursuing the analysis of multi-allelic codominant microsatellites to  
416 investigate the processes underlying the successful establishment of non-indigenous species.

### 417 **Acknowledgements**

418 We are grateful to J. Christensen, R. Collin, N. Cuvelier, X. de Montaudouin, P. Dolmer, E. Foucher,  
419 K. Halanych, H. Hummel, C. Lemaire, J.L. Martin, J. Olsen, P.G. Sauriau, D. Thielges, and divers of  
420 the “Marine Operations Department” of the Roscoff Biological Station for providing specimens of  
421 *Crepidula fornicata*. FV is grateful to D. McGlashan for his major contribution to the field sampling  
422 work in US and for preliminary results that helped to design this global analysis. We thank the  
423 Biogenouest® Genomics core facility for its technical support. We are also thankful to C. Caron, R.  
424 Whitlock, P.A. Gagnaire, T. Broquet and the University of Oslo (Bioportal) for help in computing.

425 This work was funded by the Agence Nationale de la Recherche (Project Hi-Flo, ANR-08-BLAN-  
426 0334). FR acknowledges a PhD grant from the Centre National pour la Recherche Scientifique and the  
427 Conseil Régional de Bretagne (SEA-FLO ARED project).

## 428 **References**

429 Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect  
430 molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, **9**, 323.

431 Barrett SCH, Colautti RI, Eckert CG (2008) Plant reproductive systems and evolution during  
432 biological invasion. *Molecular Ecology*, **17**, 373-383.

433 Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology and*  
434 *Evolution*, **23**, 38-44.

435 Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridizing populations.  
436 *Heredity*, **57**, 357-376.

437 Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population  
438 structure. *Proceedings of the Royal Society of London, Series B*, **263**, 1619-1626.

439 Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous  
440 Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions,  
441 CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

442 Bierne N (2010). The distinctive footprints of local hitchhiking in a varied environment and global  
443 hitchhiking in a subdivided population. *Evolution*, **64**, 3254-3272.

444 Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome  
445 scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044-2072.

446 Blanchard M (1995) Origine et état de la population de *Crepidula fornicata* (Gastropoda  
447 Prosobranchia) sur le littoral français. *Haliotis*, **24**, 75-86.

448 Blanchard M (1997) Spread of the slipper limpet *Crepidula fornicata* (L. 1758) in Europe. Current  
449 state and consequences. *Scientia Marina*, **61**, 109-118.

450 Bonin A, Pompanon F, Taberlet P (2005) Use of amplified fragment length polymorphism (AFLP)  
451 markers in surveys of vertebrate diversity. *Molecular Evolution: Producing the Biochemical Data, Part*  
452 *B*, **395**, 145-161.



453 Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci  
454 for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology*  
455 *and Evolution*, **23**, 773-783.

456 Bradbury IR, Hubert S, Higgins B, Borza T, Bowman S, Paterson IG, Snelgrove PVR, Morris CJ,  
457 Gregory RS, Hardie DC, Hutchings JA, Ruzzante DE, Taggart CT, Bentzen P (2010) Parallel adaptive  
458 evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proceedings*  
459 *of the Royal Society B-Biological Sciences*, **277**, 3725-3734.

460 Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD (2010) Genome-wide  
461 analysis of a long-term evolution experiment with *Drosophila*. *Nature*, **467**, 587-U111.

462 Caballero A, Quesada H, Rolan-Alvarez E (2008) Impact of amplified fragment length polymorphism  
463 size homoplasy on the estimation of population genetic diversity and the detection of selective loci.  
464 *Genetics*, **179**, 539-554.

465 Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced  
466 polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided  
467 populations. *Genetical Research*, **70**, 155-174.

468 Coart E, Van Glabeke S, Petit RJ, Van Bockstaele E, Roldan-Ruiz I (2005) Range wide versus local  
469 patterns of genetic diversity in hornbeam (*Carpinus betulus* L.). *Conservation Genetics*, **6**, 259-273.

470 Collin R (2001) The effects of mode of development on phylogeography and population structure of  
471 North Atlantic *Crepidula* (Gastropoda : Calyptraeidae). *Molecular Ecology*, **10**, 2249-2262.

472 Coyer JA, Hoarau G, Pearson G, Mota C, Juterbock A, Alpermann T, John U, Olsen JL (2011)  
473 Genomic scans detect signatures of selection along a salinity gradient in populations of the intertidal  
474 seaweed *Fucus serratus* on a 12 km scale. *Marine Genomics*, **4**, 41-49.

475 Dupont L, Jollivet D, Viard F (2003) High genetic diversity and ephemeral drift effects in a successful  
476 introduced mollusc (*Crepidula fornicata* : Gastropoda). *Marine Ecology-Progress Series*, **253**, 183-  
477 195.

478 Dupont L, Richard J, Paulet YM, Thouzeau G, Viard F (2006) Gregariousness and protandry promote  
479 reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of  
480 larval paternity. *Molecular Ecology*, **15**, 3009-3021.

481 Dupont L, Viard F (2003) Isolation and characterization of highly polymorphic microsatellite markers  
482 from the marine invasive species *Crepidula fornicata* (Gastropoda: Calyptraeidae). *Molecular Ecology*  
483 *Notes*, **3**, 498-500.

484 Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing  
485 STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**,  
486 359-361.

487 Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the  
488 software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.

489 Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and  
490 structuration. *Trends in Ecology and Evolution*, **23**, 347-351.

491 Facon B, Genton BJ, Shykoff J, Jarne P, Estoup A, David P (2006) A general eco-evolutionary  
492 framework for understanding bioinvasions. *Trends in Ecology and Evolution*, **21**, 130-135.

493 Fitzpatrick BM, Fordyce JA, Niemiller ML, Reynolds RG (2012) What can DNA tell us about  
494 biological invasions? *Biological Invasions*, **14**, 245-253.

495 Foll M, Gaggiotti O (2008) A Genome-scan method to identify selected loci appropriate for both  
496 dominant and codominant markers: a bayesian perspective. *Genetics*, **180**, 977-993.

497 Gagnaire PA, Minegishi Y, Zenboudji S, Valade P, Aoyama J, Berrebi P (2011) Within-population  
498 structure highlighted by differential introgression across semipermeable barriers to gene flow in  
499 *Anguilla marmorata*. *Evolution*, **65**, 3413-3427.

500 Gagnaire PA, Normandeau E, Cote C, Hansen MM, Bernatchez L (2012) The genetic consequences of  
501 spatially varying selection in the panmictic American eel (*Anguilla rostrata*). *Genetics*, **190**, 725-  
502 U703.

503 Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological invasions. *Annual*  
504 *Review of Marine Science*, **2**, 367-393.

505 Goldstien SJ, Dupont L, Viard F, Hallas PJ, Nishikawa T, Schiel DR, Gemmell NJ, Bishop JDD  
506 (2011) Global phylogeography of the widely introduced North West Pacific ascidian *Styela clava*.  
507 *Plos One*, **6**, e16755.

508 Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of*  
509 *Heredity*, **86**, 485-486.

510 Hare MP, Avise JC (1996) Molecular genetic analysis of a stepped multilocus cline in the American  
511 oyster (*Crassostrea virginica*). *Evolution*, **50**, 2305-2315.

512 Herborg LM, Weetman D, Van Oosterhout C, Hanfling B (2007) Genetic population structure and  
513 contemporary dispersal patterns of a recent European invader, the Chinese mitten crab, *Eriocheir*  
514 *sinensis*. *Molecular Ecology*, **16**, 231-242.

515 Hilbish TJ, Sasada K, Eyster LS, Pechenik JA (1999) Relationship between rates of swimming and  
516 growth in veliger larvae: genetic variance and covariance. *Journal of Experimental Marine Biology*  
517 *and Ecology*, **239**, 183-193.

518 Hoagland KE (1984) Use of molecular-genetics to distinguish species of the gastropod genus  
519 *Crepidula* (Prosobranchia, calyptraeidae). *Malacologia*, **25**, 607-628.

520 Holland BR, Clarke AC, Meudt HM (2008) Optimizing automated AFLP scoring parameters to  
521 improve phylogenetic resolution. *Systematic Biology*, **57**, 347-366.

522 Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for  
523 dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**,  
524 1801-1806.

525 Jeffreys H (1961) The theory of probability (3rd ed.) Oxford University Press, New York.

526 Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data.  
527 *Bioinformatics*, **27**, 3070-3071.

528 Jombart T, Pontier D, Dufour AB (2009) Genetic markers in the playground of multivariate analysis.  
529 *Heredity*, **102**, 330-341.

530 Koehn RK, Newell RIE, Immermann F (1980) Maintenance of an aminopeptidase allele frequency  
531 cline by natural selection. *Proceedings of the National Academy of Sciences of the United States of*  
532 *America-Biological Sciences*, **77**, 5385-5389.

533 Le Cam S, Pechenik JA, Cagnon M, Viard F (2009) Fast versus slow larval growth in an invasive  
534 marine mollusc: Does paternity matter? *Journal of Heredity*, **100**, 455-464.

535 Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local  
536 adaptation. *Molecular Ecology*, **21**, 1548-1566.

537 Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386-  
538 391.

539 Lehe R, Hallatschek O, Peliti L (2012) The rate of beneficial mutations surfing on the wave of a range  
540 expansion. *PLoS computational biology*, **8**, e1002447.

541 Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*,  
542 **17**, 183-189.

543 Li JR, Li HP, Jakobsson M, Li S, Sjodin P, Lascoux M (2012) Joint analysis of demography and  
544 selection in population genetics: where do we stand and where could we go? *Molecular Ecology*, **21**,  
545 28-44.

546 Makinen HS, Cano M, Merila J (2008) Identifying footprints of directional and balancing selection in  
547 marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. *Molecular*  
548 *Ecology*, **17**, 3565-3582.

549 Narum SR, Hess JE (2011) Comparison of F-ST outlier tests for SNP loci under selection. *Molecular*  
550 *Ecology Resources*, **11**, 184-194.

551 Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine  
552 fishes: identifying adaptive variation in space and time. *Molecular Ecology*, **18**, 3128-3150.

553 Nunes VL, Beaumont MA, Butlin RK, Paulo OS (2011) Multiple approaches to detect outliers in a  
554 genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient.  
555 *Molecular Ecology*, **20**, 193-205.

556 Pelc RA, Warner RR, Gaines SD (2009) Geographical patterns of genetic structure in marine species  
557 with contrasting life histories. *Journal of Biogeography*, **36**, 1881-1890.

558 Pennings PS, Hermisson J (2006) Soft sweeps III: The signature of positive selection from recurrent  
559 mutation. *Plos Genetics*, **2**, 1998-2012.

560 Prentis PJ, Wilson JRU, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in  
561 invasive species. *Trends in Plant Science*, **13**, 288-294.

562 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus  
563 genotype data. *Genetics*, **155**, 945-959.

564 Proestou DA (2006) Isolation and characterization of microsatellite markers in the Atlantic slipper  
565 shell *Crepidula fornicata* for use in paternity analysis. *Molecular Ecology Notes*, **6**, 437-439.

566 Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics*, **160**,  
567 1179-1189.

568 Przeworski M, Coop G, Wall JD (2005) The signature of positive selection on standing genetic  
569 variation. *Evolution*, **59**, 2312-2323.

570 Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature*, **394**, 69-  
571 72.

572 Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what  
573 empirical studies reveal about the conditions that promote adaptive evolution. *Genetica*, **112**, 183-198.

574 Riquet F, Ballenghien M, Tanguy A, Viard F (2011) *In silico* mining and characterization of 12 EST-  
575 SSRs for the invasive slipper limpet *Crepidula fornicata*. *Marine Genomics*, **4**, 291-295.

576 Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions.  
577 *Trends in Ecology and Evolution*, **22**, 454-464.

578 Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure.  
579 *Molecular Ecology Notes*, **4**, 137-138.

580 Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for  
581 Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.

582 Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, Grosberg RK, Hastings A,  
583 Holt RD, Mayfield MM, O'Connor MI, Rice WR (2007) Ecological and evolutionary insights from  
584 species invasions. *Trends in Ecology and Evolution*, **22**, 465-471.

585 Simon-Bouhet B, Garcia-Meunier P, Viard F (2006) Multiple introductions promote range expansion  
586 of the mollusc *Cyclope neritea* (Nassariidae) in France: evidence from mitochondrial sequence data.  
587 *Molecular Ecology*, **15**, 1699-1711.

588 Smaal AC, Kater BJ, Wijsman J (2009) Introduction, establishment and expansion of the Pacific  
589 oyster *Crassostrea gigas* in the Oosterschelde (SW Netherlands). *Helgoland Marine Research*, **63**, 75-  
590 83.

591 Spitze K (1993) Population-structure in *Daphnia obtusa* – Quantitative genetic and allozymic  
592 variation. *Genetics*, **135**, 367-374.

593 Steiner CC, Weber JN, Hoekstra HE (2007) Adaptive variation in beach mice produced by two  
594 interacting pigmentation genes. *Plos Biology*, **5**, 1880-1889.

595 Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics:  
596 finding the genes underlying ecologically important traits. *Heredity*, **100**, 158-170.

- 597 Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology.  
598 *Trends in Ecology and Evolution*, **18**, 94-101.
- 599 Storey J (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society*,  
600 *Series B*, **64**, 479-498.
- 601 Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence.  
602 *Molecular Ecology*, **14**, 671-688.
- 603 Teshima KM, Coop G, Przeworski M (2006) How reliable are empirical genomic scans for selective  
604 sweeps? *Genome Research*, **16**, 702-712.
- 605 Tice KA, Carlson DB (2011) Can AFLP genome scans detect small islands of differentiation? The case  
606 of shell sculpture variation in the periwinkle *Echinolittorina hawaiiensis*. *Journal of Evolutionary*  
607 *Biology*, **24**, 1814-1825.
- 608 Vekemans X (2002) AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Génétique et  
609 Ecologie Végétale, Université Libre de Bruxelles, Belgium.
- 610 Voisin M, Engel CR, Viard F (2005) Differential shuffling of native genetic diversity across  
611 introduced regions in a brown alga: Aquaculture vs. maritime traffic effects. *Proceedings of the*  
612 *National Academy of Sciences of the United States of America*, **102**, 5432-5437.
- 613 Vos P, Hogers R, Bleeker M, Reijans M, Vandeele T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper  
614 M, Zabeau M (1995) AFLP - A new technique for DNA-fingerprinting. *Nucleic Acids Research*, **23**,  
615 4407-4414.
- 616 Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure.  
617 *Evolution*, **38**, 1358-1370.
- 618 Whitlock R, Hipperson H, Mannarelli M, Butlin RK, Burke T (2008) An objective, rapid and  
619 reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular*  
620 *Ecology Resources*, **8**, 725-735.
- 621 Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of  
622 *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology*, **14**, 611-619.
- 623 Wood HM, Grahame JW, Humphray S, Rogers J, Butlin RK (2008) Sequence differentiation in  
624 regions identified by a genome scan for local adaptation. *Molecular Ecology*, **17**, 3123-3135.

625 Zayed A, Whitfield CW (2008) A genome-wide signature of positive selection in ancient and recent  
626 invasive expansions of the honey bee *Apis mellifera*. *Proceedings of the National Academy of Sciences*  
627 *of the United States of America*, **105**, 3421-3426.

628 Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA  
629 markers. *Molecular Ecology*, **8**, 907-913.

630

### 631 **Author contributions**

632 FR designed and performed molecular experiments, analyzed the data and wrote the paper. CDT  
633 developed AFLP protocols, analyzed the data and wrote the paper. MB performed part of the  
634 microsatellite genotyping. NB analyzed the data and wrote the paper. FV designed the work,  
635 performed sampling, analyzed data and wrote the paper.

### 636 **Data accessibility**

637 Data files have been deposited at DRYAD: XXXX

638 **Figures**

639 Fig. 1 Distribution (black line) and sampling locations of *Crepidula fornicata*: native populations are  
640 numbered 1 to 7 (black dots) and introduced populations are numbered 8 to 22 (white dots).

641 Fig. 2 Correspondence analysis conducted with 17 microsatellite markers. Only the barycenter of each  
642 population is shown. Native populations are numbered 1 to 7 (black dots), and introduced populations  
643 are numbered 8 to 22 (white dots), as in Fig. 1.

644 Fig. 3 Individual Bayesian assignment proportions (K=2) determined using STRUCTURE with  
645 microsatellite and AFLP markers for the 22 populations ( $N_{\text{individuals(microsatellites)}}=683$ ,  $N_{\text{individuals(AFLP)}}=579$ ). The dotted black line separates population samples from the native and  
646 introduction ranges and the solid vertical black lines separate population samples. The two clusters are  
647 distinguished by white and gray: each individual is depicted as a vertical bar with colors distinguishing  
648 its membership to the two clusters. For each population sample, the percentage of individuals with an  
649 assignment to the white cluster higher than 50% is given in parentheses.  
650

651 Fig. 4 Genome-scan analyses on AFLP markers. A-B: global analyses (over all 22 populations), C-D:  
652 summary of 231 pairwise comparisons. Identified outliers are numbered from 1 to 12 in the four  
653 figures. Black-labeled outliers were consistently identified by both methods, whereas those in gray  
654 were detected by one method only.

655 A- Outliers detected using Dfdist on 22 populations. Plot shows  $F_{ST}$  values according to  
656 heterozygosity of the 327 AFLP loci analyzed. Black lines represent the 99<sup>th</sup> quantile  
657 estimated from simulations and the gray line, the median. Outliers exceeding the 99<sup>th</sup> quantile  
658 are shown in black.

659 B- Outliers detected using BayeScan on 22 populations. Plot shows  $F_{ST}$  values (y-axis) according  
660 to the posterior odds (PO) of the 327 AFLP loci analyzed. The vertical solid and dashed lines  
661 indicate strong ( $\log_{10}(\text{PO})=1$ ,  $p(\alpha \neq 0) > 0.91$ ) and decisive ( $\log_{10}(\text{PO})=2$ ,  $p(\alpha \neq 0) > 0.99$ ) evidence  
662 for selection, respectively.



663 C- Outliers detected in the pairwise comparisons. Number of pairwise comparisons that detected  
664 outliers confirmed by both software programs (in black) and by Dfdist only (in gray). \*:  
665 outliers detected when all populations are included in the analysis.

666 D- Schematic diagram showing outliers in pairwise comparisons. Native populations are  
667 underlined. Outliers are labeled on the lines. Strong outliers are represented by black lines;  
668 outliers that are noteworthy (detected by Dfdist but with “*substantial*” or “*barely worth*  
669 *mentioning*” according to the Jeffrey Interpretation scale in BayeScan) are shown in gray.

670

671

672

673 Table 1 Genetic diversity of the population samples (named after the locality in which they were sampled) for 17 microsatellite markers and 327 AFLP markers. N:  
674 number of individuals genotyped,  $N_{all}$ : average number of alleles per locus,  $A_r$ : allelic richness,  $H_e$ : expected heterozygosity,  $F_{is}$ : fixation index and, PLP: proportion of  
675 polymorphic loci (with 95% criteria),  $H_j$ : Nei's gene diversity and  $B_r$ : band richness. For geographic locations of sampled populations, see Fig. 1.

Population	Country - State	Genetic diversity estimates based on 17 microsatellite markers					Genetic diversity estimates based on AFLP markers			
		N	$N_{all}$	$A_r$	$H_e$	$F_{is}^{(1)}$	N	PLP	$H_j$	$B_r$
Native range		202	24.47	7.28	0.656	0.168	151	30.3	0.102	1.38
Nahant	USA – Massachussetts	32	10.94	6.89	0.630	0.137	27	30.3	0.099	1.34
Fairhaven	USA – Massachussetts	23	9.65	7.07	0.647	0.160	23	28.7	0.101	1.36
Long Island	USA – New York	32	12.65	7.56	0.652	0.154	31	30.9	0.101	1.36
Somers Point	USA – New Jersey	31	11.76	7.44	0.649	0.181	12	30.6	0.124	1.43
Chesapeake	USA – Virginia	13	8.06	7.40	0.633	0.099	13	28.1	0.114	1.40
Fort Pierce	USA – Florida - East coast	53	14.47	7.28	0.640	0.124	29	30.3	0.101	1.35
Longboat Key	USA – Florida - West coast	18	9.18	7.32	0.639	0.201	16	33.0	0.111	1.41
Introduced ranges		481	22.47	6.88	0.644	0.154	428	31.7	0.102	1.37
Mud Bay	USA – Washington	32	11.00	7.05	0.645	0.159	29	31.8	0.102	1.35
Tjarno	Sweden	32	10.94	7.02	0.669	0.075	20	32.7	0.109	1.38
Limfjord	Denmark	30	9.41	6.67	0.640	0.195	32	28.4	0.100	1.34
Sylt Island	Germany	34	10.35	6.65	0.635	0.196	28	33.6	0.107	1.37
Yerseke	Netherlands	33	9.76	6.59	0.645	0.234	32	30.0	0.105	1.36
Lawrenny	United Kingdom	32	10.65	6.86	0.654	0.200	27	28.1	0.104	1.36
Portsmouth	United Kingdom	35	12.29	7.25	0.645	0.151	30	34.3	0.116	1.41
Canvey Island	United Kingdom	32	11.35	7.23	0.646	0.172	28	30.6	0.108	1.38
Gravelines	France	32	11.29	7.15	0.647	0.119	31	33.3	0.101	1.37
Port-en-Bessin	France	31	10.41	6.81	0.606	0.185	24	27.5	0.092	1.30
Bay of Morlaix	France	32	10.35	6.92	0.635	0.146	28	31.5	0.104	1.37
Bourgneuf	France	30	10.47	6.96	0.651	0.146	25	33.3	0.123	1.44
Fouras	France	32	10.12	6.68	0.627	0.072	32	31.5	0.105	1.39
Arcachon	France	32	10.12	6.62	0.627	0.070	31	34.3	0.107	1.38
Sète	France	32	9.94	6.74	0.661	0.154	31	34.3	0.100	1.36

<sup>(1)</sup> p-values associated with HWE < 0.0001 in each case

676

677 Table 2 AFLP marker analysis: primer pair combination (with ABI dyes used) and criteria for retaining AFLP loci for data analyses. The number of individuals (N)  
 678 and the percentage of replicate samples (Nrep) are indicated with the number of peaks initially analyzed (initial markers), the error rate estimation (mismatch error,  $\epsilon_{1,0}$   
 679 and  $\epsilon_{0,1}$ ) and the number and proportion of polymorphic markers used for further analyses.

ABI Dyes	Selective <i>Eco</i> RI Primers (5' - 3')	Selective <i>Mse</i> I Primers (5' - 3')	N	Nrep (%)	Initial markers	Mismatch error %	$\epsilon_{1,0}$	$\epsilon_{0,1}$	Number of markers used as polymorphic markers	% of markers used for genotyping
NED	GACTGCGTACCAATTC <b>ACG</b>	GATGAGTCCTGAGTAA <b>AGC</b>	606	12.96	187	1.26	10.73	0.02	110	58.82
PET	GACTGCGTACCAATTC <b>AAC</b>	GATGAGTCCTGAGTAA <b>ACC</b>	594	12.81	96	2.96	8.13	0.07	37	38.54
FAM	GACTGCGTACCAATTC <b>CCTG</b>	GATGAGTCCTGAGTAA <b>CAG</b>	597	12.57	101	2.48	8.75	0.13	85	84.16
VIC	GACTGCGTACCAATTC <b>CCTC</b>	GATGAGTCCTGAGTAA <b>CTC</b>	597	12.39	116	1.72	7.86	0.14	95	81.90
			total:	12.68	500	2.11	8.87	0.09	327	65.40
			SD:	0.25		0.76	1.30	0.05		

680

681

682 Table 3 Genetic structure within and among introduced and native ranges.  $F_{ST}$  estimates and probabilities (exact test) for the 3 datasets (see text).

	17microsatellite markers		327 AFLP markers		AFLPs without outliers	
	$F_{ST}$	p-value	$F_{ST}$	p-value	$F_{ST}$	p-value
All population samples (22 samples)	0.020	< 0.001	0.024	< 0.001	0.015	< 0.001
Native range (1 group) vs. Introduced range (1 group)	0.004	< 0.001	0.006	< 0.001	0.003	0.004
Native range (7 samples)	0.026	< 0.001	0.038	< 0.001	0.020	< 0.001
Native range without LBK <sup>(1)</sup> (6 samples)	0.012	< 0.001	0.020	< 0.001	0.017	< 0.001
Introduction range (15 samples)	0.015	< 0.001	0.014	< 0.001	0.011	< 0.001
European introduction range (14 samples)	0.016	< 0.001	0.015	< 0.001	0.012	< 0.001

683 (1) LBK stands for the Longboat Key population (no. 7 in Fig. 1).

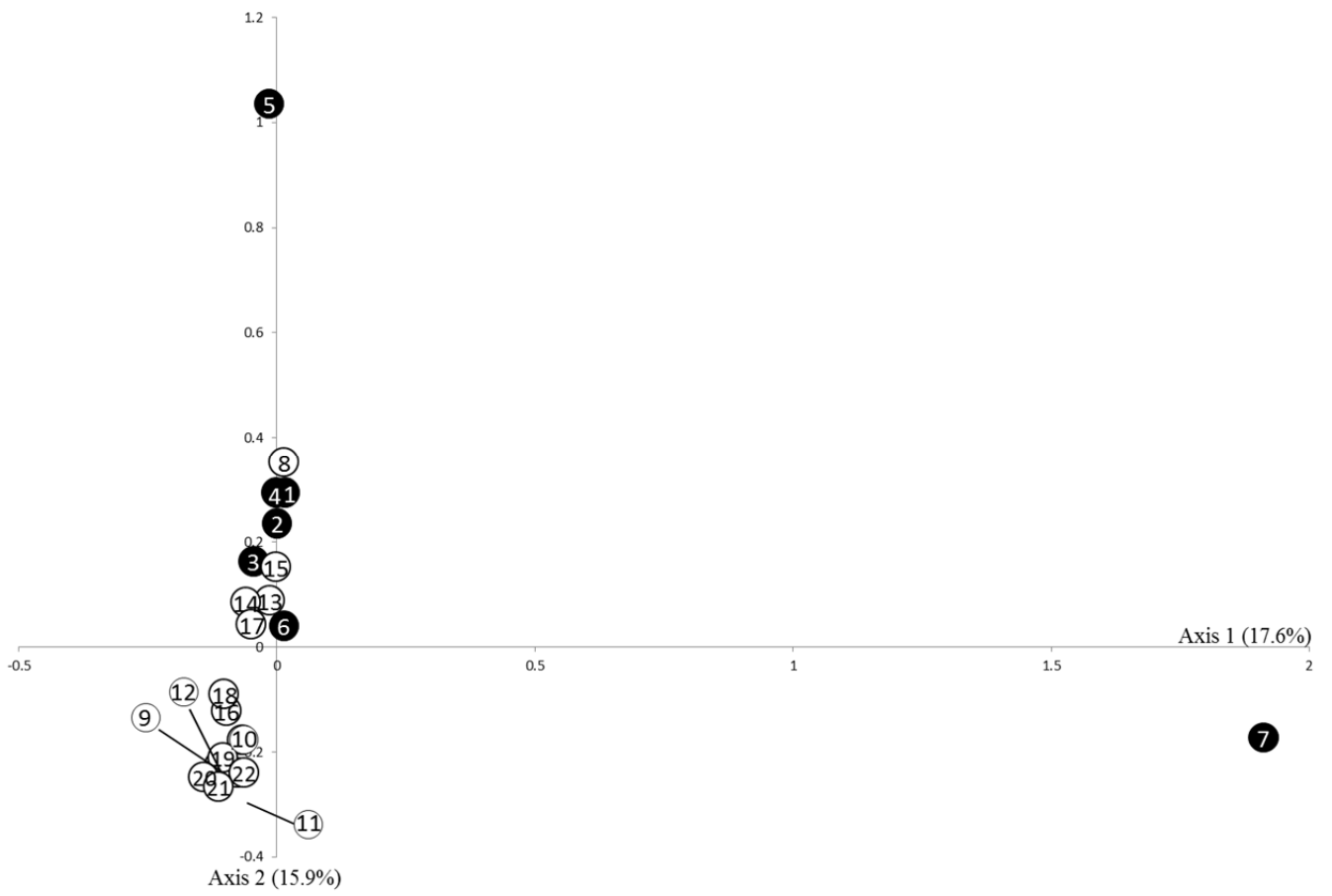
684

685

686

687





690

691

Fig. 2

**Native populations**

**Introduced populations**

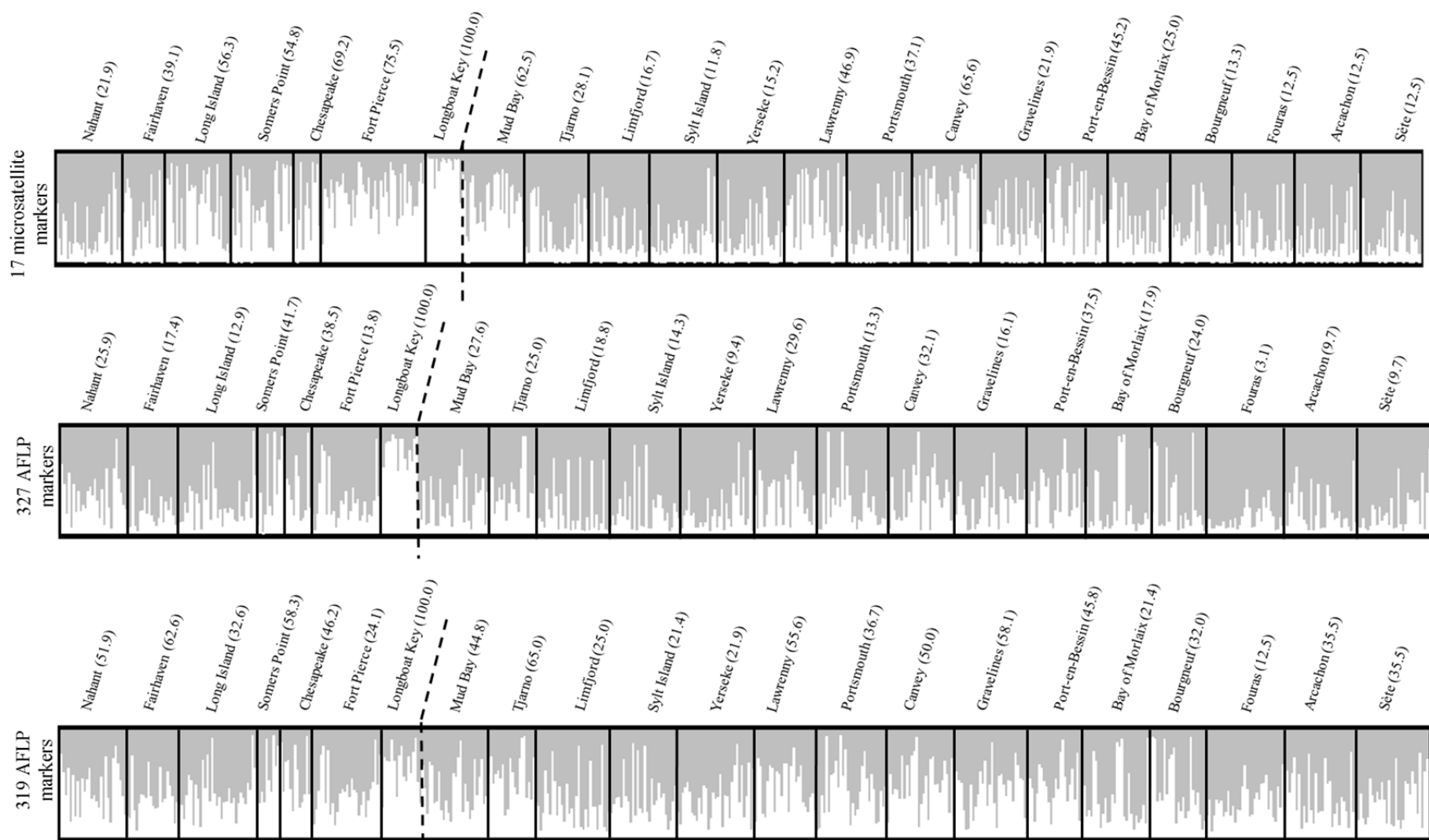


Fig. 3

692  
693





## Supporting information

0 Table S1 Pairwise population genetic differentiation (pairwise  $F_{ST}$  estimates and probability values for exact tests) A- for 17 microsatellite markers (dataset 1, see text) and B-  
 1 for AFLP markers: upper diagonal for 327 AFLPs (dataset 2; see text) and lower diagonal for 319 AFLPs (dataset 3; see text). Probability values for exact tests were  
 2 corrected for multiple comparisons: ns: non-significant ( $p>0.05$ ), \* ( $p<0.05$ ), \*\* ( $p<0.01$ ), \*\*\* ( $p<0.001$ ). Population samples are numbered from 1 to 22 (see Fig. 1 in main  
 3 text): 1: Nahant, 2: Fairhaven, 3: Long Island, 4: Somers Point, 5: Chesapeake, 6: Fort Pierce, 7: Longboat Key, 8: Mud Bay, 9: Tjarno, 10: Limfjord, 11: Sylt Island, 12:  
 4 Yerseke, 13: Lawrenny, 14:Portsmouth, 15: Canvey, 16: Gravelines, 17: Port-en-Bessin, 18: Bay of Morlaix, 19: Bourgneuf, 20: Fouras, 21: Arcachon, 22: Sète.  
 5 Comparisons among native populations are shaded in light gray, among introduced populations in dark gray and between native and introduced populations in white.

A-

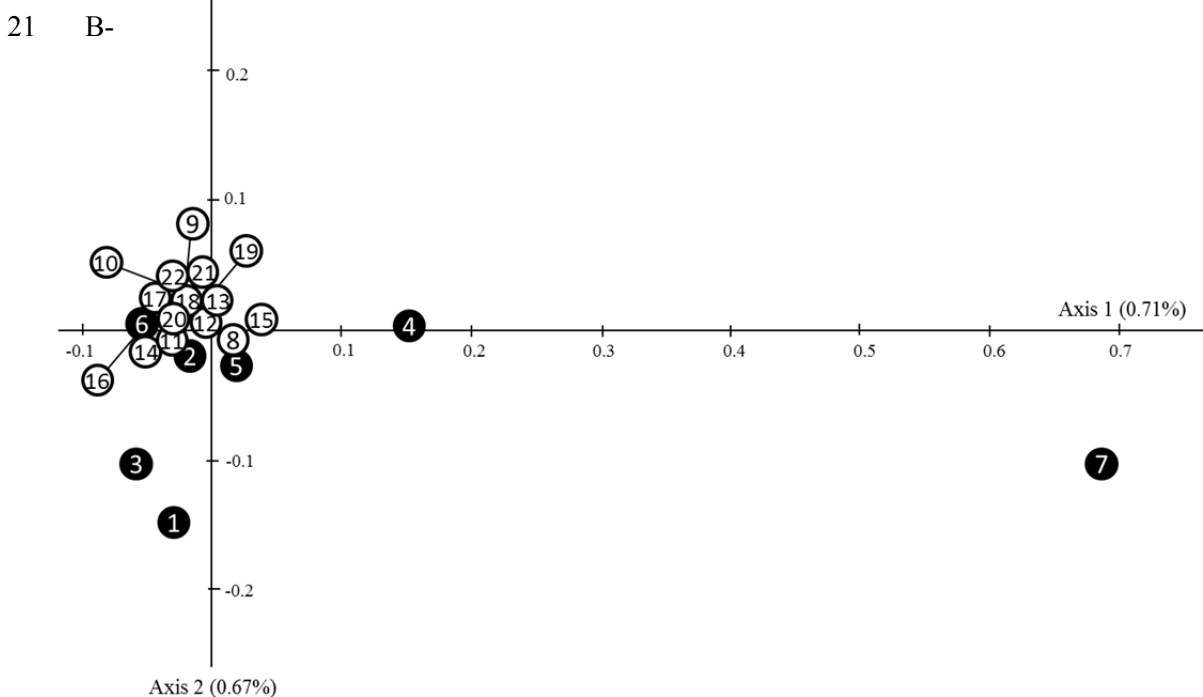
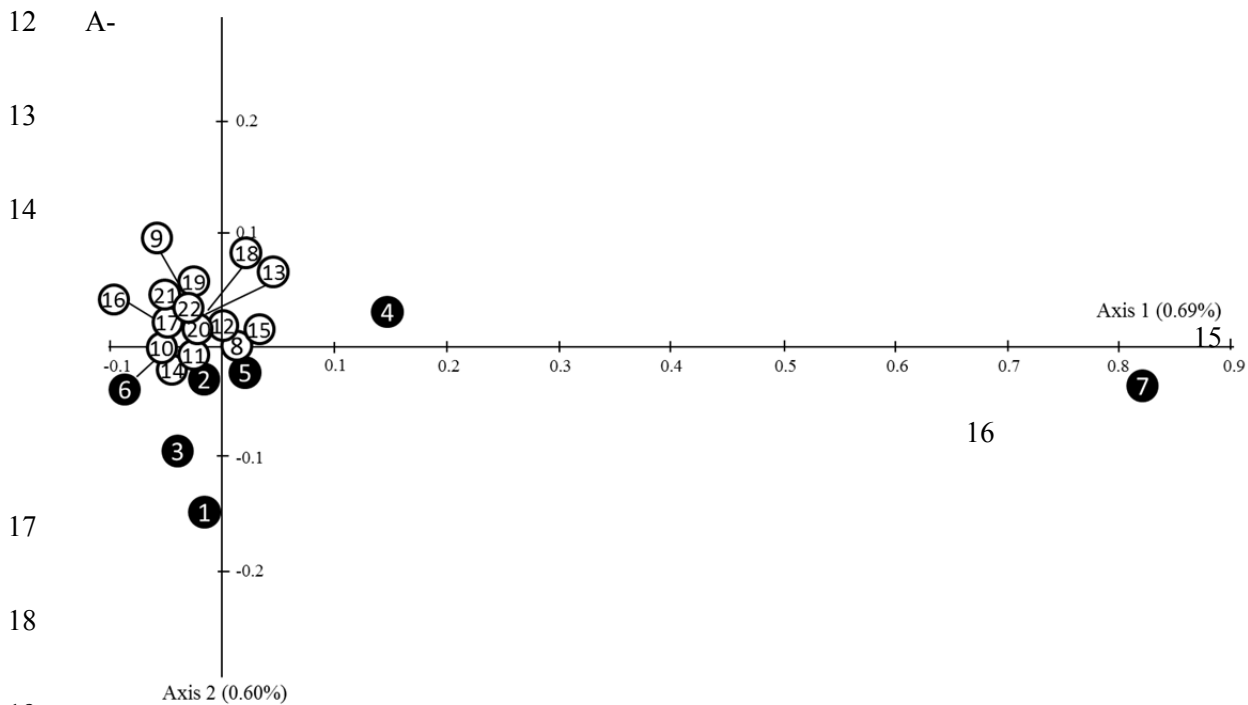
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
1	0.013***	0.014***	0.032***	0.022***	0.019***	0.083***	0.014***	0.024***	0.025***	0.028***	0.029***	0.026***	0.024***	0.017***	0.019***	0.017***	0.027***	0.028***	0.019***	0.021***	0.017***	
2		-0.005n.s.	0.006***	0.007***	0.011***	0.071***	0.001**	0.008***	0.006***	0.008***	0.014***	0.006***	0.001***	0.001***	0.000***	0.006***	0.008***	0.007***	0.002***	0.006***	0.000***	
3			-0.001n.s.	0.006***	0.011***	0.083***	-0.003*	0.004***	0.008***	0.008***	0.012***	0.005*	0.002***	0.003***	0.003n.s.	0.004*	0.001***	0.002***	0.000***	0.005***	0.001***	
4				0.008***	0.010***	0.090***	0.004***	0.006***	0.015***	0.008***	0.010***	0.004***	0.006***	0.006***	0.004***	0.010***	0.004***	0.005***	0.005***	0.009***	0.008***	
5					0.011***	0.088***	0.006***	0.014***	0.020***	0.012***	0.010***	0.003***	0.011***	0.006***	0.008***	0.009***	0.002***	0.015***	0.008***	0.010***	0.011***	
6						0.072***	0.009***	0.008***	0.019***	0.011***	0.008***	0.007***	0.013***	0.007***	0.009***	0.016***	0.010***	0.013***	0.008***	0.009***	0.007***	
7							0.081***	0.081***	0.086***	0.087***	0.090***	0.078***	0.082***	0.067***	0.084***	0.098***	0.099***	0.097***	0.091***	0.094***	0.070***	
8								0.004***	0.012***	0.010***	0.004***	0.002***	0.003***	0.005***	0.001***	0.006***	0.004***	0.007***	0.005***	0.006***	0.003***	
9									0.006***	0.004***	0.001***	0.003***	0.003***	0.006***	0.001*	0.008***	0.004***	0.004***	-0.01n.s.	0.001*	-0.002n.s.	
10										0.003***	0.006***	0.014***	0.005***	0.013***	0.005*	0.016***	0.010***	0.010***	0.002***	0.001*	0.001n.s.	
11											0.006***	0.005***	0.008***	0.006***	0.001n.s.	0.011***	0.006***	0.003***	0.003***	-0.002*	0.003***	
12												0.006***	0.005***	0.008***	0.003***	0.020***	0.004***	0.007***	0.007***	0.002***	0.004***	
13													0.003***	0.001***	0.001n.s.	0.008***	0.006***	0.010***	0.004***	0.005***	0.003***	
14														0.006***	0.000**	0.005***	0.007***	0.008***	0.004***	0.006**	0.001***	
15															0.002***	0.005***	0.008***	0.005***	0.005***	0.007***	0.003***	
16																0.002*	0.002***	-0.003n.s.	-0.003*	-0.003*	-0.004n.s.	
17																	0.008***	0.011***	0.003***	0.009***	0.008***	
18																		0.005***	0.001***	0.000***	0.003***	
19																			0.002n.s.	-0.002*	-0.001*	
20																					-0.004n.s.	
21																						-0.003n.s.

6 B-

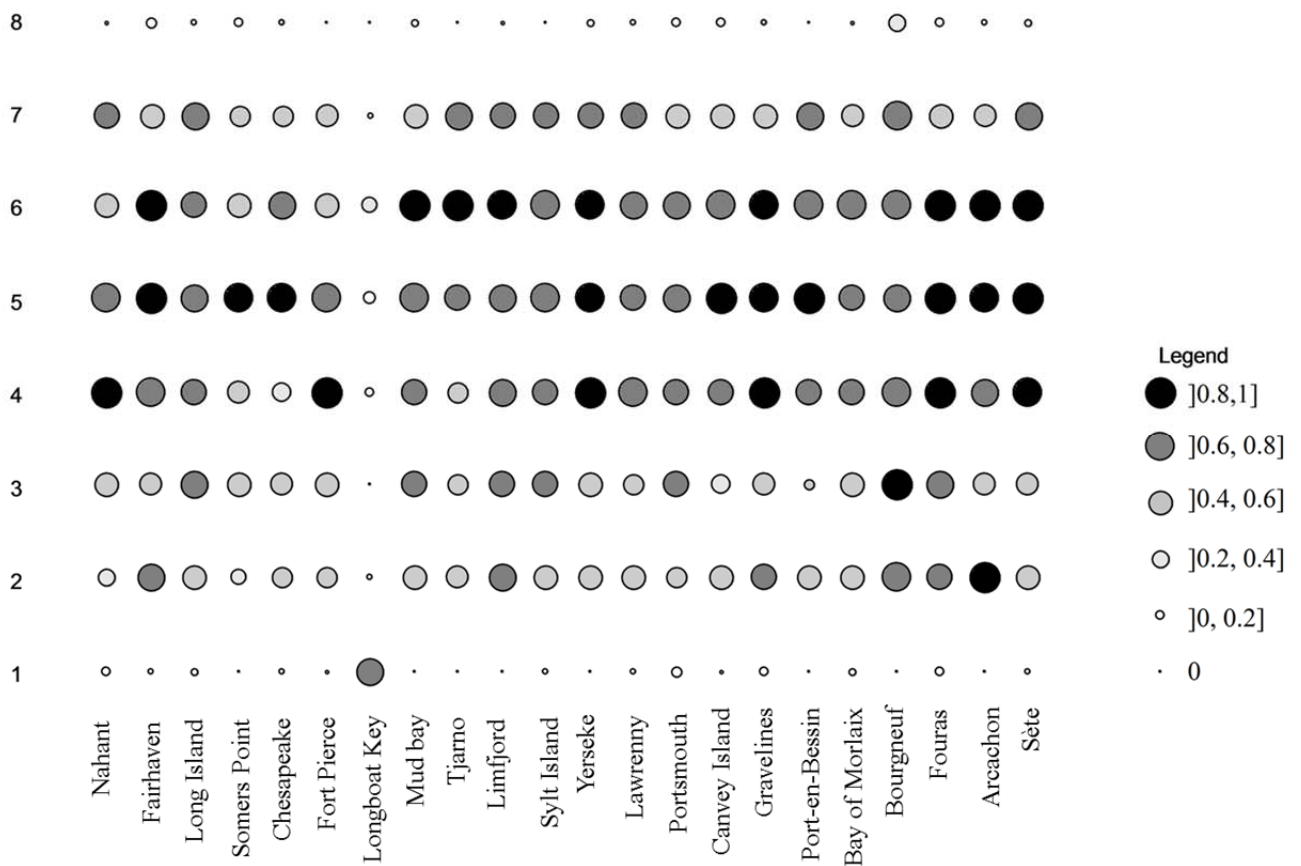
7

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1		0.016**	0.012**	0.034**	0.028***	0.022***	0.092***	0.021***	0.018**	0.025***	0.010*	0.013**	0.020***	0.017***	0.016n.s.	0.014**	0.035***	0.017*	0.028***	0.035***	0.031***	0.016**
2	0.009**		0.002n.s.	0.042***	0.026***	0.013**	0.106***	0.001n.s.	0.026***	0.005n.s.	-0.005n.s.	-0.004n.s.	0.012**	0.005n.s.	0.010n.s.	0.002n.s.	0.030***	0.002n.s.	0.021*	-0.004n.s.	0.003n.s.	-0.002n.s.
3	0.010*	0.000n.s.		0.027*	0.013*	0.007n.s.	0.087***	0.007*	0.015*	0.001**	-0.001n.s.	0.003n.s.	0.013**	0.006n.s.	0.004n.s.	0.002n.s.	0.036***	-0.0003n.s.	0.013n.s.	0.010**	0.010*	0.004n.s.
4	0.034***	0.037**	0.025*		-0.007n.s.	0.033**	0.050***	0.026**	0.023*	0.042**	0.028**	0.030**	0.019**	0.031**	0.015n.s.	0.027**	0.047***	0.032*	0.023*	0.050***	0.042***	0.038**
5	0.023***	0.023***	0.010n.s.	0.000n.s.		0.025**	0.058***	0.018**	0.005n.s.	0.036***	0.022**	0.018**	0.010n.s.	0.027**	0.008n.s.	0.018**	0.048***	0.013n.s.	0.019*	0.039***	0.023***	0.026**
6	0.025***	0.008n.s.	0.004n.s.	0.033**	0.021***		0.080***	0.022***	0.031***	0.026***	0.012*	0.008*	0.021***	0.014*	0.016**	0.008*	0.041***	0.007n.s.	0.032***	0.024***	0.016**	0.011**
7	0.037***	0.038***	0.030***	0.013***	0.018***	0.022***		0.097***	0.071***	0.110***	0.087***	0.094***	0.082***	0.091***	0.078***	0.087***	0.104***	0.084***	0.099***	0.119***	0.103***	0.100***
8	0.016***	0.001n.s.	0.006*	0.023*	0.017**	0.017***	0.034***		0.022***	0.005n.s.	0.022n.s.	0.002n.s.	0.006n.s.	0.013**	0.009n.s.	0.014**	0.046***	0.007n.s.	0.017*	0.014***	0.008*	0.009*
9	0.010*	0.022***	0.012*	0.019*	0.003n.s.	0.022***	0.020***	0.021***		0.030***	0.008n.s.	0.019***	0.012*	0.020**	0.016*	0.019**	0.034***	0.011*	0.010n.s.	0.042***	0.022***	0.027***
10	0.019***	0.005n.s.	0.010*	0.037***	0.034***	0.022***	0.046***	0.006n.s.	0.029***		0.002n.s.	0.004n.s.	0.013**	0.013**	0.014*	0.017**	0.046***	0.009n.s.	0.016*	0.008*	0.006n.s.	0.006n.s.
11	0.008*	0.000n.s.	0.000n.s.	0.026**	0.022**	0.010*	0.028***	0.004n.s.	0.007n.s.	0.003n.s.		-0.004n.s.	0.005n.s.	0.005n.s.	0.005**	0.002n.s.	0.024***	-0.001n.s.	0.010n.s.	0.008*	0.002n.s.	0.000n.s.
12	0.012***	0.000n.s.	0.000n.s.	0.024**	0.011*	0.007*	0.026***	0.001n.s.	0.014**	0.004n.s.	0.000n.s.		0.011**	0.007n.s.	0.013*	0.001n.s.	0.037***	0.003n.s.	0.015*	0.004n.s.	0.001***	-0.003n.s.
13	0.016***	0.012**	0.009*	0.017*	0.011n.s.	0.017***	0.030***	0.004n.s.	0.011*	0.011*	0.004n.s.	0.009*		0.017***	0.010*	0.012**	0.030***	0.008n.s.	0.012n.s.	0.029***	0.016**	0.012**
14	0.016**	0.004n.s.	0.004n.s.	0.028**	0.025***	0.013**	0.039***	0.012**	0.019**	0.013*	0.006n.s.	0.007*	0.017**		0.016**	0.011*	0.037***	0.006n.s.	0.016*	0.019***	0.013**	0.006n.s.
15	0.016**	0.008n.s.	0.005n.s.	0.016n.s.	0.008n.s.	0.016**	0.032***	0.009*	0.014*	0.014*	0.006n.s.	0.013**	0.009*	0.017**		0.012*	0.034***	0.009n.s.	0.013n.s.	0.016**	0.018**	0.012*
16	0.012**	0.003n.s.	0.000n.s.	0.021**	0.012*	0.007*	0.021***	0.013***	0.013*	0.017***	0.002n.s.	0.003n.s.	0.010**	0.012**	0.012*		0.030***	0.000n.s.	0.020*	0.006*	0.004n.s.	0.001n.s.
17	0.027***	0.028***	0.027***	0.041***	0.045***	0.032***	0.047***	0.041***	0.033***	0.041***	0.200***	0.033***	0.032***	0.036***	0.029***	0.026***		0.033***	0.053***	0.049***	0.043***	0.024***
18	0.014*	0.001n.s.	0.000n.s.	0.032*	0.013n.s.	0.004n.s.	0.035***	0.008*	0.010*	0.009n.s.	0.000n.s.	0.002n.s.	0.008n.s.	0.007n.s.	0.001n.s.	0.000n.s.	0.030***		0.017*	0.009*	0.000n.s.	0.002n.s.
19	0.020*	0.017n.s.	0.012n.s.	0.010n.s.	0.006n.s.	0.025***	0.023***	0.015**	0.000n.s.	0.016*	0.007n.s.	0.012n.s.	0.001n.s.	0.011n.s.	0.009n.s.	0.016*	0.040***	0.013n.s.		0.025**	0.019*	0.021*
20	0.031***	0.000n.s.	0.005n.s.	0.043***	0.033***	0.020***	0.047***	0.013***	0.034***	0.007*	0.006n.s.	0.005n.s.	0.025***	0.015***	0.014***	0.006*	0.040***	0.006n.s.	0.023*		0.004n.s.	0.001n.s.
21	0.019***	0.004n.s.	0.003n.s.	0.030***	0.016**	0.006n.s.	0.033***	0.006*	0.016***	0.006n.s.	0.000n.s.	0.000n.s.	0.013**	0.010*	0.013*	0.003n.s.	0.039***	0.000n.s.	0.014n.s.	0.001n.s.		0.000n.s.
22	0.013***	0.000n.s.	0.000n.s.	0.033**	0.022***	0.007*	0.031***	0.009*	0.025***	0.004n.s.	0.000n.s.	0.000n.s.	0.013***	0.005n.s.	0.011*	0.001n.s.	0.022***	0.000n.s.	0.016n.s.	0.000n.s.	0.000n.s.	

8 Fig. S1 Correspondence analyses conducted with A-327 AFLPs (dataset 2; i.e. original AFLP dataset; see main  
 9 text) and B- 319 AFLPs (dataset 3; see main text; i.e. without eight outliers detected in the original AFLP  
 10 dataset). Only the barycenter of each population is presented here. Native populations (black dots) are numbered  
 11 from 1 to 7 and introduced populations (white dots) are numbered from 8 to 22 according to Fig. 1.



22 Fig. S2 AFLP allele frequencies in each sampled population of *Crepidula fornicata* (x-axis) for the eight outlier  
 23 loci (numbered from 1 to 8 – see Fig. 4 - along the y-axis). For each bi-allelic locus, the size and color of the dot  
 24 varies with the class of allele frequencies.



---

**Deuxième partie :**

**Dérive, Dispersion et Invasion**

---



Parmi les nombreuses caractéristiques liées au succès d'introduction d'une espèce non-indigène, ses capacités de dispersion et son cycle de vie sont considérés comme des facteurs majeurs car participant à l'expansion de l'espèce depuis son (ses) point(s) d'introduction primaire. Les espèces à cycle benthopélagique ont de ce point de vue, des caractéristiques remarquables du fait de la présence d'une phase larvaire pélagique. Chez de nombreuses espèces marines, dont la crépidule, la phase larvaire est généralement le vecteur majeur, voire unique, de la dispersion naturelle. La durée de vie de cette phase dispersive est très variable d'une espèce à l'autre (voir les exemples donnés dans l'encadré 3 tiré de Shanks 2009), mais également au sein d'une même espèce. En combinant observations de terrain, expérimentations et modélisations, Rigal (2009) a montré au cours de sa thèse que la durée de vie larvaire pélagique chez *C. fornicata* varie de 2 à 7 semaines.

Après avoir étudié au cours du chapitre précédent les forces évolutives agissant lors de la dispersion assistée par l'homme (sélection post-introduction dans un nouvel environnement), nous nous attacherons ici à étudier plus en détail les processus de dispersion naturelle dans l'aire d'introduction. La dispersion a déjà fait l'objet d'une thèse dans le laboratoire d'accueil (L. Dupont 2004): l'objectif était alors de s'intéresser aux flux de gènes et à la relation entre dispersion potentielle et dispersion « réalisée » notamment par le couplage entre l'utilisation de modèles couplés biologie-physique et d'études de génétique des populations. Ce travail se situait donc principalement à une échelle régionale et abordant des processus spatiaux. Ces études avaient en particulier montré que malgré des capacités de dispersion importantes liées à une durée de vie larvaire de plusieurs semaines, il existait des limites aux flux de gènes, par exemple entre baies adjacentes. Par ailleurs, cette thèse avait mis en évidence des structures à micro-échelles difficilement explicables chez la crépidule compte tenu des grandes tailles de populations et fortes capacités de dispersion. La corrélation entre la durée de vie larvaire et la distance de dispersion est difficile à prédire que ce soit au niveau interspécifique (Shanks 2009 ; Encadré 3) ou intra-spécifique (*e.g.* chez la crépidule, de 2 à 7 semaines, Rigal 2009) du fait des caractéristiques environnementales mais également pour des raisons méthodologiques. En particulier, les courants de marée peuvent dans certaines zones, à l'instar de la zone étudiée, conduire à des rétentions importantes des larves à proximité des populations adultes parentales. De plus, il existe des limites méthodologiques, aussi bien avec les approches de génétique des populations (structure génétique difficile à détecter chez ces espèces fortement dispersives) qu'avec l'utilisation de modèles bio-physiques de dispersion (*e.g.* paramétrage difficile du fait de la complexité des comportements larvaires).

Au cours de cette thèse, nous nous sommes intéressés à un autre aspect de la relation entre micro-évolution et cycle de vie en étudiant principalement les effets de la dérive génétique au cours du cycle de vie de la crépidule, en se plaçant à une échelle locale (*i.e.* un site d'étude au sein d'une baie que nous appellerons par la suite une « population ») et en privilégiant des études temporelles. En effet, la fécondité et la taille des populations (limitant les risques d'extinction suite à la fondation de

**Encadré 3- Durée de la phase larvaire : *Estimates of propagule duration and dispersal distance (Shanks 2009)***  
Distances de dispersion estimées par des données expérimentales (E), par l'expansion des espèces invasives (I), par l'observation de dispersion larvaire (O) et par le déplacement de larves taguées (T).

Organism (data source)	Dispersing stage	Propagule duration	Dispersal distance	References
<b>Algae</b>				
<i>Sargassum</i> spp. (E)	spores	25 days	1 m	Deysher and Norton (1982); Kendrick and Walker (1995)
<i>Grateloupia doryphora</i> (I)	spores		10 km	Villalard-Bohnsack and Harlin (2001)
<i>Sarcothalia crispate</i> (E)	spores		900 m	Opazo and Otaiza (2007)
Crustose corallines (E)	spores		900 m	Opazo and Otaiza (2007)
<i>Fucus serratus</i> (I)	spores		4 m	Arrontes (2002)
<i>Macrocystis pyrifera</i> (E)	spores	32 h	200 m	Reed <i>et al.</i> (2006)
<b>Cnidaria</b>				
<i>Allopora californica</i> (E)	demersal planulae	minutes	0.5 m	Fritchman (1974); Ostarello (1976)
<i>Athenaria actinia</i> (O)	egg	minutes	1 m	Nyholm (1940)
<i>Alcyonium siderium</i> (E)	demersal planulae	4.5 days	3 m	Sebens (1983)
<i>Favia fragum</i> (E)	planulae	4 min	3 m	Carlson and Olson (1993)
<i>Heliopora coerulea</i> (E)	planulae	6 h	350 m	Harii <i>et al.</i> (2002); Harii and Kayanne (2003)
<i>Tubularia larynx</i> (O)	actinula	<5 min pelagic, crawling 1 day	centimeters to meters	Pyefinch and Downing (1949)
<b>Sponges</b>				
<i>Scopolina lophyropoda</i> (E)	larvae	5–10 min	0.5–2 m	Uriz <i>et al.</i> (1998)
<i>Halichondria moorei</i> (O)	larvae	60 h	2 m	Bergquist and Sinclair (1968)
<i>Stylopsis</i> spp. (O)	crawling larvae	1–4 days	meters	Ayling (1980)
<i>Chondropsis</i> spp. (O)	crawling larvae	1–4 days	meters	Ayling (1980)
<b>Tunicates</b>				
<i>Molgula pacifica</i> (O)	adhesive benthic egg	36 h	<1 m	Young <i>et al.</i> (1988)
<i>Trididemnum solidum</i> (O)	tadpole larvae	15 min	<200 m	van Duyl <i>et al.</i> (1981)
<i>Ecteinascidia turbinata</i> (O)	tadpole larvae	81.5 s	6.87 m	Bingham and Young (1991)
<b>Bryozoans</b>				
<i>Crisia denticulate</i> (T)	larvae	<2 h	10–20 m	Pemberton <i>et al.</i> (2007)
<i>Tubulipora tuba</i> (E)	larvae	<2 h	1 m	Knowlton and Jackson (1993)
<b>Molluscs</b>				
<i>Crepidula fornicata</i> (I)	larvae	2–3 weeks	28.5 km	Coe (1949); Dommasnes and Scham (1973)
<i>Rapana venos</i> (I)	larvae	14–17 days	100 km	Zolotarev (1996); Harding and Roger (1999)
<i>Philine</i> spp. (I)	larvae	30–40 days	260 km	Hansen and Ockelmann (1991); Cadien and Ranasinghe (2003)
<i>Dendropoma corallinaceum</i> (O)	crawling larvae	1–5 days	<1 m	Hughes (1978)
<i>Venerupis philippinarum</i> (I)	larvae	3–4 weeks	20 km	Bourne (1982); Quayle (1964); Strathmann (1987)
<i>Musculista senhousia</i> (I)	larvae	16 days	116 km	Willan (1987); Semenikhina <i>et al.</i> (2008)
<i>Perna viridis</i> (I)	larvae	2–3 weeks	33 km	Benson <i>et al.</i> (2001); Fajans and Baker (2005)
<i>Mya arenaria</i> (I)	larvae	10–35 days	35 km	Strathmann (1987); Zolotarev (1996)
<i>Drupella cornus</i> (T)	larvae	3 weeks	<100 m	Johnson <i>et al.</i> (1993)
<i>Adalaria proxima</i> (E)	larvae	1–3 days	100 m	Lambert <i>et al.</i> (2003)
<b>Polychaetes</b>				
<i>Spirorbis spirillum</i> (E)	larvae	15 min	<500 m	Dirnberger (1993)
<i>Terebrasabella heterouncinata</i> (I)	crawling larvae	12 h	15 m	Kuris and Culver (1999); Culver and Kuris (2000)
<i>Marenzelleria viridis</i> (I)	larvae	8 weeks	55–64 km	Boudouresque (1994); Bochert (1997)
<b>Crustaceans</b>				
<i>Paralithodes camtschaticus</i> (I)	larvae	90 days	32 km	Kuzmin <i>et al.</i> (1996); Epelbaum <i>et al.</i> (2006)
<i>Balanus glandula</i> (I)	larvae	2–4 weeks	85 km	Schwindt (2007)
<i>Cancer magister</i> (T)	larvae	3–4 months	500 km	Shanks (unpubl. data)
<b>Echinoderms</b>				
<i>Asterias amurensis</i> (I)	larvae	40–50 days	22.5 km	Buttermore <i>et al.</i> (1994); Nozais <i>et al.</i> (1997)
<i>Acanthaster planci</i> (I)	larvae	14 days	49–81 km	Moran <i>et al.</i> (1992)
<b>Fishes</b>				
<i>Pterois volitans</i> (I)	larvae	25–40 days	48.5 km	Ruiz-Carus <i>et al.</i> (2006)
<i>Sebastes melanops</i> (T)	larvae	83–174 days	<120 km	Miller and Shanks (2004)
<i>Amphiprion percula</i> (T)	larvae	10–13 days	100–500 m	Almany <i>et al.</i> (2007)
<i>Chaetodon vagabundus</i> (T)	larvae	29–48 days	100–500 m	Almany <i>et al.</i> (2007)
<i>Pomacentrus amboinensis</i> (T)	larvae	9–12 days	100–200 m	Jones <i>et al.</i> (1999)



nouvelles populations et réduisant théoriquement les effets de la dérive génétique) peuvent aussi « favoriser » le succès d'une introduction biologique. La fécondité de la crépidule est très élevée mais la mortalité est également probablement très importante dans les premiers stades du cycle de vie, comme chez tous les invertébrés marins (Rumrill 1990) avec de possibles effets de dérive génétique s'exprimant au cours du développement larvaire (mortalité pré-recrutement) et des premiers stades de la vie benthique (mortalité post-recrutement).

Au cours de cette partie, je dresserai d'abord un rapide bilan des connaissances acquises sur la dynamique démographique et sexuelle de la population étudiée (Chapitre 1). Par la suite, je présenterai sous forme d'articles, trois études (Chapitre 2 à 4) réalisées afin d'étudier différents aspects relatifs aux relations entre dérive génétique et cycle benthopélagique chez *Crepidula fornicata*.



## Chapitre 1 - Dynamique démographique, sexuelle et reproductive d'une population de *Crepidula fornicata* en baie de Morlaix

---

*Remarque préliminaire : ce chapitre dresse l'état des connaissances concernant la dynamique démographique et sexuelle d'une population de *Crepidula fornicata* située en baie de Morlaix, population que j'ai étudiée au cours de ma thèse. Les données présentées ont été obtenues (et en grande partie analysées) au cours de thèses précédemment réalisées dans l'équipe (Le Cam (2009), Rigal (2009), Leroy (2011)) auxquelles s'ajoutent des données que j'ai acquises au cours ma thèse. Ces connaissances reposent sur :*

*(1) un suivi benthique (i.e. individus juvéniles et matures se développant au fond) qui a été initié en septembre 2004 au cours des travaux de thèse de Sabrina Le Cam (2009) et poursuivi par Adélaïde Sibeaux au cours d'un travail de stage. J'ai poursuivi ce suivi en 2010 et bénéficié pour le réaliser de l'aide d'Emeline Fonteneau, que j'ai encadré lors de son stage de 2ème année d'IUT en Génie Biologique, ainsi que des données récoltées lors de Travaux Pratiques, que j'ai supervisés en 2010 et 2011, par des étudiants de l'unité d'enseignement « Dynamique et génétique des populations » du master d'Océanographie et Environnements Marins (UPMC).*

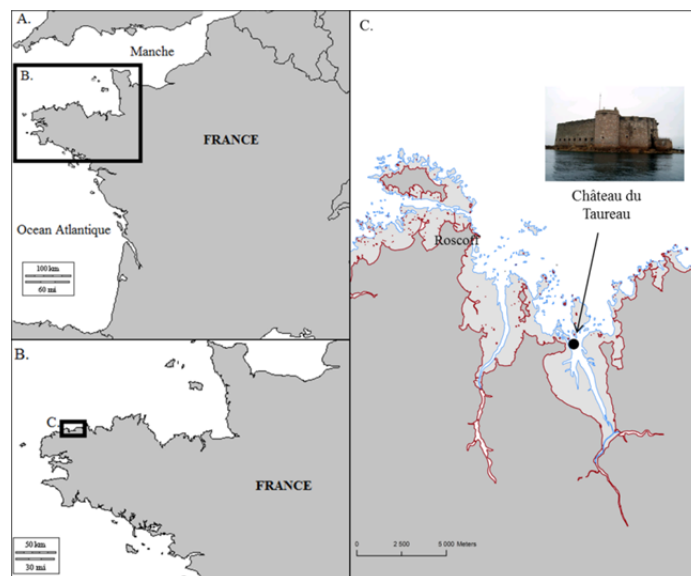
*(2) un suivi de la population pélagique qui a été initié en janvier 2005 au cours de travaux de thèse de François Rigal (2009), et poursuivi au cours de la thèse de Fanny Leroy (2011).*

*Les données que j'ai obtenues seront intégrées dans un article sur l'évolution temporelle de la structure sexuelle en fonction de la dynamique de recrutement en préparation par Sabrina Le Cam et coll. Il m'a paru important de synthétiser les résultats majeurs issus de ces deux suivis car ils sont non seulement à l'origine d'échantillons que j'ai utilisés, mais également pour partie des questions auxquelles j'ai tenté de répondre au cours de ma thèse.*

### **I. *Crepidula fornicata* en baie de Morlaix : une installation certaine mais une prolifération modérée**

La baie de Morlaix se situe dans le Finistère Nord et occupe une superficie de 5300 ha, structurée par les estuaires de La Penzé et de Morlaix (Fig. 1). Cette baie est le lieu d'activité conchylicole où sont cultivées huîtres et moules depuis plus d'une soixantaine d'années, activité qui a probablement facilité l'introduction de *Crepidula fornicata* dans la baie (Marteil 1963, Blanchard 1995). Signalée pour la première fois dans les années 60 (Marteil 1963, Blanchard 1995), la présence de ce gastéropode en baie de Morlaix semble anecdotique jusqu'aux années 90. Des populations plus importantes sont ensuite observées sans toutefois atteindre les niveaux de densité répertoriées en rade de Brest, en baie de Saint-Brieuc ou encore en baie du Mont-Saint-Michel où les stocks de *C. fornicata* sont estimés respectivement à 210 000, 250 000 et 160 000 tonnes (Blanchard et Erhold 1999, Hamon 1996). Dans les baies où l'espèce présente un caractère « invasif », les densités sont très élevées, par exemple entre 1300 et 4770 ind.m<sup>-2</sup> en baie de Marennes-Oléron (de Montaudouin *et al.* 1999) ou entre 200 et 2000 ind.m<sup>-2</sup> en baie du Mont-St-Michel (Viard *et al.* 2006). Par comparaison, la population benthique de *C. fornicata* en baie de Morlaix est caractérisée par une répartition très

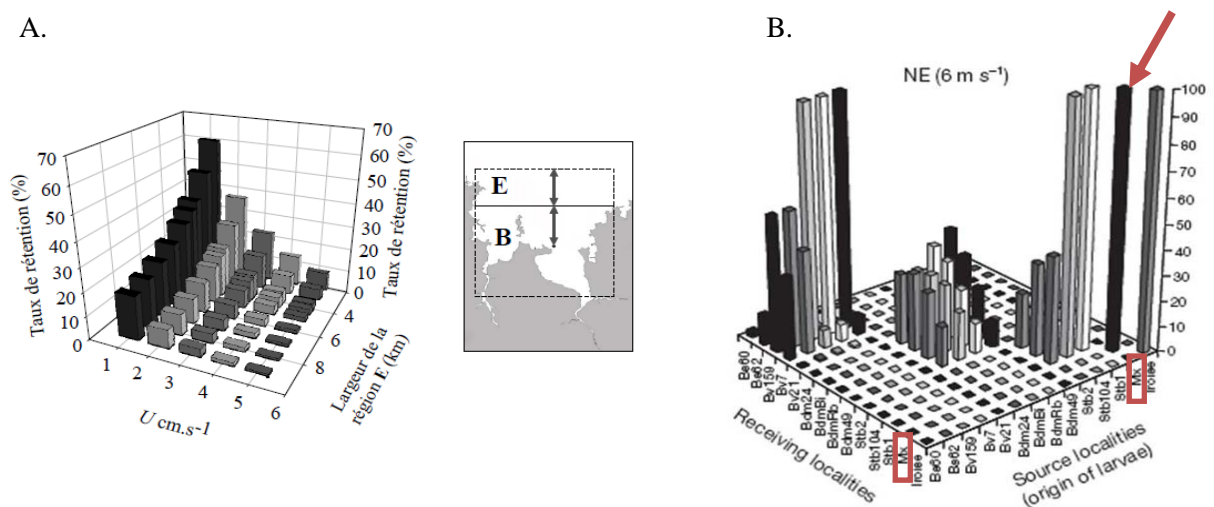
hétérogène et peu dense des individus (Rigal *et al.* 2010). Un site seulement, situé aux environs du Château du Taureau (Fig. 1) montre une population bien établie caractérisée par des densités de l'ordre de 20 à 200 ind.m<sup>-2</sup> (estimation basées sur des comptages réalisés en plongée le long de transects ; Dupont 2004). Par ailleurs, les densités locales de larves de *C. fornicata* sont également faibles (20 - 80 ind m<sup>-3</sup>, Rigal 2009, Rigal *et al.* 2010) comparativement à d'autres baies (*e.g.* 2000 larves. m<sup>-3</sup> en baie du Mont-Saint-Michel, Lasbleiz 2003). Elles sont variables selon les points de la baie : les plus fortes abondances larvaires sont observées dans la proximité immédiate de la population du Château du Taureau. C'est en ce point que les points d'échantillonnage benthique et pélagique ont été réalisés depuis 2004.



**Fig. 1 Localisation du site (« Château du Taureau ») du suivi benthopélagique réalisé en baie de Morlaix (Bretagne Nord) - tirée de Rigal (2009)**

La baie de Morlaix présente la particularité d'abriter une population stable mais caractérisée par une prolifération limitée de crépidules par rapport à d'autres baies. Ce n'est pas un cas unique ; en baie d'Arcachon, de Montaudouin *et al.* (2001) font des observations similaires, ce qu'ils expliquent par une diversité de traits hydro-sédimentaires et la présence d'herbiers de *Zostera*. En baie de Morlaix, l'hypothèse la plus probable pour expliquer cette moindre prolifération de crépidules est liée aux caractéristiques hydro-dynamiques de la baie qui favoriseraient l'export de larves tout en limitant l'apport de larves depuis les baies voisines. En effet, cette baie est caractérisée par d'importantes amplitudes de marées (*i.e.* baie mégatidale) et présente une forme en « V » (*i.e.* elle est graduellement plus profonde du fond vers le large de la baie). Ce type de configuration favoriserait l'export de particules (Dame et Allen 1996). A l'aide d'un modèle analytique de dispersion larvaire, Rigal *et al.* (2010) montrent en effet que les forts courants de marées, dominants dans cette baie, favorisent l'exportation de larves de *Crepidula fornicata* hors de la baie ainsi qu'un faible taux de rétention des

larves au sein de la baie, non compensés par des apports extérieurs en larves (Fig. 2A). Ce modèle à l'échelle de la baie donne des résultats en accord avec ceux d'une étude antérieure menée à une échelle régionale (Dupont *et al.* 2007b). Dans cette dernière étude, un modèle couplé biologie-physique lagrangien simulant la dispersion des larves depuis leur point d'émission avait montré que la totalité des larves présentes en baie de Morlaix ne pouvait pas provenir des baies voisines (baies de Brest et de St-Brieuc ; Fig. 2B). Ces différents résultats suggèrent que le renouvellement de la population est majoritairement sous la dépendance de la reproduction locale et que la population suivie est relativement « fermée à l'immigration » depuis des populations localisées dans des baies voisines.



**Fig. 2 Résultats des modèles de dispersion larvaire de *Crepidula fornicata* en Baie de Morlaix et Manche/Iroise développés respectivement par A- Rigal *et al.* (2010) et B- Dupont *et al.* (2007b).**

**A-**Résultat du modèle analytique présenté par Rigal *et al.* (2010) et construit pour tester l'influence de la marée sur le transport larvaire de *C. fornicata* en baie de Morlaix. Le taux de rétention au sein de la baie est inféré pour différentes valeurs de courant résiduel ( $U$ ). La zone B représente les points d'émission des larves et la région E représente la zone externe de la baie où les larves sont exportées depuis B par la marée.

**B-** Résultat du modèle bio-physique lagrangien (chaque larve est une particule dont la dispersion est modélisée) présenté par Dupont *et al.* (2007b) et construit pour estimer la dispersion larvaire de *C. fornicata* après 21 jours entre différents baies localisées en Manche et Mer d'Iroise. La baie de Morlaix (Mlx) est encadrée en rouge et le pourcentage de larves reçues (immigrants ou produites localement) pointée par la flèche rouge. Le modèle intègre des forçages par le vent sous deux conditions de vents dominants dans la région étudiée. La figure présente le cas avec de vents de Nord-Est. Des résultats similaires concernant la baie de Morlaix ont été obtenus avec des vents de sud-ouest ainsi que sans vent. Dans tous les cas les larves présentes en baie de Morlaix sont issues de la baie et non pas des autres utilisées comme sources potentielles, notamment les baies voisines de St-Brieuc et de la Rade de Brest. La situation est très différente dans d'autres baies (par exemple dans la baie de St-Brieuc qui reçoit des larves depuis la baie du Mont-St-Michel).

## II. Présentation des méthodes d'acquisition des données des suivis benthiques et pélagiques

### A. Etude de la population benthique

Une trentaine de chaînes sont prélevées mensuellement à partir d'un point d'échantillonnage fixe, situé à proximité du Château du Taureau où se trouvent la population de crépidules la plus importante de la baie (coordonnées GPS : 48 °40'18''N, 3°53'167''Fig. 1). L'échantillonnage est réalisé le long d'un transect de 50m dans la direction ouest, par les plongeurs de la Station Biologique de Roscoff, afin de prévenir l'endommagement de chaîne et/ou la perte d'individus. Chaque chaîne est ensuite enveloppée dans une compresse de gaze, ce qui permet de préserver sa structure intacte. Elle est conservée dans de l'éthanol 96% en vue d'analyses morphologiques et génétiques ultérieures.

Pour chaque échantillon mensuel, un minimum de 300 individus et un minimum de 30 chaînes sont analysés en déterminant :

- le nombre d'individus par chaîne et la position de chaque individu au sein de la chaîne ;
- le stade sexuel de l'individu (juvénile, mâle, individu en transition sexuelle ou femelle); pour les femelles, la présence d'une ponte, *i.e.* de capsules ovigères dans la cavité palléale de la femelle, est également noté.
- la taille de l'individu (*i.e.* la longueur droite), mesurée à l'aide d'un pied à coulisse (précision de 0,1mm).

Ces différents paramètres sont détaillés dans l'encadré 4. La tête de chaque individu, ainsi que les pontes, sont numérotées (cf. encadré 4) et préservées individuellement dans un tube eppendorf contenant de l'éthanol à 96%, en vue d'analyses génétiques ultérieures. La coquille de chaque individu est également numérotée (même numéro que les têtes préservées dans l'éthanol) et conservée. Les individus de septembre 2004 à décembre 2007 ont été analysés par Sabrina Le Cam (2009), ceux de janvier 2008 à juin 2009 par Adélaïde Sibeaux, et j'ai étudié les suivants jusqu'en décembre 2011.

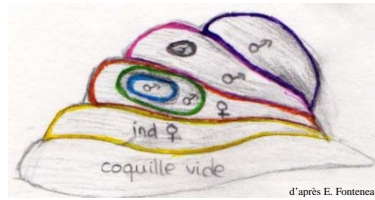
Afin de caractériser la structure démographique de la population, une décomposition modale des distributions des fréquences de taille a été réalisée pour chaque mois d'échantillonnage. Dans un premier temps, un histogramme de fréquences de taille est réalisé pour chaque échantillon en utilisant un intervalle de classe de 4mm. Chaque histogramme est ensuite lissé à l'aide d'une moyenne mobile d'ordre 3 (Frontier et Pichod-Viale 1991) afin de tenir compte du biais de mesure aux bornes de classe. La normalité de la distribution est soumise à un test de Kolmogorov-Smirnov à l'aide du logiciel Minitab 1.3.2, et le rejet de l'hypothèse nulle ( $H_0$  : « la distribution de fréquences de taille suit une loi normale ») justifie alors une décomposition modale. Sous l'hypothèse d'une corrélation positive entre la taille et l'âge, les différents « pics » observés sur les histogrammes de fréquences de taille devraient représenter des groupes d'individus issus d'une même cohorte (*i.e.* groupe d'individus

## Encadré 4 - numérotation des individus et analyses morphométriques

Pour chaque chaîne de crépidules :

1. Les individus sont numérotés selon leur position dans la chaîne :

A chaque individu est attribué un numéro et un code-position permettant de définir avec exactitude la position de l'individu dans la chaîne. Par exemple, les individus seront notés de la manière suivante pour la chaîne schématisée ci-dessous :



- 1. B1-1
- 2. B1-2
- 3. B1-3
- 4. B1-4
- 5. B2-1 <2
- 6. B2-2

7. B3-1 <3

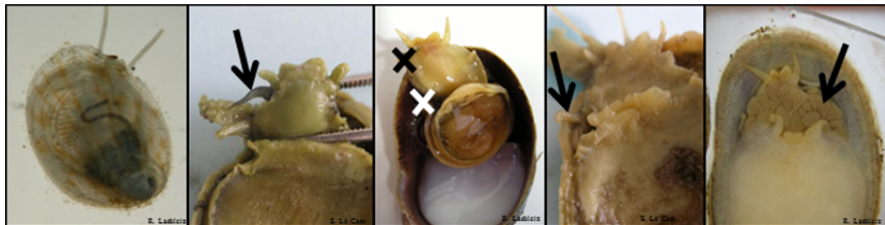
Fig. 1 : Représentation schématique d'une chaîne de crépidules et la numérotation des individus

Le numéro de l'individu est indiqué par le premier chiffre en gras et de la même couleur que sur la représentation schématique de chaîne de crépidule. Le premier individu se situe à la base de la chaîne, le deuxième individu au-dessus du premier individu, et ainsi de suite jusqu'au dernier individu situé au sommet de la chaîne principale. Les individus des chaînes satellites sont ensuite numérotés, en commençant par le premier individu situé à la base de la chaîne satellite.

Ce chiffre est suivi du code position de chaîne-individu dans la chaîne. Les individus de la chaîne principale sont notés B1- X, avec l'ordre d'empilement indiqué par le X. Les individus situés sur une chaîne satellite sont notés de la même manière mais avec le code de chaîne B2-X, B3-X... suivi du symbole < (« sur ») et le numéro de l'individu sur lequel il est fixé.

2. Le stade sexuel de chaque individu est déterminé

Les critères utilisés pour déterminer le stade sexuel de chaque individu sont les suivants :



Juvénile:  
pas de caractère sexuel visible, généralement inférieur à 10mm

Mâle:  
présence d'un pénis

Individu en transition sexuelle:  
absence de pénis et de papille génitale

Femelle:  
présence d'une papille génitale

Femelle en ponte:  
présence de sacs ovigères

3. La coquille de chaque individu est mesurée

A l'aide d'un pied à coulisse, chaque coquille est mesurée le long du plus grand axe, entre l'apex et l'avant de la coquille.



Fig. 2 Mesure d'une coquille de crépidule

issus du même évènement de sédentarisation au sein de la population), pour lesquels il est possible d'ajuster une distribution gaussienne caractérisée par sa moyenne et son écart-type. L'ajustement par une courbe polymodale, résultant des différentes gaussiennes identifiées, est réalisé grâce au package MIXDIST sous le logiciel R (<http://www.r-project.org/>).

Afin de caractériser la structure sexuelle de la population, pour chaque mois, le sex-ratio, défini par le nombre de femelles divisé par le nombre de femelles et de mâles, est calculé. Un test binomial est ensuite effectué sous R afin de tester s'il existe un écart significatif de la valeur observée du sex-ratio à la valeur théoriquement attendue (sex-ratio théorique de 0.4 car biaisé en faveur du premier sexe chez les organismes protandres ; Allsop et West 2004). Ces analyses ont été effectuées par Sabrina Le Cam.

## B. Analyse de la population pélagique

Le suivi pélagique est réalisé de manière bimensuelle à proximité du Château du Taureau (coordonnées GPS : 48°39'448''N, 3°52'343''W ; Fig. 1). Ces échantillonnages ont été réalisés à l'étable de pleine mer de morte-eau avec l'aide de l'équipage des navires Mysis, puis Néomysis de la Station Biologique de Roscoff. Les prélèvements sont effectués au filet à plancton de type WP2 d'une maille de 200 $\mu$ m sous forme de traits verticaux fond-surface. Le volume d'eau filtré à travers le filet à plancton est estimé à l'aide d'un débitmètre fixé au filet à plancton. Des profils de température et de salinité sont également effectués à l'aide d'une sonde CTD (seabird SBE19+). Les échantillons récoltés sont fixés à l'éthanol (96%) avant d'être triés.



**Fig. 3 Larves de *Crepidula fornicata***

La larve de crépidule est identifiée sans ambiguïté par rapport aux autres larves de gastéropodes par quelques critères morphologiques caractéristiques (*e.g.* présence d'une ouverture droite de la coquille, comme indiqué par la flèche) (Werner 1955, Thiriot-Quiévreux et Scheltema 1982, Lima et Lutz 1990).



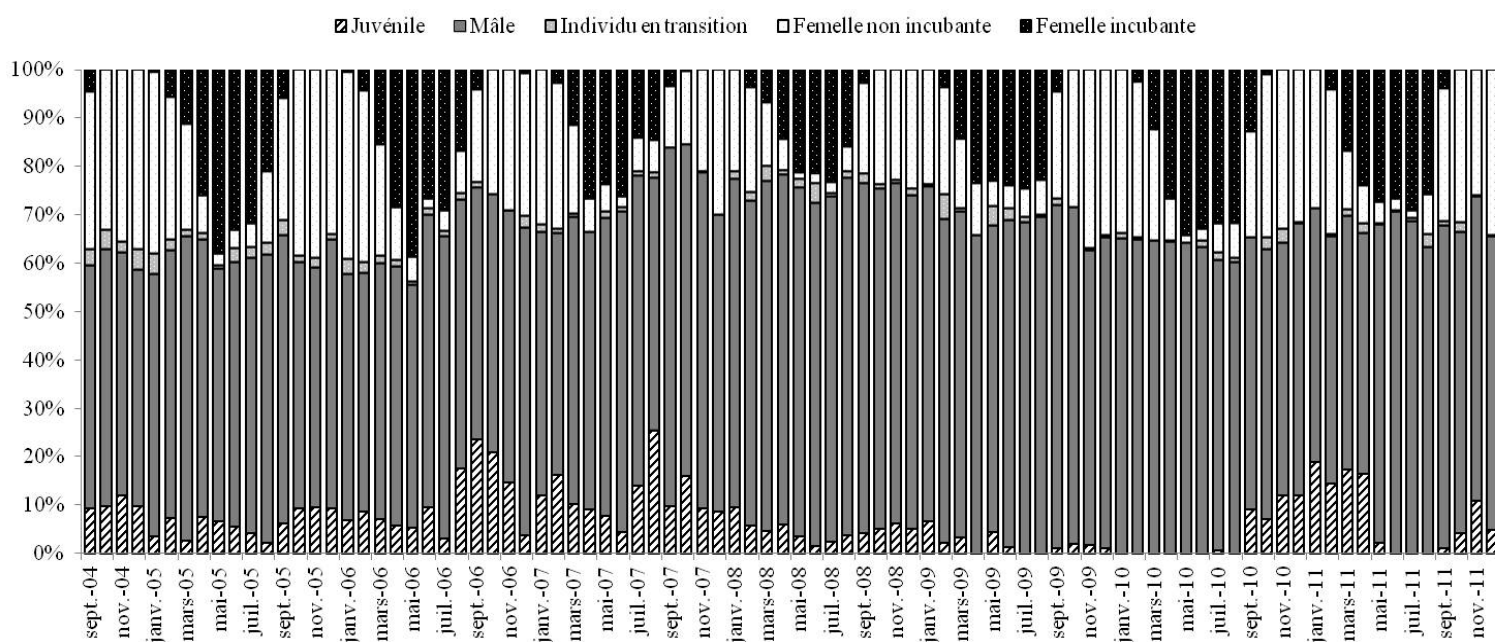
Les larves de *C. fornicata* (Fig. 3) sont identifiées, comptées, prélevées à l'aide d'une pipette pasteur et mesurées à l'aide d'un oculaire micrométrique. L'abondance larvaire (nombre de larves par mètre-cube) est ensuite estimée. Les échantillonnages, tri des larves et analyses d'abondance ont été réalisés par F. Rigal (2005-2007) et F. Leroy (2008-2010) à l'exception des mesures de taille des larves utilisées (cf. Chapitre 4) que j'ai réalisées au cours de cette thèse.

### III. Synthèse des résultats

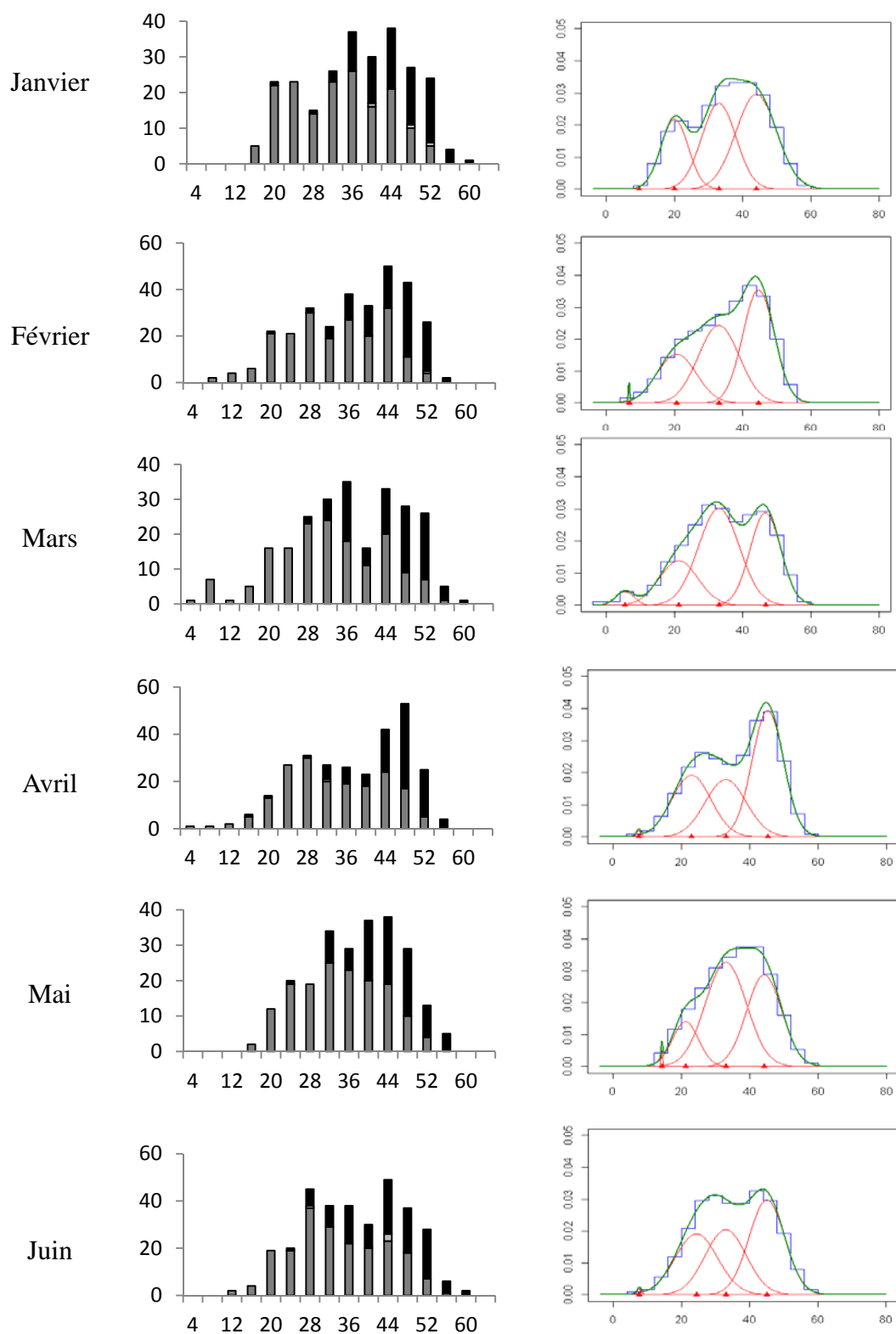
#### A. Description de la structure d'âge et identification des périodes de recrutement

Les échantillons obtenus mensuellement de septembre 2004 à décembre 2011 ont été étudiés, ce qui représente 88 mois échantillonnés et 29126 individus. Pour chacun des échantillons, l'hypothèse nulle de normalité des distributions des classes de taille a été rejetée au seuil de 1% permettant la réalisation des analyses de décompositions modales. Les résultats obtenus pour chacun des échantillons sont présentés en Annexes 7 et 8. Etant relativement similaires d'une année sur l'autre, seuls les résultats de 2010 sont présentés dans la figure 4. Différents points ressortent de ces analyses :

- trois ou quatre composantes modales sont identifiées ; la composante modale regroupant les individus les plus grands est systématiquement numérotée comme 4<sup>ème</sup> composante modale.



**Fig. 5 Proportion relative des différentes catégories d'individus benthiques (du bas vers le haut : juvéniles, mâles, individus en transition sexuelle, femelles –avec ou sans ponte) des échantillons prélevés mensuellement lors du « suivi benthique » de *C. fornicata* réalisé au Château du Taureau entre septembre 2004 et décembre 2011.**



**Fig. 4 Evolution de la structure de taille des individus échantillonnés mensuellement au Château du Taureau en 2010.**

A gauche est présentée pour chaque mois la distribution des effectifs au sein de chaque classe de taille de 4mm. Les mâles sont indiqués en gris foncé, les femelles en noir, les individus en transition sexuelle en blanc et les juvéniles en hachurés. La décomposition modale associée est donnée à droite : la distribution des fréquences de taille est représenté par les histogrammes en bleu, les courbes rouges représentent chacune des composantes gaussiennes et la courbe verte est la résultante polymodale de l'ensemble des gaussiennes.

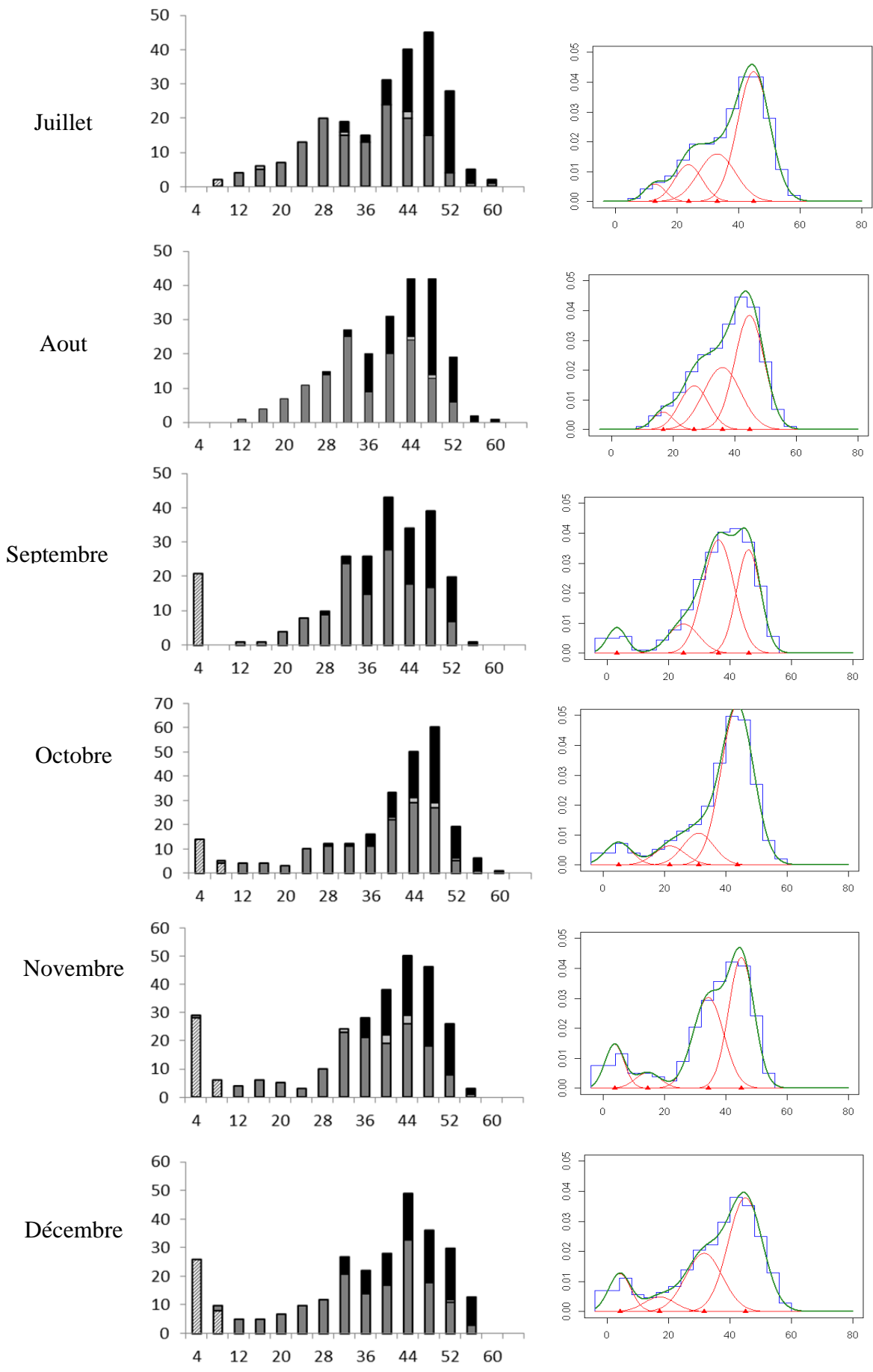


Fig. 4 (suite)

- la première composante modale n'est pas présente tous les mois. Le mode de la gaussienne et l'écart-type sont faibles : elle regroupe les recrues de l'année (*i.e.* juvéniles et très jeunes mâles) arrivées à la fin de l'été - début de l'automne. Les premières recrues sont observées en septembre 2010 (Fig. 4). La figure 5, qui indique la proportion relative des différentes catégories d'individus benthiques, montrent clairement l'arrivée des recrues au cours du temps. Cependant, cette première modale regroupe également des jeunes mâles, établis dans la population depuis plusieurs semaines. La période de recrutement a alors été déterminée avec plus de précision à l'aide d'histogrammes de fréquence de taille portant sur les individus de 10 mm et moins, défini comme la taille correspondant à l'apparition des premiers mâles, établis depuis plusieurs semaines dans la population (Dupont 2004). Ces histogrammes sont présentés dans la figure 6 pour l'ensemble des années du suivi. Ils montrent l'arrivée de juvéniles de très petites tailles, proches de la taille à la métamorphose chaque année en fin d'été et permettent d'identifier une période de 2-3 mois correspondant à l'évènement majeur de sédentarisation des recrues de l'année. Le déplacement du mode de cette composante modale dans les mois suivant la fin de l'été montre la croissance des juvéniles et petits mâles.

- les deux ou trois autres composantes modales sont beaucoup plus difficiles à distinguer sur les histogrammes bruts ; elles sont caractérisées par des écart-types importants. Sous l'hypothèse d'une relation linéaire parfaite entre la taille et l'âge, ces différentes composantes indiquent théoriquement la présence de trois ou quatre groupes d'individus issus du même évènement de sédentarisation au sein de la population (*i.e.* cohorte au sens stricte). Cependant *C. fornicata* est une espèce longévive, pouvant vivre 6 à 8 ans environ (Le Gall 1980, Coum 1979). Avec un recrutement annuel, on s'attendrait à identifier davantage de cohortes *sensus stricto* que le nombre observé. Cette différence s'explique par (1) une forte variabilité de croissance entre individus du même âge montrée par des observations expérimentales, des estimations d'âge basées sur le comptage de stries d'arrêt de croissance ainsi que par des analyses isotopiques (Dupont 2004, Richard 2005, Le Gall 1980, Coum 1979), (2) une mauvaise corrélation entre la taille et l'âge pour des âges croissants ; un résultat classique chez les mollusques et (3) le changement de sexe des individus compliquent certainement la relation entre l'âge et la taille. Il n'existe pas de modèles de croissance corrigé (par exemple modèle de von Bertalanffy) pour la population étudiée.

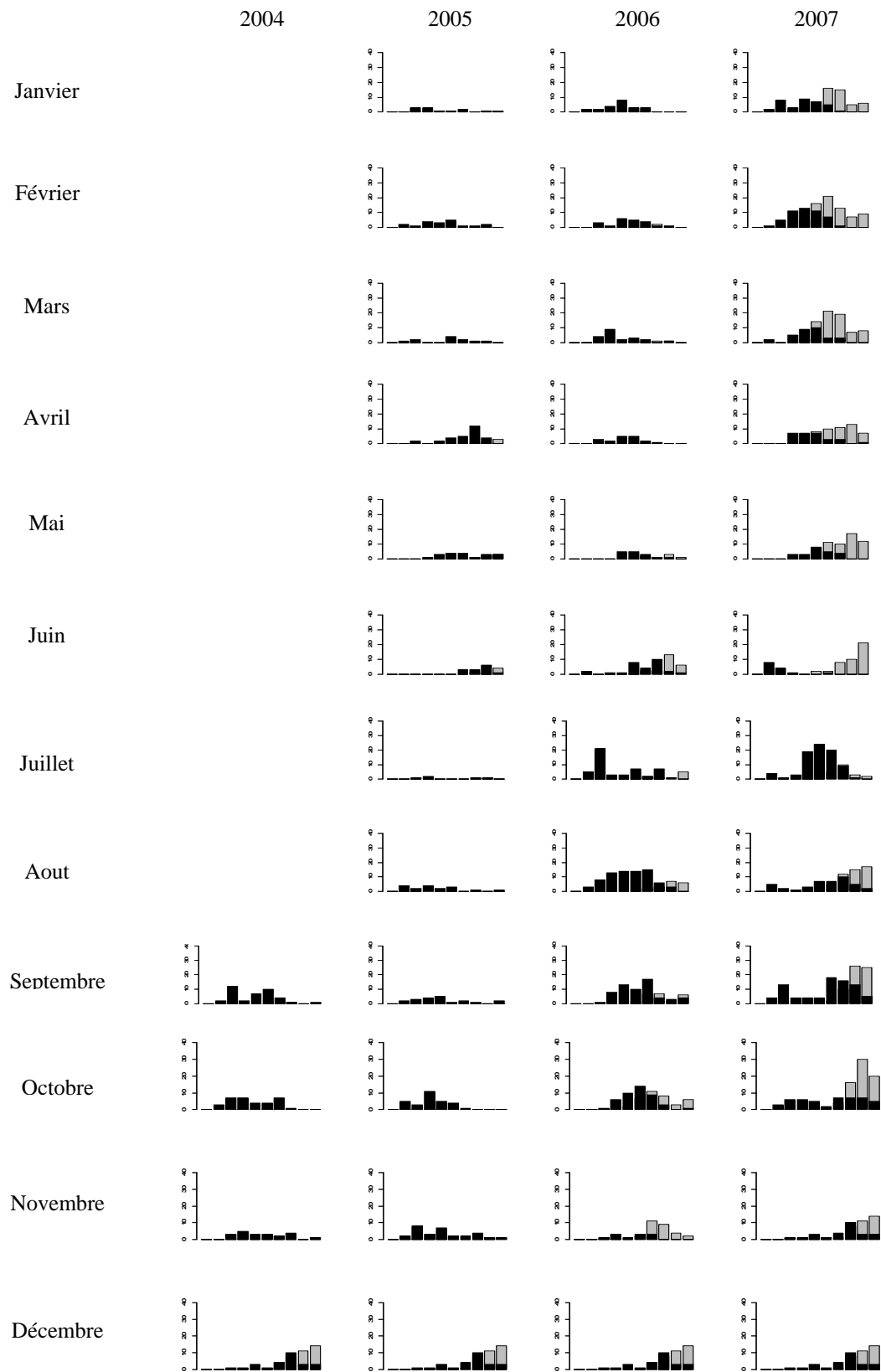
Nous ne pouvons donc identifier avec certitude que la première cohorte dans cette population.

## B. Structure d'âge, structure sexuelle et dynamique de changement de sexe

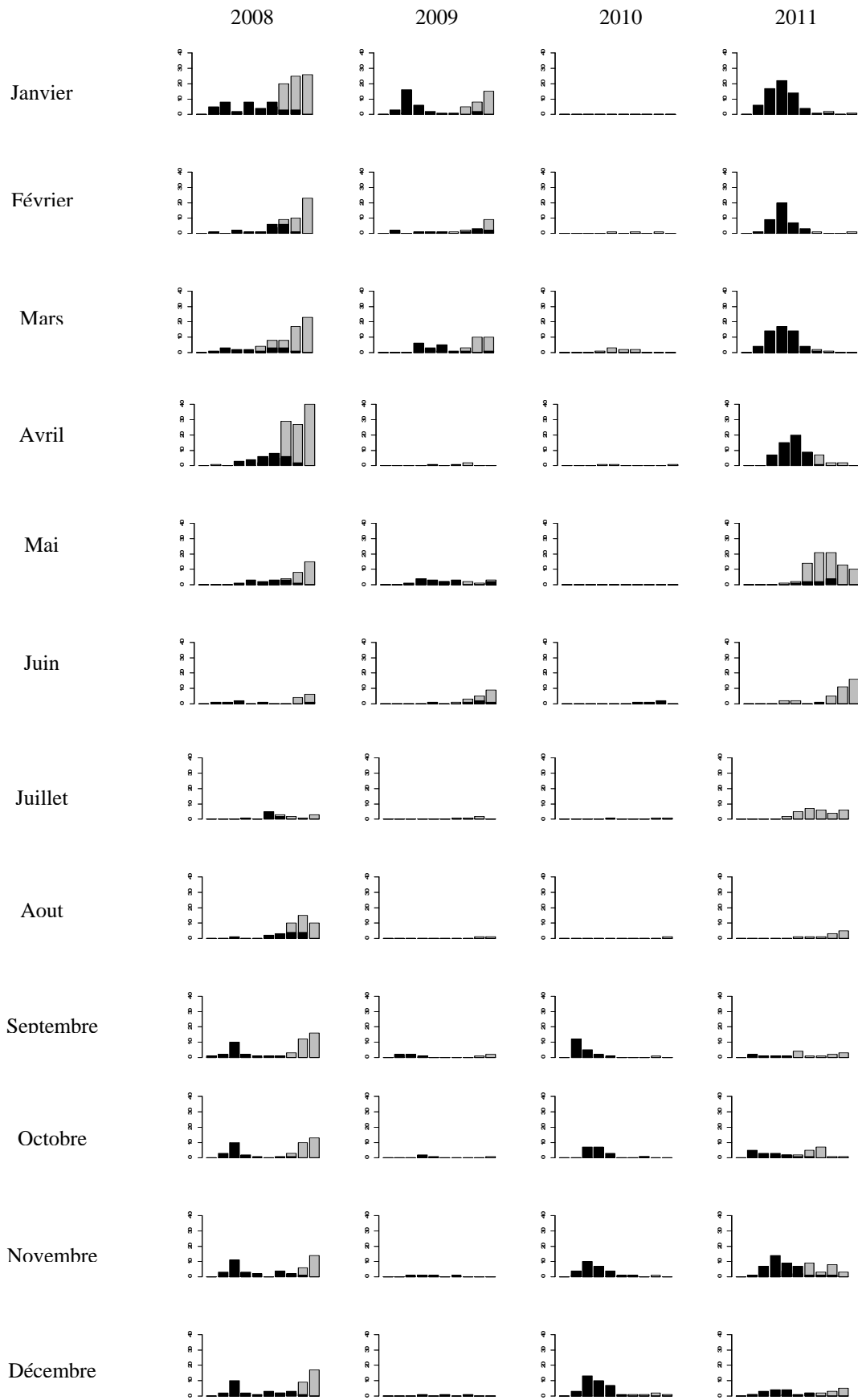
Comme attendu chez une espèce protandre et montré par des observations antérieures (Dupont 2004, Le Cam 2009), la composante modale la plus élevée (numéro 4) est majoritairement composée de femelles. La figure 4 met en évidence plusieurs points :

- la distribution des sexes n'est pas aléatoire entre les différents groupes d'âge identifiés mais il existe un chevauchement des tailles entre les différents stades sexuels, avec par exemple, des mâles aussi grands que des femelles : il n'existe alors pas de relation stricte entre la taille et le sexe des individus ;
- des femelles en ponte quelles que soient leurs tailles : la reproduction des femelles n'est donc là non plus, fonction de la taille ;
- une taille très variable des individus en transition sexuelle : il n'y a pas de taille spécifique au changement de sexe (Le Cam 2009).

Au sein de la population benthique étudiée, les proportions de mâles, individus en transition et femelles (incubantes ou non) varient au cours du temps (Fig. 5). On note que la proportion de mâles est la plus importante quel que soit le mois étudié, variant de 48.8% (décembre 2004) à 74.2% (septembre 2007). La proportion d'individus en transition sexuelle est toujours très faible (<5.0%) et il ne semble pas y avoir de période dans l'année où cette catégorie sexuelle est plus présente. La proportion de femelles varie au cours du temps de 15.3% (octobre 2007) à 43.7% (mai 2006).

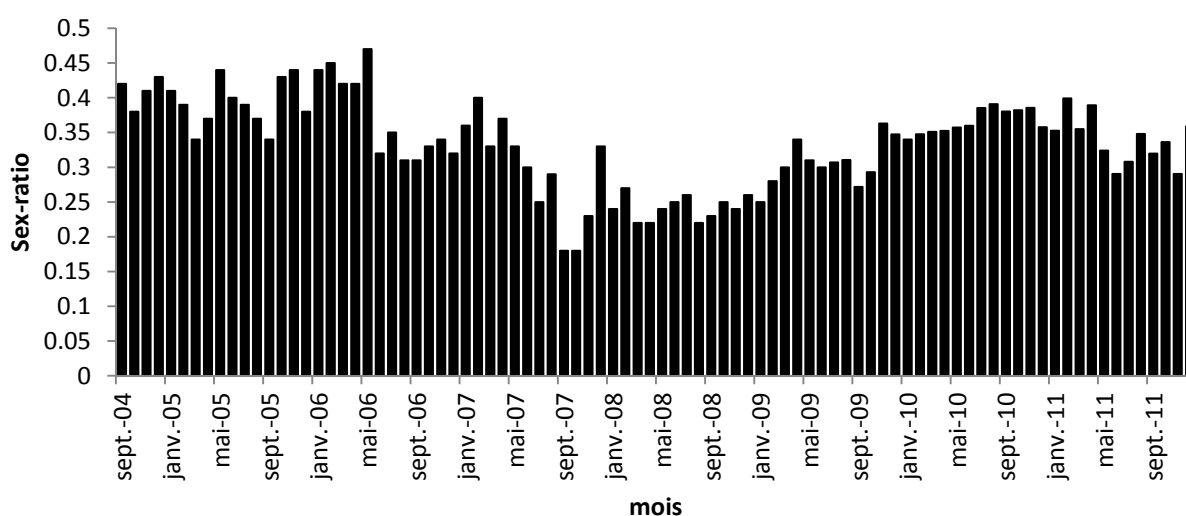


**Fig. 6** Distribution des fréquences de taille des individus de moins de 10mm échantillonnés mensuellement au Château du Taureau entre septembre 2004 et décembre 2011. Un pas de classe de taille d'1 mm est utilisé (abscisse) et les histogrammes correspondent aux effectifs cumulés des juvéniles (figurés en noir) et des mâles (figurés en gris). Les données de septembre 2004 à décembre 2007 sont extraites de la thèse de Sabrina Le Cam (2009).



**Fig. 6 (suite)**

Des variations de sex-ratio sont observées (Fig. 7) au cours de la période de suivi. Le Cam *et al.* (en préparation) ont mis en relation ses variations avec la dynamique de recrutement dans la population étudiée : les recrues se transformant rapidement en mâles, la proportion de mâles au sein de la population augmente transitoirement ; du fait du déterminisme en partie social du changement de sexe (sous l'influence du sex-ratio de la chaîne), le sex-ratio est ensuite rééquilibré par le changement de sexe de mâles. Les forts recrutements observés en 2006 et 2007 ont davantage affecté la composition de la population benthique et se sont accompagnés par un changement plus durable de la structure sexuelle de la population avec des écarts significatifs du sex-ratio aux proportions attendues de 40% de femelles jusqu'à la fin 2007 avant un « rééquilibrage » progressif (Fig. 7).



**Fig. 7 Evolution du sex-ratio de la population échantillonnée mensuellement au Château du Taureau entre septembre 2004 et décembre 2011.**

### C. Cycle annuel de reproduction

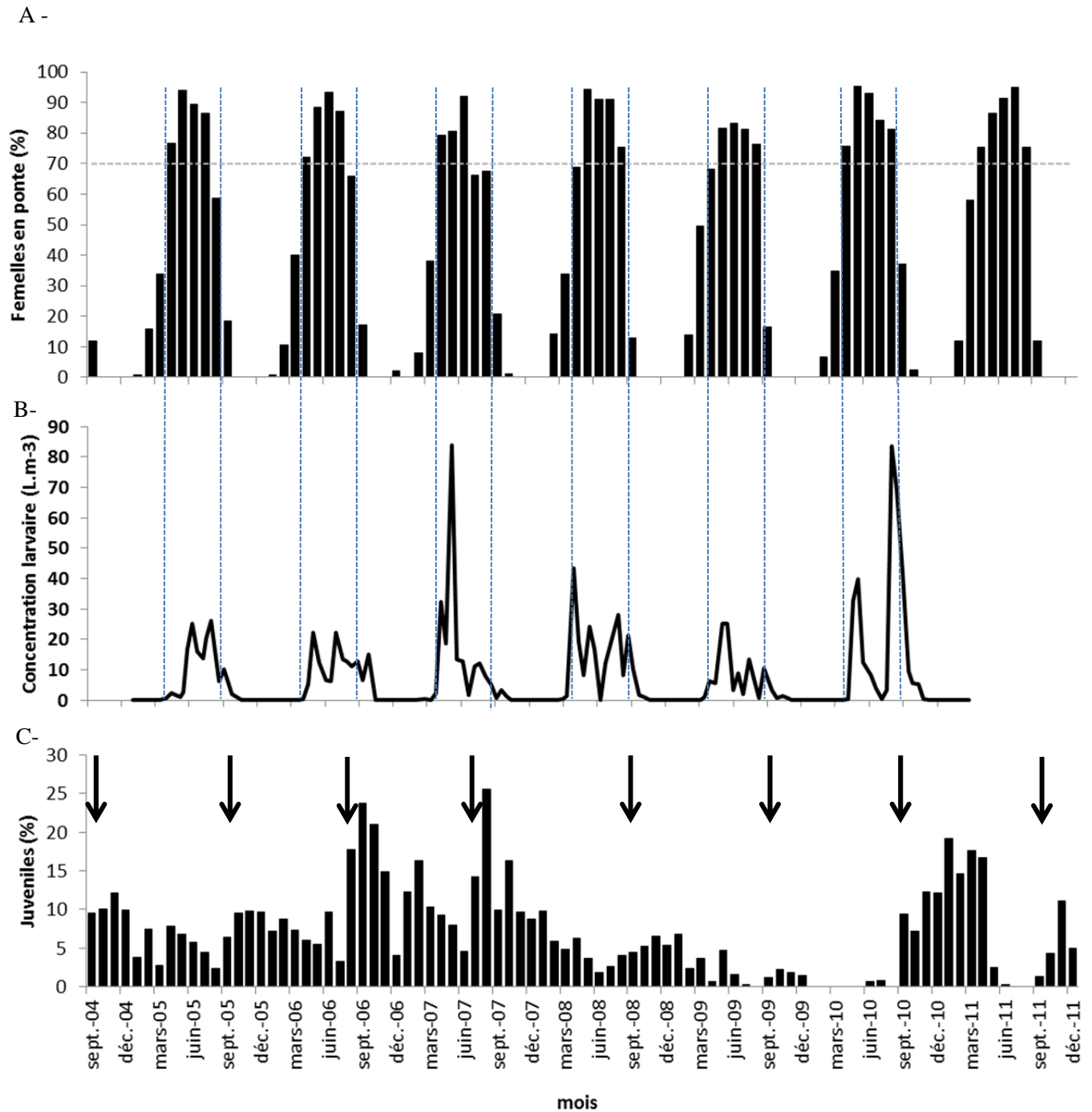
La proportion de femelles montrant des pontes est très variable au cours de l'année : elle varie de 0 à 95.3% sur l'ensemble du suivi (Fig. 5). On note toutefois une longue période de reproduction avec des femelles présentant des pontes sur plus de huit mois de l'année (Fig. 8A). Elle débute généralement en février, exception faite pour 2005 où elle débute en janvier, pour se terminer généralement en septembre – octobre. Il existe une période assez longue où on observe plus de 70% de femelles avec des pontes. Cette période s'étale sur trois (en 2007), quatre (en 2005, 2006, 2008, 2009) ou cinq (en 2010, 2011) mois. Ce schéma est très régulier d'une année sur l'autre (Fig. 8A).

Ces périodes de présence de ponte incubées par les femelles concordent avec la présence de larves dans la colonne d'eau telle que révélée par les suivis pélagiques. En effet, une longue période de présence larvaire est observée, s'étalant de mars à octobre (Fig. 8B). Les dates d'apparition des larves



dans la colonne d'eau suivent l'apparition des premières pontes observées dans la population benthique, avec un petit décalage d'un à deux mois.

Les périodes de présence de larves dans la colonne d'eau ne sont en revanche que partiellement en accord avec les périodes d'arrivée de recrues dans la population benthique (Fig. 8C). Comme nous l'avons vu précédemment (III. A-), un évènement majeur de sédentarisation est observé en fin de saison de présence larvaire, environ deux mois après le pic de reproduction (Fig. 8C). Par ailleurs, l'intensité des évènements de sédentarisation semble très variable d'une année sur l'autre et pas toujours en accord avec les concentrations larvaires mesurés *in situ* (par exemple, voir l'année 2006 dans la Fig. 8). En particulier, on observe des pics importants de recrues en 2006, 2007, 2010 mais très peu de recrues en 2009 (Fig. 8C). En automne-hiver, la présence de nombreux petits mâles (Fig. 5) traduit l'acquisition de caractères sexuels des recrues de l'année. Sans exclure totalement leur participation, il est probable que ces jeunes mâles ne soient pas des contributeurs importants de la reproduction ayant lieu au printemps suivant leur arrivée. Les analyses de paternité réalisées sur des femelles collectées *in situ* montrent que les pères sont généralement de grande taille, et même de taille supérieure à la taille moyenne des mâles de la population (Dupont *et al.* 2006).



**Fig. 8 Evolution comparée A- du pourcentage de femelles incubantes (présence de pontes), B- de la concentration larvaire et C- du pourcentage de juvéniles dans la population échantillonnée mensuellement au Château du Taureau (suivis benthiques et pélagiques).**

Les données d'abondance larvaire de sont pas disponibles en 2011.

Dans la figure 8A, la ligne pointillée marque le seuil au-delà duquel plus de 70% des femelles sont incubantes. Les lignes pointillées verticales (entre 8A et 8B) indiquent la période pendant laquelle plus de la moitié des femelles montrent des pontes. Les flèches verticales (8C) marquent le début de la période de recrutement annuel définie sur la base d'une analyse de taille et type (juvéniles ou jeunes mâles) des individus présentant une taille maximale de 10 mm (cf. Fig. 6).

#### **IV. Conclusions : des conditions idéales pour une loterie familiale ?**

Une longue saison de reproduction, des femelles pouvant se reproduire plusieurs fois par saison de reproduction (Richard *et al.* 2006 ; temps entre deux incubations de l'ordre d'un mois (Chipperfield 1951, Barranger 2011)) et plusieurs mois pendant lesquels plus de 70% des femelles présentent une ponte, suggèrent d'une part que la majorité des femelles participent à la reproduction et d'autre part, que les mêmes femelles vont se reproduire plusieurs fois le long de la saison de reproduction. Par ailleurs, les analyses de paternité réalisées sur la crépidule ont montré que (1) la reproduction se fait entre individus localisés au sein d'une même chaîne (Dupont *et al.* 2006, Le Cam *et al.* 2009), (2) que les paternités sont toujours multiples (les larves émises par une femelle sont un mélange de demi-sœurs et de pleines-sœurs) et (3) que ce sont les mêmes mâles qui contribuent aux différentes pontes produites par une femelle donnée au sein d'une saison de reproduction (Barranger 2011). Il faut toutefois noter que le nombre de pères effectifs est limité car 1 voire 2 pères dominant en général dans l'assignation des paternités (*i.e.* paternité biaisée). Ces éléments suggèrent que le pool reproducteur est stable au cours du temps.

Avec une stabilité du pool reproducteur associée à une longue saison d'émission de larves dans le milieu, une homogénéité génétique entre échantillons larvaires au cours de la saison de reproduction est attendue. Cependant, les suivis pélagiques ont mis en évidence de faibles concentrations de larves qui peuvent être expliquées par une forte mortalité des premiers stades larvaires associée à une importante exportation des larves hors de la baie (Rigal *et al.* 2010). Par ailleurs, chaque année, on observe des femelles avec des pontes ainsi que des larves dans la colonne d'eau pendant huit mois de l'année or un seul évènement majeur de recrutement est observé en fin de saison. Ceci laisse à penser qu'une part non négligeable de la production de larves par les femelles de la baie ne recrute pas localement (soit parce que les larves meurent en étant émises à une période où les conditions environnementales sont peu propices à leur développement soit parce qu'elles sont exportées hors de la baie, Rigal 2009). Compte-tenu de ces observations, les nouvelles recrues ne seraient alors pas représentatives de l'ensemble du pool larvaire au moins dans ses premiers stades de développement.

De plus, si la mortalité s'exprime de façon différentielle entre les différentes fratries larvaires, on pourrait être ici dans des conditions propices pour observer une importante variance des succès reproducteurs individuels et entre familles. Ce type de fonctionnement a été décrit par Hedgecock (1994a) sous le terme de « sweepstakes reproductive hypothesis » (que je traduirai par le terme « loterie familiale ») : les individus recrutés en phase benthique seraient issus d'une petite fraction de la génération parentale avec d'importants effets de dérive génétique entre le pool larvaire et le pool benthique.

L'observation directe des larves et leur suivi *in situ* pour, par exemple, quantifier la mortalité est une approche difficile à conduire (Levin 2006). Dans ce contexte, la génétique des populations offre des outils et un cadre théorique pour analyser les processus qui contrôlent la diversité génétique de la population. Analyser la composition génétique d'échantillons larvaires émis à différents moments au cours de la saison de reproduction en regard de celle de la population benthique (juvéniles et adultes) permettrait d'une part d'examiner la stabilité (ou les fluctuations) génétique(s) du nuage larvaire au cours de la saison de reproduction et d'autre part de tester d'éventuels effets de loterie familiale. De plus, si une fraction réduite du pool reproducteur contribue effectivement à la génération suivante, des effets de dérive génétique devraient être observables par l'analyse de la composition génétique des recrues d'une génération à l'autre. Par ailleurs, dans une population dans laquelle on fait l'hypothèse d'immigration limitée, il est plus facile d'estimer la taille efficace des populations (Waples 2005 mais voir Wang et Whitlock 2003). Néanmoins, la structure sexuelle au sein de la population favorise une reproduction entre différentes classes d'âge : les femelles, plus âgées (se retrouvant majoritairement dans la dernière composante modale) se reproduisent avec des mâles un peu plus jeunes (répartis davantage dans la troisième composante modale). Il en résulte un brassage inter-générationnel qui pourrait tamponner les changements génétiques attendus (Dupont *et al.* 2007a). Pour trancher entre ces deux hypothèses, il serait intéressant d'analyser la stabilité des pools de recrues au cours du temps : l'analyse de la structure d'âge de la population benthique échantillonnée à proximité du Château du Taureau a permis d'identifier sans ambiguïté la cohorte représentant les recrues de l'année dans la population. Dans une population relativement isolée telle qu'en baie de Morlaix, la dynamique des changements des fréquences alléliques entre générations va être influencée par (1) le nombre d'individus participant effectivement à la reproduction et (2) les événements démographiques agissant localement notamment sur les juvéniles (*i.e.* nouvelle génération). L'intensité du renouvellement de la population (*i.e.* arrivée des recrues) étant variable d'une année sur l'autre, avec notamment quatre années particulières (un fort recrutement en 2006, 2007 et 2010 et un très faible recrutement en 2009), on peut également se demander si ces variations d'intensité du recrutement laisseront des traces visibles sur la composition génétique de la nouvelle génération.

## Chapitre 2 - Dynamique de recrutement et diversité génétique des juvéniles de *Crepidula fornicata* en baie de Morlaix

---

### I. Contexte de l'étude

La dérive génétique est un processus clé dans l'évolution des populations : elle influence le rythme de perte de diversité génétique et de fixation des allèles, l'importance de la consanguinité, mais également l'efficacité de la sélection en maintenant dans la population des allèles favorables (Berthier *et al.* 2002). Dans le cas d'introductions biologiques, ce dernier mécanisme pourrait alors faciliter les réponses adaptatives des espèces introduites dans un nouvel environnement (*e.g.* Tsutsui *et al.* 2000).

Dans les populations naturelles, tous les individus ne participent pas aux événements de reproduction, et ce pour diverses raisons (comportementales (*e.g.* homogamie), écologiques (*e.g.* fragmentation du milieu), ou encore inégalité du sex-ratio), si bien qu'en général, le nombre des individus de la population qui vont effectivement contribuer à la génération suivante et qui vont ainsi influencer l'intensité de la dérive génétique (*i.e.* la taille efficace de la population,  $N_e$ ) n'est pas égale à la taille réelle de la population ( $N_c$ ). En milieu marin, la taille efficace des populations a souvent été estimée à des valeurs très inférieures à la taille réelle des populations (Hauser et Carvalho 2008), variant d'un facteur  $10^{-1}$  (Poulsen *et al.* 2006) à  $10^{-8}$  (Laurent et Planes 2007) par rapport à la taille réelle de la population. La variance des succès reproducteurs individuels est une des raisons principalement avancées pour expliquer une telle différence entre taille réelle et taille efficace de la population. Dans un environnement tel que le milieu côtier, où les conditions fluctuent rapidement (*e.g.* courant, vent), et pour les espèces présentant une forte fécondité et une forte mortalité des premiers stades de vie, seule une petite partie de la population peut contribuer à la génération suivante (hypothèse dite de « sweepstakes », Hedgecock 1994a) : on s'attend alors à une large variance du succès reproducteur et une répartition non aléatoire de cette variance entre familles.

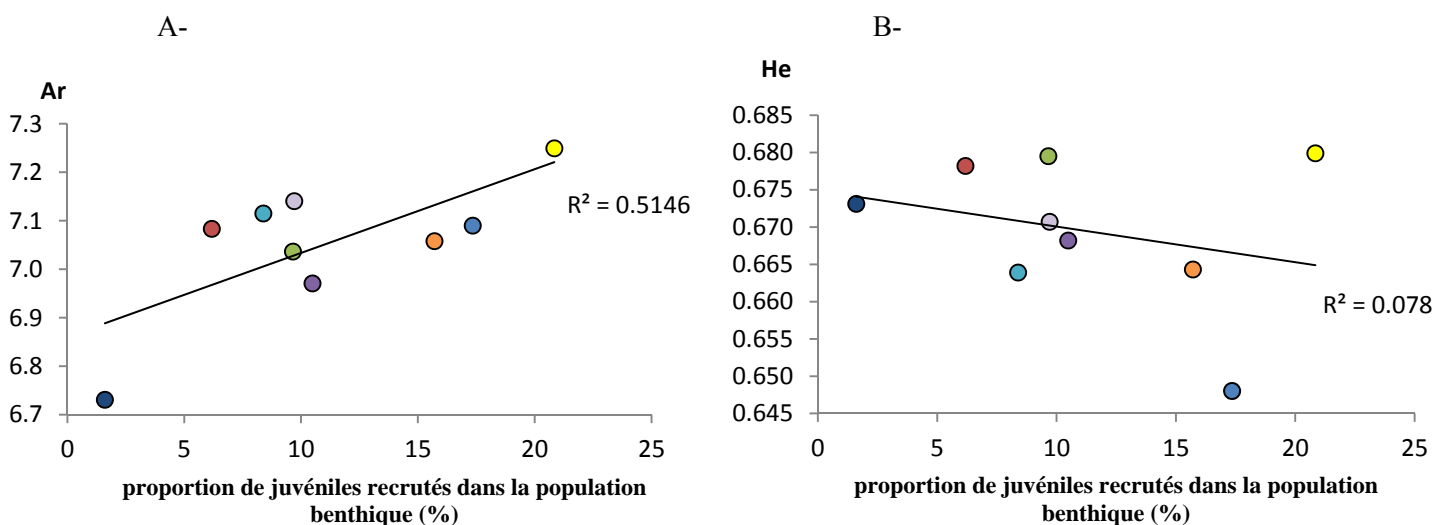
Ces traits de vie sont caractéristiques du gastéropode *Crepidula fornicata*. Ce mollusque présente en outre d'autres caractéristiques qui pourraient encore réduire la taille efficace de la population : le sex-ratio est inégal (*c.a.* 60-70% d'individus mâles), la reproduction se fait au sein d'une chaîne ce qui limite le nombre de partenaires potentiels aux individus de la chaîne, et bien que la paternité multiple soit possible (pouvant alors augmenter la taille efficace de la population, Sugg et Chesser 1994), il a été montré qu'un (ou deux) père(s) contribue(nt) majoritairement à la ponte (Le Cam *et al.* 2009, Barranger 2011) avec une contribution relativement stable au cours du temps (Barranger 2011), si bien que les larves émises par une même mère au cours du temps ont un nombre limité de pères.

Dans cette étude, nous avons analysé les effets de dérive génétique éventuels associés au recrutement dans une population de *C. fornicata* en analysant 1) la diversité génétique de 850 juvéniles échantillonnés sur 9 années consécutives (de 2002 à 2010, 681 individus) et sur sept mois de l'année 2010 (de Septembre 2010 à mars 2011, 242 individus), 2) la structure génétique temporelle de

ces individus et 3) en estimant la taille efficace de la population par six méthodes différentes. L'ensemble des individus a été génotypé à l'aide de 15 locus microsatellites.

## II. Résultats principaux de l'étude (détaillés dans l'article 3)

Dans un premier temps, nos résultats ont montré une forte diversité génétique, variable d'une année sur l'autre : la richesse allélique varie entre 6.73 en 2009 et 7.25 en 2006 et l'hétérozygotie attendue varie entre 0.648 en 2003 et 0.680 en 2005 et 2006. Cette variation temporelle de la diversité génétique reflète partiellement l'intensité du recrutement observé chaque année. En effet, le pourcentage de juvéniles dans la population benthique varie de 1.60% en 2009 à 20.84% en 2006, et une corrélation significative est observée entre la richesse allélique et la proportion de juvéniles recrutés au sein de la population benthique ( $p < 0.05$ , Fig. 1A), suggérant que plus le nombre de recrues est important plus la diversité génétique est forte. Mais aucune corrélation significative n'est observée entre l'hétérozygotie attendue et la proportion de juvéniles recrutés au sein de la population benthique ( $p > 0.05$ , Fig. 1B). Par ailleurs, la corrélation entre richesse allélique et pourcentage de juvéniles recrutés n'est plus significative quand on enlève de l'analyse les deux échantillons extrêmes, celui de 2006 (fort recrutement) et celui de 2009 (faible recrutement). Ainsi il semble que seuls des événements « extrêmes » puissent influencer le niveau de diversité génétique observé des recrues.



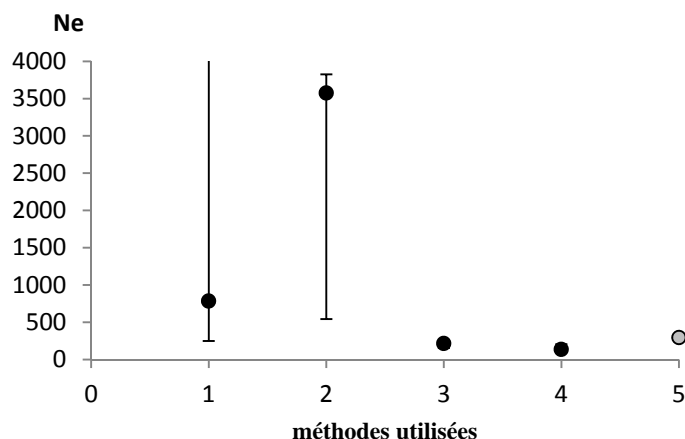
**Fig. 1** Corrélation entre la proportion de juvéniles recrutés dans la population benthique et deux indices de diversité génétique : A- la richesse allélique ( $p < 0.05$ ), B- l'hétérozygotie attendue ( $p > 0.05$ ). Chaque année est symbolisée par une couleur différente et cette couleur est identique entre les deux graphes.

De la même manière, une forte diversité génétique est observée entre les échantillons mensuels du recrutement de 2010 ( $A_r$ : 8.21 (en novembre 2010) – 8.8 (en septembre 2010);  $H_e$ : 0.652 (en novembre 2010) – 0.690 (en septembre 2010)). Aucune corrélation significative n'est observée entre ces deux indices de diversité génétique et le pourcentage mensuel de juvéniles recrutés dans la population benthique, ce dernier variant entre 7.23% (en octobre 2010) et 19.12% (en janvier 2011).

Les effets de dérive génétique entre deux épisodes de recrutement sont indiqués par une hétérogénéité génétique temporelle observée sur l'ensemble des échantillons d'une année sur l'autre ( $F_{ST} (2002-2010)=0.001$ ,  $p<0.01$ ), mais également d'un mois sur l'autre ( $F_{ST} (Sept.2010-Mars2011)=0.003$ ,  $p<0.01$ ). Les analyses des  $F_{ST}$  par paires confirment l'existence de différences génétiques entre plusieurs évènements de recrutement, en particulier, interannuel. Sous l'hypothèse que les effets d'immigration sont limités dans la population étudiée (ce qui est plausible vu des résultats antérieurs obtenus à l'aide de simulation de dispersion larvaire), cette structure génétique temporelle peut être interprétée comme résultant de variance des succès reproducteurs individuels et de processus de mortalité pendant la phase de dispersion. Tous les individus ne contribuent pas de la même manière à la nouvelle génération et seule une petite proportion des reproducteurs, différente au cours des années mais également au cours de la saison de reproduction, assure le renouvellement de la population (Hedgecock 1994a).

Afin d'avoir une estimation du nombre de reproducteurs efficaces dans la population, la taille efficace a été estimée par six méthodes : quatre d'entre elles sont des méthodes temporelles basées sur le changement des fréquences alléliques entre générations et donc estiment la taille efficace de la population sur une période de temps (ici entre 2002 et 2010) ; les deux autres méthodes estiment la taille efficace à partir d'un échantillon (méthodes basées sur l'analyse des déséquilibres de liaison et la reconstruction d'apparentés) et estiment la taille efficace du nombre d'individus de la génération précédente.

La taille efficace estimée entre 2002 et 2010 (séparée par quatre générations) est variable selon les méthodes utilisées (Fig. 2), ces différentes méthodes utilisant des propriétés mathématiques différentes ainsi que différentes hypothèses. Combinant ces quatre méthodes à partir des équations de Waples et Do (2010), la taille efficace de la population est estimée à 295,5.



**Fig. 2** Valeurs et intervalles de confiance de la taille efficace estimée à partir de quatre méthodes temporelles (en noir) entre 2002 et 2010 et estimateur de la taille efficace combinant ces quatre méthodes (en gris).

La taille efficace estimée à partir du déséquilibre de liaison est très variable d'une année sur l'autre et, à l'exception de trois années (2006, 2007, 2010), aucun effet de dérive génétique n'est mis en évidence (*i.e.* aucune réduction de la taille efficace par rapport à la taille réelle de la population). Au contraire, la méthode basée sur la reconstruction d'apparentés montre une taille efficace davantage similaire d'une année sur l'autre, variant de 54 (en 2009) à 207 (en 2006). Combinant ces deux méthodes à partir des équations de Waples et Do (2010), la taille efficace de la population est estimée à 442,3.

En combinant les deux types d'approches en un seul estimateur pour augmenter la précision (Waples et Do 2010), la taille efficace de la population a été estimée à *c.a.* 300, soit environ dix fois à cent fois moins que la taille réelle de la population estimée par Dupont (2004). Ce ratio ( $N_e/N_c$ ) est du même ordre de grandeur que la moyenne des tailles efficaces estimée sur 83 études par Palstra et Ruzzante (2008). Il est cependant plus important que la moyenne reportée pour les espèces marines, la reproduction au sein d'une chaîne de crépidule et entre générations (Dupont *et al.* 2007a) pourrait « tamponner » les effets de dérive génétique chez cette espèce.

*Des résultats complémentaires, non inclus dans l'article 3, sont présentés dans l'annexe 9 (analyses d'apparentements pour chaque échantillon), l'annexe 10 (distributions des indices de diversité génétique par rapport aux valeurs attendues) et l'annexe 11 (isolement par le temps).*

### **III. Article 3: « Inter-annual recruitment is accompanied by significant genetic drift in the invasive gastropod *Crepidula fornicata* »**

« Des effets de dérive génétique accompagnent les épisodes annuels de recrutement des juvéniles chez le gastéropode invasif *Crepidula fornicata* »

Florentine Riquet, Sabrina Le Cam, Emeline Fonteneau et Frédérique Viard – en préparation pour *Journal of Heredity*



**Inter-annual recruitment is accompanied by significant genetic drift in the invasive gastropod**

***Crepidula fornicata***

Florentine Riquet<sup>1,2</sup> – Sabrina Le Cam<sup>1,2,3</sup> Emeline Fonteneau<sup>1,2</sup> – Frédérique Viard<sup>1,2</sup>

<sup>1</sup> UPMC Univ Paris 06, UMR 7144, Team Diversity and Connectivity in Coastal Marine Landscapes, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>2</sup> CNRS, UMR 7144, Lab. Adaptation and Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>3</sup> present address: UMR INRA – Agrocampus ESE, 65 rue de Saint Brieuc, 35042 Rennes Cedex

Correspondence: F. Riquet: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France; Tel: +33298292544, Fax: +33298292324; e-mail: [riquet@sb-roscoff.fr](mailto:riquet@sb-roscoff.fr)

F. Viard: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France; Tel: +33298292312, Fax: +33298292324; e-mail: [viard@sb-roscoff.fr](mailto:viard@sb-roscoff.fr)

Running title: Genetic drift and effective size of an invasive species

Keywords: *Crepidula fornicata*, mollusk, benthopelagic life cycle, protandry, effective number of breeders, juveniles, sweepstakes

## Abstract

The population effective size is a central parameter in evolutionary biology influencing for instance the efficiency of selection and thus the adaptive potential. It has often been shown to be lower than the census size in particular in marine species with estimation sometimes orders of magnitude lower than the census size. Many marine species display high fecundity and pronounced early mortality in fluctuating environmental conditions. Consequently, a large set of families may eventually fail to reproduce offspring resulting into genetic drift from one generation of newborns to the other. Familial lottery events may be enhanced by particular reproductive traits that may further reduce the effective number of breeders. This is exemplified by the protandrous invasive species *Crepidula fornicata*. Individuals are forming small perennial breeding groups with biased sex-ratio and skewed paternity. We here investigated the importance of genetic drift over time using 15 microsatellite markers and juveniles sampling (850 individuals) obtained from a nine-year survey of annual recruitment of *C. fornicata* in the bay of Morlaix (Brittany, France), a bay little influenced by immigrants. We showed that the genetic diversity of juveniles varied over the nine study years. Part of the inter-annual genetic variation could be explained by variation in the intensity of the annual episodes of recruitment. Temporal genetic changes were observed over the study timeframe suggesting nonrandom variance in the reproductive success of different breeding groups across years. A focus on the extended 2010 recruitment (seven months) also showed temporal genetic changes over the recruitment season although much less pronounced. We estimated the contemporary population effective size using six different methods and found that it was one or two orders smaller than the census size. These results suggest that early-stages mortality and variance in reproductive success have an effect in the study species although being limited as compared to other marine species.

## Introduction

Genetic drift is influencing the rates of genetic diversity loss and allele fixation, the efficiency of selection processes and the adaptive responses to environmental changes (Frankham 1995; Berthier *et al.* 2002; Hare *et al.* 2011). Although difficult to demonstrate (e.g. Riquet *et al.* submitted<sup>1</sup>), such responses to selection has been hypothesized to be an important process to explain invasion success of non-indigenous species, especially in those species primarily introduced with a large genetic diversity on which selection on pre-existing variation may be particularly efficient (Lee 2002; Sax *et al.* 2007; Barrett & Schluter 2008; Lawson-Handley *et al.* 2011). Genetic drift is thus a critical parameter for understanding population evolutionary changes, including on short-time scale. It depends on the effective size of a population ( $N_e$ ), defined as the size of an ideal population in which genetic drift occurs at the same rate as in the study population (Wright 1931). As compared to conservation biology or fisheries management (e.g. Chapman *et al.* 2002; Hauser *et al.* 2002; Hoarau *et al.* 2005), temporal genetic changes and effective size have seldomly been investigated for non-indigenous marine species despite their evolutionary significance for understanding the success of invaders (Blanchet 2012).

Whatever the methods used, estimates of the effective size ( $N_e$ ) have been shown to be often lower than the census size of the population ( $N_c$ ) (Frankham 1995; Hauser *et al.* 2002; Yu *et al.* 2004; Palstra & Ruzzante 2008). Low  $N_e/N_c$  ratios are a particularly commonplace in marine species with  $N_e$  sometimes orders of magnitude smaller than  $N_c$  (see Hauser & Carvalho 2008 for a review). Many factors are likely to influence  $N_e$ , in particular variance in reproductive success, uneven sex-ratio, variation in population size (Hedrick 2000, Turner *et al.* 2006). A common explanation for low  $N_e/N_c$  ratio in marine species is unequal reproductive success among families. This mechanism may be particularly efficient for species with high fecundity and a free-floating larval stage experiencing high mortality (type III survivorship curve) in a fluctuating environment (particularly nearshore; Siegel *et al.* 2008). For those species, the combination of high fecundity, larval mortality and nonrandom variance in reproductive success may be viewed as a familial lottery or “sweepstakes” as coined by Hedgecock (1994), in which a large set of families will eventually fail to reproduce offspring. A

---

<sup>1</sup> Première Partie – Chapitre 2

reduction of the effective population size as compared to the census size is one among other predictions of the sweepstake reproductive success (SRS) hypothesis (Hedgcock & Pudovkin 2011).

Life history traits and mating systems are important component of the SRS and, more generally, of the reduction of the ratio of  $N_e$  to  $N_c$ . In addition to the specific properties of marine species mentioned above, the effect of unequal reproductive success among families may be further enhanced in species forming small perennial breeding groups with mating opportunities limited to a few partners. Such a situation is exemplified in the invasive marine mollusk *Crepidula fornicata*. Native from the Eastern costs of North America, this marine gastropod has been introduced in Europe since the late 19<sup>th</sup> century and reached nowadays large densities in several European bays (Blanchard 2009). This long-lived species (6-8 years, Coum 1979; Le Gall 1980; Blanchard 1995) displays a benthic-pelagic life cycle alternating a free-swimming larval stage (2 to 7 weeks, Rigal 2009) with a sessile adult stage. With a high fecundity (5000-20000 larvae released per female at each reproduction event, Richard *et al.* 2006) and high mortality in the early life-stages (Pechenik & Levine 2007), this species has a typical type III survivorship curve and is prone to nonrandom variance in reproductive success as described in other marine invertebrates (Hedgcock & Pudovkin 2011). Hedrick (2005) demonstrated theoretically that the effective size  $N_e$  is expected to be small when there are a few successful breeders ( $N_b$ ) in a population. In *C. fornicata*, several biological characteristics may reduce  $N_b$ . Firstly, the mollusk shows a gregarious behavior and the reproductive individuals have a limited number of mates available. Males and females are piling up on each other constituting perennial stacks of two to 30 individuals which will not move around. Because the species is protandrous (i.e. males change into females according to several parameters including their age (size) and the stack sex-ratio ; Coe 1938; Hoagland 1978), the two sexes are not randomly distributed within a stack: females (older individuals) are at the basis whereas males (younger individuals) are at the top of the stack. Fertilization being internal, the males are fertilizing females closed to them and within the stack they occupy. Consequently, a given female may receive sperms from different males but paternity analyses (e.g. Dupont *et al.* 2006; Proestou *et al.* 2008; Le Cam *et al.* 2009) showed that one (or two) is (are) usually dominant whatever the number of males in the stack. Each set of larvae released by a given female (releasing larvae 3 to 5 times a year; Richard *et al.* 2006) has thus a limited number of fathers.

In addition, it has been shown that the paternity contribution is stable over time (Barranger 2011). Temporal polyandry is therefore unlikely to balance the effect of a reduced effective number of fathers (Pearse & Anderson 2009). Secondly, as expected (Charnov 1982; Charnov & Bull 1989) and observed (e.g. Allsop & West 2004) in many protandrous species, the sex-ratio at the stack and population level is biased towards males further increasing the possibility of reduction of the effective size (Nunney 1993; Caballero 1994).

The aim of this study was to document the importance of genetic drift of newborns from one year to the other. Reproduction is occurring each year during spring and summer. Recruits (i.e. juveniles) that follow the dispersal and metamorphosis of the larvae settle massively each year during 2-4 months in late summer – early fall. We compared juvenile pools of *C. fornicata* that settle during the main annual recruitment event over a nine-year survey (2002-2010). For the reproductive season of 2010, we also sampled all the juveniles found over the full reproductive period (from September 2010 to March 2011) in order to look at temporal genetic changes occurring over a very short time scale. Altogether, we based our analysis on 850 juveniles genotyped over the nine years at 15 microsatellite loci. The survey was carried out in one population located in the invasive European range of *C. fornicata* (bay of Morlaix; Brittany, France), unlikely to be strongly influenced by immigration from other colonized bays (Rigal *et al.* 2010, Riquet *et al.* in preparation<sup>2</sup>). We also attempted to estimate short-term (i.e. contemporary) effective size. Estimating  $N_e$  is challenging and the use of several methods is recommended (Waples 2005; Fraser *et al.* 2007; Hare *et al.* 2011). We employed six methods, some of them based on the analysis of the change in allele frequency across temporal samples (so-called “temporal” methods; Waples 1989; Wang 2001; Beaumont 2003; Wang & Whitlock 2003) and other using only one sample and analyzing linkage disequilibrium (LD; Waples & Do 2010) or identifying half and full sibs (Wang 2009). Application of these methods generally requires analyzing a given cohort and replicating samplings or to randomize the sampling across cohorts (Jorde & Ryman 1995; Waples & Yokota 2007; Waples & Do 2010). Because cohorts older than one year cannot be distinguished in the study species (Le Cam 2009) and because we were primarily interested in inter-annual genetic variation of recruits due to a reduced effective number of

---

<sup>2</sup> Deuxième Partie – Chapitre 3

breeders, we analyzed juveniles only. We tested the following hypothesis: 1) temporal changes of genetic diversity is occurring from one recruitment episode to the other, 2) the importance of genetic drift event is explained by variation in the intensity of recruitment, i.e. a higher wave of recruits reduce genetic diversity loss 3) the genetic drift events occurred through reduction in effective population size in support of the sweepstake reproductive hypothesis.

## **Material and Method**

### **Sampling design**

The sampling of *Crepidula fornicata* was conducted in the bay of Morlaix (Brittany, France) at the GPS coordinates 48°40'18''N, 3°53'17''W. In this study site, as part of a long-term sexual and demographic survey, the benthic phase of *C. fornicata* has been monitored since 2004 based on a monthly sampling. The sampling of c.a. 30 full stacks (min. 300 individuals) per month is done by scuba diving to preserve the stack composition, along a 50m transect. Juveniles are quite easy to collect as they settle preferably on stacks of adults because of chemical-cues (Pechenik & Heyman 1987; McGee & Targett 1989). In laboratory, the juveniles were isolated and preserved in 96% ethanol for subsequent genetic analyses. This monitoring allowed determining the yearly recruitment event which is occurring during late summer – early autumn. The percentage of recruits in the population as compared to the other benthic individuals was shown to be variable across years (Table 1A; Le Cam 2009; Le Cam, Riquet & Viard unpublished data) with two extreme values: 1.6% in 2009 (J09) and 21.0% in 2006. For each year except 2009, we pooled in a same ‘annual recruits’ sample all the juveniles collected during three successive months to be representative of the major recruitment episode (Table 1A). For 2009, as a very low recruitment occurred, we selected all the juveniles collected before the start of the next reproductive period (March 2010). Additional samples were available for 2002 and 2003, i.e. before the start of the monthly monitoring. Because only 4 samplings per year were made at this time, we could use only one sample for collecting recruits. For fine scale assessment of the temporal genetic variation, all the recruits of 2010 collected were analyzed (from September 2010 to March 2011; Table 1B).

### **DNA extraction and microsatellite genotyping**

Genomic DNA extraction of juveniles was performed using a Nucleospin® Multi-96 Tissue Kit (MACHEREY-NAGEL) following the manufacturer's protocol. Each individual was genotyped at 15 microsatellite loci (12 EST-SSRs: BI13YE17, CL191, CL322, DA5YM24, DA8YN14, CL270, DA1YN22, DA4YO03, CL291, CL83, DA1YL23, DA1YA23 (Riquet *et al.* 2011), and 3 anonymous markers: CfH7 (Dupont *et al.* 2006), Cf8 (Proestou 2006) and CfCA4 (Dupont *et al.* 2003)). Loci were amplified by polymerase chain reactions (PCR) following Riquet *et al.* (2011) protocols but with 2µL of stock DNA diluted 1:20. Amplification products were separated by electrophoresis on an ABI 3130xl DNA sequencer (Applied Biosystems™). Scoring of alleles was performed using GeneMapper® v. 4.0 (Applied Biosystems™) and manually checked.

### **Analyses of genetic diversity**

Allele frequencies, allelic richness ( $A_r$ , i.e. the expected number of alleles corrected for sampling size, based on a rarefaction procedure) and expected heterozygosity ( $H_e$ ) were estimated for each sample (i.e. yearly for 2002-2010 and monthly for 2010) using FSTAT 2.9.3.2 (Goudet 1995). The correlation between these estimates and the percentage of recruits observed during the monitoring was assessed by Spearman's correlation analyses using the software package R (R Development Core Team, 2011). Correspondence Analysis (CA) implemented in GENETIX (Belkhir *et al.* 1996 - 2004) were performed to graphically visualize the genetic distance between each sample of juveniles over the nine study years and over the seven-month study of 2010. The genetic structure between samples was assessed by calculating an estimate of  $F_{ST}$  (Weir & Cockerham 1984) using GENEPOP 4.1 (Rousset 2008). Exact tests for population differentiation (10,000 dememorization steps, 500 batches and 5,000 iterations per batch) were carried out to test for differences in allelic distributions among samples and between pairwise samples. To adjust the p-values for multiple tests, q-values were computed using the QVALUE package in the R software (Storey 2002). Isolation By Time (IBT) was tested over the nine years studied as well as over the seven months in 2010, using Mantel test (Mantel 1967) implemented in GENEPOP on the web (<http://genepop.curtin.edu.au/>) by correlating temporal distance (years or months between recruitment) to  $F_{ST}/1-F_{ST}$ .

## Effective size estimation

To estimate effective population size, we used two categories of methods. Firstly, we used methods in which  $N_e$  is estimated by measuring the temporal changes in allelic frequencies between two or more samples separated by one or more generations (i.e. the so called ‘temporal methods’, Waples 1989). Four methods were used: 1) a moment-based F statistics (Waples 1989), 2) a pseudo-likelihood based approach (Wang 2001), 3) a maximum likelihood coalescent-based method (Beaumont 2003) and 4) a method identical to 2) but in which migration rate is jointly estimated (Wang & Whitlock 2003), relaxing the assumption that the population is completely closed to immigration. Sampling plan II was employed (Waples 1989) as the individuals were sampled before reproduction and not replaced. Each of these methods assumes that mutation and selection are negligible in changing allelic frequencies relative to genetic drift. Sampling periods spanned nine years (2002-2010) and the slipper limpet has a generation time of c.a. 2 years. The effective size was estimated between the most temporally-spaced samples (i.e. recruits of 2002 and recruits of 2010) separated by c.a. four generations, a span time sufficiently short that the effects of mutation can be ignored (Waples 1989; Beaumont 2003) but bias due to non-discrete generations reduced (Waples & Yokota 2007). It is more difficult to validate the assumption about selection although our loci showed little evidence for being under strong selection processes in the introduced range as they did not show any outliers in a genome-scan study (Riquet *et al.* submitted<sup>3</sup>). The moment-based temporal estimates and their 95% confidence intervals (Waples 1989) were computed using NeEstimator 1.3 (Ovenden *et al.* 2007; method 1). The pseudo likelihood methods (methods 2 and 4) used is implemented in MLNe 2.3 (Wang & Whitlock 2003). We set the maximum  $N_e$  at 50 000, higher than the census size estimated by Dupont (2004) for the study population in which density was estimated to vary from 20 to 200 ind.m<sup>-2</sup>, resulting in an upper bound of the census size estimated to 10 000 individuals in the monitored site. Migration was set at zero, a condition under which MLNe estimates the likelihood estimator proposed by Wang (2001 – method 2). For the method 4 which used the MLNe software as well, samples used as “source population” were adults collected at four sites of the bay of Morlaix and of two sites from the closest adjacent bay

---

<sup>3</sup> Première partie – Chapitre 2



(Saint-Brieuc), previously genotyped in another study (Riquet *et al.* in preparation<sup>4</sup>). Finally we used a coalescent-based maximum likelihood method implemented in TMVP (method 3 - Beaumont 2003), that uses MCMC simulations to generate a posterior distribution of  $N_e$  and incorporates a Bayesian prior on the maximum  $N_e$ . We set  $N_e$  at 50 000 and used 25 000 MCMC replicates. A combined point estimator based on harmonic mean was computed across these four temporal methods, according to strategy 1 described in Appendix A of Waples and Do (2010).

Secondly, we used methods based on the analysis of one sample estimating the effective number of breeders ( $N_b$ ) that produced the sample (Waples 2005): 1) by measuring the nonrandom association of alleles at different loci within a sample (Hill 1981) using a bias correction (Waples 2006) and implemented in LDNe (Waples & Do 2008; method 5) and 2) a method based on kinship reconstruction as proposed by Wang (2009) and available in the COLONY software (Jones & Wang 2010). For method 5, a criterion was used for excluding rare alleles and balancing the precision-bias tradeoff due to these rare alleles. Given our sample size, we chose to exclude alleles with frequency below 0.02 following recommendations by Waples & Do (2010). Like for temporal estimates, the estimates from the two methods were combined into a joint point estimate per sample (Waples & Do 2010).

Finally, to get an overall estimate of effective size with higher precision, a joint point estimate, based on all single-sample and temporal estimates, was computed (Waples & Do 2010). The single-sample estimators provide an estimate of the effective number of breeders the previous year. To avoid combining inferences over different timeframe, we computed new single-sample and temporal estimates over comparable timeframes and also excluded the 2009 sample with only 17 individuals genotyped. Two timeframes could be used (2002/2006 and 2003/2007). Computations were done for both: we combined the estimates computed using (1) temporal methods applied to the recruits of 2002 and 2006 and single-sample methods applied to the samples of 2003 and 2007 and (2) temporal methods applied to the recruits of 2003 and 2007 and single-sample methods applied to the samples of 2004 and 2008. These computations were done following recommendations and equations provided in Waples and Do (2010) and shown in Appendix S1.

---

<sup>4</sup> Deuxième Partie -Chapitre 3

## Results

*Genetic diversity of juveniles over nine years and over the 2010 recruitment event.* A total of 850 juveniles were genotyped over 15 highly polymorphic microsatellites (the total number of alleles per locus varied from 2 (locus CL291) to 68 (locus DA8YN14) over 850 individuals genotyped). Estimates of genetic diversity assessed for each sample are summarized in table 1. Taking into account sample size, genetic diversity varied across the nine study years with  $A_r$  ranging from 6.73 (J09) to 7.25 (J06). The same holds with  $H_e$ , ranging from 0.648 (J03) to 0.680 (J05 and J06; Table 1A). The most genetically diverse pool of recruits was observed in 2006. A significant correlation was observed between  $A_r$  and the percentage of yearly recruits ( $r= 0.72$ ,  $p<0.05$ ). No correlation was observed between  $H_e$  and the percentage of new recruits ( $r= -0.27$ ,  $p>0.1$ ). Over the 2010 recruitment period,  $A_r$  and  $H_e$  ranged from 8.21 (November) to 8.84 (September) and 0.652 (November) to 0.690 (September), respectively (Table 1B). Neither  $A_r$  nor  $H_e$  showed a correlation with the percentage of recruits settled in the population ( $p>0.1$  in all tests).

*Genetic structure of juveniles between years and between months for the 2010 recruitment event.* Significant temporal genetic change was observed over the nine study years ( $F_{ST-all\ samples}=0.001$ ,  $p<0.001$ ). In the Correspondence Analysis (CA), juveniles of 2006 and 2007 stand out from juveniles of the seven other years, which clustered together (Fig. 1A). This singular pattern observed in 2006 and 2007 is ascertained with  $F_{ST}$  pairwise estimates (Table 2): significant genetic differentiation was observed between J07 and all the other samples as well as between J06 and most of the other samples (all but J08 and J09). No isolation by time was noticed across the nine-year sampling ( $p$ -Mantel=0.22). Over the 2010 recruitment period, two months (September and October) stand out in the CA (Fig.1 B), but this pattern was only partially confirmed by  $F_{ST}$  pairwise estimates (Table 2) as significant genetic differentiation was only observed between October and all the other months. No isolation by time was shown when analyzing the seven study months ( $p$ -Mantel=0.19).

*Effective size estimated using the nine-year sampling.* Temporal changes in allelic frequencies over the nine study years were used to estimate effective population size (2002/2010 in Table 3). Over all methods, point estimates of  $N_e$  ranged from 136.4 (method 4) to 3576.0 (method 2). Methods 1 and 2 overlapped in confidence intervals as well as methods 3 and 4 but methods 1-2 and 3-4 did not (or

slightly for method 3 and 2) overlap (Table 3). The largest confidence intervals were obtained with the moment-based method (method 1) which an infinite value for the upper bound. Adding intermediate samples between J02 and J10 (i.e. J04, J06 and J08 standing for the first, second and third generations) changed slightly the estimates of  $N_e$  (method 2:  $N_e = 3012.8$ ,  $IC = [264.3, 4741.3]$ , method 3:  $N_e = 204.3$ ,  $IC = [150.6, 264.5]$ , method 4:  $N_e = 90.5$ ,  $IC = [75.8, 111.6]$ ). By combining the four temporal methods, the point estimate was  $N_e = 295.5$ . Using the single-sample method based on linkage disequilibrium analysis (method 5), in six out the nine study year,  $N_e$  estimates were negative (Table 4). In this case, the interpretation is that there is no evidence for any disequilibrium caused by genetic drift due to a finite number of parents. The lower bound of confidence intervals remained however useful providing information about plausible lowest limits of  $N_e$  (Fraser *et al.* 2007; Waples & Do 2010). Considering the two single-sample methods and over all samples, estimates of  $N_e$  ranged from 54 (in J09 using method 6) to an infinite value (in J02 to J05, J08 and J09 using method 5, Table 4). Confidence intervals did not overlap for 7 out of 9 samples between these two methods. By combining these two methods, point estimates were found to be variable across years and combining all samples over nine-years to get an overall effective size estimate lead to an estimated value of 442.3. This value is however not comparable to the temporal estimate computed between 2002 and 2010 as corresponding to slightly different timeframe (2001-2009 for the single sample method). For combining temporal and single-sample methods, other samples were used to apply the two methods to the same timeframe. The point estimate across temporal methods alone were  $N_e = 267.5$  and  $N_e = 250.9$  when considering 2002-2006 and 2003-2007 respectively. Over the same two timeframes, the point estimates for single-sample methods were  $N_e = 1111.2$  and  $N_e = 300.9$  respectively, thus similar for 2003-2007 but three times higher for 2002-2006 (details are given in Appendix S1). Combining the two categories of methods by taking into account the estimates variance, the point estimate were  $N_e = 273.4$  for 2002-2006 and  $N_e = 290.9$  for 2003-2007, thus very similar values (Appendix S1).

## Discussion

The genetic analysis of juveniles of *Crepidula fornicata* recruited annually over nine years in the bay of Morlaix (Brittany, France) showed variations in genetic diversity and shifts in allelic

frequencies over time. In this relative closed bay (Rigal *et al.* 2010, Dupont *et al.* 2007b), the temporal genetic changes observed are likely to reveal variance in adult reproductive success over time. *C. fornicata* display high fecundity and pronounced mortality in the early stages, as observed in many other marine species notably because of the variability in ocean conditions (Siegel *et al.* 2008; Hedgecock & Pudovkin 2011). Many individuals are likely to eventually fail to reproduce, because only a random subset of larvae will survive dispersal, metamorphosis and settlement, i.e. the so-called sweepstake reproductive success hypothesis (Hedgecock 1994). The effective number of breeders may thus be reduced as compared to the number of reproductive adults. The temporal genetic changes observed from one year to another was however more or less marked across years suggesting that the extent of this familial lottery may vary over years, producing more or less genetic drift. Interestingly, the recruitment success was also variable from one year to the other ( $\chi^2$ -test:  $p < 0.001$ ). Three years (2003, 2006 and 2007; Table 1) showed particularly important recruitment with a higher percentage of juveniles than the mean value (11.6%). One year (2009) was conversely associated to a very low recruitment with less than 2% of the benthic individuals being recruits over the recruitment period (Table 1). Using the proportion of juveniles in the benthic samples as a proxy for settlement success, we observed a significant correlation with the estimate of allelic richness, suggesting that when more juveniles are recruited, they are also more genetically diverse. This correlation was however not evidenced with gene diversity. Moreover the correlation was mainly due to two years (2006 and 2009) during which recruitment events were very different from the other years (very important in 2006 and very low in 2009): the correlation was no more significant when J06 and J09 were discarded from the analysis ( $r = 0.47$ ,  $p = 0.78$ ). This result suggests that extreme (low or high) recruitment events may influence the genetic diversity in juveniles but that slight inter-annual changes are not strong enough to impact the genetic diversity in the study population. The same holds over a shorter time scale, i.e. comparing intra-annual variation of juveniles settled in 2010-2011. As expected in a species with marked episode of recruitment, the settlement dynamics was highly variable over time with a population composed by 7.23% (October 2010) to 19.12% (January 2011) of recruits. No variation in genetic diversity indices was observed despite this variation in the monthly proportion of juveniles ( $\chi^2$ -test:  $p < 0.001$ ). Shift in allelic frequencies observed from one year to the other was also not observed

from one month to the other in 2010 with the exception of the juveniles sampled in October and to a lesser extent in November (CA, Fig. 1B). The first two months of the 2010 recruitment appeared slightly different from the other studied months suggesting that recruits from the beginning of the recruitment season underwent important mortality or/and were produced by a different set of parents (Gosselin & Qian 1997; Hunt & Scheibling 1997).

Temporal genetic changes in population closed to migration are largely driven by reduction in effective population size as compared to census population size. We used several methods to estimate short-term  $N_e$ . The four temporal methods used differ in mathematical properties and assumptions (review in Luikart *et al.* 2010) and revealed some inconsistencies among methods.  $N_e$  estimates were lower in methods 3 and 4 than in methods 1 and 2. Larger CIs in  $N_e$  estimates was revealed by methods 1 and 2, suggesting more uncertainty. Sample size is influencing both precision and biases of temporal estimates (Waples 1989). The 2002 sample is limited in size with 32 individuals. Method 1 is corrected for the effect of sample size; the others are not but are expected to be less sensitive to sample size (Palstra & Ruzzante 2008). Based on a review of 83 studies computing empirical estimates of  $N_e$ , Palstra & Ruzzante (2008) showed that for temporal methods, the precision of the estimates is increased with marker polymorphism. This effect was shown to be higher than the increased in precision due to an increased sample size. We used a set of 15 microsatellites which should help to increase the precision despite the low sample size in some of the samples. However, with many rare alleles, method 1 is expected to show a decrease in precision as compared to the other temporal methods (Waples 1989, Pastra & Ruzzante 2008). Point estimates for which CIs are large should be carefully interpreted; point estimates by methods 3 and 4 appeared from this point of view more reliable. They also showed more similar point estimate values when changing slightly the timeframe (Table 3). One assumption in most methods estimating  $N_e$  is that migration can be neglected. Genetic differentiation from source populations of migrants can strongly influenced  $N_e$  estimation (Fraser *et al.* 2007). The study population was expected to be under limited influence of immigrants. However, method 2 and method 4, which are identical except that method 4 included migration from a “source population”, showed inconsistencies. Highest values of  $N_e$  were estimated when ignoring gene flow. In the case of low genetic differentiation in source populations, as demonstrated in *C. fornicata* in the

study area (Dupont *et al.* 2007b), migration might compensate for genetic drift effect in short-term, resulting in an overestimation of  $N_e$  if migration is ignored.  $N_e$  was also estimated using single-sample methods. Here again some inconsistencies were observed. Method 5 (Waples & Do 2008), based on measuring a nonrandom association of alleles at different loci (Hill 1981) corrected for biases, showed the estimates the most different from all the other estimates obtained with temporal or single-sample methods. The  $N_e$  estimates were often negative, meaning that the results can be explained by sampling error only (i.e. without any genetic drift; Waples & Do 2010). Only three samples (J06, J07 and J10) showed finite estimates of  $N_e$ . This is a surprising finding as Waples & Do (2010) showed that the precision of the linkage disequilibrium method is higher than temporal methods in most circumstances (large number of generations between samples). This result is likely to be explained by the large effective size for which this method, as other, is more challenging to use. Less precision of  $N_e$  estimates are expected when mean value increased (Palstra & Ruzzante 2008) as well when sampling errors, technical errors and biases are not minimized (Hare *et al.* 2011). Lower bounds of CIs remained however informative about plausible low limits for  $N_e$ . Results from method 6 appeared more congruent with temporal methods, in particular 3 and 4. Departure from different assumptions may differentially affect some of the estimates. To increase precision, Waples & Do (2010) proposed several strategies to combine several estimates derived from temporal and single-sample methods, taking into account appropriate weights for each estimate. Combining point estimates across single-sample and temporal methods give congruent results over the two study timeframe, with  $N_e$  estimated at 273.4 between 2002 and 2006 and at 290.9 between 2003 and 2007. This value were also much closed to the estimate obtained over the whole study period with temporal methods only ( $N_e = 295.5$ ). Despite large CIs for some methods, altogether, the estimates were of the same order (except for method 5) and the combined estimates similar across various computations, with a value close to 300.

Considering the field estimated census size (Dupont 2004) and the above  $N_e$  estimate, the  $N_e/N_c$  ratio is around 0.30 when using the lower bound of the estimated census size ( $N_{c- \min} = 1000$ ) and around 0.03 when using the upper bound of the estimated census size ( $N_{c- \max} = 10\ 000$ ). The effective size of the population was thus estimated to be at most one or two orders of magnitude less than the

number of potential breeders ( $N_e = 1000-10000$ ). This ratio is thus far to be orders of magnitude smaller than the census population size as shown in other marine species (Hare *et al.* 2011 and references here in). By analyzing  $N_e / N_c$  ratios of different species characterized by different life history traits, Palstra & Ruzzante (2008) showed that life history traits such as high fecundity and high mortality in the early stage, which are traits generating high variance in reproductive success, may result in a reduction of 14% in effective size relative to the census size. This is a value of the same order than the one observed in this study. In *C. fornicata*, at the stack level, i.e. within a breeding group (Dupont *et al.* 2006, Proestou *et al.* 2008, Le Cam *et al.* 2009), individuals from several distinct generations co-occur and mate. This particular age-sex structure may buffer genetic drift at the population level (Dupont *et al.* 2007a). The skewed paternity and the limited number of mates, conversely to our expectations, do not seem to balance strongly this ‘buffer effect’ and the relatively low genetic drift observed in our study may be attributed to the stacking behavior and protandry of *C. fornicata*.

## **Conclusion**

Temporal genetic changes and a reduced effective population size relative to the census size were observed in the study populations of *C. fornicata*. They are signatures left by sweepstake events (Hedgecock 1994). This footprint was however weak as compared to values reported in other studies investigating sweepstake events in marine species (i.e.  $N_e/N_c \ll 0.01$ , Hedgecock & Pudovkin 2011). Estimating effective size is challenging especially in marine species with larval dispersal and high effective population size because of the risks of biases due to sampling bias and gene flow from distant populations. Sweepstake events should thus also be tested using other predictions such as the loss of genetic diversity between larvae or juveniles and adults (e.g. Li & Hedgecock 1998) or higher relatedness within larval or juvenile cohorts as compared to adults (e.g. Planes *et al.* 2002). All these testable predictions need to study at once the different life-stages, including larvae before and after dispersal, when genetic drift is expected to be the more severe.

## **Acknowledgment**

We are grateful to the divers of “Marine Operations Department” of the Biological Station of Roscoff for providing stacks of *Crepidula fornicata*. We thank the Biogenouest® Genomics core facility for its technical support. This work was funded by the Agence Nationale de la Recherche (Project Hi-Flo, ANR-08-BLAN-0334) and the AXA Research Fund (project ‘Marine Aliens and Climate Change’). FR acknowledges a PhD grant from the Centre National pour la Recherche Scientifique and the Conseil Régional de Bretagne (SEA-FLO ARED project).

## References

- Allsop DJ, West SA (2004) Sex-ratio evolution in sex changing animals. *Evolution*, **58**, 1019-1027.
- Barranger A (2011) Système de reproduction et dérive génétique chez un gastéroode protandre *Crepidula fornicata*. Mémoire de Master Océanographie et Environnements Marins - Université Pierre et Marie Curie.
- Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38-44.
- Beaumont MA (2003) Estimation of population growth or decline in genetically monitored populations. *Genetics*, **164**, 1139-1160.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).
- Berthier P, Beaumont MA, Cornuet JM, Luikart G (2002) Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. *Genetics*, **160**, 741-751.
- Blanchard M (1995) Origine et état de la population de *Crepidula fornicata* (Gastropoda Prosobranchia) sur le littoral français. *Halictis*, **24**, 75-86.
- Blanchard M (2009) Recent expansion of the slipper limpet population (*Crepidula fornicata*) in the Bay of Mont-Saint-Michel (Western Channel, France). *Aquatic Living Resources*, **22**, 11-19.
- Blanchet S (2012) The use of molecular tools in invasion biology: an emphasis on freshwater ecosystems. *Fisheries Management and Ecology*, **19**, 120-132.
- Caballero A (1994) Developments in the prediction of effective population-size. *Heredity*, **73**, 657-679.
- Chapman RW, Ball AO, Mash LR (2002) Spatial homogeneity and temporal heterogeneity of red drum (*Sciaenops ocellatus*) microsatellites: Effective population sizes and management implications. *Marine Biotechnology*, **4**, 589-603.
- Charnov EL (1982) The theory of sex allocation. *Monographs in population biology*, **18**, 1-355.
- Charnov EL, Bull JJ (1989) Non-fisherian sex-ratios with sex change and environmental sex determination. *Nature*, **338**, 148-150.



- Coe WR (1938) Conditions influencing change of sex in mollusks of the genus *Crepidula*. *Journal of Experimental Zoology*, **77**, 401-424.
- Coum A (1979) La population de *Crepidula fornicata* (L, 1758) en rade de Brest - Ecologie et dynamique. *Thèse de doctorat. Université Bretagne Occidentale, France*.
- Dupont L (2004) Invasion des côtes françaises par le mollusque exotique *Crepidula fornicata*: contribution de la dispersion larvaire et du système de reproduction au succès de la colonisation.
- Dupont L, Bernas D, Viard F (2007a) Sex and genetic structure across age groups in populations of the European marine invasive mollusc, *Crepidula fornicata* L. (Gastropoda). *Biological Journal of the Linnean Society*, **90**, 365-374.
- Dupont L, Ellien C, Viard F (2007b) Limits to gene flow in the slipper limpet *Crepidula fornicata* as revealed by microsatellite data and a larval dispersal model. *Marine Ecology Progress Series*, **349**, 125-138.
- Dupont L, Jollivet D, Viard F (2003) High genetic diversity and ephemeral drift effects in a successful introduced mollusc (*Crepidula fornicata* : Gastropoda). *Marine Ecology-Progress Series*, **253**, 183-195.
- Dupont L, Richard J, Paulet YM, Thouzeau G, Viard F (2006) Gregariousness and protandry promote reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of larval paternity. *Molecular Ecology*, **15**, 3009-3021.
- Frankham R (1995) Effective population-size adult-population size ratios in wildlife - a review. *Genetical Research*, **66**, 95-107.
- Fraser DJ, Hansen MM, Ostergaard S, Tessier N, Legault M, Bernatchez L (2007) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology*, **16**, 3866-3889.
- Gosselin LA, Qian PY (1997) Juvenile mortality in benthic marine invertebrates. *Marine Ecology-Progress Series*, **146**, 265-282.
- Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485-486.
- Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, Ruegg K, Palstra F (2011) Understanding and estimating effective population size for practical application in marine species management. *Conservation Biology*, **25**, 438-449.
- Hauser L, Adcock GJ, Smith PJ, Ramirez JHB, Carvalho GR (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11742-11747.
- Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, **9**, 333-362.
- Hedgecock D (1994) Does variance in reproductive success limit effective population sizes of marine organisms?
- Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bulletin of Marine Science*, **87**, 971-1002.

- Hedrick P (2000) Genetics of populations - 2nd ed. *Boston, Jones and Bartlett Publishers, Inc.*,
- Hedrick P (2005) Large variance in reproductive success and the  $N_e/N$  ratio. *Evolution*, **59**, 1596-1599.
- Hill WG (1981) Estimation of effective population-size from data on linkage disequilibrium. *Genetical Research*, **38**, 209-216.
- Hoagland KE (1978) Protandry and evolution of environmentally-mediated sex change - study of mollusca. *Malacologia*, **17**, 365-391.
- Hoarau G, Boon E, Jongma DN, Ferber S, Palsson J, Van der Veer HW, Rijnsdorp AD, Stam WT, Olsen JL (2005) Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa* L.). *Proceedings of the Royal Society B-Biological Sciences*, **272**, 497-503.
- Hunt HL, Scheibling RE (1997) Role of early post-settlement mortality in recruitment of benthic marine invertebrates. *Marine Ecology-Progress Series*, **155**, 269-301.
- Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, **10**, 551-555.
- Jorde PE, Ryman N (1995) Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics*, **139**, 1077-1090.
- Lawson-Handley LJ, Estoup A, Evans DM, Thomas CE, Lombaert E, Facon B, Aebi A, Roy HE (2011) Ecological genetics of invasive alien species. *Biocontrol*, **56**, 409-428.
- Le Cam S (2009) Grégarité, changement de sexe et polyandrie: modalités de la reproduction chez une espèce invasive *Crepidula fornicata*. *Thèse de doctorat. Université Pierre et Marie Curie, Paris*.
- Le Cam S, Pechenik JA, Cagnon M, Viard F (2009) Fast versus slow larval growth in an invasive marine mollusc: does paternity matter? *Journal of Heredity*, **100**, 455-464.
- Le Gall P (1980) Etude expérimentale de l'association en chaîne et son influence sur la croissance et la sexualité chez la crépidula *Crepidula fornicata*. *Université de Caen*.
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, **17**, 386-391.
- Li G, Hedgecock D (1998) Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1025-1033.
- Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, **11**, 355-373.
- Mantel N (1967) Detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-&.
- McGee BL, Targett NM (1989) Larval habitat selection in *Crepidula* (L) and its effect on adult distribution patterns. *Journal of Experimental Marine Biology and Ecology*, **131**, 195-214.

- Nunney L (1993) The influence of mating system and overlapping generations on effective population size. *Evolution*, **47**, 1329-1341.
- Ovenden JR, Peel D, Street R, Courtney AJ, Hoyle SD, Peel SL, Podlich H (2007) The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*). *Molecular Ecology*, **16**, 127-138.
- Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428-3447.
- Pearse DE, Anderson EC (2009) Multiple paternity increases effective population size. *Molecular Ecology*, **18**, 3124-3127.
- Pechenik JA, Heyman WD (1987) Using KCL to determine size at competence for larvae of the marine gastropod *Crepidula fornicata* (L). *Journal of Experimental Marine Biology and Ecology*, **112**, 27-38.
- Pechenik JA, Levine SH (2007) Estimates of planktonic larval mortality using the marine gastropods *Crepidula fornicata* and *C. plana*. *Marine Ecology-Progress Series*, **344**, 107-118.
- Planes S, Lecaillon G, Lenfant P, Meekan M (2002) Genetic and demographic variation in new recruits of *Naso unicornis*. *Journal of Fish Biology*, **61**, 1033-1049.
- Proestou DA (2006) Isolation and characterization of microsatellite markers in the Atlantic slipper shell *Crepidula fornicata* for use in paternity analysis. *Molecular Ecology Notes*, **6**, 437-439.
- Proestou DA, Goldsmith MR, Twombly S (2008) Patterns of male reproductive success in *Crepidula fornicata* provide new insight for sex allocation and optimal sex change. *Biol Bull*, **214**, 192-200.
- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Richard J, Huet M, Thouzeau G, Paulet YM (2006) Reproduction of the invasive slipper limpet, *Crepidula fornicata*, in the Bay of Brest, France. *Marine Biology*, **149**, 789-801.
- Rigal F, Viard F, Ayata SD, Comtet T (2010) Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? *Biological Invasions*, **12**, 3171-3186.
- Rigal F (2009) Dynamique spatio-temporelle du nuage larvaire du gastéropode introduit *Crepidula fornicata* au sein d'une baie mégatidale, la baie de Morlaix (France). Thèse de doctorat. Université Pierre et Marie Curie, Paris.
- Riquet F, Ballenghien M, Tanguy A, Viard F (2011) *In silico* mining and characterization of 12 EST-SSRs for the invasive slipper limpet *Crepidula fornicata*. *Marine Genomics*, **4**, 291-295.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour*, **8**, 103-106.
- Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, Grosberg RK, Hastings A, Holt RD, Mayfield MM, O'Connor MI, Rice WR (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22**, 465-471.

- Siegel DA, Mitarai S, Costello CJ, Gaines SD, Kendall BE, Warner RR, Winters KB (2008) The stochastic nature of larval connectivity among nearshore marine populations. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 8974-8979.
- Storey J (2002) A direct approach to false directory rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479-498.
- Turner TF, Osborne MJ, Moyer GR, Benavides MA, Alo D (2006) Life history and environmental variation interact to determine effective population to census size ratio. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 3065-3073.
- Turner TF, Wares JP, Gold JR (2002) Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics*, **162**, 1329-1339.
- Wang J (2009) A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, **18**, 2148-2164.
- Wang JL (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genetical Research*, **78**, 243-257.
- Wang JL, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, **163**, 429-446.
- Waples RS (1989) A generalized-approach for estimating effective population-size from temporal changes in allele frequency. *Genetics*, **121**, 379-391.
- Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, **14**, 3335-3352.
- Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167-184.
- Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753-756.
- Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, **3**, 244-262.
- Waples RS, Yokota M (2007) Temporal estimates of effective population size in species with overlapping generations. *Genetics*, **175**, 219-233.
- Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.
- Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 0097-0159.
- Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, Li WH (2004) Nucleotide diversity in gorillas. *Genetics*, **166**, 1375-1383.

## **Figures and Tables**

**Fig. 1** Correspondence analyses conducted with 15 microsatellite markers used to analyze A- recruits sampled from 2002 to 2010, B- recruits sampled along the 2010 recruitment events. Only the barycenter of each population is shown. Samples are labeled according to table 1.

**Table 1** Genetic diversity indices of the study samples using 15 microsatellite markers for: A- juveniles sampled from 2002 to 2010 (with the number of months (N months) used to collect them) and B- juveniles sampled from September 2010 to March 2011.

N: number of individuals successfully genotyped,  $N_{all}$ : actual mean number of alleles per locus,  $A_r$ : allelic richness,  $H_e$ : expected heterozygosity.

A-

year	N months	month of the 1st sampling	labels	recruits (%)	N	$N_{all}$	$A_r$	$H_e$
2002	1	October	J02	8.39	32	11.20	7.11	0.664
2003	1	September	J03	17.35	41	11.47	7.09	0.648
2004	3	September	J04	10.49	93	15.47	6.97	0.668
2005	3	October	J05	9.65	71	14.53	7.04	0.680
2006	3	August	J06	21.01	135	19.60	7.25	0.680
2007	3	July	J07	15.71	144	18.60	7.06	0.664
2008	3	November	J08	6.18	75	15.40	7.08	0.678
2009	4	September	J09	1.60	17	8.53	6.73	0.673
2010	3	September	J10	10.05	73	16.13	7.14	0.671

B-

month	recruits (%)	N	$N_{all}$	$A_r$	$H_e$
Sep-10	10.53	21	10.27	8.84	0.690
Oct-10	7.23	18	9.27	8.69	0.674
Nov-10	12.23	34	11.13	8.21	0.652
Dec-10	12.10	29	10.47	8.29	0.664
Jan-11	19.12	62	14.67	8.76	0.678
Feb-11	14.60	40	11.67	8.23	0.655
Mar-11	17.65	38	12.40	8.58	0.660

**Table 2** Pairwise comparison of the genetic structure (pairwise  $F_{st}$  estimates and probability values for exact tests) for A- juveniles sampled from 2002 to 2010 and B- juveniles sampled from September 2010 to March 2011. Probability values for exact tests were corrected for multiple comparisons: ns: non-significant ( $p>0.05$ ), \* ( $p<0.05$ ), \*\* ( $p<0.01$ ), \*\*\* ( $p<0.001$ ). Significant values are highlighted in bold.

A-

	J02	J03	J04	J05	J06	J07	J08	J09
J03	-0.003n.s.							
J04	-0.001n.s.	-0.001n.s.						
J05	0.001 n.s.	<b>0.002**</b>	<b>0.002**</b>					
J06	<b>0.002***</b>	<b>0.002*</b>	<b>0.002**</b>	<b>0.002***</b>				
J07	<b>0.002***</b>	<b>0.002**</b>	<b>0.000**</b>	<b>0.003***</b>	<b>0.003***</b>			
J08	-0.001n.s.	0.000 n.s.	-0.001 n.s.	<b>-0.001***</b>	0.000 n.s.	<b>0.000**</b>		
J09	0.002 n.s.	0.005 n.s.	-0.003 n.s.	0.001 n.s.	0.002 n.s.	<b>0.002**</b>	-0.001n.s.	
J10	<b>0.004*</b>	0.002 n.s.	<b>0.001*</b>	<b>0.004***</b>	<b>0.001***</b>	<b>0.001***</b>	0.001 n.s.	0.002 n.s.

B-

	Sep 10	Oct 10	Nov 10	Dec 10	Jan 11	Feb 11
oct-10	0.002n.s.					
nov-10	<b>0.007*</b>	<b>0.004**</b>				
déc-10	0.004 n.s.	<b>0.015***</b>	0.003 n.s.			
janv-11	0.004 n.s.	<b>0.012***</b>	0.004 n.s.	-0.002 n.s.		
févr-11	0.001 n.s.	<b>0.007***</b>	0.002 n.s.	-0.002 n.s.	0.001 n.s.	
mars-11	0.003 n.s.	<b>0.004*</b>	0.003 n.s.	-0.003 n.s.	0.004 n.s.	-0.004 n.s.

**Table 3** Effective size estimates using four temporal methods and their confidence intervals (CI) computed over the whole study (2002/2010) and over two reduced timeframe (2002/2006 and 2003/2007) for comparison with single-sample estimates (see Material and Methods). A combined estimate across temporal methods is provided for each timeframe.

	Method 1 Waples (1989)		Method 2 Wang (2001)		Method 3 Beaumont (2003)		Method 4 Wang & Whitlock (2003)		Combined estimate
	$N_e$	CI	$N_e$	CI	$N_e$	CI	$N_e$	CI	$N_v$
2002/2010	783.7	[247.6, $\infty$ [	3576.0	[541.8, 3825.0]	215.0	[154.3, 259.9]	136.4	[95.8, 204.5]	295.46
2002/2006	1893.1	[219.0, $\infty$ [	4414.0	[2533.6, 8532.6]	186.1	[54.6, 290.9]	113.3	[88.6, 155.6]	267.48
2003/2007	394.2	[140.1, $\infty$ [	1093.5	[304.1, 2415.1]	165.8	[55.2, 268.3]	154.8	[118.7, 210.7]	250.90



**Table 4** Effective size point estimates using single-sample methods for each sample and their confidence intervals (CI). Combined point estimates across methods are indicated for each sample.

Sample (year for the effective number of breeders)	Method 5 (Waples & Do 2008)		Method 6 (Wang 2009)		Combined point estimates
	$N_e$	CI	$N_e$	CI	$N_e$
J02 (2001)	-128.5	[218.1, $\infty$ [	55	[34, 96]	192.31
J03 (2002)	-22.2	[-38.1, $\infty$ [	62	[41, 95]	-69.17
J04 (2003)	-8163.4	[597.7, $\infty$ [	132	[99, 178]	268.34
J05 (2004)	-2122	[583.3, $\infty$ [	110	[79, 155]	232.03
J06 (2005)	2259.5	[676.0, $\infty$ [	186	[146, 239]	343.71
J07 (2006)	6872.2	[732.3, $\infty$ [	207	[84, 265]	401.89
J08 (2007)	-840.2	[1308.5, $\infty$ [	118	[29, 164]	274.56
J09 (2008)	-81	[-215.6, $\infty$ [	54	[29, 156]	324.00
J10 (2009)	676.6	[248.8, $\infty$ [	111	[80, 161]	190.71

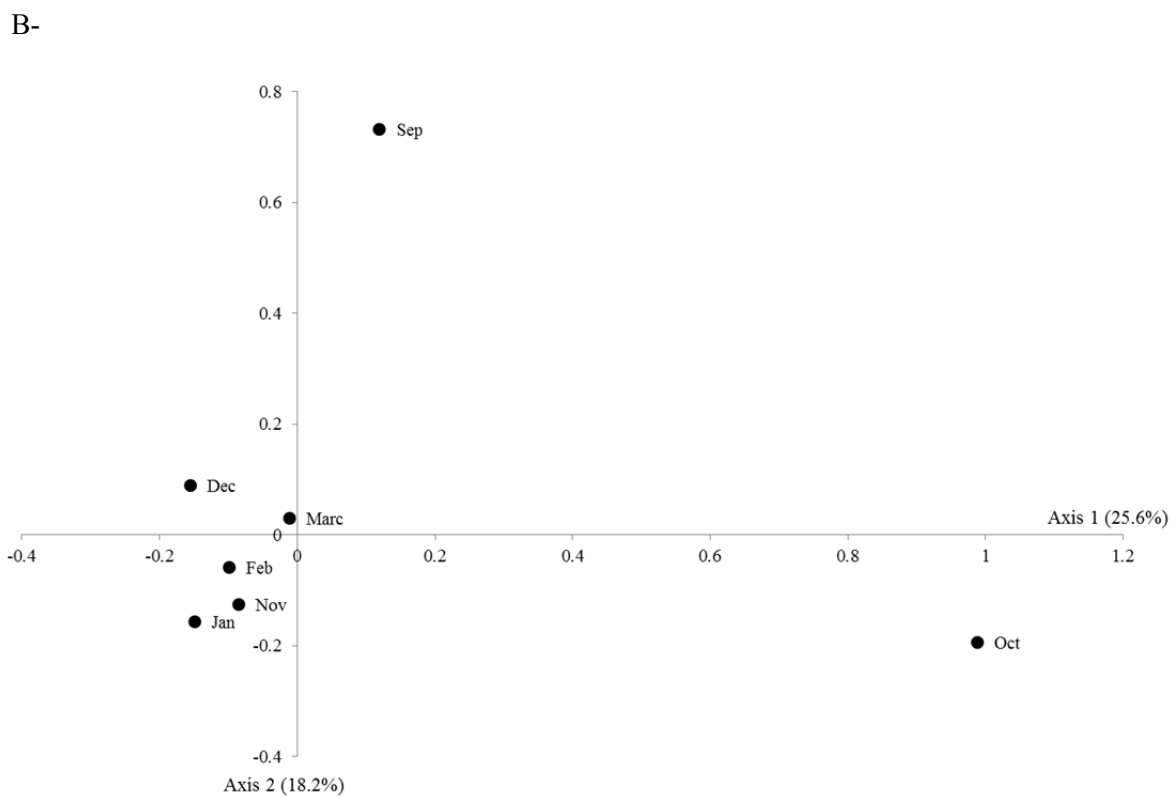
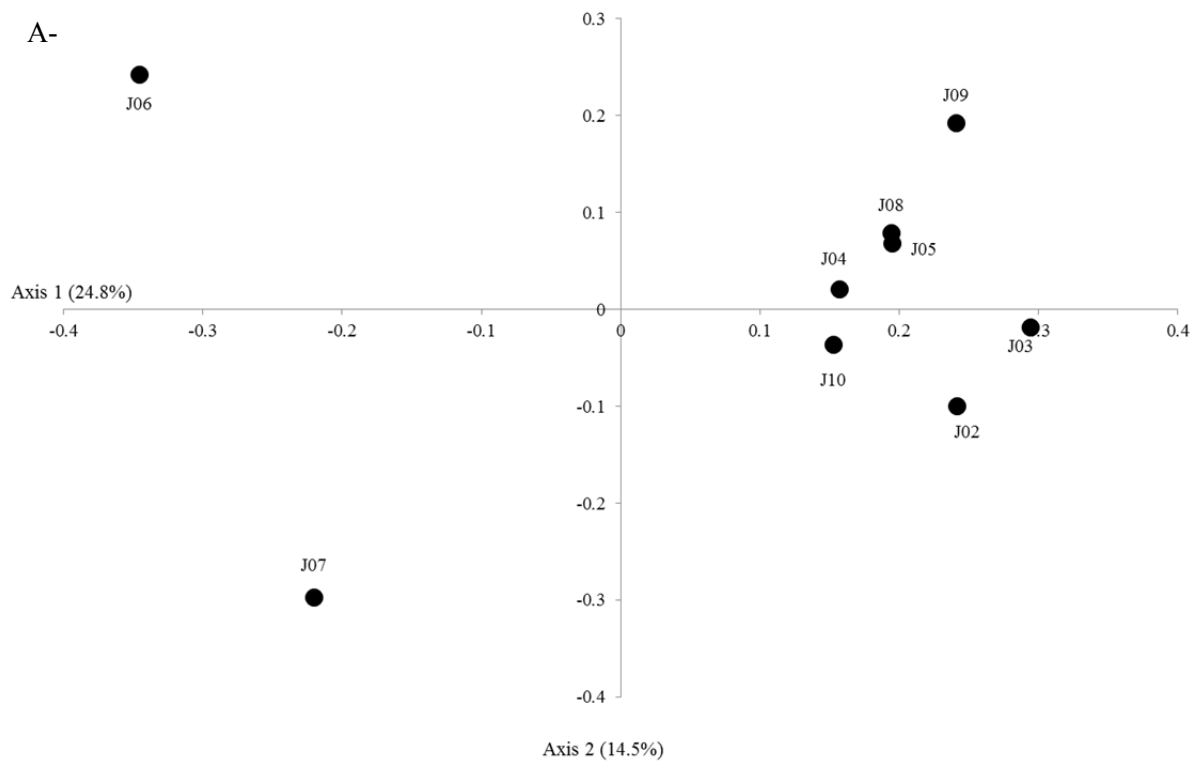


Fig. 1

**Appendix S1 Details of computations done for combining estimates of  $N_e$  across single-sample (SS) and temporal (T) methods.**

(formula and methods from Waples & Do 2010 – Appendix A)

Firstly, an overall point estimate for all single samples is obtained through a weighted harmonic mean. The computation of weights first needs the calculation of the variance of  $N_{e(SS)}$  over the timeframe of interest (here 2003-2007 and 2004-2008).

The variance is (equation A4 in Waples & Do 2010):

$$(1) \quad \text{Var}_{LD}(\hat{N}_e) \approx \frac{2N_e^2}{n} \left[ 1 + \frac{3N_e}{S} \right]^2.$$

with  $n$ : total number of independent allelic combinations in the sample (provided by equation (1) of Waples & Do 2010),  $S$ : the number of individuals sampled and  $N_e$ : harmonic mean of  $N_e$  unweighted harmonic mean over the years of interest.

Then, the weights for each sample are computed following the equation (A2 in Waples & Do 2010):

$$(2) \quad W_1 = \frac{1/V_1}{\sum_{i=1}^j 1/V_i}, W_2 = \frac{1/V_2}{\sum_{i=1}^j 1/V_i}, \dots, W_j = \frac{1/V_j}{\sum_{i=1}^j 1/V_i}.$$

with  $V_i$ , the variance previously estimated using formula (1).

Finally, a point estimate over the study period is computed as a weighted harmonic mean using the formula (equation A3 in Waples & Do 2010):

$$(3) \quad \tilde{N}_{e(\text{Weighted})} = \frac{1}{\sum_{i=1}^j W_i / \hat{N}_{e(i)}}.$$

Using our dataset, we obtained:

sample	Combined estimate across method 5 and 6	n	S	J03 – J07		J04 – J08	
				var <sub>LD</sub>	W	var <sub>LD</sub>	W
J03	-69.17	10475	41	3.45E-09	0.133		
J04	268.34	15029	93	2.40E-09	0.191	1.44E-09	0.181
J05	232.03	13901	71	2.60E-09	0.176	1.56E-09	0.167
J06	343.71	19983	135	1.81E-09	0.254	1.08E-09	0.241
J07	401.89	19377	144	1.86E-09	0.246	1.12E-09	0.233
J08	274.56	14748	75			1.47E-09	0.178
Combined $N_{e(SS)}$				1112.2		300.9	

So to summarize, the combined estimates with the two categories of methods are:

(values for  $N_{e(T)}$ , which is simply an unweighted harmonic mean over the 4 temporal methods, are given in Table 3 of the main text)

Time frame	$N_{e(T)}$	$N_{e(SS)}$
2002-2006	267.5	1111.2
2003-2007	250.9	300.9

We then combined estimates of  $N_e$  across single-sample (SS) and temporal (T) methods for the two different timeframe 2002-2006 and 2003-2007. The combination is based on a weighted harmonic mean, the weights being defined as a function of the variances.

For the single sample estimate the equation is given above (1; using harmonic mean of S and n over the study period and the combined estimate computed above) and for temporal estimates, the variance is

$$(4) \quad \text{Var}_T(\hat{N}_e) \approx \frac{2N_e^2}{n'} \left[ 1 + \frac{2N_e}{tS} \right]^2.$$

with  $n'$ : the number of independent alleles for temporal methods (provided by equation (5) of Waples & Do 2010),  $t$ : the time in generations elapsed between the two study dates,  $S$  the sample size (harmonic mean).

The values of the parameters, variances and weights are given below with the final combined estimates for the two timeframes:

	2002-2006	2003-2007
sample size (s)	49.8	49.8
n	13108.6	13108.6
n'	188.7	188.7
generation (t)	2.0	2.0
$\text{var}_{(N_{eSS})}$	869202.2	5050.1
$\text{var}_{(N_{eT})}$	30769.5	24317.4
$W_{SS}$	0.034	0.828
$W_T$	0.966	0.172
$N_{e(SS+T)}$	273.4	290.9

## Chapitre 3 – L’hypothèse de loterie familiale chez *Crepidula fornicata* : comparaison de la diversité génétique au cours des différentes étapes du cycle de vie

---

### I. Contexte de l’étude

Chez les invertébrés marins à cycle benthopélagique, la phase larvaire est une étape clé du cycle de vie assurant la dispersion, souvent sur de longues distances, et assurant le renouvellement des populations benthiques à l’issue de la métamorphose des larves. Cette étape est néanmoins sujette à d’importantes fluctuations environnementales et de forts taux de mortalité. Ces événements peuvent conduire à une forte variance du succès reproducteur individuel et à une diminution de la taille efficace des populations. Pour décrire ces processus, Hedgecock (1994a) a proposé le terme de « sweepstakes » (traduit ici par « loterie familiale »). De nombreuses études ont montré la vraisemblance de ces loteries familiales chez les espèces marines présentant une phase larvaire (voir pour une revue Hedgecock et Pudovkin 2011). Plusieurs prédictions sont proposées pour valider l’existence de loterie familiale. En particulier, parce que tous les individus ne contribuent pas de la même manière au renouvellement de la population, on s’attend entre autre à une perte de diversité génétique chez les larves et les juvéniles comparée à celle du pool reproducteur ainsi qu’une importante variance génétique entre groupes larvaires issus de différents épisodes de reproduction (*i.e.* cohortes larvaires). De plus, dans des cas particulièrement marqués de loterie familiale, on s’attend également à observer des larves apparentées au sein d’une cohorte. Un élément majeur dans l’étude du « sweepstakes » est donc de pouvoir comparer les étapes avant et après dispersion, et donc non seulement les recrues (juvéniles) mais également les larves. Les études focalisées sur les larves des invertébrés marins sont cependant encore peu nombreuses (mais voir Domingues *et al.* 2011 pour un exemple récent avec des larves de crabes). En effet, leur échantillonnage et leur analyse génétique reste difficile à réaliser.

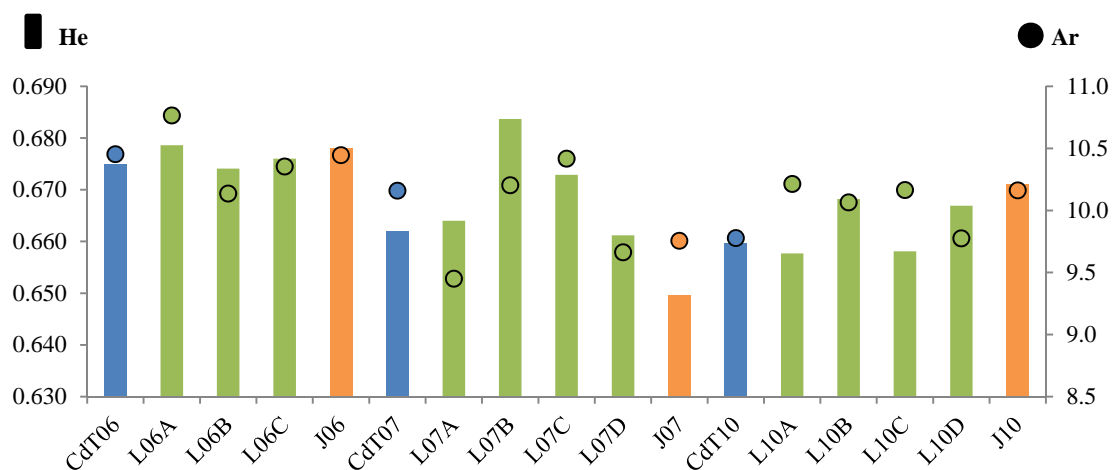
Grâce à des suivis pélagiques d’une population de *Crepidula fornicata*, réalisés dans l’équipe depuis plusieurs années, nous avons pu d’une part documenter les moments de présence des larves dans le milieu et avoir accès à un échantillonnage de jeunes larves (caractérisées par des tailles proches de la taille d’émission). Ceci nous a permis de tester différentes prédictions associées à l’hypothèse de loterie familiale chez le gastéropode *Crepidula fornicata*. Nous avons notamment comparé la composition génétique de plusieurs échantillons larvaires obtenus au cours d’une saison de reproduction entre eux ainsi qu’avec celle des juvéniles et des adultes de la même population. Cet échantillonnage en baie de Morlaix a été réalisé sur trois années, en 2006, 2007 et 2010. Afin de pouvoir interpréter les résultats de différences génétiques entre cohortes larvaires, plusieurs échantillons de populations adultes situées dans la même baie ainsi que dans deux populations de deux

baies adjacentes à la baie de Morlaix (baie de Saint-Brieuc et rade de Brest) ont également été intégrés à l'analyse. Au total, 1780 individus (869 adultes, 696 larves et 215 juvéniles) ont été génotypés sur 16 marqueurs microsatellites.

## II. Résultats principaux de l'étude (détaillés dans l'article 4)

Les larves et juvéniles montrent une structure génétique moins forte avec les adultes de la baie qu'avec les adultes des baies voisines, suggérant un recrutement de larves émises par les adultes locaux (*i.e.* intra-baie) et ce, malgré les fortes capacités de dispersion potentielle de l'espèce. Ces résultats vont dans le même sens que des résultats antérieurs ayant proposés que la baie de Morlaix soit exportatrice de larves et relativement « fermée » à l'immigration (Rigal *et al.* 2009, Dupont *et al.* 2007b).

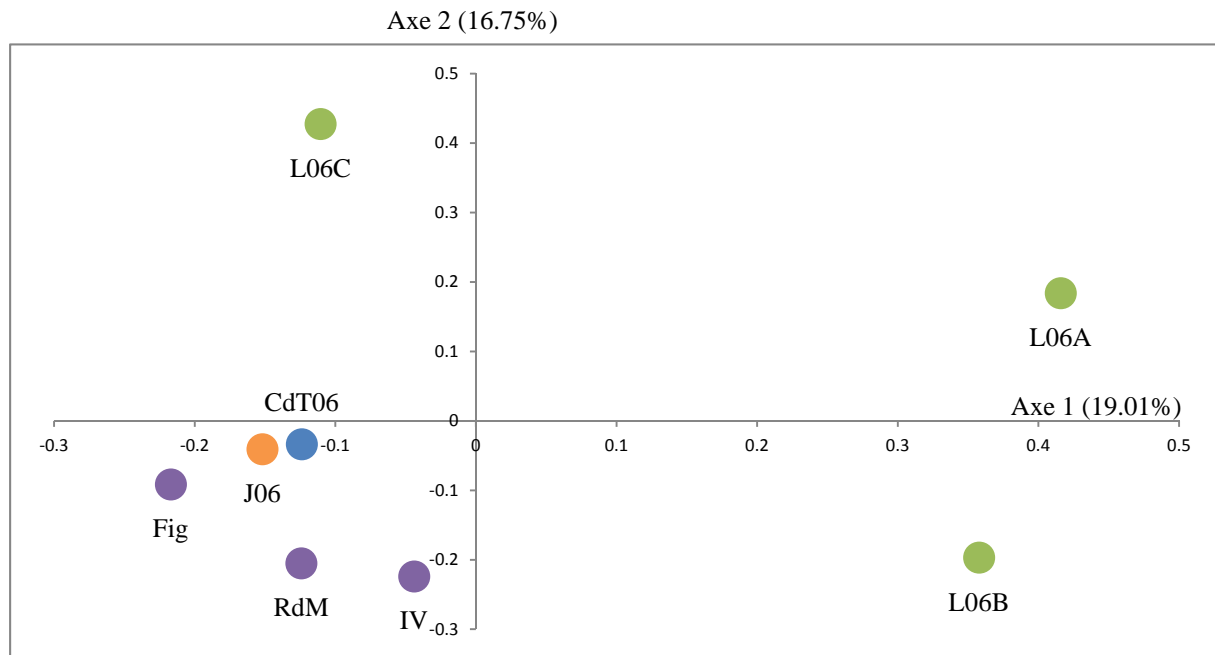
Contrairement à une des prédictions de l'hypothèse de loterie familiale, une forte diversité génétique a été observée, et ce quel que soit le stade de vie étudié (Fig. 1).



**Fig. 1** Diversité génétique (richesse allélique et hétérozygotie attendue) estimée par 16 marqueurs microsatellites pour différents échantillons du cycle de vie de *C. fornicata*. Les adultes (CdTOX) sont figurés en bleu, les larves (LOX) en vert et les juvéniles (JOX) en orange. L'année d'échantillonnage est indiquée par 06, 07 ou 10 dans le nom des échantillons et pour les larves, une lettre (A-D) différencie les échantillons au cours de la saison de reproduction, l'échantillon A étant systématiquement le premier de la saison.

En revanche, d'autres prédictions du « sweepstakes » sont vérifiées. En particulier, une structure génétique significative entre cohortes larvaires échantillonnées au cours de la saison de reproduction, a été observée. Cette structure génétique est supérieure à celle observée entre différentes populations d'adultes échantillonnées dans la même baie (Fig. 2 pour un exemple en 2006). Les larves sont également génétiquement différentes des juvéniles et des adultes. Comme attendu compte-tenu de la dynamique de recrutement de l'espèce, les différences génétiques entre juvéniles et adultes sont

quelquefois moins marquées qu'avec les larves ; les juvéniles sont en effet constitués de plusieurs cohortes juxtaposées localement.



**Fig. 2** Résultat de l'AFC réalisée avec les échantillons prélevés en baie de Morlaix en 2006. Trois cohortes larvaires (points verts), un échantillon des juvéniles recrutés (point orange) dans la population adulte cible (point bleu) et trois échantillons de trois localités où la crépidule est présente (en moindre densité) dans la baie ont été analysés avec 16 locus microsatellites. Cette AFC permet de visualiser les différences génétiques entre les trois échantillons de larves et les populations juvéniles et adultes du site étudié.

Enfin, des analyses d'apparentement ont mis en évidence que des larves, de petites tailles (436µm en moyenne, proches de la taille à l'émission), étaient pleine- ou demi-sœurs. Des juvéniles apparentés (demi-frères uniquement) ont également été observés, mais en une plus faible proportion comparée aux échantillons larvaires. La structure familiale observée chez les larves est « diluée » au cours de la phase de dispersion. Cependant, identifier des juvéniles apparentés pourrait suggérer l'existence de migration collective de larves issues d'un même évènement de reproduction, un résultat inattendu chez l'espèce étudiée dont la phase de dispersion peut durer jusqu'à sept semaines.

Ainsi, à l'exception de la diversité génétique stable entre les différents stades du cycle de vie, les résultats obtenus sont compatibles avec l'hypothèse de variations des pools reproducteurs et de variance de succès de reproducteur individuel au cours de la saison de reproduction. Ils sont ainsi en accord avec l'hypothèse d'une loterie familiale dans la population étudiée.

*Des résultats complémentaires, non inclus dans l'article 4, sont présentés dans l'annexe 12 (nombre d'individus apparenté en fonction du nombre de runs réalisé par le logiciel COLONY (Jones et Wang 2010)) et, l'annexe 13 (données d'apparentement de dix simulations de 100 plein-frères, 100 demi-frères et 100 individus non-apparentés avec un taux d'erreur de génotypage de  $10^{-3}$  réalisées par le logiciel Coancestry (Wang 2011)).*

**III. Article 4: « Genetic variance and kin-related individuals within the larval pool support the sweepstakes reproductive success hypothesis in the gastropod *Crepidula fornicata* »**

« Des différences génétiques et la présence d'individus apparentés dans le nuage larvaire sont en accord avec l'hypothèse de loterie familiale chez le gastéropode *Crepidula fornicata*»

Florentine Riquet, Thierry Comtet et Frédérique Viard – en préparation pour *Heredity*



**Genetic variance and kin-related individuals within the larval pool support sweepstakes reproductive success in the gastropod *Crepidula fornicata***

Florentine Riquet<sup>1,2</sup> – Thierry Comtet<sup>1,2</sup> – Frédérique Viard<sup>1,2</sup>

<sup>1</sup> UPMC Univ Paris 06, UMR 7144, Team Diversity and Connectivity in Coastal Marine Landscapes, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>2</sup> CNRS, UMR 7144, Lab. Adaptation and Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France

Correspondence: F. Riquet: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France; Tel: +33298292544, Fax: +33298292324; e-mail: [riquet@sb-roscoff.fr](mailto:riquet@sb-roscoff.fr)

F. Viard: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France; Tel: +33298292312, Fax: +33298292324; e-mail: [viard@sb-roscoff.fr](mailto:viard@sb-roscoff.fr)

Running title: Sweepstakes in *C. fornicata*

Keywords: *Crepidula fornicata*, gastropod, invasive species, relatedness, larval dispersal, sweepstakes

Abstract:

In many marine invertebrates, dispersal over large distance may be achieved during a pelagic larval phase. Although such a dispersal mechanism should result in high gene flow over large spatial scale, a significant genetic structure is commonly observed at a fine scale. This pattern has notably been attributed to variance in reproductive success across families. To examine such hypothesis, larvae released at the first stage of dispersal should be compared with the post-dispersal stages, i.e. benthic recruits and adults. Such data are however scarce in the literature because of the difficulty to sample and analyze larvae of minute size. In the present study, we investigated the genetic diversity of larvae of the highly dispersive and invasive marine gastropod *Crepidula fornicata*. The sampling of 3-4 larval pools were conducted in the field over the reproductive season and repeated three years. The genetic composition of the larval pools, obtained with 16 microsatellites loci, was compared to those of recruits and adults sampled in the same site and years, with reference to adults from adjacent bays. Altogether, 1780 larvae, juveniles and adults were examined. In contrast to expectations from the sweepstake hypothesis, no reduction of the genetic diversity of the larvae as compared to the parental pool was observed. Important temporal variations of the genetic composition of the larvae released at different time of the same reproductive season was however noticed. The genetic differences among the different batches of larvae were found to be higher than the variation observed among adults within the study bay or even with adults from an adjacent bay. Finally, we identified full and/or half sibs within the larval samples and juveniles. Altogether, our results support the occurrence of sweepstakes in the study species although with variable intensity across years. This study advocates for increasing temporal surveys over the whole life-cycle for a comprehensive understanding of the sweepstakes mechanisms in marine invertebrates.

## Introduction

Dispersal is a major evolutionary driver by modifying the genetic structure across populations. Weak genetic structures over large distances are expected for species with a long dispersal potential (e.g. in terrestrial species, Marthinsen *et al.* 2009, or in marine species, Shulman & Birmingham 1995). This assumption is particularly strong in marine invertebrates which display a larval stage (more than 70% of marine invertebrates (Mileikovski 1971)): with limited or no movements as adults, it is essentially the free-swimming larval phase that ensures dispersion. With a typical life span of a few weeks in the water column (Shanks 2009), larvae may disperse over large distances and contribute to high gene flow between populations as shown by several meta-analyses (e.g. Kinlan & Gaines 2003; Shanks 2009; Selkoe & Toonen 2011). Yet, significant genetic structure at a very local scale has also been observed. This puzzling pattern was first described by Johnson & Black (1982) who referred to an “apparently genetic chaotic patchiness” pattern. High genetic structure over very fine spatial scales was then documented in many marine invertebrate species such as bivalves (Hedgecock 1994; David *et al.* 1997), echinoderms (Banks *et al.* 2007) or crustaceans (Marino *et al.* 2010).

One of the main hypotheses that have been proposed to explain chaotic genetic patchiness is the “sweepstakes reproductive success” (SRS) hypothesis as coined by Hedgecock (1994): many marine species exhibit high fecundity and high early life-stages mortality, making them prone to high variance in reproductive success among individuals and in larval survival among families. The release of larvae in fluctuating environmental conditions may be viewed as a lottery in which a large random set of larvae will eventually die and thus a large set of individuals will eventually fail to effectively reproduce (Hedgecock 1994; Hedrick 2005). As reviewed by Hedgecock & Pudovkin (2011), this process leads to testable predictions in terms of genetic diversity, structure and relatedness from one generation to the other. First, the effective size of the population should be smaller than the census size, as the reproducing population is limited to a fraction of the whole population (Hauser *et al.* 2002). Second, if only a few individuals effectively reproduce, one would expect a reduced genetic diversity in offspring from a single reproduction event relative to the adults (e.g. Li & Hedgecock 1998). In addition, a shift in allelic frequencies from one reproductive event to the next is expected, as a result of the small, random and different set of individuals that succeed in reproducing at each event

(e.g. Lee & Boulding 2007). Besides, Hedgecock & Pudvokin (2011) pointed out that in cases of severe sweepstakes, full and half sibs are expected to be identified in the pool of dispersing larvae.

Studies that tested these predictions did not always reach to a consensus concerning the plausibility of the SRS hypothesis. For instance, Flowers *et al.* (2002) did not find significant differentiation between several cohorts of the sea urchin *Strongylocentrotus purpuratus*. Conversely, Moberg & Burton (2000) observed temporal genetic changes between recruits and adults in another sea urchin, *S. franciscanus*, giving support to the existence of sweepstake events. The lack of support to SRS hypothesis may be explained by the methods used, the fact that only some of the predictions were tested or that the intensity of SRS may be variable over time (i.e. “*the right place, but the wrong time*” as stated by Hedgecock & Pudovkin 2011). Frequency and level of SRS events may indeed vary from one year to the other, as exemplified by the important variation in the estimates of effective size of a population of *Crassostrea gigas* in Dabob Bay (WH, USA) that varied from c.a. 50 to 500 individuals over decades (Hedgecock & Pudovkin 2011). Altogether, there are numerous evidences for SRS to be widespread in the marine world (reviewed in Hedgecock & Pudvokin 2011).

The precise biological processes that lead to SRS are however still to be examined. An essential component when investigating the SRS hypothesis is to study pre- and post-dispersal stages, for instance in marine invertebrates larvae recently released by adults should be compared to recruits and adults. Studying pelagic eggs or larvae directly turns out to be extremely difficult notably due to their minute size (Levin 2006). Recruits are thus often favored to test SRS hypothesis predictions (e.g. Liu & Ely 2009, Christie *et al.* 2010 for recent studies ; but see for instance Li & Hedgecock 1998; Domingues *et al.* 2011 for studies on larval stages). Although important to study, recruits are providing an incomplete picture of the mechanisms behind SRS as they are post-dispersal stages. The effect of instantaneous genetic drift created by sweepstake reproductive events may be hidden by gene flow homogenizing larval cohorts in the water column during their development or by immigration of larvae from distant populations. In addition recruits may undergo post-settlement mortality (e.g. post-settlement selection as exemplified by Johnson & Black (1982)), modifying the genetic composition relative to that of the larval pool and, in some cases, also shading sweepstakes signatures.

In the present study, we examined the predictions of the SRS hypothesis by investigating the genetic composition of the larval, juvenile and adult pools of one emblematic invasive mollusc, the slipper limpet *Crepidula fornicata*. Our aim was to examine pre- and post-dispersal stages, over the whole reproductive season. Native to the northwestern Atlantic, *C. fornicata* is well established in the northeastern Pacific and in Europe, where it has been introduced in the late 19<sup>th</sup> and during the 20<sup>th</sup> centuries. In Europe, it is a major invader in some bays where it occurs at large densities (e.g. Blanchard 2009). This long-lived species (living up to 10 years) displays a typical marine benthopelagic life cycle alternating a long free-swimming larval stage (2-7 weeks) and a sessile adult stage. This mollusk is characterized by high fecundity with females releasing on average 5000-7500 swimming veliger larvae, several times per reproductive season (Richard *et al.* 2006). Fertilization is internal and females are brooding larvae before they hatch at a shell length of c.a. 400  $\mu\text{m}$  (Pechenik & Lima 1984) and metamorphose and settle at a length of c.a. 1000  $\mu\text{m}$  on a hard substrate (often on *C. fornicata* shells when present). *C. fornicata* is protandrous (i.e. changes sex from male to female; Coe 1938) and displays a stacking behavior, with piled-up individuals constituting stacks of two to 50 individuals with females at the base of the stack, males on the top. Reproduction occurs between males and females located within a given stack, each stack forming a local breeding group (Dupont *et al.* 2006; Proestou *et al.* 2008; Le Cam *et al.* 2009). The large population size, the long reproductive period, the high fecundity as well as the long larval phase with pulses of larvae released several times per reproductive season make this species well suited to investigate the SRS mechanisms. We analyzed the genetic diversity of the larval pool (3 to 4 samples per reproductive season) compared with juveniles (i.e. recruits of the year) and adults of *C. fornicata* sampled within a bay. The same sampling design was repeated in 2006, 2007 and 2010 for sake of comparison. These investigations were carried out in the bay of Morlaix (Brittany, France) where the species is established for decades. The population is exporting larvae but seems to receive a few immigrants from adjacent bays (Dupont *et al.* 2007b, Rigal *et al.* 2010). Following the SRS predictions, we hypothesized that the genetic diversity of the larval pool is reduced compared to the benthic individuals (both juveniles and adults) and that larvae display important genetic variation over time (both within and among reproductive seasons) as compared to juveniles and adults.

## Material and Methods

### Sampling design

*Larvae and benthic individuals in bay of Morlaix.* In the bay of Morlaix (Brittany, France), a monitoring over seven years (2005-2011) in the site “Chateau du Taureau” (CDT in Fig. 1) reported (i) the presence of brooding females from February to October with a peak on April-June with more than 70% of females brooding eggs (Le Cam 2009; F. Riquet & F. Viard, unpublished data), (ii) the presence of larvae from February to November (Rigal 2009; Leroy 2011) and (iii) a major recruitment event during late summer-early fall (Le Cam 2009, F. Riquet & F. Viard, unpublished data). Larvae, juveniles and adults of *Crepidula fornicata* were sampled in this site in 2006, 2007 and 2010 (table 1A). To be representative of the long reproductive period, larvae of *C. fornicata* were collected three to four times each year (from April to September, table 1A). They were sampled using a WP2 plankton net with a 200 $\mu$ m mesh size, in a 5-min oblique tow (from 10m depth to the surface). In the laboratory, larvae of *C. fornicata* were preserved in 96% ethanol; c.a. 60 larvae per date were then sorted and measured using a dissecting microscope with an ocular micrometer. In 2006, larval samples were separated by more than seven weeks (i.e. the maximum pelagic larval duration for *C. fornicata* as estimated by Rigal (2009)). In 2007 and 2010, samples were separated by c.a. 30 days (table 1A), a span time longer than the pelagic larval duration estimated for *C. fornicata* in the bay of Morlaix at the time of sampling (Rigal 2009). Besides, most of these larvae were small, with a mean size of 436.2 $\mu$ m  $\pm$ 71.7 (N=602 larvae), corresponding to larvae released recently. The different larval samples thus belonged to different cohorts (i.e. different reproduction events) and were likely released by local adults (Rigal *et al.* 2010). To sample the benthic phase, entire stacks of *C. fornicata* were collected along a 50m transect by scuba diving (12m depth at low tide). Adults were collected each year in April, the peak of the breeding period and matching as putative parents of the sampled larvae. Juveniles were collected during the major recruitment event (i.e. late summer in 2006 and 2007 and early winter in 2010). Altogether, the 215 juveniles had a mean size of 4.8mm  $\pm$ 1.7 (range: 1.0mm (in J06) - 8.8mm (in J06)). After collection benthic individuals (i.e. juveniles, males and females) were preserved in 96% ethanol. Labels of larva, juvenile and adult samples are given in table 1A.

*Adults from neighboring sites.* *C. fornicata* is established in other sites in the bay of Morlaix, although at much lower density and with lower larval abundance (Rigal *et al.* 2010). Because we cannot discard that larvae collected at Château du Taureau (CdT) could have been released in other part of the bay, adults from three other sites in the bay of Morlaix were examined using samples collected in 2006 and 2010. In addition, two sites in the two neighboring bays where *C. fornicata* is reported (bay of Brest and bay of Saint-Brieuc) were also examined using sampling available for 2002-2003 and 2010 (Fig. 1 and table 1B). The sampling procedures in 2002-2003 and 2006 are detailed in Dupont *et al.* (2007a) and Rigal *et al.* (2010) for the bays of Saint Brieuc and Brest and for the bay of Morlaix, respectively. Adults sampled in 2010 (i.e. the three bays) were collected by scuba diving at the same location than in 2002 and 2006. Individuals were preserved in 96% ethanol. Labels of these samples are given in table 1B.

#### **DNA extraction and microsatellite genotyping**

Except for the samples L06A and part of L06B, DNA extraction of larvae, adults and juveniles was performed using a Nucleospin® Multi-96 Tissue Kit (MACHEREY-NAGEL) following the manufacturer's protocol. Genomic DNA was extracted from (1) the whole larvae, (2) tissues and shell of juveniles less than 3mm in length, (3) whole tissues without the shell for juveniles above 3mm in length and (4) 20 mg of adults. Larval DNA was eluted in 60µL of elution buffer and DNA of adults and juveniles in 100µL. For the samples L06A and part of L06B, larval DNA was extracted using the Higuchi's (1989) method. Each individual was then genotyped at 16 microsatellite loci, 12 EST-SSRs: BI13YE17, CL191, CL322, DA5YM24, DA8YN14, CL270, DA1YN22, DA4YO03, CL291, CL83, DA1YL23, DA1YA23 (Riquet *et al.* 2011), and 4 anonymous markers: CfH7 (Dupont *et al.* 2006), Cf8 (Proestou 2006), CfCA4 and CfCA2, (Dupont & Viard 2003). Loci were amplified by polymerase chain reactions (PCR) following protocols detailed in Riquet *et al.* (2011), but with 2µL of stock DNA diluted 1:5 for larvae and 2µL of stock DNA diluted 1:20 for juveniles (adults: 1µL of stock DNA diluted 1:100). Amplification products were separated by electrophoresis on an ABI 3130xl DNA sequencer (Applied Biosystems™). Alleles were scored using GeneMapper® v. 4.0 (Applied Biosystems™) and manually checked. A total of 1780 individuals (869 adults, 696 larvae and 215

juveniles) were genotyped. A part of the whole dataset (8.4%) was genotyped twice, leading to the same results.

### **Data analyses**

Allele frequencies, the average number of alleles ( $N_{\text{all}}$ ) and allelic richness ( $A_r$ ) were estimated for each sample using FSTAT 2.9.3.2 (Goudet 1995). The software GENEPOP 4.1 (Rousset 2008) was used to estimate expected heterozygosity ( $H_e$ ), fixation index ( $F_{IS}$ ) and to test for departure from Hardy–Weinberg equilibrium in each sample (10,000 dememorization steps, 500 batches and 5,000 iterations per batch). To adjust the p-values for multiple tests, q-values were computed using the QVALUE package in the R software (Storey 2002). Indices of genetic diversity were compared among larvae, adults and juveniles as well as among adults from the three bays using a permutation procedure implemented in FSTAT (5,000 permutations).

Correspondence Analyses (CAs) implemented in GENETIX (Belkhir *et al.* 1996 - 2004) were used to picture the overall genetic distance between samples. This multivariate technique depicts how individuals are associated according to their individual multi-locus genotype. The genetic structure between samples was assessed by calculating an estimate of  $F_{ST}$  (Weir & Cockerham 1984) using GENEPOP 4.1. Exact tests for population differentiation (10,000 dememorization steps, 500 batches and 5,000 iterations per batch) were carried out to test for differences in allele distributions among samples and between pairwise samples.

We looked for clues supporting the hypothesis of larval collective dispersal, i.e. pools of larvae released from a local breeding group, here a stack, remaining aggregated during dispersal. The relatedness level between larvae, between juveniles and between adults was examined by computing a relatedness coefficient within each sample using the approach described in Wang (2007) and implemented in COANCESTRY 1.0.0.1 (Wang 2011). Briefly, this approach is based on a likelihood method to infer the relatedness coefficient between two individuals knowing allele frequencies. A third individual is used as a control to reduce the chance of genes identical by state to be mistakenly identified as identical by descent. COANCESTRY allows for inbreeding in the parental generation. Prior to the estimation of relatedness coefficients, simulations were performed with COANCESTRY



in order to 1) test for the robustness of the results taking into account different genotyping error rates and 2) assess variation in relatedness coefficients for three categories targeted in the analysis, namely full sibs, half sibs and unrelated individuals. For these simulations, population allele frequencies are needed: they were estimated using all the adults genotyped in the bay of Morlaix. Simulations remained robust for a large range of genotyping errors (Fig. S2) and we chose error rates of  $10^{-3}$  for all loci. Simulations also showed that the distributions of the relatedness coefficients overlap between the three categories. As we wanted to be conservative in the identification of full and half sibs, these simulations were used to define thresholds for which the pair of individuals could unambiguously be assigned as half-sibs, full-sibs or unrelated individuals (Fig. S3). Relatedness coefficients and their 95% confidence intervals were estimated with a bootstrap procedure (500 bootstraps). Differences in relatedness between samples were tested using a bootstrap method (10,000 bootstraps). In addition to the pairwise approach implemented in COANCESTRY, we used a more global approach, expected to be more powerful (Wang 2004; Wang 2009; Hedegcok & Pudovkin 2011; Wang 2012), to assign each individual to sib groups. We used the software COLONY 2.0.0.1 (Wang 2004; Jones & Wang 2010) which implements a maximum likelihood approach for partitioning pairs of individuals in three groups (full sibs, half sibs or unrelated) according to the population allele frequencies with groups modified until obtaining the best configuration (i.e. with the maximum likelihood). Like in the previous analysis, allele frequencies were estimated using all the adults sampled in the bay of Morlaix (2006, 2007 and 2010). The full-likelihood analysis with a high likelihood precision was used and allele frequencies were not updated during runs. We used a genotyping error as recommended by Wang (2004) and chose the same error rate than in COANCESTRY ( $10^{-3}$  for all loci). To check the reliability of results, for each sample, at least three runs (up to twelve runs for larval samples) were carried out with different seed numbers. Only pairwise relationships consistently inferred across runs (i.e.  $p > 0.90$  in each run) were retained for further discussions. The distribution of the pairwise relatedness coefficients were plotted against size characteristics (mean and difference) for larvae and juveniles.

## Results

*Genetic diversity across life cycle.* Significant departures from Hardy-Weinberg equilibrium were found in each sample, including larvae (e.g. L06A), juveniles (e.g. J07) or adults (e.g. CdT10). Seven out of the 16 loci showed higher heterozygote deficiencies that may suggest the presence of null alleles. All the analyses were performed with and without these seven loci. A strong correlation was observed between indices of diversity estimated using all the loci or only nine loci (Supporting Information –Fig. S1). The results presented are those based on the 16 loci. Table 1 provides the estimates of genetic diversity for each sample. Whatever the year and month within a year, the larval and juvenile samples displayed a high genetic diversity, similar to that of adults (table 1A): no significant difference was evidenced between larvae, adults and juveniles either in allelic richness ( $p=0.99$ ) or in expected heterozygosity ( $p=0.98$ ). This genetic diversity was also very close to the values observed for adults from the bay of Saint-Brieuc (table 1B;  $p=0.84$ ,  $p=0.53$  for  $A_r$  and  $H_e$ , respectively). Conversely, adults from Brest showed a significant reduced allelic richness as compared to the adults from the bays of Morlaix and Saint-Brieuc ( $p=0.004$ ; table 1B), which was not supported when considering the expected heterozygosity ( $p=0.16$ ).

*Genetic structure between larval samples and comparison with the genetic structure between adult populations.* Results of the genetic structure analyses were similar with the 16 loci or when excluding 7 loci with putative null alleles (see above). They are described with the analyses made using 16 loci. Within each year, a high genetic variation among larval cohorts was observed. This is featured by the Correspondence Analysis (CA) with larval samples scattered over the first two axes (Fig. 2). This high heterogeneity was further assessed by  $F_{ST}$  estimates which showed a significant genetic differentiation between the larval samples collected within a year (2006:  $F_{ST}=0.005$ ,  $p<0.001$ ; 2007:  $F_{ST}=0.003$ ,  $p<0.001$ , 2010:  $F_{ST}=0.0002$ ,  $p<0.001$ ). This also holds in pairwise comparison for the samples collected in 2006 and 2007 (table 2A and B). A slightly different pattern was observed in 2010 with a genetic homogeneity between the three samples collected in the second part of the reproductive season (i.e. L10B-D in Fig. 2C, Table 2C). The three larval samples were however significantly different from the first larval sample (L10A; Fig 2C and Table 2C).

Adults from the bay of Morlaix clustered together (Fig. 2A, C) and  $F_{ST}$  estimates showed a genetic differentiation among sites within the bay (2006:  $F_{ST}=0.002$ ,  $p<0.001$ , 2010:  $F_{ST}=-0.0005$ ,  $p<0.001$ ). But the difference is lower than among larval samples collected the same year over the reproductive season. When considering pairwise comparisons (Table 2), no significant difference was evidenced between sites within the bay of Morlaix (in 2006: Table 2A, in 2010: Table 2C). Spatial genetic differences among adult populations are significant at the regional scale: the genetic differentiation among the three bays ( $F_{ST-3bays}=0.006$ ,  $p<0.001$ ) was higher than within a bay (Table 2) and among larval samples. CA further showed that larvae clustered with adults of the same bay along the first axis, pointing out a higher genetic similarity of larvae to the adults of the same bay than to adults of adjacent bays. This was also supported by pairwise  $F_{ST}$  estimates (Table 2).

*Family structure across life cycle.* Within each year, one or two larval or juvenile samples (L06B, L07A, J07 and J10) exhibited significantly higher values of mean relatedness coefficient than other samples. The reproductive pool (adults) exhibited a similar mean relatedness coefficient over the three years examined (0.025-0.027;  $p>0.05$ ). Pairwise relatedness coefficients (i.e. between individuals) were highly variable, ranging from 0 to 0.69. Similar ranges were observed among adults (CdT06: 0-0.38; CdT07 and CdT10: 0-0.36) but the ranges were variable among larval or juvenile samples (Fig. 3A). Based on the relatedness coefficient, most of the individual pairs could be assigned as unrelated individuals, with a mean of 79.3% over the samples, a minimum of 68.0% in J07 and a maximum of 82.6% in L10B (Fig. 3B). The frequency of pairs that could not be assigned as being unrelated or related ranged from 10.7% (L06A) to 19.9% (J07) across the samples, with a mean over samples of 13.1% (Fig. 3B). A slight proportion was made of unambiguous kin related individuals (half- or full sibs), ranging from 5.6% in L10C to 12.8% in L07A (mean over samples of 7.6%) (Fig. 3B). Six samples (L07B, L07D, L10B, J06, J07 and J10) showed unambiguously pairs of full sibs, which occurred at frequencies less than 1% (Fig. 3B). The same was observed with the COLONY software: three pairs of full sibs (all in the larval sample L10B) and 38 pairs of half sibs were inferred within larval samples (Fig. 4). The three pairs of full-sibs are composed of three larvae forming a trio of sister larvae. Interestingly, when analyzing the relationships over the whole reproductive period (i.e.

analyzing at once all the larval samples collected the same year), 19, 23 and 33 pairs of half-sibs were additionally identified in 2006, 2007 and 2010, respectively. Half-sibs were also observed in juvenile samples with 12 pairs (four in 2006, five in 2007 and three in 2010). No full-sibs were identified among juveniles. Note that no full or half sibs was identified in adult samples.

For larvae and juveniles, the individual pairwise relatedness was examined as a function of their mean size and pairwise size difference. In figure 5, the pairwise relatedness of every sample analyzed is pictured. The highest values of larval pairwise relatedness were observed when the mean size of the dyads was closed to c.a. 400 $\mu$ m (i.e. size at hatching, Fig. 5A) and when size differences tended to be smaller than 50 $\mu$ m (Fig. 5B). In juveniles, we also observed that most closely related individuals displayed low differences in size, with no trend for the mean size (Fig. 5C-D). Some pairs of relatives however also displayed large differences in size (e.g. relatedness of 0.38 with a size difference of 400 $\mu$ m, Fig. 5A-B).

## Discussion

In the present study of the slipper limpet *Crepidula fornicata*, we compared the genetic composition of 3-4 larval cohorts collected during one reproductive season between them and with that of juveniles and adults from the same site. One prediction of the sweepstakes reproductive success (SRS) hypothesis is that larvae and juveniles should have less diversity as compared to adults, which was not detected here. This prediction is hardly tested (Hedgecock & Pudovkin 2011) and, when tested, less diversity in larvae or juveniles relative to the adults is hardly found (e.g. Ruzzante *et al.* 1996 but see Hedgecock *et al.* 2007). In some early allozymes studies, results were consistent with this SRS prediction (Hedgecock & Pudovkin 2011) but apparently more difficult to obtain with highly polymorphic microsatellites. Conversely, in agreement with other predictions of SRS, in each of the three study years, we observed significant genetic differences between larval pools. To interpret genetic differences between larval cohorts however requires using genetic differences between adults as a reference. A weak genetic structure was observed between adults collected at different sites within a single bay (e.g. sites within the bay of Morlaix or sites within the bay of Brest). Most of the differences at the adult stage were observed between bays. In particular, the bay of Brest showed large

genetic differences with the two other bays. These results are in agreement with a previous study which investigated the genetic structure among populations of *C. fornicata* at a regional scale using four microsatellite loci (Dupont *et al.* 2007b). The strong genetic divergence between the bay of Brest and the other bays has been explained by the semi-enclosed nature of this bay which may impede gene flow (Richard *et al.* 2006; Dupont *et al.* 2007b). A few exceptions to the inter-bay differences and within-bay homogeneity were however noticed. Two sites of the bay of Saint-Brieuc showed genetic differentiation, as previously observed in Dupont *et al.* (2007b). Non-significant genetic differences between some sites of Morlaix and Saint-Brieuc (e.g. between Fig10 in Morlaix and SB1\_10 in St-Brieuc) (Table 2C) were also observed and might be explained by episodic long-distance dispersal events between these two bays. In the bay of Morlaix, Rigal *et al.* (2010) has shown that larvae can be exported outside the bay. Some larvae released in Morlaix could thus disperse towards the bay of St-Brieuc, a pattern however not supported by the outcome of a 2D-Lagrangian model of larval dispersal in the western Channel (Dupont *et al.* 2007b). The relative genetic homogeneity between sites of different bays does not oppose to the high larval genetic heterogeneity observed in this study as they can be the outcome of processes operating at different scales: gene flow operating at regional spatial scales through larval dispersal vs. variance in adults' reproductive success and larval mortality operating over shorter temporal and local scale. A similar scenario was proposed by Christie *et al.* (2010) when examining temporal and spatial genetic variation in a coral-reef fish.

Within the bay of Morlaix, the temporal variation in allele frequencies observed between larval pools within a reproductive season was found to be higher than the spatial genetic variation among adults sampled in four sites of this bay. This may be caused either by fluctuations in source of larvae, some of them being immigrants from adjacent bays (e.g. Selkoe *et al.* 2006) and/ or by locally-produced larvae originating from different groups of adults. The former hypothesis can be ruled out as 1) the studied larvae are all very young, with size close to the size at hatching, making them unlikely to come from adjacent bays and 2) as pointed above, the study bay was shown to receive little immigrants from the adjacent bays (Dupont *et al.* 2007b; Rigal *et al.* 2010). This pattern was however not fully expected as Richard *et al.* (2006) showed that a given female of *C. fornicata* can reproduce

several times per reproductive season. In our study site, more than 70% of brooding females were observed over several months per reproductive season (Le Cam 2009). Similar data (i.e. long reproductive period) have been obtained in other populations of *C. fornicata* within its native and introduced ranges (Hoagland 1978; Thieltges *et al.* 2004; Richard *et al.* 2006). These monitoring data suggest that many females can contribute to several larval cohorts over the reproductive period. Furthermore, the paternal reproductive pool was also expected to be similar over time as females store and use the sperm transferred by several males for several months (e.g. Le Cam *et al.* 2009 and references therein). Despite these reproductive characteristics, we did observe important changes in allelic frequencies of the study larval samples suggesting that there is large variance in individual reproductive success.

The significant genetic differentiation between larval samples observed over the reproductive season is the first clue supporting the SRS hypothesis (Hedgcock 1994): although all individuals may theoretically contribute to different larval pools, individuals that effectively contribute vary from one reproductive event to the other. This result is in agreement with a previous study (Riquet *et al.* in prep<sup>1</sup>) which examined genetic changes of juveniles recruited over nine years and showed a reduced effective size relative to the census size in the same study population. Similar observations were documented in Pacific oyster (*Crassostrea gigas*) larvae by Li & Hedgcock (1998) who demonstrated a shift in allelic frequencies over the reproductive season, especially for early and late samples that differed from the other samples. Furthermore, in the present study, the intensity of SRS events varied from one year to the other. The variance in contribution of the reproductive pool to the larval pool seemed less important in 2010 as compared to 2006 and 2007. In 2010, no significant genetic differentiation was observed among the three last larval samples, suggesting that a part of the reproductive pool contributed “equally” to these larval pools. Hedgcock & Pudovkin (2011) pointed out large variations in the effective size of the Pacific oyster population in Dabob Bay (WH, USA) that they attributed to the stochasticity in oceanic conditions. Less fluctuating environmental conditions and a larger reproductive pool during the 2010 summer may explain a slighter genetic drift during the

---

<sup>1</sup> Deuxième Partie – Chapitre 2

dispersal phase and a weaker signal of sweepstake events relative to those of the two other studied years.

Some of the larvae and juveniles were either half- or full-sibs. A higher proportion of relatives was however expected with COLONY based on COANCESTRY results (5.6 to 12.8%). The full and half sibs that we retained with the Colony software were those pairs that were consistently observed in every run (up to 12 runs for larval samples). This conservative approach reduced the number of dyads identified from c.a. 7.3% pairs per sample in one run to less than 1% after considering every run. Although it may have underestimated the number of relatives, it greatly improved the reliability of relatedness inferred by Colony. In addition to genetic differentiation between temporally distinct larval supply events, another expected signature of sweepstakes (Hedgecock & Pudovkin 2011) was thus documented by the presence of close relatives within larval and juvenile pools (but not in adults) over the three studied years. Examining relatedness within invertebrate larval pools, an important prediction of the SRS hypothesis, has been done to our knowledge only once by Domingues *et al.* (2011) studying the megalopea larvae of a crab species. These authors however did not find any evidences of family relatedness. Finding related offspring in young larvae is not only compatible with SRS but also gives support to some form of collective dispersal (i.e. concerted migration of related offspring). Such collective dispersal can have some influence on a commonly reported pattern in marine invertebrates, the so-called “chaotic genetic patchiness”. Using analytical models and simulations, Broquet *et al.* (in revision) indeed showed that, when migration is not limited and in the absence of overlapping generations, such collective dispersal may facilitate chaotic genetic patchiness.

The variation in relatedness coefficients across stages (i.e. from larvae to juveniles) reflect a family-structure that is diluted once larvae have been dispersed and metamorphosed. Juveniles are indeed composed of a series of cohorts that have settled during the reproductive season and the genetic signals of SRS are weakened (Hedgecock & Pudvokin 2011). For instance, Planes & Lenfant (2002) suggested that cohorts of recruits of the marine fish *Diplodus sargus* were actually made up by different families assembling during the larval development or after the settlement. The presence of kin-related juveniles in our study also suggests that some kin related larvae can disperse together after

their release in the water column until settlement, which is surprising considering that the larvae can disperse up to seven weeks in the study area. It may be due to larvae trapped in the same water mass (Siegel *et al.* 2003) and/or to an active behavior of larvae. The existence of aggregation of siblings during dispersal has been mainly documented in fishes (e.g. Selkoe *et al.* 2006; Miller-Sims *et al.* 2008; Buston *et al.* 2009; Pujolar *et al.* (2009); Christie *et al.* (2010)) but also shown in some marine invertebrates (e.g. Veliz *et al.* (2006); Taris *et al.* (2009)). In many cases, for fishes, active larval behavior has been documented to be based on auditory, olfactory or other cues has been shown (e.g. Kingsford *et al.* 2002; Montgomery *et al.* 2006). For instance, Pujolar *et al.* (2009) calculated genetic relatedness among *Anguilla anguilla* glass eels and found c.a. 1 to 3% of close relatives in new arrival waves plus a majority of unrelated individuals. Planes *et al.* (2002) examined relatedness of settling larvae of *Naso unicornis* in the reefs of Moorea and showed that larvae of the same spawning dates were related, suggesting a recruitment of siblings.

As expected with the assumption of collective dispersal, related individuals often showed a similar size thus supporting the idea that they originate from a single event of reproduction. Based on our data, we cannot ascertain the mean duration of kin aggregation during the dispersive phase since the larvae are close to the size at hatching (i.e. 400 $\mu$ m). The few pairs of related individuals displaying differences in size can be explained by two different hypothesis: (1) variance in larval growth within a given offspring as evidenced by Le Cam *et al.* (2009) and (2) that the same adults reproduced several time over the reproductive period which is consistent with the reproductive biology of the studied species and that by chance their larvae were sampled. In addition that some relatives showed large difference in size (up to 400-500  $\mu$ m; Fig. 5) suggested that the larvae released after several reproductive events and separated by 3-5 weeks all stayed within the bay. Rigal *et al.* (2010) using an analytical model however showed low retention rates in the bay of Morlaix. Additional sampling of larvae of different size (age) is needed to monitor the time during which collective dispersal is acting and its intensity in the wild. Ideally, a single larval cohort should be sampled over the course of its development. This is however very difficult to carry out with species, like *C. fornicata*, which have



larvae with a long-development time and because sampling of large (old) larvae is hard (Rigal *et al.* 2010).

## **Conclusion**

Our results support a number of SRS predictions in the invasive gastropod *Crepidula fornicata* with notably genetic differences among larval cohorts and the existence of kin related individuals in larvae and juveniles. This study however showed that the footprint is variable across years advocating for temporal surveys in similar studies. We also demonstrated the plausibility of collective dispersal during the first stage of larval dispersal, a pattern hardly investigated in marine invertebrates. We however focused on very early stages of larval dispersal. The mechanisms and the importance of such correlated dispersal of kin individuals deserve further examination in marine invertebrates.

## **Acknowledgment**

We are grateful to Sabrina Le Cam and the divers of “Marine Operations and Monitoring” Department of the Station Biologique of Roscoff for providing adults and juveniles sampling and data, as well as François Rigal, Fanny Leroy and the crew of the Neomysis (“Marine Operations Department” - the Roscoff Biological Station) for providing larvae. We thank the Biogenouest® Genomics core facility for its technical support. We are also thankful to E. Billard for help in genotyping and to J. Wang for advices for computation. This work was funded by the Agence Nationale de la Recherche (Project Hi-Flo, ANR-08-BLAN-0334), the Interreg IVa Marinexus programme and the AXA Research Funds (Marine Aliens and Climate Change project). FR acknowledges a PhD grant from the Centre National pour la Recherche Scientifique and the Conseil Régional de Bretagne (SEA-FLO ARED project).

## **References**

Banks SC, Piggott MP, Williamson JE, Bove U, Holbrook NJ, Beheregaray LB (2007) Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. *Ecology*, **88**, 3055-3064.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Blanchard M (2009) Recent expansion of the slipper limpet population (*Crepidula fornicata*) in the Bay of Mont-Saint-Michel (Western Channel, France). *Aquatic Living Resources*, **22**, 11-19.

Buston PM, Fauvelot C, Wong MYL, Planes S (2009) Genetic relatedness in groups of the humbug damselfish *Dascyllus aruanus*: small, similar-sized individuals may be close kin. *Molecular Ecology*, **18**, 4707-4715.

Christie MR, Johnson DW, Stallings CD, Hixon MA (2010) Self-recruitment and sweepstakes reproduction amid extensive gene flow in a coral-reef fish. *Molecular Ecology*, **19**, 1042-1057.

Coe WR (1938) Conditions influencing change of sex in mollusks of the genus *Crepidula*. *Journal of Experimental Zoology*, **77**, 401-424.

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

Domingues CP, Creer S, Taylor MI, Queiroga H, Carvalho GR (2011) Temporal genetic homogeneity among shore crab (*Carcinus maenas*) larval events supplied to an estuarine system on the Portuguese northwest coast. *Heredity*, **106**, 832-840.

Dupont L, Bernas D, Viard F (2007a) Sex and genetic structure across age groups in populations of the European marine invasive mollusc, *Crepidula fornicata* L. (Gastropoda). *Biol J Linn Soc*, **90**, 365-374.

Dupont L, Ellien C, Viard F (2007b) Limits to gene flow in the slipper limpet *Crepidula fornicata* as revealed by microsatellite data and a larval dispersal model. *Marine Ecology Progress Series*, **349**, 125-138.

Dupont L, Richard J, Paulet YM, Thouzeau G, Viard F (2006) Gregariousness and protandry promote reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of larval paternity. *Mol Ecol*, **15**, 3009-3021.

Dupont L, Viard F (2003) Isolation and characterization of highly polymorphic microsatellite markers from the marine invasive species *Crepidula fornicata* (Gastropoda : Calyptraeidae). *Mol Ecol Notes*, **3**, 498-500.

Flowers JM, Schroeter SC, Burton RS (2002) The recruitment sweepstakes has many winners: Genetic evidence from the sea urchin *Strongylocentrotus purpuratus*. *Evolution*, **56**, 1445-1453.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485-486.

Hauser L, Adcock GJ, Smith PJ, Ramirez JHB, Carvalho GR (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11742-11747.

Hedgcock D (1994) Does variance in reproductive success limit effective population sizes of marine organisms? In: *Genetics and Evolution of Aquatic organism* (ed. Beaumont AR). pp. 122-135, Chapman and Hall, London.

Hedgcock D, Launey S, Pudovkin AI, Naciri Y, Lapegue S, Bonhomme F (2007) Small effective number of parents ( $N_b$ ) inferred for a naturally spawned cohort of juvenile European flat oysters *Ostrea edulis*. *Marine Biology*, **150**, 1173-1182.

- Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bulletin of Marine Science*, **87**, 971-1002.
- Hedrick P (2005) Large variance in reproductive success and the  $N_e/N$  ratio. *Evolution*, **59**, 1596-1599.
- Higuchi R (1989) Simple and rapid preparation of samples for PCR. *Erlich HA, editor. PCR technology, principles and applications for DNA amplification. New York: Stockton Press.*, 31-38.
- Hoagland KE (1978) Protandry and evolution of environmentally-mediated sex change - study of mollusca. *Malacologia*, **17**, 365-391.
- Johnson MS, Black R (1982) Chaotic genetic patchiness in an inter-tidal limpet, siphonaria sp. *Marine Biology*, **70**, 157-164.
- Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, **10**, 551-555.
- Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: A community perspective. *Ecology*, **84**, 2007-2020.
- Kingsford MJ, Leis JM, Shanks A, Lindeman KC, Morgan SG, Pineda J (2002) Sensory environments, larval abilities and local self-recruitment. *Bulletin of Marine Science*, **70**, 309-340.
- Le Cam S (2009) Grégarité, changement de sexe et polyandrie: modalités de la reproduction chez une espèce invasive *Crepidula fornicata*. *Thèse de doctorat. Université Pierre et Marie Curie, Paris.*,
- Le Cam S, Pechenik JA, Cagnon M, Viard F (2009) Fast versus slow larval growth in an invasive marine mollusc: does paternity matter? *Journal of Heredity*, **100**, 455-464.
- Lee HJ, Boulding EG (2007) Mitochondrial DNA variation in space and time in the northeastern Pacific gastropod, *Littorina keenae*. *Molecular Ecology*, **16**, 3084-3103.
- Leroy F (2011) Influence des conditions trophiques sur le développement larvaire de l'espèce invasive *Crepidula fornicata*. Conséquences sur sa capacité d'expansion. *Thèse de doctorat. Université Pierre et Marie Curie, Paris.*,
- Levin LA (2006) Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology*, **46**, 282-297.
- Li G, Hedgecock D (1998) Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1025-1033.
- Liu J-X, Ely B (2009) Sibship reconstruction demonstrates the extremely low effective population size of striped bass *Morone saxatilis* in the Santee-Cooper system, South Carolina, USA. *Molecular Ecology*, **18**, 4112-4120.
- Marino IAM, Barbisan F, Gennari M, Giomi F, Beltramini M, Bisol PM, Zane L (2010) Genetic heterogeneity in populations of the Mediterranean shore crab, *Carcinus aestuarii* (Decapoda, Portunidae), from the Venice Lagoon. *Estuarine Coastal and Shelf Science*, **87**, 135-144.
- Marthinsen G, Wennerberg L, Solheim R, Lifjeld JT (2009) No phylogeographic structure in the circumpolar snowy owl (*Bubo scandiacus*). *Conservation Genetics*, **10**, 923-933.

- Mileikovski S (1971) Types of larval development in marine bottom invertebrates, their distribution and ecological significance - re-evaluation. *Marine Biology*, **10**, 193-213.
- Miller-Sims VC, Gerlach G, Kingsford MJ, Atema J (2008) Dispersal in the spiny damselfish, *Acanthochromis polyacanthus*, a coral reef fish species without a larval pelagic stage. *Molecular Ecology*, **17**, 5036-5048.
- Moberg PE, Burton RS (2000) Genetic heterogeneity among adult and recruit red sea urchins, *Strongylocentrotus franciscanus*. *Marine Biology*, **136**, 773-784.
- Montgomery JC, Jeffs A, Simpson SD, Meekan M, Tindle C (2006) Sound as an orientation cue for the pelagic larvae of reef fishes and decapod crustaceans. **51**, 143-196.
- Pechenik JA, Lima GM (1984) Relationship between growth, differentiation and length of larval life for individually reared larvae of the marine gastropod, *Crepidula fornicata*. *Biological Bulletin*, **166**, 537-549.
- Planes S, Lecaillon G, Lenfant P, Meekan M (2002) Genetic and demographic variation in new recruits of *Naso unicornis*. *Journal of Fish Biology*, **61**, 1033-1049.
- Planes S, Lenfant P (2002) Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. *Molecular Ecology*, **11**, 1515-1524.
- Proestou DA (2006) Isolation and characterization of microsatellite markers in the Atlantic slipper shell *Crepidula fornicata* for use in paternity analysis. *Molecular Ecology Notes*, **6**, 437-439.
- Proestou DA, Goldsmith MR, Twombly S (2008) Patterns of male reproductive success in *Crepidula fornicata* provide new insight for sex allocation and optimal sex change. *Biological Bulletin*, **214**, 192-200.
- Pujolar JM, Maes GE, Volckaert FAM (2006) Genetic patchiness among recruits in the European eel *Anguilla anguilla*. *Marine Ecology-Progress Series*, **307**, 209-217.
- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Richard J, Huet M, Thouzeau G, Paulet YM (2006) Reproduction of the invasive slipper limpet, *Crepidula fornicata*, in the Bay of Brest, France. *Marine Biology*, **149**, 789-801.
- Rigal F (2009) Dynamique spatio-temporelle du nuage larvaire du gastéropode introduit *Crepidula fornicata* au sein d'une baie mégatidale, la baie de Morlaix (France). *Thèse de doctorat. Université Pierre et Marie Curie, Paris.*
- Rigal F, Viard F, Ayata SD, Comtet T (2010) Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? *Biological Invasions*, **12**, 3171-3186.
- Riquet F, Ballenghien M, Tanguy A, Viard F (2011) *In silico mining* and characterization of 12 EST-SSRs for the invasive slipper limpet *Crepidula fornicata*. *Marine Genomics*, **4**, 291-295.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour*, **8**, 103-106.

- Ruzzante DE, Taggart CT, Cook D (1996) Spatial and temporal variation in the genetic composition of a larval cod (*Gadus morhua*) aggregation: Cohort contribution and genetic stability. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 2695-2705.
- Selkoe KA, Gaines SD, Caselle JE, Warner RR (2006) Current shifts and kin aggregation explain genetic patchiness in fish recruits. *Ecology*, **87**, 3082-3094.
- Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology-Progress Series*, **436**, 291-305.
- Shanks AL (2009) pelagic larval duration and dispersal distance revisited. *Biological Bulletin*, **216**, 373-385.
- Shulman M, Bermingham E (1995) Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. *Evolution*, **49**, 897-910.
- Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology-Progress Series*, **260**, 83-96.
- Storey J (2002) A direct approach to false directory rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479-498.
- Taris N, Boudry P, Bonhomme F, Camara MD, Lapegue S (2009) Mitochondrial and Nuclear DNA Analysis of Genetic Heterogeneity Among Recruitment Cohorts of the European Flat Oyster *Ostrea edulis*. *Biological Bulletin*, **217**, 233-241.
- Thieltges DW, Strasser M, van Beusekom JEE, Reise K (2004) Too cold to prosper - winter mortality prevents population increase of the introduced American slipper limpet *Crepidula fornicata* in northern Europe. *Journal of Experimental Marine Biology and Ecology*, **311**, 375-391.
- Veliz D, Duchesne P, Bourget E, Bernatchez L (2006) Genetic evidence for kin aggregation in the intertidal acorn barnacle (*Semibalanus balanoides*). *Molecular Ecology*, **15**, 4193-4202.
- Wang J (2004) Sibship reconstruction from genetic data with typing errors. *Genetics*, **166**, 1963-1979.
- Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetics Research*, **89**, 135-153.
- Wang J, Santure AW (2009) Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, **181**, 1579-1594.
- Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, **11**, 141-145.
- Wang J (2012) Computationally efficient sibship and parentage assignment from multilocus marker data. *Genetics*, **191**, 183-94.
- Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

## Figures and Tables

**Fig. 1** Sampling locations. A- Study area in Brittany with the bay of Brest (sites BrK and BrZ) and the bay of Saint-Brieuc (sites SB1 and SB104) and B- the bay of Morlaix (sites Le Figuier (Fig), Ile Verte (IV), RdM: Rivière de Morlaix and Château du Taureau (CdT)).

**Fig. 2** Correspondence analyses conducted with 16 microsatellite markers on samples (larvae, adults and juveniles) collected in A- 2006 (numbered 1 to 8), B- 2007 (numbered 9 to 14) and C- 2010 (numbered 15 to 27).

The life-cycle stages are featured as follows: adults with circles, larvae with squares and juveniles with triangles. Samples are numbered chronologically as follow: 1. Fig06, 2. IV06, 3. RdM06, 4. CdT06, 5. L06A, 6. L06B, 7 L06C, 8. J06, 9. CdT07, 10. L07A, 11. L07B, 12. L07C, 13. L07D, 14. J07, 25. BrK10, 16 BrZ10, 17. SB1\_10, 18. SB104\_10, 19. Fig10, 20. IV10, 21. RdM10, 22. CdT10, 23. L10A, 24. L10B, 25. L10C, 26. L10D, 27. J10. In C, the samples are colored as follows: bay of Brest: dark grey, bay of Saint Brieuc: light grey, bay of Morlaix: white.

**Fig. 3** Pairwise relatedness coefficients for each sample (adults, larvae and juveniles). A- Pairwise relatedness distribution in each sample. The mean value per sample is given in the bracket below the sample code (\* stands for samples showing a mean value statistically different from the others). B- Frequency of pairs of individuals distributed over different categories defined according to the simulations results (see text and Fig. S3).

**Fig. 4** Kin-related individuals (full and half sibs) identified by the software Colony in each samples of larvae and juveniles. The number of half sibs identified when comparing samples collected over the same reproductive period is indicated below the figure.

**Fig. 5** Distributions of the mean size (A, C) or size difference (B, D) in larvae (A, B) and juveniles (C, D) as a function of pairwise relatedness between individuals. Relationships identified by Colony were featured by grey dots for half sibs and white dots for full sibs. Note that larvae of the samples L06A and part of L06B were not measured and thus not pictured here. Vertical lines indicate the limit of the five relationship categories: 1-unrelated individuals, 2- unrelated or half sibs, 3- half sibs, 4 -half or full sibs and 5-full sibs.

Table 1 Sample definition and genetic diversity indices obtained across 16 microsatellite loci: A- adults, larvae and juveniles sampled in the targeted site (Château du Taureau; bay of Morlaix; Fig. 1) and B- adults sampled in other locations of the same bay and neighboring bays.

N: number of individuals,  $N_{all}$ : average number of alleles per locus,  $A_r$ : allelic richness,  $H_e$ : expected heterozygosity,  $F_{IS}$ : fixation index.

A.

Year and type of sampling	life cycle stage	date of sampling	labels	N	$N_{all}$	$A_r$	$H_e$	$F_{IS}^*$
<b>2006</b>								
benthic	adults	18-Apr	CdT06	88	16.38	10.45	0.675	0.130
pelagic	larvae	all pools	L06	(total 2006) 180	21.13	10.79	0.679	0.265
pelagic	larvae	4-May	L06A	56	15.00	10.77	0.679	0.239
pelagic	larvae	4-Jul	L06B	60	14.06	10.14	0.674	0.373
pelagic	larvae	29-Sep	L06C	64	14.88	10.35	0.676	0.184
benthic	juveniles	19-Sep	J06	72	15.63	10.45	0.678	0.097
<b>2007</b>								
benthic	adults	19-Apr	CdT07	88	16.19	10.16	0.662	0.082
pelagic	larvae	all pools	L07	(total 2007) 241	20.81	10.32	0.672	0.205
pelagic	larvae	12-Apr	L07A	54	12.56	9.45	0.664	0.396
pelagic	larvae	11-May	L07B	62	14.19	10.20	0.684	0.134
pelagic	larvae	8-Jun	L07C	62	14.69	10.42	0.673	0.132
pelagic	larvae	22-Aug	L07D	63	13.50	9.66	0.661	0.193
benthic	juveniles	20-Aug	J07	81	14.44	9.76	0.650	0.235
<b>2010</b>								
benthic	adults	16-Apr	CdT10	48	12.44	9.78	0.660	0.164
pelagic	larvae	all pools	L10	(total 2010) 275	20.44	10.24	0.663	0.137
pelagic	larvae	5-May	L10A	66	14.50	10.21	0.658	0.175
pelagic	larvae	9-Jun	L10B	65	13.69	10.07	0.668	0.147
pelagic	larvae	4-Aug	L10C	72	14.44	10.17	0.658	0.121
pelagic	larvae	2-Sep	L10D	72	14.31	9.78	0.667	0.108
benthic	juveniles	19-Jan-11	J10	62	13.94	10.16	0.671	0.074

\*: p-values associated with test for Hardy-Weinberg equilibrium is less than 0.0001 in each case

B.

Bay	Populations (labels)	GPS coordinates		date of sampling	Labels	N	Nall	Ar	He	F <sub>IS</sub> *
		latitude N	longitude W							
<b>2002-2003</b>										
Bay of Brest	Keraliou (BrK)	48° 22' 33''	04° 25' 69''	Jul 2003	BrK03	48	9.00	7.71	0.661	0.088
	Rozegat (BrZ)	48° 19' 30''	04° 31' 70''	Jul 2003	BrZ03	53	9.25	7.69	0.654	0.099
Bay of St Brieuc	Saint Brieuc (SB1)	48° 42' 62''	02° 50' 90''	Aug 2002	SB1_02	32	11.13	10.00	0.677	0.207
	Saint Brieuc (SB104)	48° 37' 37''	02° 45' 41''	Aug 2002	SB104_02	32	10.75	9.62	0.660	0.099
Bay of Morlaix	Château du Taureau (CdT)	48°40'185"	3°53'162"	1-Oct 2002	CdT02	47	13.44	10.31	0.684	0.129
<b>2006</b>										
Bay of Morlaix	Figuier (Fig)	48°40'580"	3°56'192"	25-Jul	Fig06	29	11.50	10.52	0.669	0.060
	Rivière de Mlx (RdM)	48°39'487"	3°52'343"	20-Jul	RdM06	49	12.31	9.52	0.646	0.078
	Ile Verte (IV)	48°42'435"	3°54'772"	26-Jul	IV06	34	11.00	9.59	0.675	0.099
	Château du Taureau (CdT)	48°40'185"	3°53'162"	18-Apr	CdT06	88	16.38	10.45	0.675	0.130
<b>2010</b>										
Bay of Brest	Keraliou (BrK)	48° 22' 33''	04° 25' 69''	12-Apr	BrK10	46	8.69	7.55	0.659	0.133
	Rozegat (BrZ)	48° 19' 30''	04° 31' 70''	12-Apr	BrZ10	41	8.44	7.26	0.642	0.184
Bay of St Brieuc	Saint Brieuc (SB1)	48° 42' 62''	02° 50' 90''	28-Apr	SB1_10	45	12.88	10.09	0.658	0.153
	Saint Brieuc (SB104)	48° 37' 37''	02° 45' 41''	28-Apr	SB104_10	47	12.94	10.11	0.645	0.206
Bay of Morlaix	Figuier (Fig)	48°40'580"	3°56'192"	10-Jun	Fig10	47	13.13	10.22	0.668	0.117
	Rivière de Mlx (RdM)	48°39'487"	3°52'343"	17-Jun	RdM10	47	12.81	10.19	0.689	0.103
	Ile Verte (IV)	48°42'435"	3°54'772"	22-Jun	IV10	48	12.63	9.92	0.672	0.177
	Château du Taureau (CdT)	48°40'185"	3°53'162"	16-Apr	CdT10	48	12.44	9.78	0.660	0.164

\*: p-values associated with HWE < 0.0001 in each case



Table 2 Pairwise comparison of the genetic structure (pairwise  $F_{ST}$  estimates and probability values for exact tests) for samples collected in A- 2006, B- 2007 C- 2010 and D- 2002-3 (adults only). Labels are described in Table 1. Probability values for exact tests were corrected for multiple comparisons: ns: non-significant ( $p>0.05$ ; in bold), \* ( $p<0.05$ ), \*\* ( $p<0.01$ ), \*\*\* ( $p<0.001$ ).

A-

	Fig06	IV06	RdM06	CdT06	L06A	L06B	L06C
IV06	<b>0.005n.s.</b>						
RdM06	<b>-0.002n.s.</b>	<b>0.006n.s.</b>					
CdT06	<b>0.001n.s.</b>	<b>0.001n.s.</b>	<b>0.002n.s.</b>				
L06A	0.002**	0.004**	0.004***	0.000*			
L06B	0.006***	0.010***	0.007***	0.006***	0.003***		
L06C	0.007**	0.008***	0.012***	0.003***	0.006***	0.007***	
J06	0.003**	0.001*	0.004***	0.000*	0.002***	0.005***	0.003***

B-

	CdT07	L07A	L07B	L07C	L07D
L07A	0.004***				
L07B	0.001***	0.004***			
L07C	<b>0.001 n.s.</b>	0.007***	0.001***		
L07D	<b>0.000 n.s.</b>	0.007***	0.001***	0.002***	
J07	0.002***	0.009***	0.006***	0.003***	0.005***

C-

	BrK10	BrR10	SB1_10	SB104_10	Fig10	IV10	RdM10	CdT10	L10A	L10B	L10C	L10D
BrK10												
BrR10	<b>-0.003n.s.</b>											
SB1_10	0.015***	0.019***										
SB104_10	0.019***	0.019***	0.003***									
Fig10	0.003***	0.006***	<b>0.003n.s.</b>	0.006***								
IV10	0.009***	0.010***	0.003**	0.005***	<b>-0.001n.s.</b>							
RdM10	0.005***	0.011***	0.004***	0.005***	<b>-0.001n.s.</b>	<b>-0.002n.s.</b>						
CdT10	0.0010***	0.012***	<b>0.001n.s.</b>	0.002*	<b>0.000n.s.</b>	<b>0.002n.s.</b>	<b>0.001n.s.</b>					
L10A	0.011***	0.013***	0.003**	0.002***	0.001*	0.005***	0.002**	<b>-0.001n.s.</b>				
L10B	0.007***	0.011***	0.003**	0.004***	-0.002*	0.001***	0.000*	-0.001*	0.001**			
L10C	0.008***	0.010***	0.003***	0.002***	-0.001*	-0.001***	0.002***	-0.001**	0.002**	<b>-0.002n.s.</b>		
L10D	0.004***	0.009***	0.004***	0.006***	-0.002*	0.001***	0.000**	0.001***	0.002**	<b>-0.001n.s.</b>	<b>-0.001n.s.</b>	
J10	0.009***	0.014***	0.003*	0.005***	0.002**	0.001**	0.001**	0.004**	0.004***	<b>-0.001n.s.</b>	<b>0.001n.s.</b>	0.002*

D-

	BrK03	BrR03	SB1_02	SB104_02
BrK03				
RZ03	<b>0.000 n.s.</b>			
SB1_02	0.016 ***	0.015 ***		
SB104_02	0.014 ***	0.013 ***	0.007 ***	
CdT02	0.011 ***	0.010 ***	0.003 ***	0.005 ***

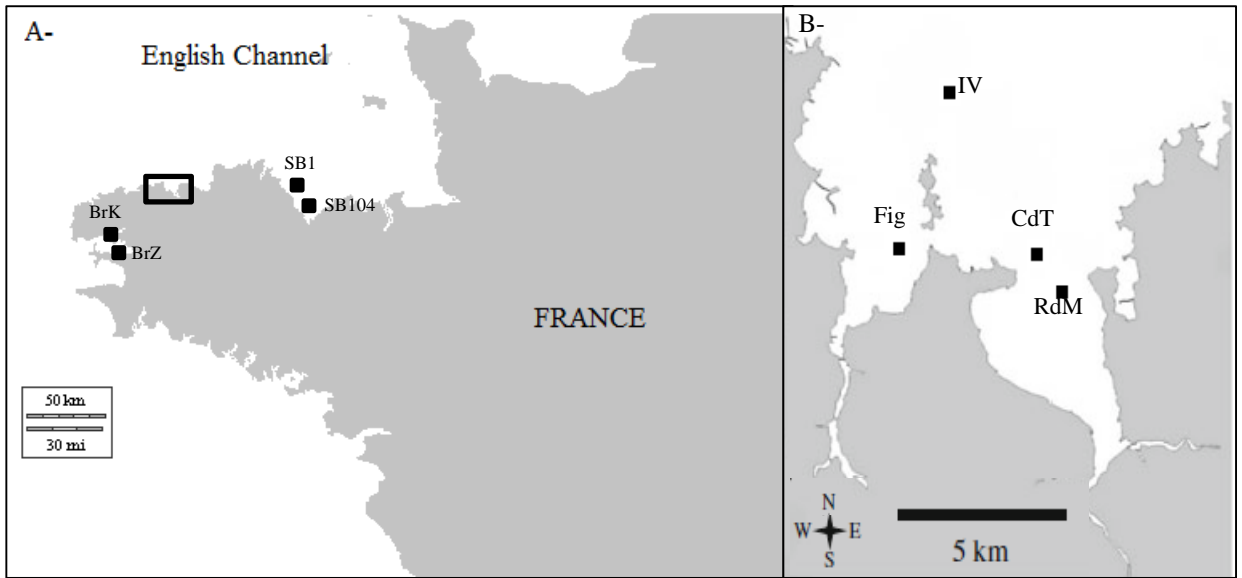
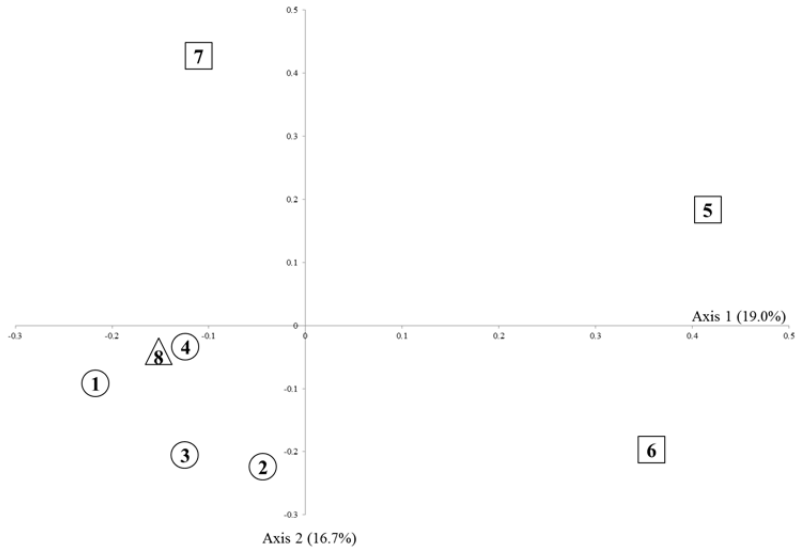
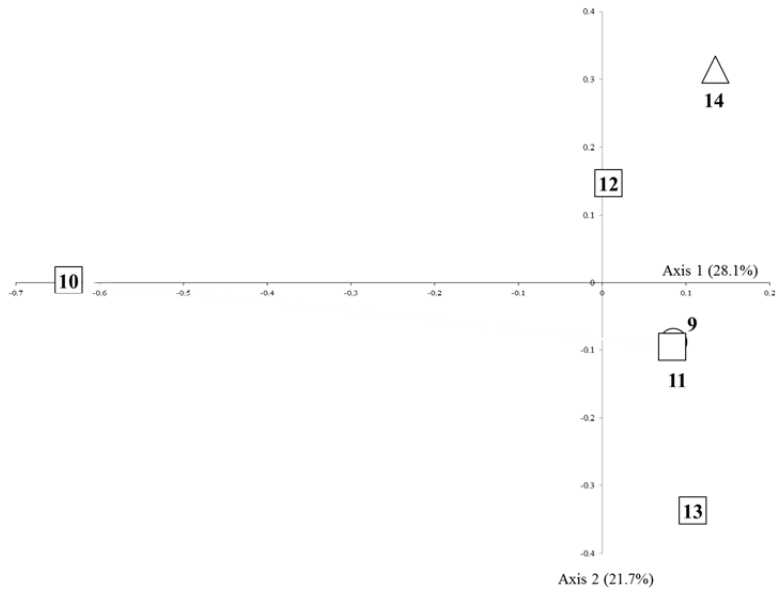


Fig.1

A.



B.



C.

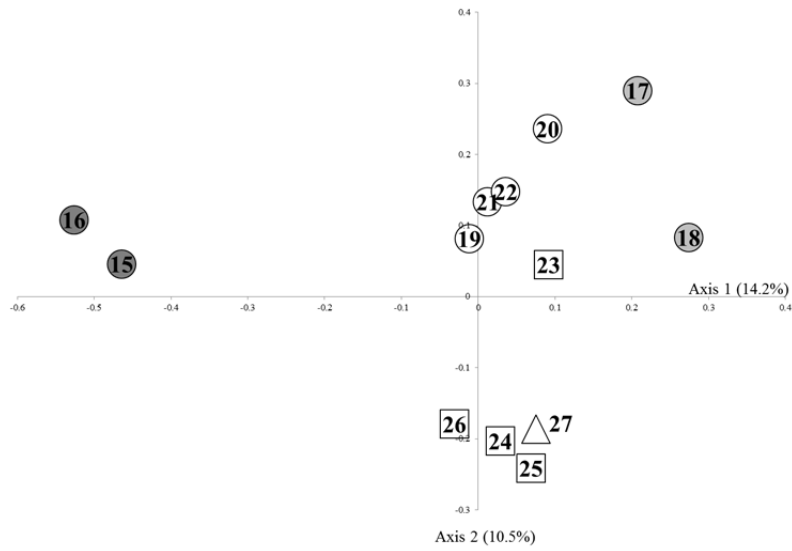


Fig. 2

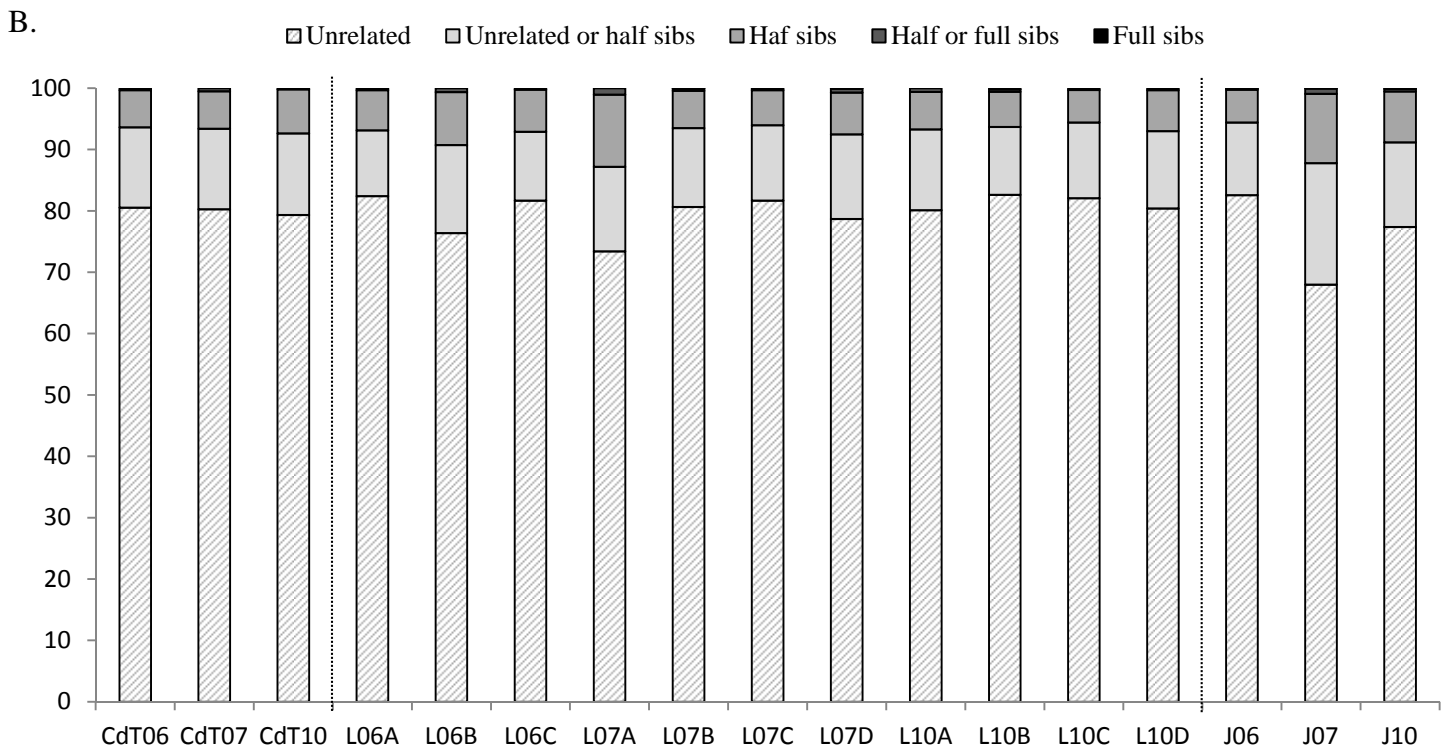
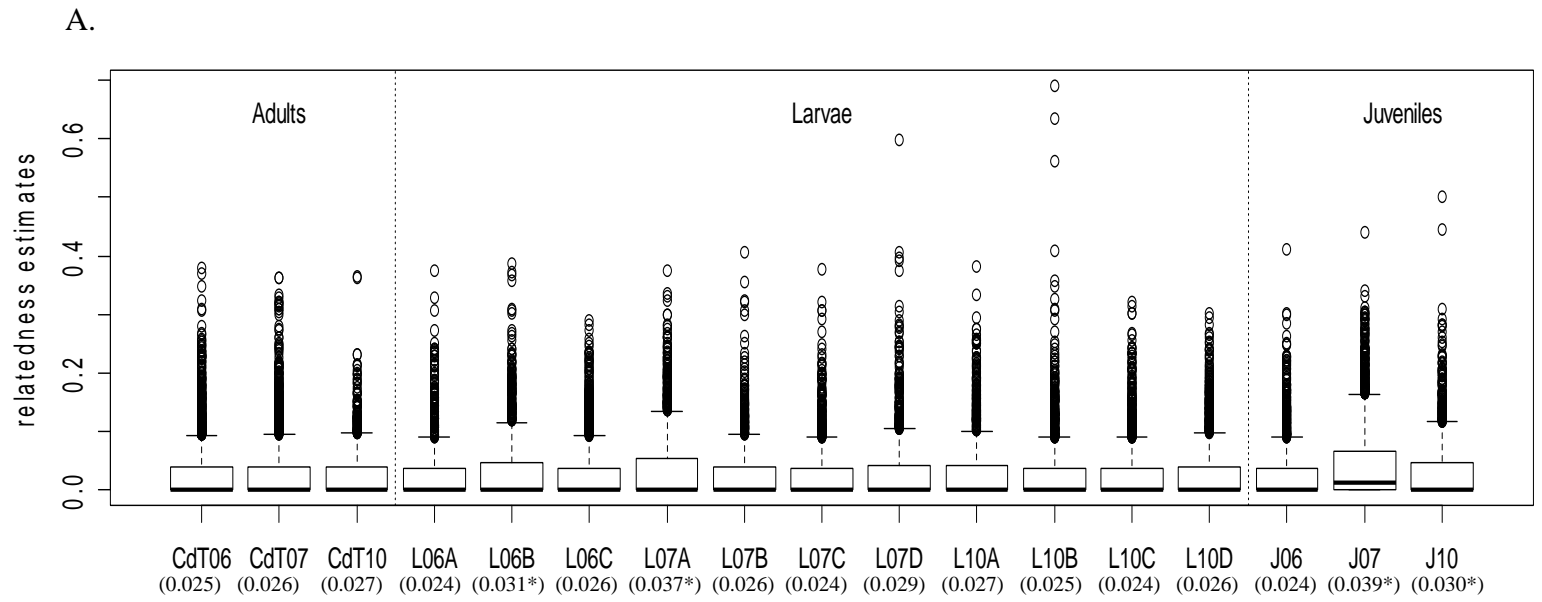


Fig. 3

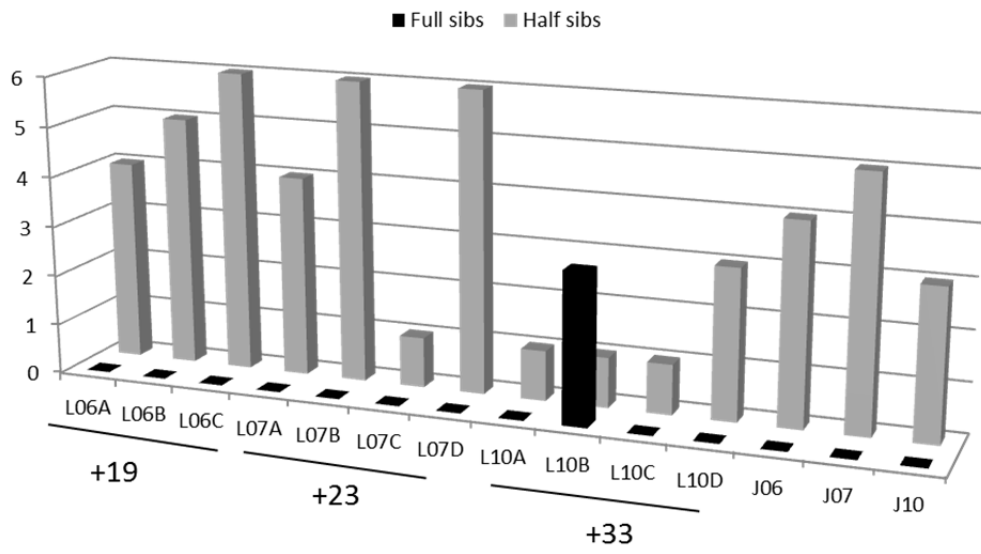


Fig. 4

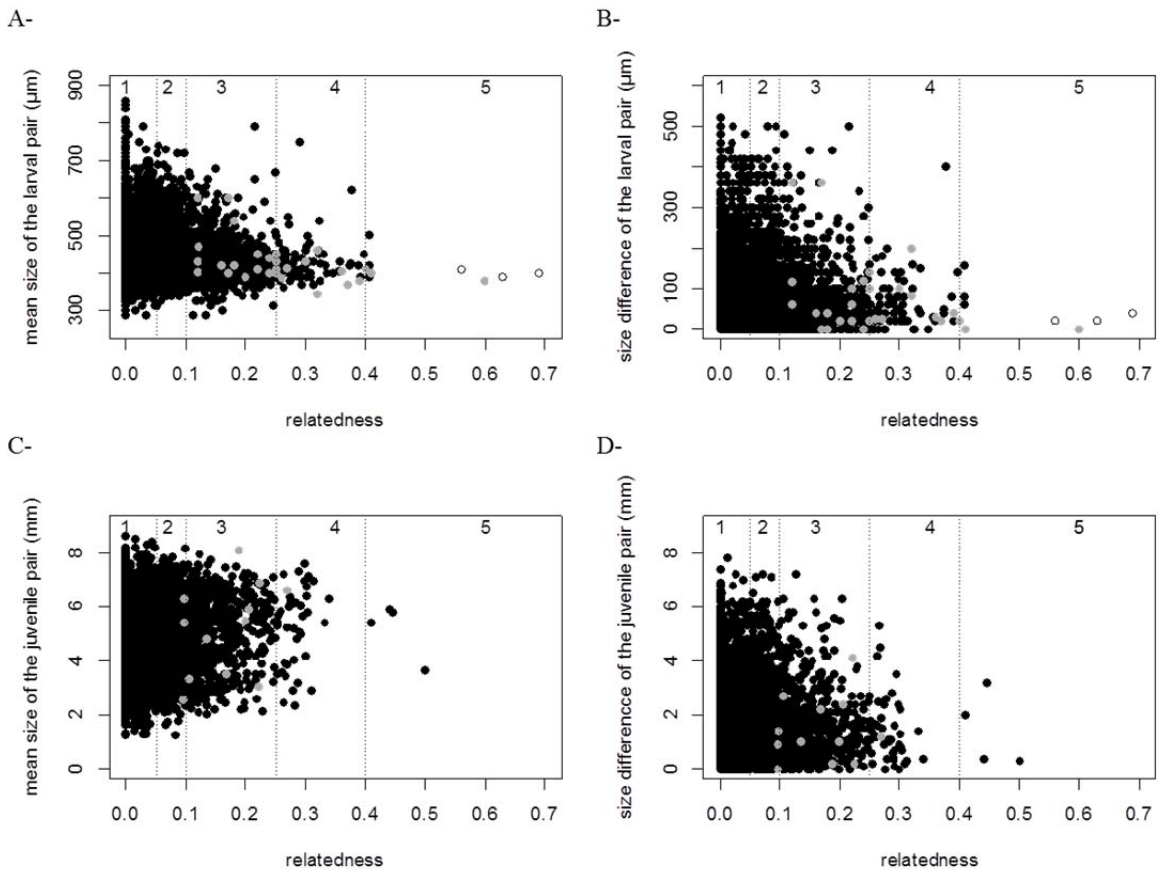


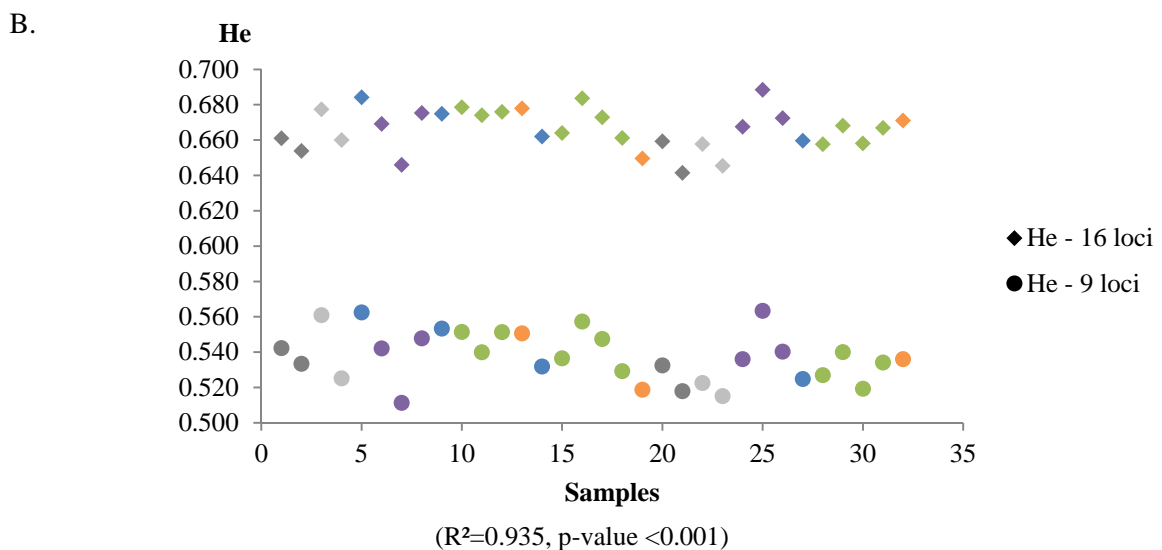
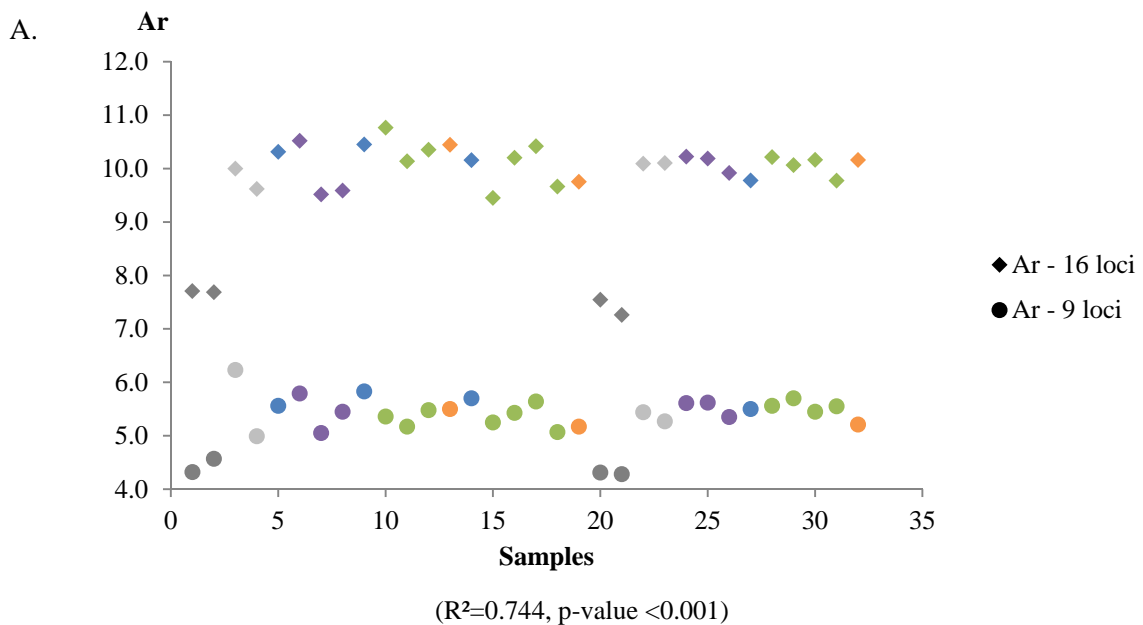
Fig. 5

## Supplementary Information

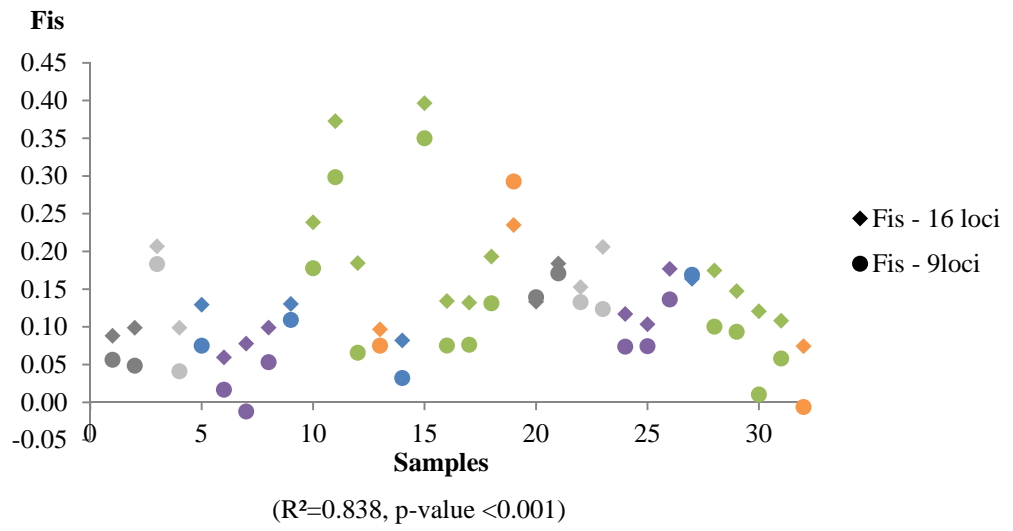
Figure S1 Comparisons of the genetic diversity assessed with 16 and 9 loci in each sample for: A- allelic richness ( $A_r$ ), B- gene diversity ( $H_e$ ) and C- fixation index ( $F_{is}$ ). Correlation coefficients between these indexes computed with 16 and 9 loci ( $R^2$ ) and associated p-values are given below the figures.

The sample code is detailed in Table 1. Samples are numbered as follow: 1. BrK03, 2.BrZ03, 3.SB1\_02, 4. SB104\_02, 5.CdT02, 6. Fig06, 7. IV06, 8. RdM06, 9. CdT06, 10. L06A, 11. L06B, 12. L06C, 13. J06, 14. CdT07, 15. L07A, 16. L07B, 17. L07C, 18. L07D, 19. J07, 20. BrK10, 21. BrZ10, 22. SB1\_10, 23. SB104\_10, 24. Fig10, 25. IV10, 26. RdM10, 27. CdT10, 28. L10A, 29. L10B, 30. L10C, 31. L10D, 32. J10.

Samples are colored as follows: bay of Brest in dark grey, bay of Saint Briec in light grey, bay of Morlaix in violet; for the site Château du Taureau, adults, larvae and juveniles are in blue, green and orange, respectively.



C.





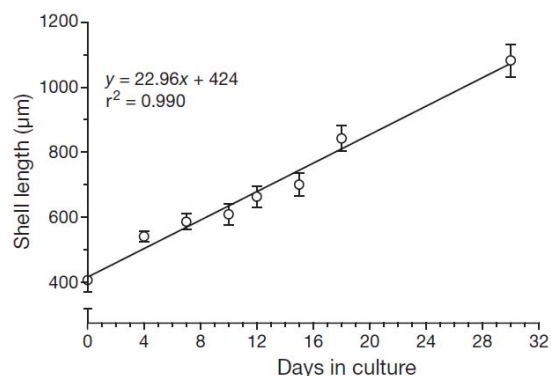
## Chapitre 4 – Distribution de l’hétérozygotie multilocus, déséquilibre d’identité et taille des larves chez *Crepidula fornicata*

### I. Contexte de l’étude

Chez *C. fornicata*, une forte mortalité des premiers stades de vie est observée (Pechenik et Levine 2007). En parallèle, des changements de fréquences alléliques sont observés entre stades (cf. article 4). On pourrait imaginer que des modifications génotypiques ont également lieu et que la forte mortalité larvaire pourrait être non aléatoire mais fonction du degré d’hétérozygotie des individus. En effet, avoir un génome hétérozygote peut conférer de multiples avantages (*e.g.* protéger des allèles délétères, ou promouvoir de la vigueur hybride). Les corrélations entre hétérozygotie et valeur sélective (« Heterozygosity-fitness correlation » ; HFC) ont fait l’objet de nombreuses études (pour revue, références et discussions, voir David 1998, Szulkin *et al.* 2010). Elles sont étudiées par des associations statistiques entre le nombre de locus hétérozygotes sur un ensemble de marqueurs (« hétérozygotie multi-locus », en anglais « multilocus heterozygosity » ou MLH) et un trait associé à la fitness. Les HFC peuvent être expliqués par des effets locaux (associations entre les locus analysés et les locus impliqués dans le trait étudié) ou par des effets globaux (consanguinité qui génère des déséquilibres d’identité à l’échelle du génome). Ces déséquilibres d’identité sont définis par des associations non-aléatoires des génotypes diploïdes c’est-à-dire un excès ou un déficit en hétérozygotie multi-locus (David 1998). Dans le cas de marqueurs neutres, tels que ceux que nous utilisons dans ce travail de thèse, nous nous intéressons à des effets globaux.

Les espèces marines sont souvent supposées panmictiques ou à grandes tailles de population ce qui rend plus difficile l’observation d’effets globaux générés par de la consanguinité. Nos résultats montrent néanmoins que des effets de dérive génétique ont lieu dans la population de crépidule étudiée en baie de Morlaix. Dans cette étude, nous avons donc cherché (1) à mettre en évidence d’éventuels déséquilibre d’identité puis (2) à vérifier si les niveaux de MLH observés pouvaient être corrélés à la taille des larves, sous l’hypothèse que ce trait soit déterminé génétiquement par de nombreux locus. En effet il a été montré que la taille des larves est un bon proxy de leur âge (Fig. 1).

**Fig. 1** Relation positive entre taille de la coquille larvaire et âge de la larve. D’après Pechenik et Levine (2007)



La même analyse a été faite chez des juvéniles et des adultes. Pour cela, nous avons utilisé le jeu de données d'une étude antérieure, celui de la Deuxième Partie – Chapitre 3 : plusieurs échantillons larvaires au cours d'une saison de reproduction (trois en 2006 et quatre en 2007 et 2010) ainsi que trois échantillons de juvéniles (en 2006, 2007 et 2010) et d'adultes (en 2006, 2007 et 2010) avaient été génotypés sur 16 locus microsatellites. De plus, à l'exception de deux échantillons larvaires de 2006, tous les individus avaient été mesurés.

## II. Résultats principaux de l'étude (détaillés dans l'article 5)

Aucun déséquilibre d'identité n'a été observé au sein des échantillons de juvéniles et d'adultes et dans cinq échantillons larvaires. En revanche, 6 échantillons larvaires montrent un déséquilibre d'identité significatif. Ainsi les déséquilibres d'identité et l'hétérozygotie multi-locus varient au sein des échantillons larvaires au cours d'une saison de reproduction.

La taille des larves varie entre les différents échantillons collectés au cours de la saison de reproduction. En effet, en 2007 et 2010 (données de taille non disponibles pour deux échantillons larvaires en 2006), les échantillons larvaires de début et fin de saison de reproduction ne sont significativement pas différents l'un de l'autre mais sont significativement différents de ceux de mi-saison (qui ne sont par ailleurs pas significativement différents l'un de l'autre). Si ce sont des larves qui ont la même distribution d'âge dans l'échantillon prélevé, des conditions favorables à la croissance larvaire en début d'été (*e.g.* température plus élevée) pourraient expliquer ces différences. La taille des juvéniles de 2010 est significativement différente de celle des deux autres années ; en 2010 les juvéniles sont plus petits suggérant un recrutement plus récent. La taille des adultes n'est pas significativement différente d'une année sur l'autre.

Finalement aucune corrélation entre taille des individus et la valeur individuelle de MLH n'a été détectée, ne soutenant pas l'hypothèse d'HFC en utilisant ce trait. Il est possible que la taille des larves soit non seulement un mauvais indicateur de fitness mais également un mauvais indicateur de l'âge des larves. En effet, des études antérieures ont montré qu'au sein d'une même descendance, on pouvait observer des larves à croissance rapide et des larves à croissance lente (Pechenik et Levine 2007, Le Cam *et al.* 2009). On ne peut pas non plus exclure que la qualité du génotypage et la présence d'allèles nuls aient pu biaiser les résultats.

*Des résultats complémentaires, non inclus dans l'article 5, sont présentés dans l'annexe 14 (transformation des données larvaires pour le modèle linéaire utilisé).*

**III. Article 5: « Identity disequilibrium but no heterozygosity-fitness correlation in the marine gastropod *Crepidula fornicata* »**

« Des déséquilibres d'identité au sein du pool larvaire mais pas de corrélation hétérozygotie-fitness détectés chez le gastéropode marin *Crepidula fornicata*»

Florentine Riquet, Thomas Broquet, Thierry Comtet et Frédérique Viard

## **Identity disequilibrium but no heterozygosity-fitness correlation in the marine gastropod**

### ***Crepidula fornicata***

Florentine Riquet<sup>1,2</sup> - Thomas Broquet<sup>1,2</sup> - Thierry Comtet<sup>1,2</sup> - Frédérique Viard<sup>1,2</sup>

<sup>1</sup> UPMC Univ Paris 06, UMR 7144, Team Diversity and Connectivity in Coastal Marine Landscapes, Station Biologique de Roscoff, 29680 Roscoff, France

<sup>2</sup> CNRS, UMR 7144, Lab. Adaptation and Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France

Correspondence: F. Viard: Lab. Adaptation & Diversity in the Marine Environment, Station Biologique de Roscoff, 29680 Roscoff, France; Tel: +33298292312, Fax: +33298292324; e-mail: [viard@sb-roscoff.fr](mailto:viard@sb-roscoff.fr)

Running title: HFCs in *C. fornicata*?

Keywords: *Crepidula fornicata*, gastropod, heterozygosity-fitness-correlation, multi-locus heterozygosity, heterozygote deficiencies

#### Abstract

Heterozygosity at fitness loci confers advantages, especially in fluctuating environments. This effect can be revealed by neutral marker loci if there is a genome-wide correlation in heterozygosity across loci. Accordingly, a correlation between genome-wide heterozygosity and fitness traits, also termed Heterozygosity-Fitness-Correlations (HFCs), has long been evidenced in wild populations. Here we tested the hypothesis that high mortality in early life stages may be correlated to heterozygosity at marker loci in the slipper limpet *Crepidula fornicata*. By analyzing correlations of homozygosity among loci, we showed strong identity disequilibrium (ID) in some larval pools released during the reproductive season. No association between larval size, used as a trait of fitness, and heterozygosity was however observed, casting doubt as a possible HFC in *Crepidula fornicata*.

## Introduction

Heterozygosity at fitness loci is considered to be advantageous as, additionally to the general effect of masking deleterious recessive alleles, it shows more genetic variation to respond to the fluctuating environment encountered (Hedrick *et al.* 1976). When heterozygosity at fitness loci can be approached by estimates of heterozygosity at marker loci, then a positive correlation between heterozygosity and fitness may be observed, a pattern coined Heterozygosity-Fitness-Correlations (HFCs). HFCs have been widely examined in natural populations (review in Chapman *et al.* 2009) with a number of fitness traits reported to be associated with higher heterozygosity at individual level, such as weight (Coltman *et al.* 1998), longevity (Coltman *et al.* 1999), reproductive traits (Slate *et al.* 2000). Such hypothesis had been also investigated in marine species (e.g. in bivalve (Koehn *et al.* 1973; Koehn *et al.* 1976; Zouros *et al.* 1980), false limpet: (Johnson & Black 1982; 1984), gastropod: (Johnson *et al.* 1993). In many marine species HFC may be easier to observe in early stages, which suffer a very large mortality. Empirical evidences of HFC in these species may thus be limited by our ability to sample and analyze pelagic larvae (but see Bierne *et al.* 1998).

As pointed out by David *et al.* (1995), two main hypotheses are proposed to explain HFCs called the “local effect hypothesis” or “the general effect hypothesis”. The former hypothesis suggests that neutral loci are linked to functional loci influencing fitness; heterozygotes exhibit an advantage via dominance (i.e. effect of deleterious alleles are masked) and/or overdominance (i.e. heterozygotes exceed the performance of both homozygotes *per se*, Hansson & Westerberg 2002) and are hence selected. The latter hypothesis suggests that HFCs result rather from inbreeding than local effects (Szulkin *et al.* 2010). If Multi Locus Heterozygosity (MLH) reflects genome-wide variation then it indicates general inbreeding: more inbred individuals show both a reduced heterozygosity and reduced fitness. In this situation HFC rely on correlations between observed heterozygosity at genotyped markers and heterozygosity at functional loci, correlations resulting from the presence of Identity Disequilibrium (ID, correlation in heterozygosity and/or homozygosity across loci, Szulkin *et al.* 2010). Identity disequilibrium can be driven by population demography (e.g. genetic drift, bottleneck) and/or life-history (e.g. mating system, Szulkin *et al.* 2010, Bierne *et al.* 2000).

The present study aimed at examining ID and HFCs in the slipper limpet *Crepidula fornicata*, a marine gastropod that displays a typical benthic-pelagic life cycle alternating a free-swimming larval phase with a sessile phase. Using a sampling of larvae, juveniles and adults in the bay of Morlaix (Brittany, France) conducted over three years (2006, 2007, 2010), a genetic shift in allelic frequencies was identified between the three life stages considered, with deficiencies in heterozygotes observed in larvae, juveniles and adults (Riquet *et al.* in preparation<sup>1</sup>). The slipper limpet exhibits high mortality during the early stages (in pre-settlement as a larvae (Thieltges *et al.* 2004; Pechenik & Levine 2007) and post-settlement as juvenile (Richard *et al.* 2006)) as most invertebrates with a benthic-pelagic life cycle (type III survivorship curve). Consequently, the genotypic composition of the larva sample could change in a non-random way, assuming that mortality depends upon heterozygosity. Families composed of larvae in which heterozygosity is higher could be favored during the larval stage. One important fitness trait in *C. fornicata* larvae is the larval size, shown to be a good proxy of larval age (Pechenik & Levine 2007). It has been shown that the slow-growing larvae were more vulnerable to predators than fast-growing larvae (Pechenik & Levine 2007). Hence larger larvae should be the outcome of larvae that grew fast and that are more likely to survive during the larval stage (Pechenik & Levine 2007). In addition, Zouros & Foltz (1984) demonstrated in the Atlantic oyster *Crassostrea gigas* that heterozygotes were shown to exhibit faster growth than homozygotes. Therefore, larval size appeared as a good fitness trait to investigate HFCs in the slipper limpet *C. fornicata*.

We here tested for the occurrence of HFCs in the slipper limpet *C. fornicata*, first by analyzing Identity Disequilibrium in larvae, juveniles and adults, MLH being used as a proxy of the genome-wide heterozygosity, and then by examining correlation between heterozygosity and size of the individuals.

## Materials and Methods

Larvae, juveniles and adults of *Crepidula fornicata* were sampled over three years (2006, 2007 and 2010) in the same site in the bay of Morlaix (Brittany, France). Sampling and genotyping are detailed

---

<sup>1</sup> Deuxième Partie- Chapitre 3

in Riquet *et al.* (in preparation<sup>2</sup>). Briefly, sampling was conducted either by using a WP2 plankton net with a 200µm mesh size in a 5-min horizontal tow for larvae (three to four samples a year) or along a 50m transect by scuba diving (adults and juveniles). Each individual was measured either using a dissecting microscope with an ocular micrometer (larvae) or using a calliper (adults and juveniles), then separately preserved in 96% ethanol, except for the first and part of the second larval samples of 2006 that were not measured. Labels of these samples given in table 1 will be used hereafter. After DNA extraction performed using a Nucleospin® Multi-96 Tissue Kit (MACHEREY-NAGEL) following the manufacturer's protocol, all the individuals were genotyped for 16 microsatellite markers in which 12 are EST-SSRs (Riquet *et al.* 2011) and four are markers obtained from anonymous DNA libraries enriched for microsatellites (Dupont & Viard 2003; Dupont *et al.* 2006; Proestou 2006). DNA of the first and part of the second larval samples (L06A and part of L06B) were extracted using the Higuchi (1989) method and genotyped for the 16 microsatellite markers. Summary statistics (expected heterozygosity ( $H_e$ ), fixation index and tests from departure to Hardy–Weinberg equilibrium for each sample) obtained for each sample are given in table 1, kin-related and relatedness coefficient for larval samples are shown in Fig. 2 (from Riquet *et al.* in preparation<sup>3</sup>).

Individual Multi-Locus Heterozygosity (MLH) was calculated as the number of loci that were heterozygous divided by the number of loci successfully genotyped. Difference in mean MLH among samples was tested by multiple comparison test using the software package R (R Development Core Team, 2011). Four other individual heterozygosity estimates (standardized heterozygosity based on the mean expected heterozygosity (Coltman *et al.* 1999), standardized heterozygosity based on the mean observed heterozygosity (Amos 2005), Internal Relatedness (Amos *et al.* 2001) and Homozygosity by locus (Aparicio *et al.* 2006)) were calculated using the R function GENHET (Coulon 2010). The five individual heterozygosity estimates were highly correlated (Supplementary information – Fig. S1) and using either estimate produced equivalent results. We used MLH in all analyses.

---

<sup>2</sup> Deuxième Partie : Chapitre 3

<sup>3</sup> Deuxième Partie- Chapitre 3

The strength of ID was measured for each sample by calculating the parameter  $g_2$  using the software RMES (David *et al.* 2007). This parameter measures the excess of double heterozygotes at two loci relative to the expectation under a random association, standardized by average heterozygosity (Szulkin *et al.* 2010). To test whether the parameter  $g_2$  was significantly different than zero, hence significant correlation indicating the presence of ID, genotypes were resampled and tested 1000 times (David *et al.* 2007). P-values for each sample were corrected using the FDR method implemented in QVALUE (Storey 2002).

Linear models were constructed to test for a relationship between MLH and size using the software package R (R Development Core Team, 2011). Samples were included as an additional factor in the model. The R Box-Cox function in the MASS library (R Development Core Team, 2011) was used to identify a suitable power transformation of the data.

## Results

Altogether 602 larvae were measured. Their size ranged from 250 $\mu$ m (in L07A) to 900 $\mu$ m (in L07C), with a mean size of 436.2 $\mu$ m  $\pm$ 71.7. In 2006, no difference in length was observed between the two measured larval samples ( $p > 0.05$ , Fig. 1A, first Fig.). In 2007, L07A and L07D were significantly smaller than L07B and L07C; no difference in length was observed between L07A and L07D as well as between L07B and L07C (Fig. 1B, first Fig.). The same pattern was observed in 2010 with L10A and L10D significantly smaller than L10B and L10C and no difference observed between L10A and L10D and between L10B and L10C (Fig. 1C, first Fig.). Juvenile length ranged from 1.0 (J06) to 8.8mm (J06) and J10 showed significant differences in length relative to J06 and J07 (Fig.1D, first Fig.). Adult length ranged from 4.0 (CdT10) to 53.3mm (CdT10) and no difference in length was observed between the three years considered (Fig. 1E, first Fig.).

No individual was fully homozygote or heterozygote according to the number of loci genotyped. Mean Multi-Locus Heterozygosity (MLH) per sample ranged from 3.98 (L07A) to 6.19 (J10) with an overall mean of 5.46 (table1). Over the three study years, no difference in mean MLH was observed from one year to the other among adult samples (table 1, Fig. 1E, middle Fig.). Among juvenile samples, the mean MLH of J07 was significantly different from those of J06 and J10 (table 1, Fig. 1D, middle Fig.). Among larvae, a significant difference in MLH was observed between L06 and L10 (table 1).



Within a year, L06B and L06C differed from L06A, and L07A differed from all other samples from 2007 (Fig.1 A-B, middle Fig.). Larvae of 2006 showed a significantly lower mean MLH than adults and juveniles of the same year; larvae and juveniles had a lower MLH than adults in 2007. No difference was observed in mean MLH among samples from year 2010.

Identity Disequilibrium (ID) was found to be low for all adult and juvenile samples, and not significantly different from zero (table 1). Significant correlations in heterozygosity and/or homozygosity across loci were however observed each year for larvae (table 1). When investigating ID in each of the 11 larval samples, correlations in heterozygosity across loci were observed for six larval samples (L06A, L06B, L06C, L07A, L07D and L10A, table 1). Their observed MLH distributions compared to the expected distribution under random mating are given in Fig. S2. ID may arise in larvae due to technical problems, such as the presence of null alleles suspected in 7 out of the 16 loci, or a greater sensitivity in genotyping error (Fig. S3), but ID was still identified in L06A, L06B and L06C when discarded these 7 loci and in L06B and L10A when keeping only larvae genotyped for all the 16 loci. Larval samples with significant ID generally showed higher inter-individual relatedness (Fig. 2).

Using linear models, MLH has no effect on the size of larvae, juveniles or adults. However, some larval samples showed both relatively large sizes and high MLH (this is particularly visible in year 2007, Fig. 1B). Such differences among sample would drive a spurious HFC signal when all individuals are pooled without consideration for the fact that larvae belong to different samples.

## **Discussion**

Theory predicts that larval mortality in *Crepidula fornicata* could depend on heterozygosity at fitness loci, a hypothesis examined in the present study. Identity disequilibrium was identified in some larval samples but not in juvenile or adult samples. No association was however observed between larval size and heterozygosity, calling into question a possible HFC in *Crepidula fornicata*.

One interesting pattern that has emerged from multi-locus analysis is the presence of Identity Disequilibrium (ID). ID was inconsistent among larval samples over the same year and from one year to the other but was clearly evidenced in six of our larval samples. This provides an interesting insight about genetic associations during the larval phase, which is not identified in juveniles or adults. As

reviewed by Szulkin *et al.* (2010), ID can be observed under a variety of conditions without the need for HFCs to hold. Besides methodological considerations, ID may be generated by inbreeding and/or demographic effects, such as admixture or genetic drift. Methodological artifacts may have generated ID, but may not be the only reason that caused ID since ID was still identified when corrected for methodological noise. Despite a long-lived larval stage, individuals of *C. fornicata* might be more inbred than expected if individuals disperse less than expected. The study population receives little immigrants from adjacent bays (Rigal *et al.* 2010) and larvae recruiting in the parental population mainly originated from the focal population (Dupont *et al.* 2007), thus increasing opportunities for inbreeding. Inbred larvae should be the outcome of parents that are more related than the average. Individuals of *C. fornicata* exhibit a stacking behavior, with piled-up individuals constituting stacks of two to 20 individuals. Reproduction occurs within stacks (Dupont *et al.* 2006; Le Cam *et al.* 2009) and in general, there is one male, located above the female that contributes to the majority of the larvae within a brood (Le Cam *et al.* 2009). The probability that relatives recruited in the same stack following larval dispersion should be extremely low but deserves further investigations by genotyping several complete stacks of *C. fornicata* to examine inbreeding level within stacks. Inbred larvae may also suffer inbreeding depression occurring in the larval phase, as demonstrated in other marine invertebrates (Beaumont & Abdulmatin 1994; Hedgecock *et al.* 1995). Larvae were small (on average 436.2  $\mu\text{m}$ , close to the size of hatching (c.a. 400 $\mu\text{m}$ )) and may have been sampled before suffering inbreeding depression. Similarly, inbreeding effects may be invisible in juveniles and adults because inbred individuals died previously during larval development. The identification of ID in larval samples may also be correlated to a familial structure, as revealed by the proportion of kin-related individuals and mean relatedness in samples with high ID (Fig. 2, Riquet *et al.* (in preparation)<sup>4</sup>).

No HFCs have been demonstrated in the present study. We do not know if this pattern is a commonplace in marine species as, additionally to a substantial bias in favor of significant correlation between heterozygosity and fitness trait in the published literature resulting in null results underrepresented for all species (Hansson & Westerberg 2002), only few studies examined HFCs in marine species (but see Bierne *et al.* 1998). Based only on the heterozygosity, HFCs were identified, but

---

<sup>4</sup> Deuxième Partie – Chapitre 3

heterozygosity and size correlations were no more evidenced when considering that individuals belonged to different samples. With milder conditions, especially during the late spring and summer, larvae grew faster (Rigal 2009) which is exemplified here by larger larvae in May 2007, June 2007 and 2010 and August 2010. MLH however varied randomly from one sample to the other and no pattern arose when comparing the three study years. This suggested that heterozygosity and larval size of 2007 samples varied by chance in the same way, creating preliminary a false signal and pointing out the importance of integrating all variables in the models. The null result observed in the present study may be because of no HFCs *per se*. Relevance of the methods used or the fitness trait examined should also be questioned. First, the number of genetic markers has long been questioned to evidence HFCs in wild populations (Balloux *et al.* 2004; Slate *et al.* 2004) and it has nowadays been proved that a small panel of microsatellites is sufficient enough to greatly detect HFCs (e.g. Forstmeier *et al.* 2012). All markers will reflect the same genetic process for a given trait, although different classes of markers (e.g. functional vs. neutral loci, Olano-Marin *et al.* 2011) may generate variation of intensity in the HFCs signal (Szulkin *et al.* 2010). Besides, MLH is in general the most robust measure to be related to HFCs (Szulkin *et al.* 2010). Presence of null alleles, which is commonly reported in marine invertebrates (David *et al.* 1997; Yu & Li 2008; Lemer *et al.* 2011) may however have obscured HFCs by underestimating the heterozygosity (David 1998). ID in some larval samples was still identified when correcting for the presence of null alleles in seven suspected loci and when keeping only larvae fully genotyped but no HFCs was then identified. On another hand, we here focused on one aspect of individual fitness, i.e. the individual size, and Hoffman *et al.* (2010) suggested that focusing on several fitness traits may confer a more powerful signal of HFCs. The number of fitness traits in larvae of *C. fornicata* appeared however limited, especially because of the minute size of the larvae. In particular, their small size impeded to divide them for other analyses than genotyping (e.g. estimation of fatty acid, usually used as an estimator of quality (Leroy 2011)). Le Cam *et al.* (2009) showed that the more related were larvae within a brood, the lower variation was in growth rate, supporting previous results about the heritability of larval growth rate (Hilbish *et al.* 1999)). They also demonstrated a large variance in larval growth within a given offspring. This is not contradictory to the first result as multiple paternity occurs in *C. fornicata* (Dupont *et al.* 2006; Proestou *et al.* 2008; Le Cam *et al.*

2009), resulting in larvae that differ in degree of relatedness within broods. Consequently, although one of the rare available fitness trait available, larval size may not be the most appropriate to examine HFCs.

### **Acknowledgement**

We are grateful to Sabrina Le Cam and the divers of “Marine Operations Department” of the Biological Station of Roscoff for providing adults and juveniles of *Crepidula fornicata*, as well as François Rigal, Fanny Leroy and the crew of the Neomysis (“Marine Operations Department” - the Biological Station of Roscoff) for providing larvae of *C. fornicata*. We thank the Biogenouest® Genomics core facility for its technical support. This work was funded by the Agence Nationale de la Recherche (Project Hi-Flo, ANR-08-BLAN-0334). FR acknowledges a PhD grant from the Centre National pour la Recherche Scientifique and the Conseil Régional de Bretagne (SEA-FLO ARED project).

### **References**

Amos W, Wilmer JW, Fullard K, Burg TM, Croxall JP, Bloch D, Coulson T (2001) The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **268**, 2021-2027.

Amos W (2005) IR macroN3. IR Manual. <http://www.zoo.cam.ac.uk/zoostaff/amos/#ComputerPrograms> [accessed date April 10th 2009].

Aparicio JM, Ortego J, Cordero PJ (2006) What should we weigh to estimate heterozygosity, alleles or loci? *Molecular Ecology*, **15**, 4659-4665.

Balloux F, Amos W, Coulson T (2004) Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology*, **13**, 3021-3031.

Beaumont AR, Abdulmatin AKM (1994) Differences in morphology, survival and size between self-fertilized and cross-fertilized larvae of *Mytilus galloprovincialis*. *Journal of the Marine Biological Association of the United Kingdom*, **74**, 445-448.

Bierne N, Launey S, Naciri-Graven Y, Bonhomme F (1998) Early effect of inbreeding as revealed by microsatellite analyses on *Ostrea edulis* larvae. *Genetics*, **148**, 1893-1906.

Bierne N, Tsitroni A, David P (2000) An inbreeding model of associative overdominance during a population bottleneck. *Genetics*, **155**, 1981-1990.

Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. *Molecular Ecology*, **18**, 2746-2765.

- Coltman DW, Bowen WD, Wright JM (1998) Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **265**, 803-809.
- Coltman DW, Pilkington JG, Smith JA, Pemberton JM (1999) Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution*, **53**, 1259-1267.
- Coulon A (2010) GENHET: an easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources*, **10**, 167-169.
- David P (1998) Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity*, **80**, 531-537.
- David P, Delay B, Berthou P, Jarne P (1995) Alternative models for allozyme-associated heterosis in the marine bivalve *Spisula ovalis*. *Genetics*, **139**, 1719-1726.
- David P, Delay B, Jarne P (1997) Heterozygosity and growth in the marine bivalve *Spisula ovalis*: testing alternative hypotheses. *Genetical Research*, **70**, 215-223.
- David P, Pujol B, Viard F, Castella V, Goudet J (2007) Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, **16**, 2474-2487.
- Dupont L, Ellien C, Viard F (2007) Limits to gene flow in the slipper limpet *Crepidula fornicata* as revealed by microsatellite data and a larval dispersal model. *Marine Ecology Progress Series*, **349**, 125-138.
- Dupont L, Richard J, Paulet YM, Thouzeau G, Viard F (2006) Gregariousness and protandry promote reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of larval paternity. *Molecular Ecology*, **15**, 3009-3021.
- Dupont L, Viard F (2003) Isolation and characterization of highly polymorphic microsatellite markers from the marine invasive species *Crepidula fornicata* (Gastropoda : Calyptraeidae). *Molecular Ecology Notes*, **3**, 498-500.
- Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012) Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Molecular Ecology*, **21**, 3237-3249.
- Hansson B, Westerberg L (2002) On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology*, **11**, 2467-2474.
- Hedgecock D, McGoldrick DJ, Bayne BL (1995) Hybrid vigor in Pacific oysters: An experimental approach using crosses among inbred lines. *Aquaculture*, **137**, 285-298.
- Hedrick PW, Ginevan ME, Ewing EP (1976) Genetic-polymorphism in heterogeneous environments. *Annual Review of Ecology and Systematics*, **7**, 1-32.
- Higuchi R (1989) Simple and rapid preparation of samples for PCR. *Erlich HA, editor. PCR technology, principles and applications for DNA amplification*. New York: Stockton Press., 31-38.
- Hilbish TJ, Sasada K, Eyster LS, Pechenik JA (1999) Relationship between rates of swimming and growth in veliger larvae: genetic variance and covariance. *Journal of Experimental Marine Biology and Ecology*, **239**, 183-193.

- Hoffman JI, Hanson N, Forcada J, Trathan PN, Amos W (2010) Getting long in the tooth: a strong positive correlation between canine size and heterozygosity in antarctic fur seals *Arctocephalus gazella*. *Journal of Heredity*, **101**, 527-538.
- Johnson MS, Black R (1982) Chaotic genetic patchiness in an inter-tidal limpet, siphonaria sp. *Marine Biology*, **70**, 157-164.
- Johnson MS, Black R (1984) Pattern beneath the chaos - the effect of recruitment on genetic patchiness in an intertidal limpet. *Evolution*, **38**, 1371-1383.
- Johnson MS, Holborn K, Black R (1993) Fine-scale patchiness and genetic-heterogeneity of recruits of the corallivorous gastropod *Drupella cornus*. *Marine Biology*, **117**, 91-96.
- Koehn RK, Milkman R, Mitton JB (1976) Population-genetics of marine pelecypods .4. selection, migration and genetic differentiation in blue mussel *Mytilus edulis*. *Evolution*, **30**, 2-32.
- Koehn RK, Turano FJ, Mitton JB (1973) Population genetics of marine pelecypods II: genetic differences in microhabitats of *Modiolus demissus*. *Evolution*, **27**, 100-105.
- Le Cam S, Pechenik JA, Cagnon M, Viard F (2009) Fast versus slow larval growth in an invasive marine mollusc: does paternity matter? *Journal of Heredity*, **100**, 455-464.
- Lemer S, Rochel E, Planes S (2011) Correction method for null alleles in species with variable microsatellite flanking regions, a case study of the black-lipped pearl oyster *Pinctada margaritifera*. *Journal of Heredity*, **102**, 243-246.
- Leroy F (2011) Influence des conditions trophiques sur le développement larvaire de l'espèce invasive *Crepidula fornicata*. Conséquences sur sa capacité d'expansion. *Thèse de doctorat. Université Pierre et Marie Curie, Paris.*,
- Olano-Marin J, Mueller JC, Kempnaers B (2011) Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. *Molecular Ecology*, **20**, 4028-4041.
- Pechenik JA, Levine SH (2007) Estimates of planktonic larval mortality using the marine gastropods *Crepidula fornicata* and *C. plana*. *Marine Ecology-Progress Series*, **344**, 107-118.
- Proestou DA (2006) Isolation and characterization of microsatellite markers in the Atlantic slipper shell *Crepidula fornicata* for use in paternity analysis. *Molecular Ecology Notes*, **6**, 437-439.
- Proestou DA, Goldsmith MR, Twombly S (2008) Patterns of male reproductive success in *Crepidula fornicata* provide new insight for sex allocation and optimal sex change. *Biol Bull*, **214**, 192-200.
- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Richard J, Huet M, Thouzeau G, Paulet YM (2006) Reproduction of the invasive slipper limpet, *Crepidula fornicata*, in the Bay of Brest, France. *Marine Biology*, **149**, 789-801.
- Rigal F (2009) Dynamique spatio-temporelle du nuage larvaire du gastéropode introduit *Crepidula fornicata* au sein d'une baie mégatidale, la baie e Morlaix (France). *Thèse de doctorat. Université Pierre et Marie Curie, Paris.*,

- Rigal F, Viard F, Ayata SD, Comtet T (2010) Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? *Biological Invasions*, **12**, 3171-3186.
- Riquet F, Ballenghien M, Tanguy A, Viard F (2011) *In silico* mining and characterization of 12 EST-SSRs for the invasive slipper limpet *Crepidula fornicata*. *Marine Genomics*, **4**, 291-295.
- Slate J, David P, Dodds KG, Veenvliet BA, Glass BC, Broad TE, McEwan JC (2004) Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity*, **93**, 255-265.
- Slate J, Kruuk LEB, Marshall TC, Pemberton JM, Clutton-Brock TH (2000) Inbreeding depression influences lifetime breeding success in a wild population of red deer (*Cervus elaphus*). *Proceedings of the Royal Society of London Series B-Biological Sciences*, **267**, 1657-1662.
- Storey J (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479-498.
- Szulkin M, Bierne N, David P (2010) Heterozygosity-Fitness Correlations: a time for reappraisal. *Evolution*, **64**, 1202-1217.
- Thieltges DW, Strasser M, van Beusekom JEE, Reise K (2004) Too cold to prosper - winter mortality prevents population increase of the introduced American slipper limpet *Crepidula fornicata* in northern Europe. *Journal of Experimental Marine Biology and Ecology*, **311**, 375-391.
- Yu H, Li Q (2008) Exploiting EST databases for the development and characterization of EST-SSRs in the pacific oyster (*Crassostrea gigas*). *Journal of Heredity*, **99**, 208-214.
- Zouros E, Foltz DW (1984) Possible explanations of heterozygote deficiency in bivalve mollusks. *Malacologia*, **25**, 583-591.
- Zouros E, Singh SM, Miles HE (1980) Growth-rate in oysters - an overdominant phenotype and its possible explanations. *Evolution*, **34**, 856-867.

## Figures and Tables

Fig.1 Size distributions of each sample, the Multi-Loci Heterozygotes (MLH) distributions measured by 16 microsatellite loci of each sample and scaled from 0 to 1, and the individual MLH according to the individual size are represented for A- larvae of 2006, B- larvae of 2007, C- larvae of 2010, D- juveniles over the three study years and E- adults over the three study years. In the last figures (MLH according to size), the first sample of the year (e.g. L06A, L07A, L10A) or the first sample of the three study years (e.g. J06, CdT06) are featured by black squares, the second samples by red dots, the third samples by green triangle and the last samples by blue diamond. A boxplot featured in grey showed significant differences of the study samples to the other samples.

Fig. 2 Kin-related and relatedness coefficient estimates for larval samples sorting according to ascending relatedness coefficients. Larval samples exhibiting ID are featured in red, the other in blue.

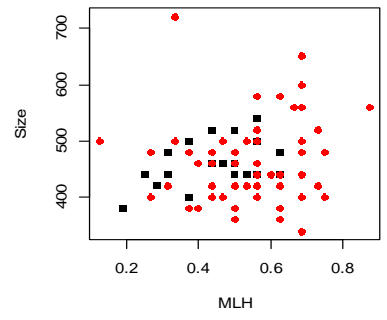
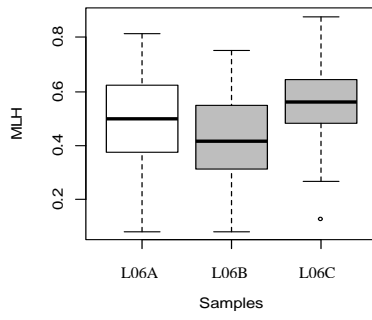
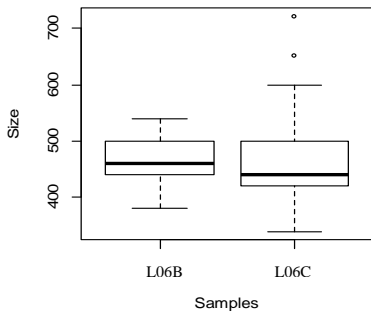


Table 1 Genetic diversity of the study samples using 16 microsatellite markers for adults, larvae and juveniles, all sampled in the same site in the bay of Morlaix. N: number of individuals successfully genotyped, He: expected heterozygosity,  $F_{IS}$ : fixation index, MLH: Multi-Loci Heterozygote,  $g_2$ : estimate of the identity disequilibrium, sd: standard deviation of  $g_2$ . Probability values for  $g_2=0$  were corrected for multiple comparisons: ns: non-significant ( $p>0.05$ ), \* ( $p<0.05$ ), \*\* ( $p<0.01$ ), \*\*\* ( $p<0.001$ ).

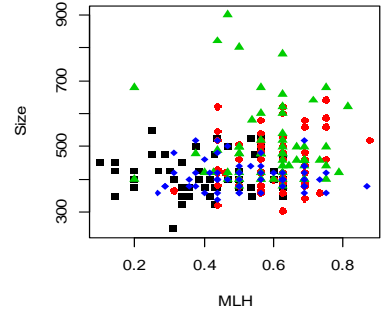
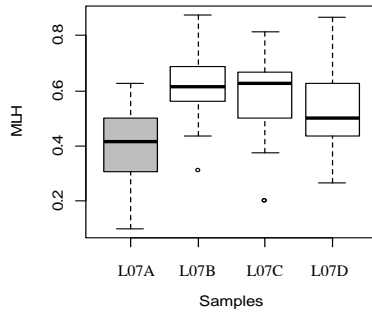
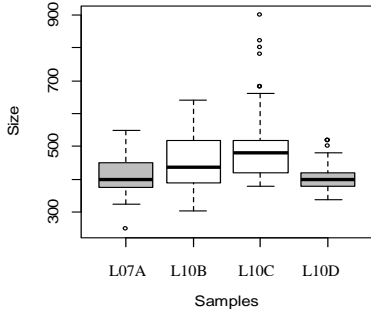
sample phase	life cycle stage	date of sampling	labels	N	He	$F_{IS}^{(1)}$	mean MLH	varMLH	$g_2$	sd
<b>2006</b>										
benthic	adults	18-Apr	CdT06	88	0.675	0.130	5.86	1.27	0.002ns	0.006
pelagic	larvae	-	L06	180	0.679	0.265	4.94	2.54	<b>0.066***</b>	0.015
pelagic	larvae	4-May	L06A	56	0.679	0.239	5.04	2.88	<b>0.085***</b>	0.030
pelagic	larvae	4-Jul	L06B	60	0.674	0.373	4.24	2.20	<b>0.065***</b>	0.028
pelagic	larvae	29-Sep	L06C	64	0.676	0.184	5.44	1.86	<b>0.024***</b>	0.015
benthic	juveniles	19-Sep	J06	72	0.678	0.097	6.12	1.10	-0.001ns	0.005
<b>2007</b>										
benthic	adults	19-Apr	CdT07	88	0.662	0.082	6.06	1.33	0.007ns	0.007
pelagic	larvae	-	L07	241	0.672	0.205	5.31	2.13	<b>0.040***</b>	0.010
pelagic	larvae	12-Apr	L07A	54	0.664	0.396	3.98	1.11	<b>0.067***</b>	0.032
pelagic	larvae	11-May	L07B	62	0.684	0.134	5.91	2.04	-0.004ns	0.006
pelagic	larvae	8-Jun	L07C	62	0.673	0.132	5.83	1.06	0.006ns	0.009
pelagic	larvae	22-Aug	L07D	63	0.661	0.193	5.31	1.49	<b>0.028***</b>	0.016
benthic	juveniles	20-Aug	J07	81	0.650	0.235	4.98	1.85	-0.001ns	0.007
<b>2010</b>										
benthic	adults	16-Apr	CdT10	48	0.660	0.164	5.47	2.05	0.004ns	0.021
pelagic	larvae	-	L10	275	0.663	0.137	5.71	1.41	<b>0.008*</b>	0.005
pelagic	larvae	5-May	L10A	66	0.658	0.175	5.40	1.80	<b>0.026***</b>	0.015
pelagic	larvae	9-Jun	L10B	65	0.668	0.147	5.67	1.55	0.009ns	0.010
pelagic	larvae	4-Aug	L10C	72	0.658	0.121	5.78	1.17	-0.001ns	0.006
pelagic	larvae	2-Sep	L10D	72	0.667	0.108	5.94	1.06	-0.003ns	0.006
benthic	juveniles	19-Jan-11	J10	62	0.671	0.074	6.19	1.75	0.012ns	0.010

<sup>(1)</sup>: p-values associated with HWE < 0.0001 in each case.

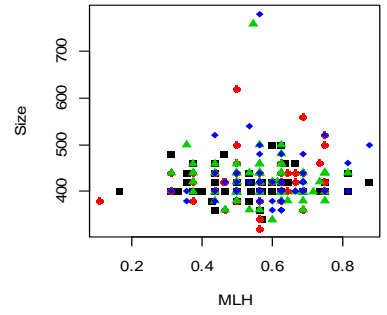
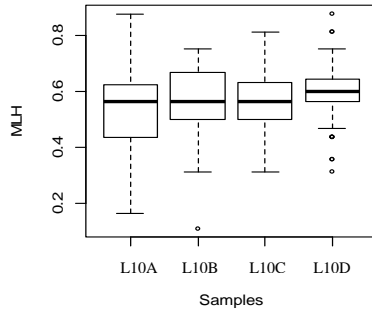
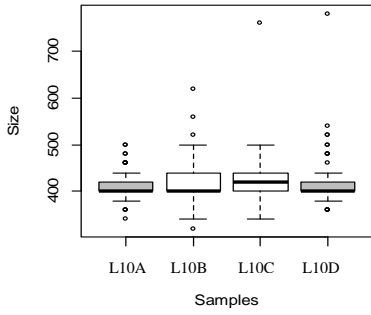
A. Larvae 2006



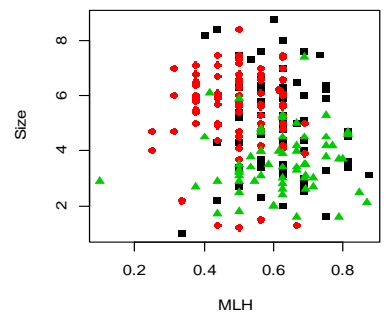
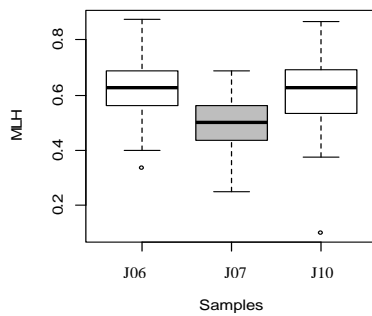
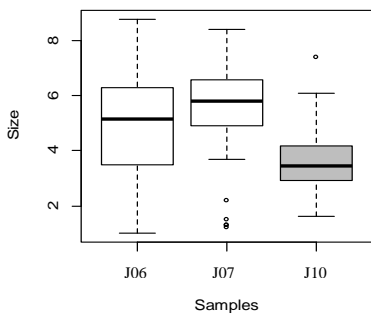
B. Larvae 2007



C. Larvae 2010



D. Juveniles



E. Adults

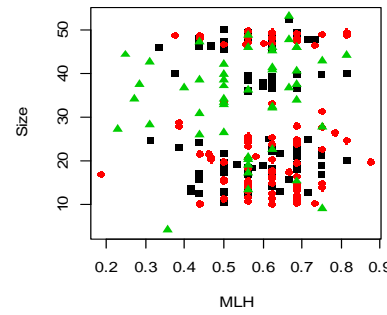
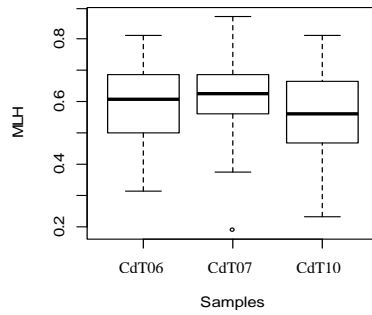
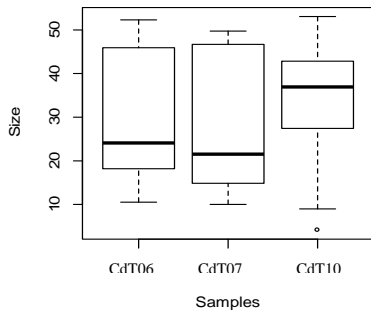


Fig. 1

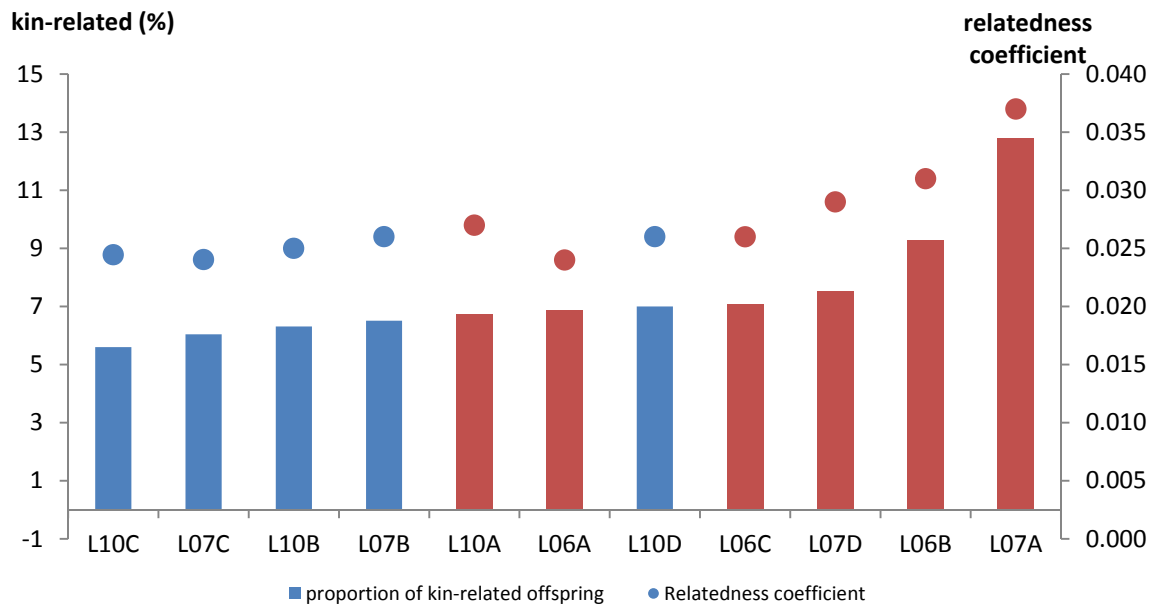


Fig. 2

## Supplementary Information

Fig. S1 Correlations between five individual heterozygosity estimates: Multi-Locus Heterozygosity (MLH), standardized heterozygosity based on the mean expected heterozygosity (Hs\_obs, Coltman *et al.* (1999), standardized heterozygosity based on the mean observed heterozygosity (Hs\_exp, Amos 2005), Internal Relatedness (IR, Amos *et al.* 2001) and Homozygosity by locus (HL, Aparicio *et al.* 2006). A strong correlation was observed with Spearman correlation coefficients higher than 0.89 or lower than -0.91 ( $r(\text{MLH-Hs\_obs}) = 0.94$ ;  $r(\text{MLH-Hs\_exp}) = 0.99$ ;  $r(\text{MLH-IR}) = -0.98$ ;  $r(\text{MLH-HL}) = -0.96$ ;  $r(\text{H\_obs-Hs\_exp}) = 0.94$ ;  $r(\text{H\_obs-IR}) = -0.91$ ;  $r(\text{H\_obs-HL}) = -0.90$ ;  $r(\text{H\_exp-IR}) = -0.98$ ;  $r(\text{H\_exp-HL}) = -0.96$ ;  $r(\text{IR-HL}) = 0.96$ , for all pairwise correlations  $p < 0.0001$ , adjusted for multiple comparisons).

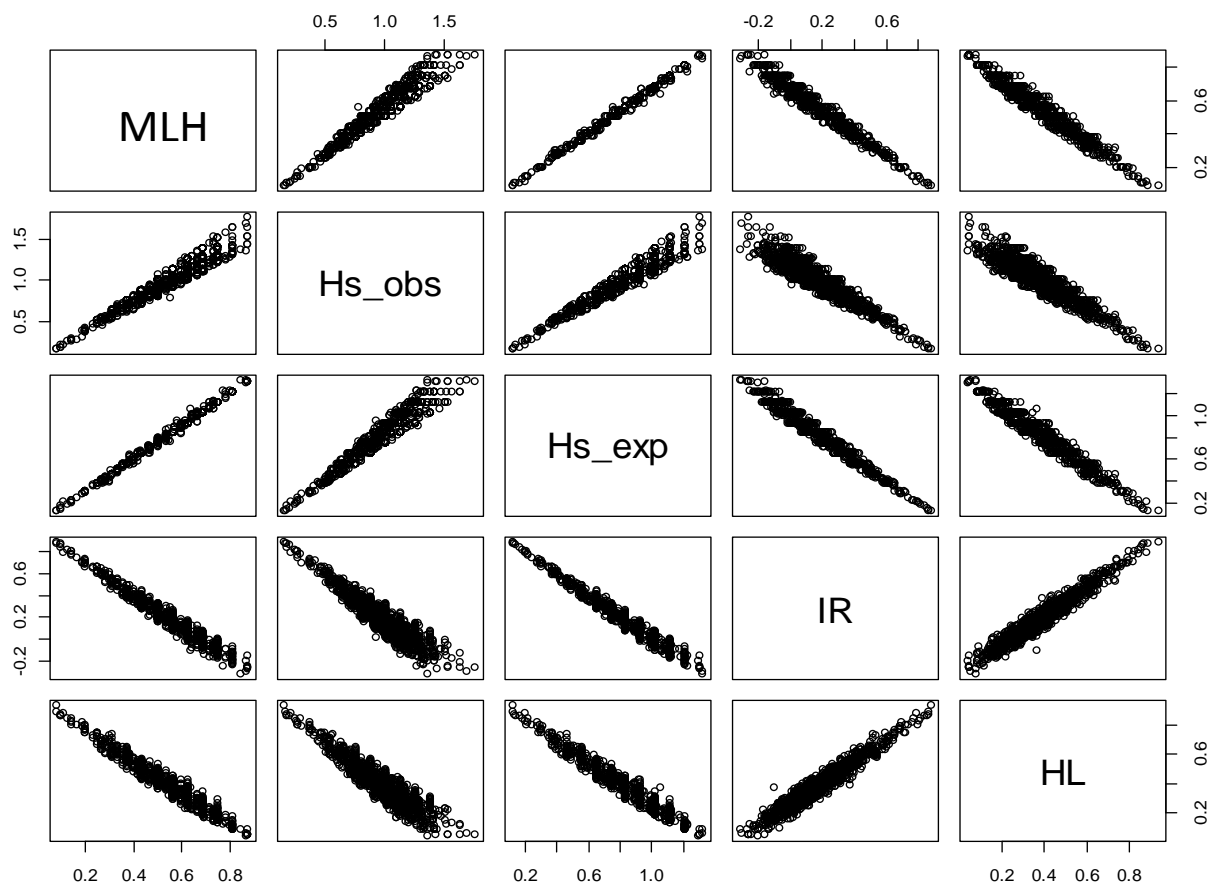


Fig. S2 Distributions of observed multi-locus heterozygosity (in striped blue) and expected multi-locus heterozygosity under random mating (in red) for the larval samples exhibiting identity disequilibrium.

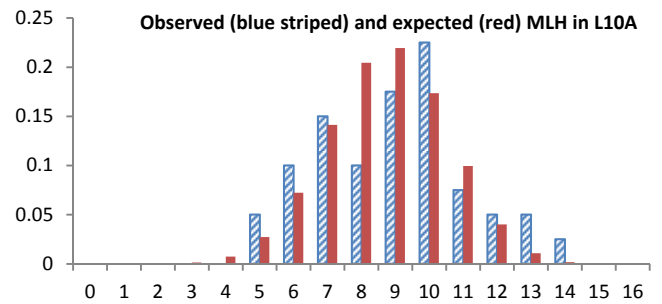
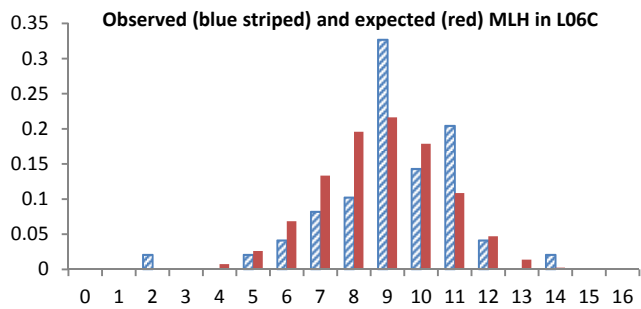
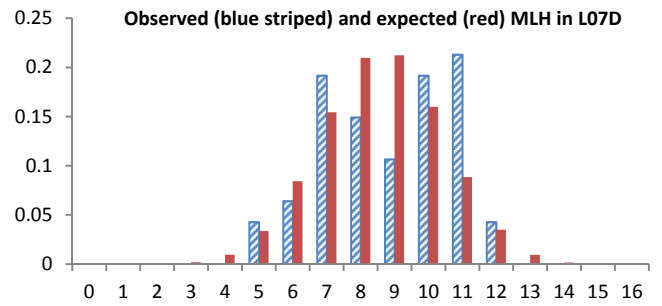
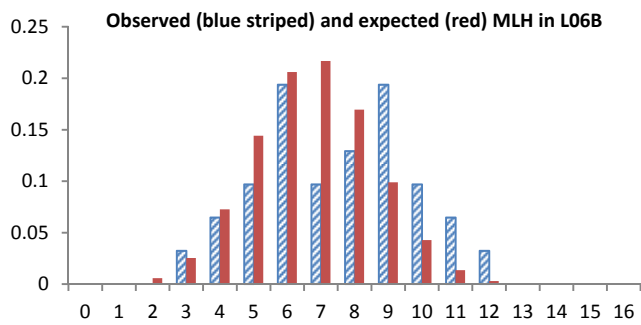
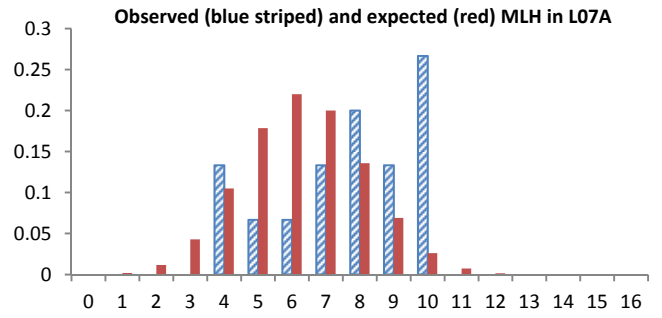
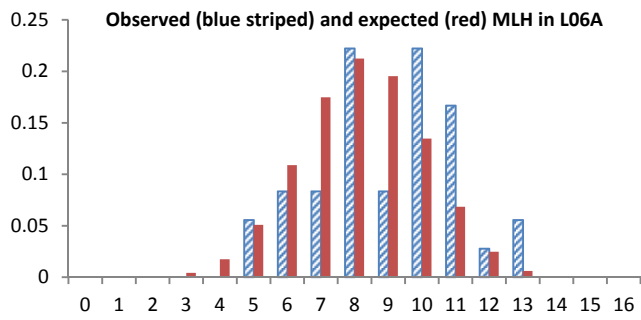


Fig. S3 MLH according to the number of genotyped loci for adults, larvae and juveniles. These graphs suggest that larvae genotyped for less than 16 loci were more easily homozygotes (i.e. lower MLH) than adults and juveniles genotype for less than 16 loci. This suggests that larvae are more sensitive to genotyping error if amplification failed for some loci.

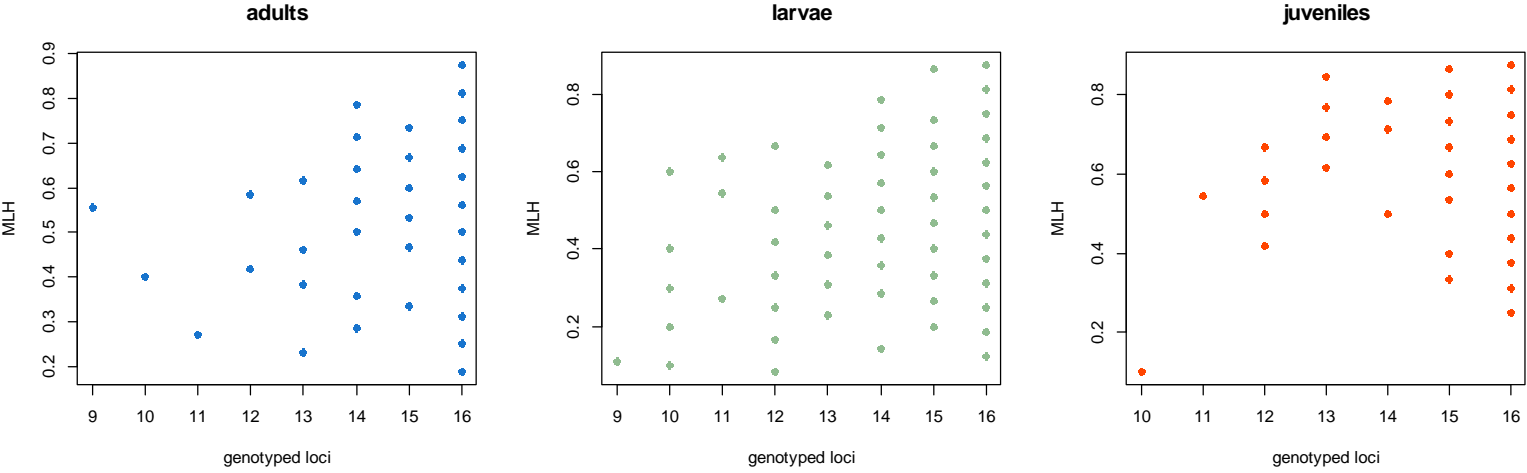


Figure S2 Distributions of the relatedness coefficients in eight simulated datasets with 100 full sibs, 100 half sibs and 100 unrelated individuals and with no genotyping error accounted (Sim1), genotyping error of 0.001 for all loci (Sim2), genotyping error of 0.005 for all loci (Sim3), genotyping error of 0.01 for all loci (Sim4), genotyping error of 0.015 for all loci (Sim5), genotyping error of 0.1 for all loci (Sim6), genotyping error of 0.15 for all loci (Sim7), genotyping error of 0.2 for all loci (Sim8).

Distributions of relatedness coefficient were significantly different between full sibs, half sibs and unrelated individuals (ANOVA,  $p < 0.001$ ). When considering full sibs only, the distribution for Sim7 and Sim8 were significantly different from the six other simulations (Sim1-Sim6) (ANOVA,  $p < 0.001$  with Sim7, Sim8 or both simulations,  $p = 0.25$  for Sim1-Sim6). No significant difference in relatedness estimates was evidenced when comparing the eight simulations for half sibs and unrelated individuals (ANOVA,  $p > 0.05$ ).

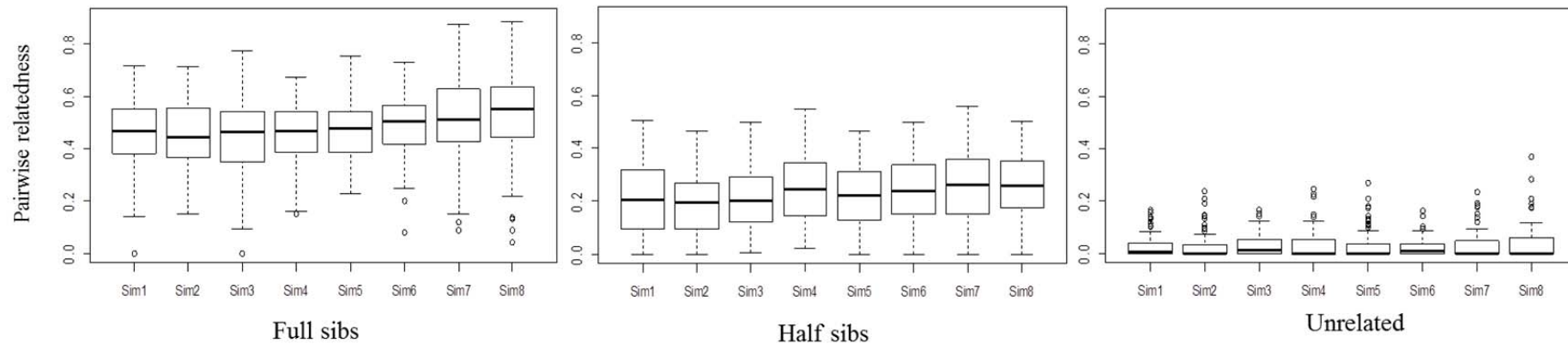
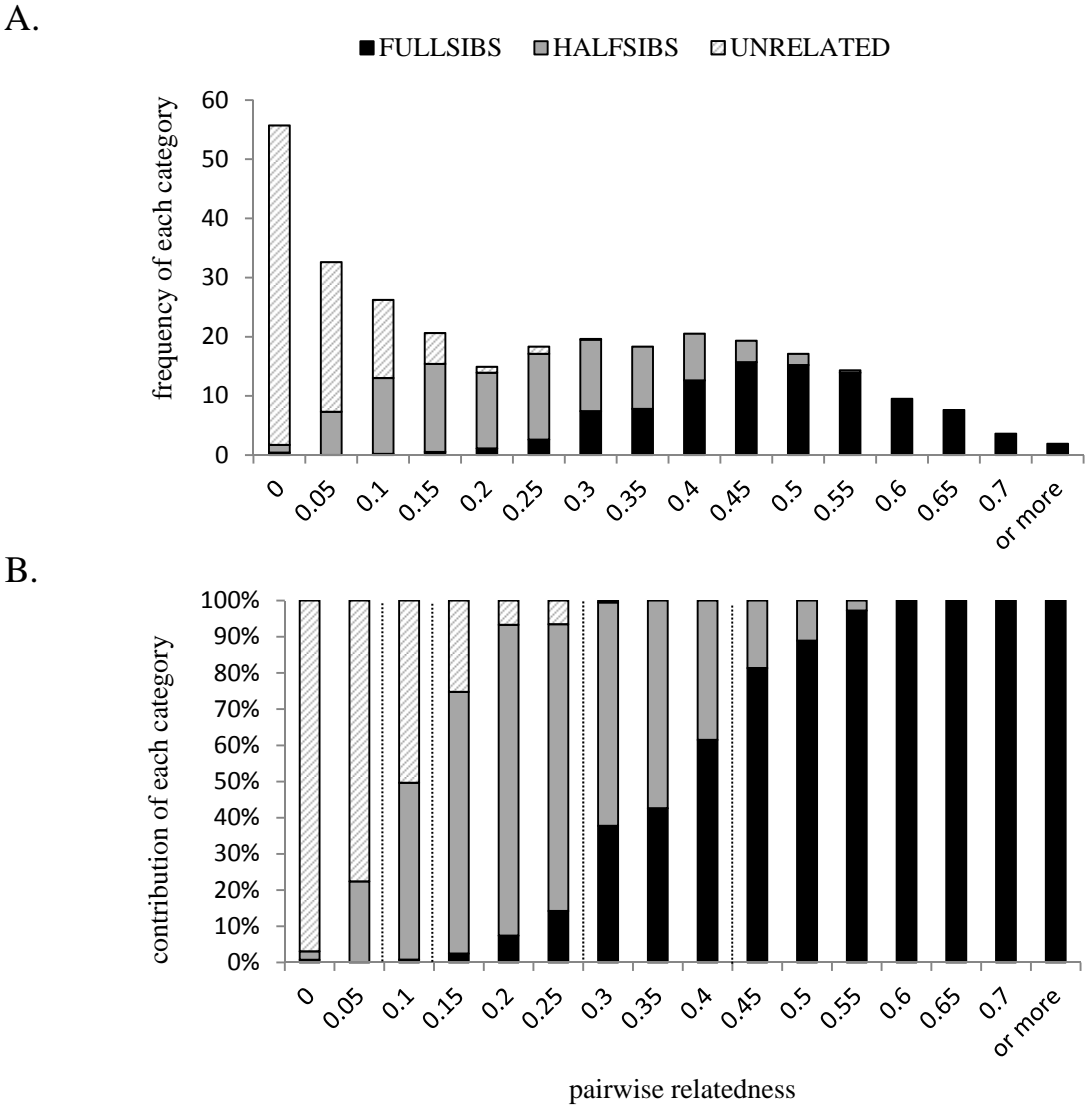


Figure S3. Results of simulations showing the distribution of pairwise relatedness coefficients for the three targeted categories, i.e. full sibs, half sibs and unrelated individuals. Ten simulations were carried out with genotyping error of  $10^{-3}$  for each locus.

For each category, no significant difference in relatedness estimates has been demonstrated across the simulations (ANOVA,  $p > 0.05$ ). Results of the 10 simulated datasets were thus pooled. The distributions of relatedness coefficients for full sibs, half sibs and unrelated individuals were significantly different (ANOVA,  $p < 0.001$ ). The graphs below show: A- the percentage of pairs of individual that are full sibs, half sibs or unrelated per relatedness size class of 0.05 and B- the contribution of full sibs, half sibs and unrelated individuals in each relatedness size class.

Thresholds for categories assignments were defined based on the latest distribution (B) by considering that a category is correctly identified if it is contributing for more than 75% of the class. Unrelated individuals were thus defined by a relatedness coefficient ranging from 0 to 0.05, half sibs by a relatedness coefficient ranging from 0.1 to 0.25 and full sibs by a relatedness estimates higher than 0.40. Individuals falling in between these intervals were not assigned to a given category (e.g. relatedness coefficient between 0.05 and 0.1 could define either half sibs or unrelated individuals).





---

## **Synthèse et Perspectives**

---



Dans le cadre de cette thèse, nous avons étudié l'importance de deux forces évolutives, la sélection et la dérive génétique, chez le gastéropode *Crepidula fornicata*, en se plaçant à deux échelles d'espace très différentes : nous avons dans un premier temps recherché des indices d'évènements de sélection dus au processus d'introduction de ce mollusque, en intégrant dans l'analyse son aire de distribution naturelle et ses aires d'introduction, et, dans un second temps, nous avons examiné l'importance des mécanismes de dérive génétique à une échelle locale (*i.e.* une baie) au sein de l'aire d'introduction. Les résultats majeurs obtenus à partir de ce travail de thèse peuvent être résumés en deux points:

- Aucun mécanisme de sélection ayant agi postérieurement à l'introduction de *C. fornicata* en Europe et sur la côte ouest des Etats-Unis n'est mis en évidence ; en revanche, huit locus outliers ont été identifiés au sein de l'aire native, pouvant être mis en relation avec des processus adaptatifs ayant évolué sur le long terme le long des côtes de la Floride;
- Différentes prédictions de l'hypothèse dite de « sweepstakes » (traduit ici par « loterie familiale ») ont été validées chez *C. fornicata*, en particulier des effets de dérive génétique, une instabilité génétique des cohortes larvaires et des juvéniles et la présence de larves et juvéniles apparentés. Ces observations suggèrent une forte variance du succès reproducteur des adultes de la population étudiée.

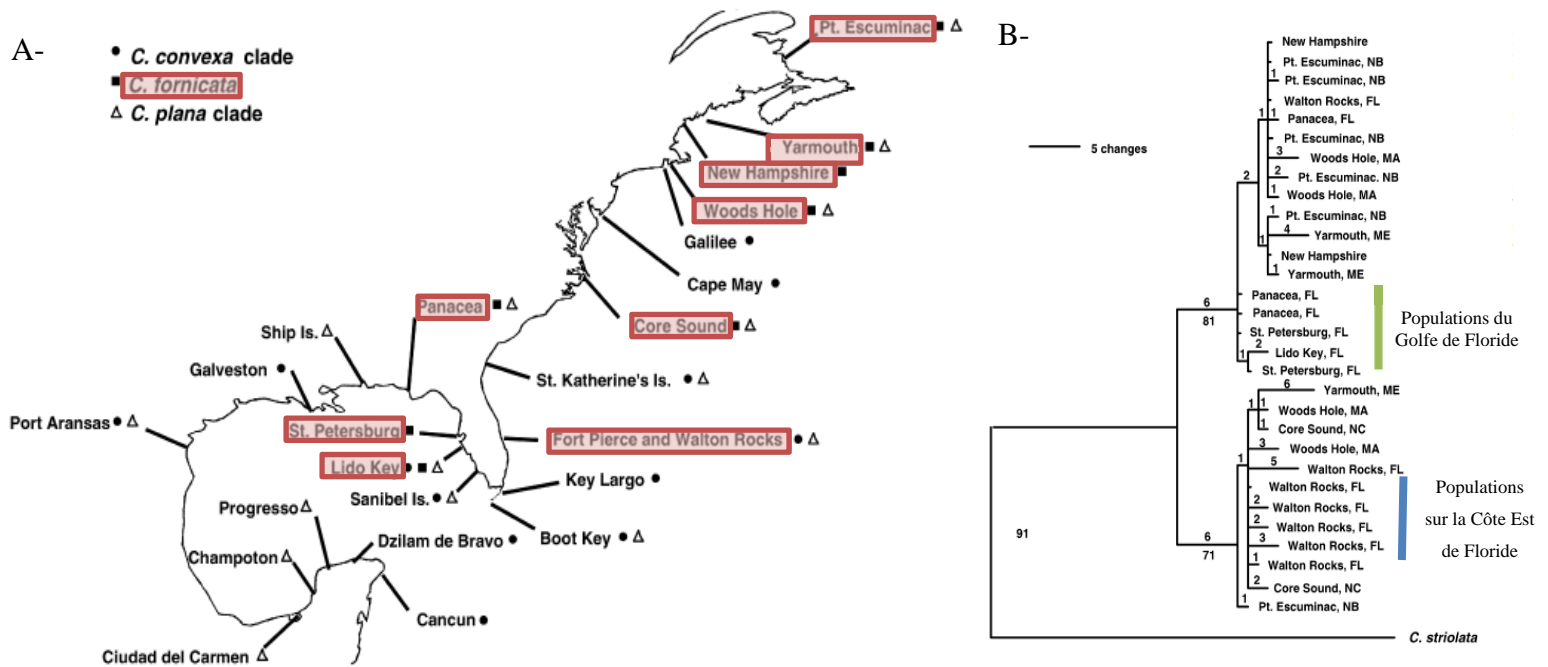
Ces résultats seront discutés ci-dessous en regard des connaissances acquises antérieurement à ce travail. Certaines des limites de ces études seront abordées et des perspectives de travail envisagées.

## **I. Processus de sélection dans l'aire de distribution naturelle de *C. fornicata* : un exemple de couplage entre barrières endogènes et exogènes ?**

L'étude basée sur une approche par scan génomique, présentée dans l'article 2, a mis en évidence une structure génétique significative élevée impliquant la population de Longboat Key (LBK) située dans le Golfe de la Floride. En effet, par exemple avec les 17 locus microsatellites, les valeurs de  $F_{ST}$  estimées entre cette population et une autre population de la côte Est américaine (de la Floride jusqu'au Massachussets) sont en moyenne de 0.081, une valeur 7 fois plus élevée que celle observée entre les autres populations de cette côte ( $F_{ST}$  moyen de 0.011). Cette structure génétique est importante en comparaison de ce qui est traditionnellement observé chez les organismes marins (*e.g.* Hauser et Carvalho 2008, Hedgecock *et al.* 2007). De plus, huit marqueurs AFLPs parmi les 344 marqueurs analysés, ont été identifiés comme des marqueurs directement ou indirectement sous sélection différenciant LBK des autres populations de l'aire d'origine.

Une telle structure génétique se situant le long de la péninsule de Floride avait déjà été mise en évidence par Collin (2001) dans le cadre d'une étude de phylogéographie comparative de 3 espèces de crépidules de la côte Est américaine. Dans cette étude, utilisant un marqueur mitochondrial (le gène codant pour la cytochrome oxydase I), peu de spécimens par population avaient été analysés mais la

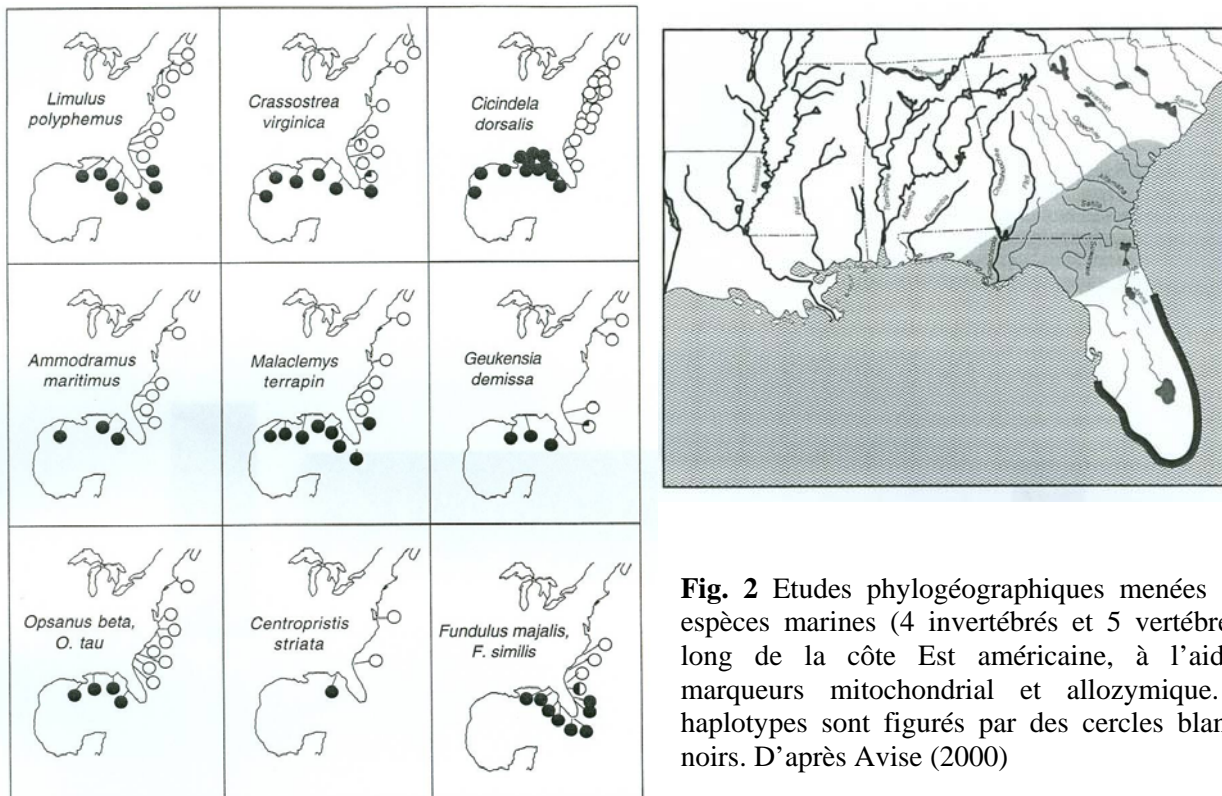
couverture géographique, notamment du Golfe de Floride, était plus importante que la nôtre (Fig. 1A). Collin (2001) avait observé une structure génétique significative ( $p=0.029$ ) le long du gradient longitudinal et la distribution des individus de Floride dans deux clades soutenus par une divergence de 2% (Fig. 1B). Les individus échantillonnés dans les 3 populations du Golfe se retrouvent dans un clade et ceux échantillonnés sur la côte Est de la Péninsule se rassemblent dans le second clade, qui regroupe majoritairement les individus des populations situés plus au nord. Par ailleurs, l'auteur a montré l'existence d'une barrière génétique le long de la côte Est de Floride, au Cap Canaveral (situé 150km au Nord de Fort-Pierce) chez deux autres espèces du même genre, *Crepidula convexa* et le complexe d'espèces *Crepidula plana* (*C. plana*, *C. depressa* et *C. atrasolea*). Cette barrière phylogéographique semble située dans la même région pour les trois espèces étudiées bien qu'elles présentent des cycles de vie différents (développement direct pour *C. convexa* et indirect pour *C. fornicata* et *C. plana*).



**Fig. 1** Localisation des échantillonnages de *C. fornicata* (en rouge), *C. plana* et *C. convexa* (A-) le long de la côte Atlantique Nord et arbre phylogénétique réalisé pour *C. fornicata* (9 localités) à partir de 640pb du COI (B-). Cap Canaveral est situé 150 km au Nord de Fort-Pierce et Longboat Key est situé près de Lido Key. D'après Collin (2001)

La singularité génétique des populations de Floride coïncide avec la barrière biogéographique bien connue le long de cette péninsule. En effet, il s'agit d'une zone de transition entre les écosystèmes marins tempérés, situés au Nord de cette barrière, et tropicaux, situés au Sud-Ouest de cette barrière (Hare et Avise 1996, Pelc *et al.* 2009). De nombreuses études phylogéographiques réalisées chez des espèces marines (vertébrés et invertébrés) de la côte Est américaine ont mis en évidence de façon

récurrente la présence de deux clades le long de cette côte, se rejoignant au niveau de la Péninsule de Floride (Fig. 2). On peut noter que la position géographique de ces barrières phylogéographiques varie selon les espèces, y compris pour des espèces présentant des cycles de vie similaires, *e.g.* *Limulus polyphemus* et *Crassostrea virginica* deux invertébrés marins au développement indirect dont la barrière génétique ne se situe au même endroit (Fig. 2)

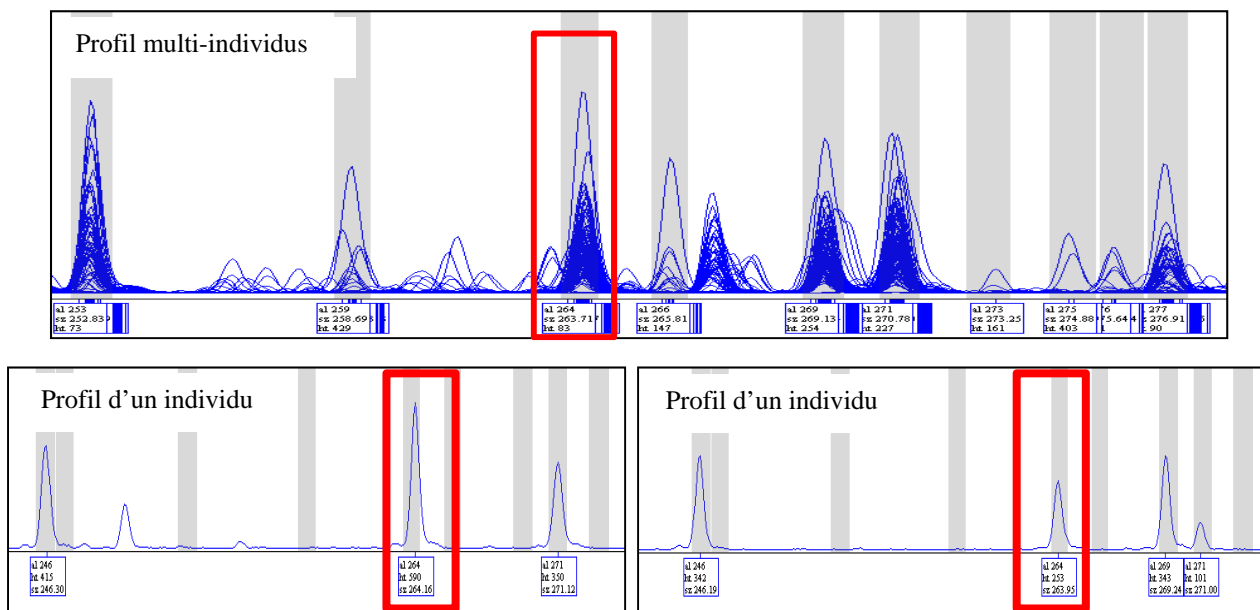


**Fig. 2** Etudes phylogéographiques menées sur 9 espèces marines (4 invertébrés et 5 vertébrés) le long de la côte Est américaine, à l'aide de marqueurs mitochondrial et allozymique. Les haplotypes sont figurés par des cercles blancs et noirs. D'après Avise (2000)

Ces « cassures » phylogéographiques, et dans notre cas l'identification de huit marqueurs vraisemblablement influencés par des processus de sélection, peuvent refléter plusieurs mécanismes. D'une part, elles pourraient être expliquées par des mécanismes d'adaptation locale de part et d'autre de la zone de transition en relation avec des gradients environnementaux. Cette hypothèse a été classiquement avancée dans la littérature des espèces marines pour expliquer des clines de fréquence allélique sur certains marqueurs et la mise en évidence d'outliers (*e.g.* effet de la salinité, Coyer *et al.* 2011). Bierne *et al.* (2011) proposent d'autres mécanismes pour expliquer ce type de patrons. Ces auteurs expliquent que l'identification des locus outliers coïncide avec une frontière environnementale contrastant deux habitats différents ; cependant le cline de fréquence allélique observé ne résulte pas directement d'adaptations locales mais d'incompatibilités génétiques intrinsèques réduisant le flux de gènes sur une partie du génome. Cette barrière endogène, *i.e.* indépendantes des facteurs environnementaux, est « piégée » au niveau d'une frontière environnementale, coïncidant avec des barrières exogènes, elles-mêmes liées aux habitats contrastés. Bierne *et al.* (2011) soulignent qu'en milieu marin, il y a *a priori* peu de barrières environnementales ; on s'attend alors à ce que les barrières génétiques endogènes soient couplées au niveau des rares barrières physiques (*e.g.* front

océanique). S'appuyant sur le travail de Barton (1979), ces auteurs soulignent que de telles barrières peuvent varier géographiquement entre des espèces présentant des capacités de dispersion différentes, et notamment à développement direct *versus* indirect ; l'échantillonnage de Collin (2001) n'est malheureusement pas assez resserré pour analyser finement la position des barrières génétiques chez les trois espèces du genre *Crepidula* étudiées.

Une perspective de ce travail serait de poursuivre le travail initié au cours de cette thèse dans l'aire de distribution naturelle de *C. fornicata*. Il s'agirait dans un premier temps de localiser à une échelle plus fine la barrière génétique observée à l'aide d'un effort d'échantillonnage plus important le long de la Péninsule de Floride chez *C. fornicata* mais également chez *C. convexa* et *C. plana*. Par ailleurs, si ce sont bien des facteurs endogènes qui sont à l'origine des patrons observés, il serait intéressant d'étudier si des isolements reproducteurs sont observés lors de croisements entre individus collectés de part et d'autre de la barrière (ou des accouplements préférentiels entre individus issus de la même « région »). Enfin, bien qu'il soit très peu probable que les outliers détectés chez *C. fornicata* soient directement soumis à sélection et que les gènes sous sélection soient difficiles à identifier (Barton et Bengtsson 1986, Bierne *et al.* 2011), il pourrait être intéressant de caractériser plus précisément les huit outliers identifiés. Il s'agirait de les séquencer et d'accéder aux régions du génome environnantes, sous l'hypothèse que les régions séquencées ne soient pas « anonymes » mais correspondent à des gènes annotés chez d'autres espèces.



**Fig. 3** Electrophorégrammes du locus outlier numéroté 4 dans l'article 2 : cet outlier pourrait facilement être isolé sur gel car il ne présente pas de pics adjacents au pic outlier.

## II. *Crepidula fornicata* : une espèce introduite avec succès hors de son aire de distribution naturelle

De nombreux scénarios évolutifs sont proposés pour expliquer le succès d'une espèce dans un environnement dans lequel elle n'a pas évolué dont les 3 suivants (non exclusifs les uns des autres):

(1) **Scénario 1** : L'introduction de génotypes déjà adaptés à ce nouvel environnement et sur laquelle la sélection va agir (Barrett et Schluter 2008). Parmi les génotypes introduits, certains mieux adaptés à ce nouvel environnement vont être favorisés. Sous ce scénario, le niveau de diversité génétique du « pool d'individus » introduits apparaît ainsi comme un facteur déterminant. Ce scénario a été particulièrement mis en avant lors d'introductions biologiques en milieu marin car une forte diversité génétique est souvent observée dans les populations introduites, notamment du fait de fortes pressions en propagules (*i.e.* nombre d'individus introduits) et de multiples introductions de différentes sources (*e.g.* Goldstien *et al.* 2011, Simon-Bouhet *et al.* 2006, Roman et Darling 2007).

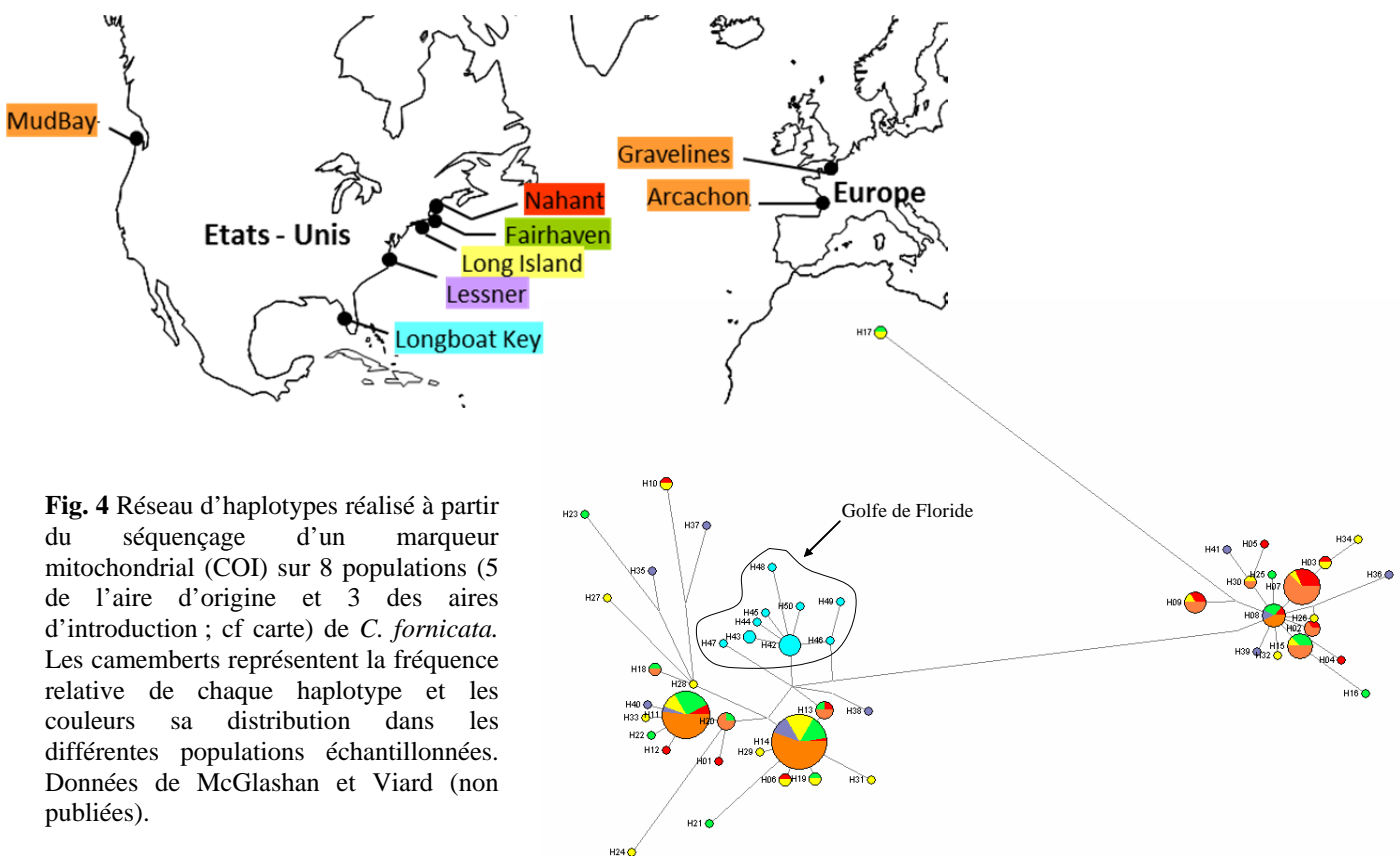
(2) **Scénario 2** : La combinaison de génotypes (en anglais « admixture ») issue de la reproduction entre individus venant de plusieurs populations sources génétiquement différenciées ; ces nouveaux génotypes peuvent être plus « avantageux » que les génotypes originaux issus de chacune des sources (*e.g.* Rius *et al.* 2012). La multiplicité d'introductions via de multiples sources souvent observée lors d'introductions biologiques marines offre des conditions favorables à la réalisation de ce scénario, chez les espèces pour lesquels les flux de gènes dans l'aire de distribution naturelle sont faibles (*e.g.* espèces à développement direct ou à phase dispersive de courte durée de vie).

(3) **Scénario 3** : la capacité des individus à moduler leurs phénotypes à un trait donné en fonction de l'environnement, *i.e.* la plasticité phénotypique (*e.g.* Richards *et al.* 2005, Sultan 2001). Pour les espèces invasives, deux hypothèses sont envisagées (Richards *et al.* 2006): les espèces qui ont réussi leur introduction présentent des caractères plus plastiques que les autres espèces (*e.g.* McDowell 2002) et/ou elles présentent une plasticité phénotypique plus importante dans l'aire d'introduction que dans l'aire native (*i.e.* évolution vers une plasticité accrue ; *e.g.* Parker *et al.* 2003). Ce dernier scénario est plus difficile à évaluer : il requiert des expérimentations en conditions contrôlées sur des individus de l'aire d'origine et de l'aire d'introduction, ce qui en pratique est souvent difficilement envisageable, pour des raisons pratiques et légales.

Rappelons par ailleurs que, comme souligné par Facon *et al.* (2006), les modifications de l'environnement dans lequel a été introduite l'espèce jouent aussi un rôle critique sur le succès de cette espèce, et notamment dans le cadre de « co-variations » entre l'environnement et l'espèce introduite.

Chez *C. fornicata*, la très faible structure génétique entre populations ayant servi de sources à l'introduction et l'absence de variations génétiques entre l'aire d'origine et les aires d'introduction, ne soutient pas l'hypothèse de succès expliqué par des processus d'admixture (scénario 2). En outre, une

forte diversité génétique et une faible structure génétique (sauf concernant le Golfe de Floride) dans les aires natives et d'introductions ainsi qu'une faible structure génétique entre les deux types d'aires ont été observées. Ces résultats sont compatibles avec un scénario d'introduction caractérisé par une forte pression en propagules et de multiples introductions de différentes sources (scénario 1). Ce scénario avait déjà été proposé lors de différentes études génétiques menées avec des marqueurs enzymatiques (Hoagland 1985, Dupont *et al.* 2003). En effet, dans ces deux études, la diversité génétique d'une population de l'aire d'introduction ou de l'aire d'origine avait été analysée en regard de plusieurs populations de l'aire d'origine ou de l'aire d'introduction (Hoagland 1985 et Dupont *et al.* 2003 respectivement). McGlashan et Viard (données non publiées) avaient également montré à l'aide d'un marqueur mitochondrial, que les populations introduites étaient génétiquement diversifiées et présentaient sur la côte Ouest comme en Europe des haplotypes observés le long de la côte Est Américaine, à l'exception de la Floride (conformément aux résultats de Collin (2001) discutés auparavant ; Fig. 4).



**Fig. 4** Réseau d'haplotypes réalisé à partir du séquençage d'un marqueur mitochondrial (COI) sur 8 populations (5 de l'aire d'origine et 3 des aires d'introduction ; cf carte) de *C. fornicata*. Les camemberts représentent la fréquence relative de chaque haplotype et les couleurs sa distribution dans les différentes populations échantillonnées. Données de McGlashan et Viard (non publiées).

Ainsi quels que soient les marqueurs utilisés, une forte diversité génétique est observée (allozymes : Hoagland 1985, Dupont *et al.* 2003, mtDNA : McGlashan et Viard (données non publiées), microsatellites : notre étude), suggérant des conditions favorables pour que des processus de sélection aient agi suite à l'introduction de *C. fornicata* en Europe et sur la côte ouest des Etats-Unis



(scénario 1). Sous l'effet de nouvelles pressions de sélection, des changements évolutifs importants et rapides, observables sur une vingtaine de générations (Prentis *et al.* 2008), sont possibles pour expliquer la réussite d'une espèce introduite dans un nouvel environnement (Sakai *et al.* 2001, Lee 2002, Roman et Darling 2007) et, ceci a été notamment observé chez les plantes (*e.g.* Barrett *et al.* 2008). Dans le cas de la crépidule, l'histoire d'introduction semble suffisamment longue pour que des processus de sélection aient pu agir. Mais utiliser la diversité génétique neutre comme proxy de la diversité génétique adaptative est discutable et discutée (*e.g.* Holderegger *et al.* 2006). Aussi rechercher des preuves de processus adaptatifs dans les populations introduites est important pour valider ce scénario. Or ce travail de thèse a montré qu'aucune signature de sélection n'était identifiée entre les aires natives et introduites par l'approche de scans génomiques.

Nos résultats ne remettent pas définitivement en cause que des mécanismes de sélection sur le polymorphisme préexistant aient pu avoir lieu, mais nous suggérons que les différentiels de sélection entre les deux environnements sont trop faibles pour laisser une trace détectable par des scans génomiques. Quasiment aucune donnée n'est encore disponible dans la littérature pour pouvoir comparer nos résultats avec d'autres situations (histoire de l'invasion, traits de vie de l'espèce etc.). En croisant dans la base de recherche Web of Knowledge, les termes « invasive species \* genome scan », « alien \* genome scan », « non indigeneous species \* genome scan » ou encore « invader\* genome scan », seule une étude correspond à une introduction biologique récente (*i.e.* dispersion à longue distance ayant moins de 200 ans) analysée par des scans génomiques. Il faut noter qu'un biais de publication est possible, les études ne montrant « aucun résultat », *i.e.* aucun mécanisme de sélection identifié, pouvant ne pas être publiées. Dans l'étude identifiée lors de cette recherche bibliographique (Wang *et al.* 2012), les auteurs se sont intéressés à l'introduction de la liane Américaine *Mikania micrantha* (Fig. 5), une plante originaire des zones tropicales d'Amérique centrale et d'Amérique du Sud, introduite notamment en Chine à la fin des années 1980 et début des années 1990.



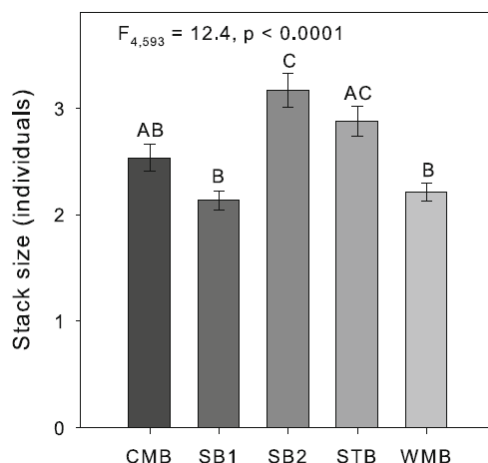
**Fig. 5** *Mikania micrantha*, la liane Américaine, originaire d'Amérique centrale et du Sud, devenue invasive en Inde, à Maurice, en Asie et dans la zone Australienne.

Parmi les 483 AFLPs analysés, 14 outliers ont été identifiés, suggérant l'action de processus sélectifs. Cependant, les scans génomiques ont été conduits sur 28 populations échantillonnées

uniquement dans l'aire d'introduction, et non entre aire native et aire introduite. On ne peut donc pas exclure que ces 14 marqueurs soient déjà soumis à sélection dans l'aire d'origine de *M. micrantha* ou qu'ils soient issus de nouvelles combinaisons de génotypes suite à d'introductions de multiples sources.

Ainsi il serait intéressant de pouvoir étendre des approches par scan génomique à d'autres modèles biologiques ayant des caractéristiques similaires en termes de diversité génétique et d'histoire d'introduction (récurrence, multiples sources). Des mollusques, tels que le gastéropode *Ocenebrillus inornatus* (Martel *et al.* 2004), des crustacés, tels que *Carcinus maenas* (Roman 2006) ou des ascidies telles que *Styela clava* (Goldstien *et al.* 2011) seraient des candidats intéressants pour de telles études.

Des processus d'adaptations consécutifs à des introductions biologiques ont été montrés chez d'autres espèces par d'autres approches (*e.g.* Kollmann et Bañuelos 2004, Leger et Rice 2003). Une autre piste de recherche serait donc de développer des approches alternatives pour l'étude de *C. fornicata*. Cibler un caractère particulier, ayant pu jouer un rôle important dans le succès de cette espèce et analyser les variations pour ce caractère entre aires native et introduites est une possibilité. La grégarité, mesurée par le nombre d'individus par chaîne, pourrait être un tel caractère. En effet, ce nombre varie entre 2,1 et 3,2 (Hoch et Cahill 2012, Fig. 6) dans l'aire d'origine tandis qu'il est de 9,8 au sein de l'aire introduite (données acquises en baie de Morlaix, Le Cam 2009, Fig. 7). Certaines des chaînes que j'ai pu analyser au cours des suivis benthiques (présentés dans le chapitre 1 de la seconde partie) atteignent un nombre de 50 individus.



**Fig. 6** Nombre moyen d'individus par chaîne calculé sur *c.a.* 300 individus prélevés sur 5 sites de l'aire d'origine de *C. fornicata*. D'après Hoch et Cahill (2012)

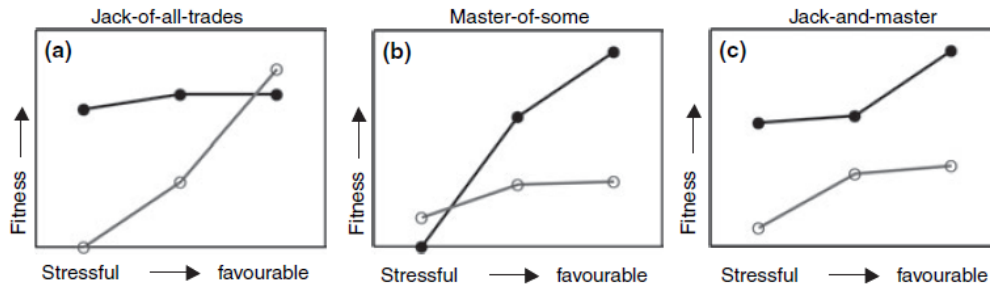


**Fig.7** Chaîne de crépidules échantillonnées au sein de l'aire d'introduction. Cette chaîne est composée d'une quinzaine d'individus. Photo de W. Thomas (SBR)

Ce trait pourrait être adaptatif dans la mesure où il traduit une très bonne résistance à la compétition intra-spécifique et où il favorise la reproduction en maximisant les chances de rencontre

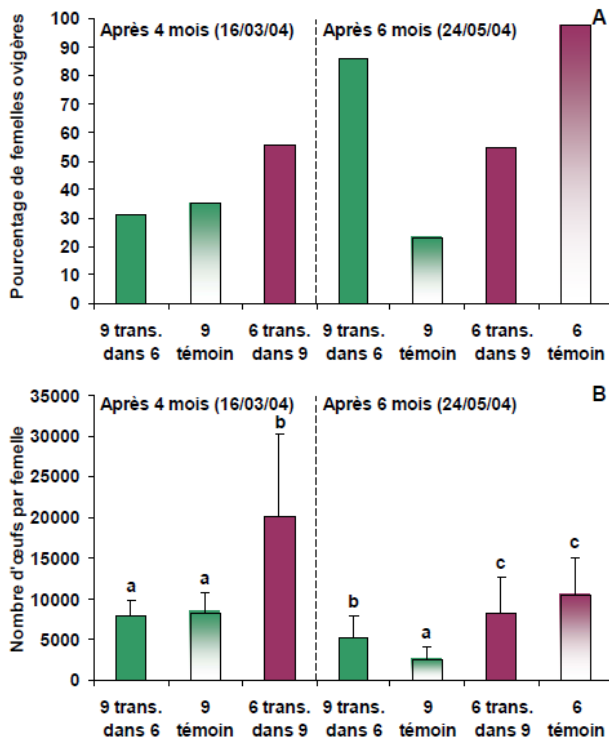
entre partenaires des deux sexes. Il n'existe cependant aucune preuve que ce caractère puisse être soumis à sélection. Il faudrait également pouvoir caractériser avec plus de précisions la moyenne et la variance de ce trait dans les aires de distribution de l'espèce. Nous avons déjà eu l'occasion de suggérer que la croissance larvaire puisse être aussi un bon candidat (variable, héritable). Une série d'étape devrait être franchie dont (1) documenter les changements de distribution de moyenne et variance de ces traits pour lesquels on suppose qu'il existe une base génétique, entre l'aire d'origine et l'aire d'introduction, (2) vérifier l'existence d'une base génétique à la différence des moyennes du trait observé en étudiant des familles issues des deux aires et élevées dans un même environnement, (3) ces traits étant possiblement déterminés par de multiples gènes, se tourner vers des approches de génétique quantitative (Lynch et Walsh 1998) par exemple comparer la différenciation génétique adaptative ( $Q_{ST}$ , Spitze 1993, Whitlock 2008) et la différenciation génétique neutre ( $F_{ST}$ , Wright 1951). Une limite à cette dernière approche est d'estimer la variance génétique additive totale qui requiert des expérimentations en conditions contrôlées (« common garden experiment » ou « expérience en jardin commun »). Une alternative à de telles expérimentations serait d'estimer le  $Q_{ST}$  via la différenciation phénotypique d'un trait particulier entre populations, *i.e.*  $P_{ST}$  (Leinonen *et al.* 2006), relié mathématiquement au  $Q_{ST}$ , mais nécessitant de connaître l'héritabilité pour ce trait. Dans le cas de la croissance larvaire une valeur d'héritabilité (égale à 0.35 au sens large) a été estimée par Hilbish *et al.* (1999). Une autre difficulté, soulignée par Keller et Taylor (2008) est que ces approches menées sur des espèces introduites nécessitent de pouvoir bien connaître le processus d'introduction et les populations sources afin que toute modification phénotypique puisse être attribuée avec confiance au processus d'introduction. Dans notre cas, nous avons pu exclure certaines zones (*e.g.* Golfe de Floride) mais il reste des incertitudes quant aux populations sources, compte tenu de la faible structure génétique dans l'aire naturelle de l'espèce.

Développer des approches phénotypiques pourrait également être pertinent pour tester le troisième scénario évolutif proposé, *i.e.* le rôle de la plasticité phénotypique. La relation entre plasticité phénotypique et succès d'invasion repose sur deux éléments: la capacité de l'individu à conserver sa fitness en dépit des contraintes environnementales (un génotype donné produit un phénotype généraliste, « Jack-of-all-trades », traduit littéralement par « passe-partout ») et la capacité de l'individu à augmenter sa fitness dans ce nouvel environnement (un génotype donné produit un phénotype spécialisé, (« Master-of-some », traduit littéralement par « maîtrise d'un trait »), voire un compromis entre un génotype produisant un phénotype généraliste ou spécialiste (« Jack and Master ») (Baker 1965, cité dans Richards *et al.* 2006 ; Fig. 8).



**Fig. 8** Normes de réaction (*i.e.* relation entre les phénotypes possibles produits par un génotype donné et la gamme de valeurs d'une variable environnementale à laquelle il est exposé) attendues pour des traits de fitness plastiques d'un génotype/espèce/population invasif (en noir) vs. non-invasif (en gris) : (a) l'individu non-natif présente de meilleures valeurs de fitness dans un environnement qui ne lui est pas favorable, ce qui lui conférerait un avantage écologique, (b) l'individu non-natif présente de meilleures valeurs de fitness dans un environnement qui lui est favorable, lui permettant alors d'atteindre des densités de populations importantes, (c) l'individu non-natif présente de meilleures valeurs de fitness dans un environnement qui ne lui est pas favorable et dans un environnement qui lui est favorable. D'après Richards *et al.* (2006)

Richard (2005) a montré d'importantes variabilités phénotypiques au sein des populations européennes de crépidules, que ce soit sur des caractères de croissance (*e.g.* la longueur droite des individus) ou de reproduction (*e.g.* nombre d'œufs lors d'un événement de reproduction) et a montré, à l'aide de transplant entre deux sites de la Rade de Brest, que des traits relatifs à la reproduction (nombre de femelles présentant une ponte et nombre d'œufs par femelle) étaient plastiques (Fig.9).



**Fig.9** Pourcentage de femelles ovigères (A) et nombre d'œufs par femelle (B) 4 mois et 6 mois après le transplant pour les femelles transplantées de l'Aulne (site 9) à Rozegat situé en Rade Brest (site 6, appelé sur la figure « 9 trans. 6 » figuré en vert), pour les femelles transplantées de Rozegat (site 6) dans l'Aulne (site 9, appelé sur la figure « 6 trans. 9 », figuré en rouge) ainsi que pour les femelles témoin dans l'Aulne (site 9, « 9 témoin », figuré dans un dégradé vert) et les femelles témoin à Rozegat (site 6, « 6 témoin », figuré dans un dégradé rouge). Les barres d'erreur représentent les écart-types associés aux valeurs moyennes et les lettres, le résultat de l'analyse *a posteriori* après l'analyse de variance : les histogrammes ne contenant pas la même lettre sont significativement différents les uns des autres ( $p < 0,05$ ). D'après Richard (2005)

Compte-tenu des résultats obtenus au cours de cette thèse avec l'absence de preuves d'évènements de sélection et les résultats de Richard (2005), il semblerait intéressant d'explorer plus avant si *C. fornicata* présente des traits plus plastiques que d'autres espèces non-invasives de l'aire native et/ou de l'aire introduite et de chercher à savoir si la plasticité de certains traits a évolué entre l'aire native et l'aire introduite. Les traits étudiés par Richard (2005) comme le nombre de crépidules par chaînes seraient des candidats intéressants pour une telle étude. La difficulté principale ici étant encore une fois notre capacité à réaliser des expérimentations sur des populations transplantées mais il n'est pas irréaliste de pouvoir étudier des descendance de première génération dans un dispositif expérimental en jardin commun. En effet, s'il est très long d'obtenir plusieurs générations, des expériences réalisées à la Station Biologique de Roscoff ont permis de déterminer des conditions d'élevage propices pour boucler un cycle de vie complet.

**Sélection et Invasion chez *C. fornicata* : trois scénarios sont supposés pour expliquer le succès d'une espèce invasive : la préadaptation, l'admixture et la plasticité phénotypique. Chez *C. fornicata*, seul le scénario d'admixture peut être écarté. Des études complémentaires sont nécessaires pour valider l'un, ou les deux, autre(s) scénario(s) proposé(s). Les perspectives les plus intéressantes seraient de développer des approches phénotypiques, que ce soit pour analyser des évolutions sur des traits phénotypiques ou caractériser la plasticité phénotypique de la crépidule. Cependant, ces études seront ardues et longues compte-tenu de l'importance de l'expérimentation dans de telles approches et des caractéristiques du modèle biologique.**

### **III. La phase larvaire : facteur de succès et moment critique du cycle de vie**

Au cours de sa thèse, François Rigal (2009) s'était intéressé à la dynamique de la phase larvaire chez la crépidule dans la population du Château du Taureau en baie de Morlaix. Ses travaux avaient permis de mettre en évidence la présence de larves pendant huit mois de l'année (avec des densités très variables au cours du temps) mais les données benthiques (thèse de Sabrina Le Cam (2009)) avaient montré un seul épisode de sédentarisation majeur à la fin de l'été – début de l'automne. Compte-tenu de ces informations, Rigal (2009) s'était interrogé sur la composition génétique du pool larvaire notamment afin de savoir s'il existait une homogénéité génétique au cours de la saison de reproduction et si les juvéniles reflétaient la diversité de ce pool larvaire. Notre étude menée sur la même population a permis d'apporter des éléments de réponse : plusieurs échantillons de larves prélevés à différents moments de la saison de reproduction montrent de fortes variations génétiques entre eux et la diversité génétique de ce pool larvaire global n'est retrouvée que partiellement chez les juvéniles indiquant finalement une forte variance des succès reproducteurs individuels. En outre nous avons montré que la taille efficace est probablement inférieure à la taille démographique. De plus, avec l'analyse des juvéniles recrutés sur neuf années consécutives, nous avons montré que la diversité

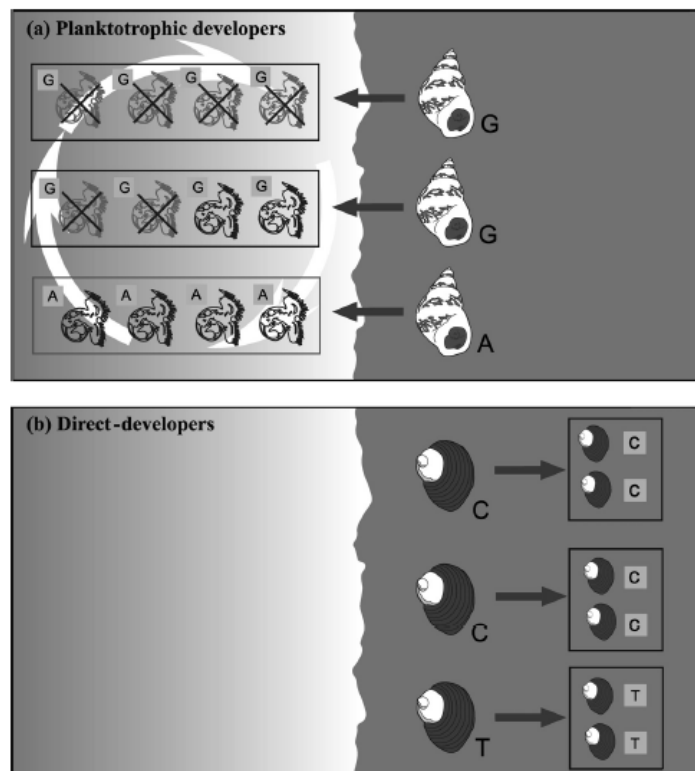
génétique peut être modulée par la dynamique de recrutement. Cependant, en analysant conjointement la dynamique du changement de sexe et la dynamique de recrutement de la population benthique de crépidules, Le Cam (2009) avait montré que le sex-ratio est lui-même modulé par la dynamique de recrutement : par exemple, un apport important de juvéniles en 2006 et 2007 s'est traduit par un fort déséquilibre du sex-ratio (excès de mâle). Ces déséquilibres de sex-ratio sont susceptibles de renforcer la réduction du nombre efficace de reproducteurs dans la population.

L'ensemble de ces résultats sont ainsi tout à fait compatibles avec un scénario de « loterie familiale » (ou « sweepstakes » selon Hedgecock (1994a)). En raison du caractère aléatoire des processus hydrodynamiques affectant la survie des larves dans la colonne d'eau, une loterie s'établie durant la phase larvaire et le recrutement conduisant à une contribution variable des individus à la génération suivante (Hedgecock 1994a). Les prédictions découlant de ce scénario ont été validées chez de nombreuses espèces (voir pour une revue Hedgecock & Pudovkin 2011). Ces derniers auteurs soulignent cependant qu'il est important et encore trop rarement fait de comprendre les mécanismes mis en œuvre tout au long du cycle. Notre étude a cherché à mener une telle approche en combinant l'analyse de phases pré- et post-dispersives (et pour ces dernières, l'étude simultanée de nouvelles recrues et adultes). En analysant les phases pré et post- dispersion (*i.e.* les larves de petites tailles et les juvéniles), nous avons vu que les effets de dérive s'exprimaient principalement au cours de la phase larvaire et l'apparentement entre larves étaient plus fort que celui estimé entre juvéniles. Par ailleurs, l'apparentement entre juvéniles suggère une migration concertée des larves. Une perspective de travail pour poursuivre dans cette voie serait de non seulement analyser les larves à l'émission, ce que nous avons fait ici, mais également des larves de différents âges et à différents moments de leur développement. Ce suivi très fin permettrait de déterminer les moments les plus critiques en termes de dérive génétique (due à la mortalité pendant la dispersion et à la variance initiale du succès reproducteur) à la base de l'hypothèse de sweepstakes. Mais le manque de « plus grandes » larves échantillonnées, *i.e.* des larves plus âgées, ne permet pas d'observer l'évolution de ces processus au cours du temps. Suivre une cohorte larvaire et échantillonner des larves de cette cohorte à des intervalles de temps réguliers permettrait d'analyser ces mécanismes à une plus fine échelle et d'observer notamment quand s'exprime la dérive génétique, comment évolue l'apparentement, etc. Cependant une telle approche apparaît techniquement difficile. Par exemple, les larves sont de petites tailles et ne peuvent être tagguées, et l'utilisation de bouées dérivantes ne prend pas en compte le comportement larvaire (*e.g.* migration verticale).

Il faudrait auparavant valider l'hypothèse que la plupart des larves échantillonnées sont des larves produites localement c'est-à-dire que l'immigration est très faible (ce qui est une hypothèse plausible dans notre étude vu leur très jeune âge mais qui, pour des larves plus âgées, peut être moins réaliste). L'utilisation de modèles couplés biologie-physique de dispersion larvaire pourrait être dans ce contexte particulièrement utile. Les modèles utilisés jusqu'à présent (*e.g.* Viard *et al.* 2006, Dupont *et al.* 2007b) sont trop imparfaits pour analyser finement les dispersions potentielles des larves mais le

développement de modèles lagrangiens de plus en plus sophistiqués (e.g. Ayata *et al.* 2009) représentent une voie prometteuse.

Par ailleurs, il serait également envisageable d'examiner les processus de loterie familiale chez une autre espèce du genre *Crepidula*, mais présentant un développement direct, par exemple *Crepidula convexa*. Les effets de loterie familiale sont attendus comme moins marqués chez de telles espèces (Hedgecock et Pudvokin 2011). Du fait d'une phase dispersive beaucoup moins importante chez *C. convexa*, une structure génétique spatiale plus forte est également attendue, ce qui a notamment été montré par Collin (2001). A titre d'exemple, Lee et Boulding (2009) ont étudié la structure génétique et temporelle de quatre espèces de littorines sur la côte Ouest Américaine. Ils ont notamment montré une structure génétique spatiale plus forte chez les espèces à cycle de développement direct, une structure génétique temporelle plus forte chez les espèces à cycle de développement indirect et une taille efficace de la population plus importante pour les espèces à cycle de développement indirect. Ces processus à l'origine de tels mécanismes sont schématisés dans la figure 10 (Lee et Boulding 2009).

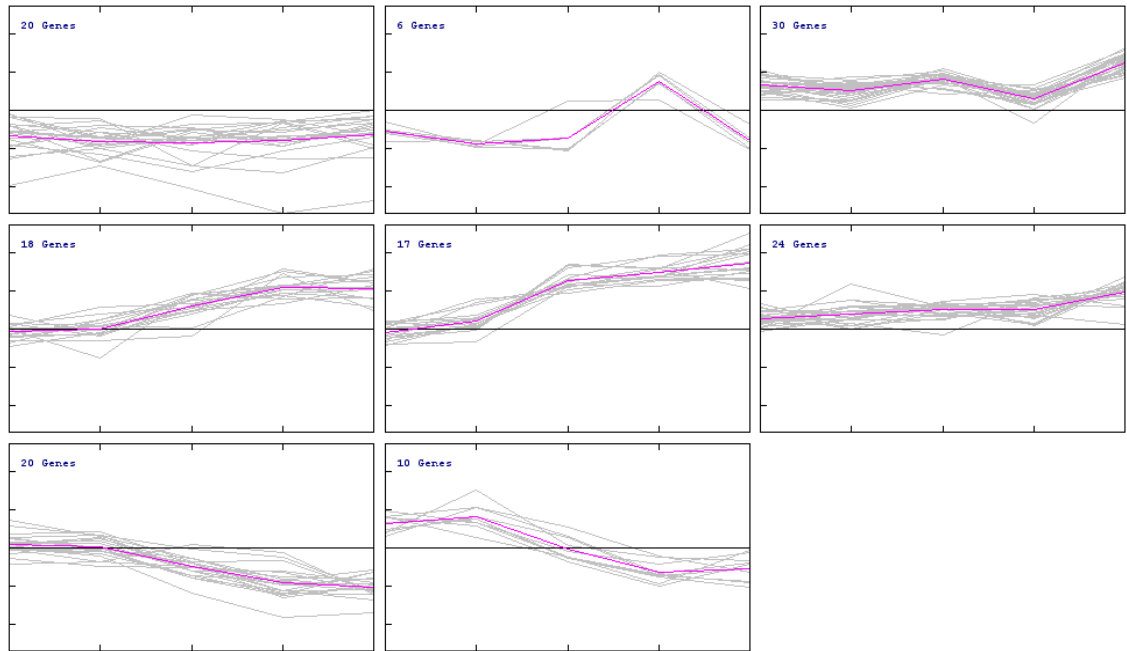


**Fig. 10** Différences du succès reproducteur entre espèces à développement indirect (a-) et espèces à développement direct (b-). La forte fécondité des espèces à développement indirect est schématisée par un nombre de larves deux fois plus importants que le nombre de juvéniles des espèces à développement direct. Les larves véligères ne survivant pas au cours de la phase planctonique sont marquées d'une croix. (a) Une forte hétérogénéité temporelle est observée quand, les larves présentant le génotype le moins commun (« A ») survivent suite à des conditions favorables lors de la phase larvaire ; la structure génétique spatiale est faible en raison de la phase larvaire dispersive. (b) Une faible hétérogénéité génétique est observée, tous les descendants vont contribuer à la génération suivante et les fréquences alléliques ne sont pas modifiées d'une génération à l'autre. D'après Lee et Boulding (2009)

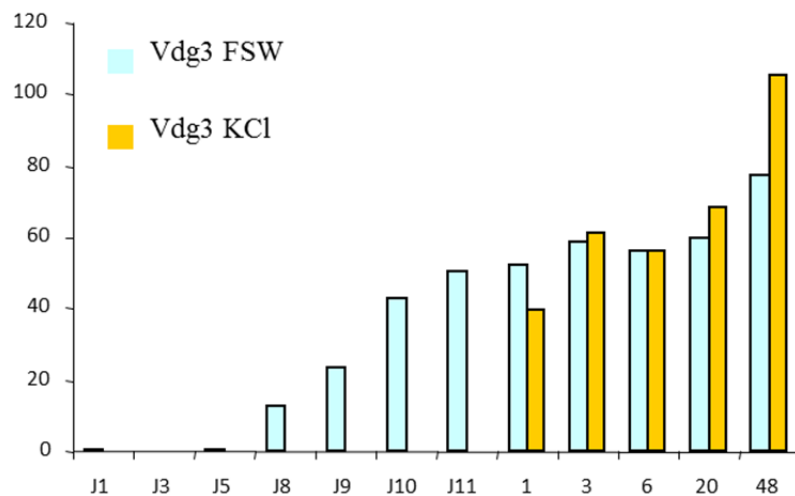
La phase larvaire est une étape clé de la dynamique démo-génétique de la crépidule mais également un trait de vie déterminant dans le succès de son introduction : cette longue phase larvaire a certainement facilité les phases d'installation et d'expansion dans l'aire d'introduction. En effet, les larves peuvent être transportées à longue distance (par exemple dans les ballasts des navires). Elles participent à la « pression en propagules » pour l'établissement initial des populations, favorisent l'expansion de l'espèce puis assurent son établissement à plus long terme en déterminant la dynamique des populations locales. L'étude menée au cours de ce travail à une échelle globale comme locale (baie de Morlaix) en rend compte, par exemple (1) les fortes diversités génétiques et fort flux de gènes observées dans l'aire naturelle sont tout à fait caractéristiques d'une espèce à cycle benthopélagique ; (2) l'absence de structure génétique entre populations naturelles et invasives renvoient à des scénarios impliquant une forte pression en propagules, et (3) les études menées à une échelle locale soulignent l'importance des processus affectant la phase larvaire dans la dynamique démo-génétique des populations.

Cependant, les processus biologiques et physiologiques opérant pendant la phase larvaire restent encore mal compris, en particulier concernant le retour des larves sur les populations adultes. La fin de la vie larvaire (la métamorphose) est un élément clé pour l'espèce étudiée puisqu'il assure le comportement grégaire de cette espèce, un caractère déterminant pour la réussite de sa reproduction (fécondation interne entre mâles et femelles d'une même chaîne), garantissant alors le maintien des populations locales. Chez *C. fornicata*, le retour des larves vers leurs congénères dans un milieu dispersif peut s'expliquer par le développement de mécanismes d'attraction conspécifiques (Thorson 1950, Pawlik 1990). La détection de ce signal par des larves compétentes (cf. encadré 5) à la métamorphose se traduit par leur métamorphose en post-larves qui se sédentarisent à proximité de leurs congénères, c'est-à-dire un habitat favorable. Des travaux antérieurs réalisés au laboratoire ont permis d'identifier un certain nombre de gènes possiblement impliqués dans le développement des larves de crépidules et l'acquisition de leur compétence à la métamorphose (1) via l'analyse d'une puce d'expression (2669 unigènes ; Fig. 12). (2) via une approche gènes candidats (*e.g.* oxyde nitrique synthase, vdg3 ; Fig. 13).





**Fig. 12** Profil d'expression de 2669 gènes de *C. fornicata* spottés sur une puce d'expression et analysés au cours du développement larvaire. Les gènes sont regroupés par profil d'expression « type ». Par exemple, dans le carré central, on note une série de gènes dont l'expression est croissante au cours du développement. Données de Taris et Viard (non publiées)



**Fig. 13** Analyse d'expression du gène candidat *vdg3* au cours du développement larvaire de *C. fornicata*. L'histogramme jaune correspond à l'expression du gène *vdg3* suite à l'induction de la métamorphose par KCL, l'histogramme bleu est l'expression du gène *vdg3* sans induction à la métamorphose, dans de l'eau de mer (FSW). En abscisse sont figurés les jours suite à l'émission larvaire (de J1 à J11) et les heures suite au test KCL (de 1 à 48h). Données de Taris et Viard (non publiées).

### Encadré 5 – Développement larvaire de *Crepidula fornicata* et induction à la métamorphose

Les différentes étapes du développement larvaire sont bien décrites chez *C. fornicata*. Elles peuvent être décomposées en deux grandes phases, caractérisées chacune par de profondes modifications morphologiques et physiologiques : la phase pré-compétente et la phase compétente à la métamorphose. Une larve pré-compétente (Fig. 14) possède les caractères larvaires et des structures juvéniles (Hadfield *et al.* 2001) mais est incapable de se métamorphoser (Pechenik et Heyman 1987). L'acquisition de la compétence (Fig. 15) marque la fin de cette 1<sup>ère</sup> phase et le début de la phase de compétence, également appelée délai à la métamorphose. Les larves compétentes vont par la suite se métamorphoser, en réponse à des stimuli spécifiques environnementaux (chimiques ou physiques).



**Fig. 14** Larve pré-compétente de *C. fornicata*



**Fig. 15** Larve de *C. fornicata* présentant des signes morphologiques de larves compétentes à la métamorphose : la coquille est évasée et le vélum se détache

Chez *Crepidula fornicata*, il n'existe pas de caractère phénotypique strictement facilement identifiable du moment de l'acquisition de la compétence. Néanmoins, Pechenik et Heyman (1987) ont montré qu'une solution de KCL 20mM dans de l'eau de mer filtrée induisait la métamorphose (caractérisée par une perte du velum) chez des larves compétentes. Ils ont alors proposé un protocole permettant de déterminer la date à laquelle les larves d'une population sont dites compétentes : dès que les larves mesurent en moyenne  $\sim 700\mu\text{m}$ , elles sont exposées à une solution d'eau de mer filtrée contenant 20mM de KCL. Au bout de 6h, si au moins 70% des larves sont métamorphosées, les larves de la population sont dites compétentes. Si moins de 70% des larves sont métamorphosées, ce test est répété jusqu'à atteindre ce seuil, seuil dit de compétence de l'ensemble des larves. Cependant, les larves peuvent également se métamorphoser spontanément, en l'absence de stimuli externes (Pechenik et Lima 1984). Il faut donc un témoin négatif pour tous les tests de compétence à la métamorphose.

Plusieurs substances naturelles et artificielles ont été mises en évidence dans l'induction de la métamorphose larvaire de *C. fornicata* tels que de l'eau conditionnée par des adultes de la même espèce (Pechenik et Lima 1984, McGee et Targett 1989), des films microbiens (Pechenik et Gee 1993) ou encore des composés halogénés (Pechenik et Heyman 1987, Pechenik 1990, Taris *et al.* 2010). Une grande variété de molécules contrôle ainsi la métamorphose et le recrutement. Cependant, les mécanismes moléculaires régissant ces étapes restent largement méconnus, notamment les bases moléculaires qui répondent à ces signaux.

Ces résultats ont été obtenus (1) en réponse à des signaux chimiques d'induction de la métamorphose et (2) en laboratoire ; il serait notamment intéressant de valider et étudier les variations d'expression de ces gènes sur des larves prélevés dans le milieu, en absence ou présence d'individus adulte et au cours de la saison de reproduction (à différentes périodes de présence de larves dans le milieu, dans des conditions environnementales diversifiées).

**La phase larvaire, une phase clé du cycle de vie de *C. fornicata* et de la réussite de l'introduction: la phase larvaire joue un rôle important dans le succès d'introduction de l'espèce et dans les processus de loterie familiale. Elle reste néanmoins difficile à étudier et les mécanismes mettant fin à la vie larvaire pélagique sont insuffisamment compris. Plusieurs perspectives émergent visant à mieux comprendre les processus liés à la phase larvaire : mettre au point une stratégie d'échantillonnage permettant d'analyser des larves d'âge différent et affiner l'étude des réponses moléculaires liées au développement et à la métamorphose larvaire.**



---

**Bibliographie**  
**(Hors articles)**

---



## A

---

- Adam W, Leloup E (1934) Sur la présence du gastéropode *Crepidula fornicata* (Linné, 1758) sur la côte belge. *Bulletin du Musée Royal d'Histoire Naturelle de Belgique/Mededeelingen van het Koninklijk Natuurhistorisch Museum van België*, **10**, 1-6.
- Allsop DJ, West SA (2004) Sex-ratio evolution in sex changing animals. *Evolution*, **58**, 1019-1027.
- Audet D, Davis DS, Miron G, Moriyasu M, Benhalima K, Campbell R (2003) Geographical expansion of a nonindigenous crab, *Carcinus maenas* (L.), along the Nova Scotian shore into the southeastern Gulf of St. Lawrence, Canada. *Journal of Shellfish Research*, **22**, 255-262.
- Avise JC (2000) Phylogeography: the history and formation of species. Harvard Univ. Press.
- Ayata SD, Ellien C, Dumas F, Dubois S, Thiébaud E (2009) Modelling larval dispersal and settlement of the reef-building polychaete *Sabellaria alveolata*: Role of hydroclimatic processes on the sustainability of biogenic reefs. *Continental Shelf Research*, **29**, 1605-1623.

## B

---

- Barranger A (2011) Système de reproduction et dérive génétique chez un gastéropode protandre *Crepidula fornicata*. *Mémoire de Master Océanographie et Environnements Marins - Université Pierre et Marie Curie*.
- Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38-44.
- Barrett SCH, Colautti RI, Eckert CG (2008) Plant reproductive systems and evolution during biological invasion. *Molecular Ecology*, **17**, 373-383.
- Barton N (1979) Dynamics of hybrid zones. *Heredity*, **43**, 341-359.
- Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridizing populations. *Heredity*, **57**, 357-376.
- Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969-980.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **263**, 1619-1626.
- Berthier P, Beaumont MA, Cornuet JM, Luikart G (2002) Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. *Genetics*, **160**, 741-751.
- Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044-2072.
- Blackburn TM, Pysek P, Bacher S, Carlton JT, Duncan RP, Jarosik V, Wilson JRU, Richardson DM (2011) A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, **26**, 333-339.
- Blanchard M (1995) Origine et état de la population de *Crepidula fornicata* (Gastropoda Prosobranchia) sur le littoral français. *Halictis*, **24**, 75-86.
- Blanchard M (1997) Spread of the slipper limpet *Crepidula fornicata* (L. 1758) in Europe. Current state and consequences. *Scientia Marina*, **61**, 109-118.
- Blanchard M (2009) Recent expansion of the slipper limpet population (*Crepidula fornicata*) in the Bay of Mont-Saint-Michel (Western Channel, France). *Aquatic Living Resources*, **22**, 11-19.
- Blanchard M, Erhold A (1999) Cartographie et évaluation du stock de crépidules (*Crepidula fornicata* L.) en baie du Mont Saint-Michel. *Halictis*, **28**, 11-20.
- Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737-3758.
- Boudouresque C-F (2008) Les espèces introduites et invasives en milieu marin. *Troisième édition. GIS Posidonie publ., Marseille*.
- Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK, Cavallisforza LL (1991) Drift, admixture, and selection in human evolution - a study with DNA polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 839-843.
- Brenchley GA, Carlton JT (1983) Competitive displacement of native mud snails by introduced periwinkles in the new-england intertidal zone. *Biological Bulletin*, **165**, 543-558.

- Broquet T, Petit EJ (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology Evolution and Systematics*, **40**, 193-216.
- Brouwer AB (1934) Verslag over de visscherij gedurende het jaar 1933. *Verslagen en Mededeelingen van de Afdeling Visscherijen*, **23**.

## C

---

- Caballero A (1994) Developments in the prediction of effective population size. *Heredity*, **73**, 657-679.
- Carlton JT (1999) Molluscan invasions in marine and estuarine communities. *Malacologia*, **41**, 439-454.
- Carlton JT (2011) The inviolate sea? Charles Elton and biological invasion in the World's oceans. in *Fifty years of invasion ecology: the legacy of Charles Elton*. Edited by David M. Richardson. Wiley-Blackwell pp. 23-33.
- Carlton JT, Cheney DP, Vermeij GJ (1982) Ecological effects and biogeography of an introduced marine species : the periwinkle, *Littorina littorea*. *Malacologia Review*, **15**, 143-150.
- Carlton JT, Geller JB (1993) Ecological roulette - The global transport of nonindigenous marine organisms. *Science*, **261**, 78-82.
- Cavalli-Sforza L (1966) Population structure and human evolution. *Proceedings of the Royal Society of London series B-Biological Sciences*, **164**, 362-379.
- Chapple DG, Simmonds SM, Wong BBM (2012) Can behavioral and personality traits influence the success of unintentional species introductions? *Trends in Ecology & Evolution*, **27**, 57-64.
- Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621-631.
- Chipperfield PNJ (1951) The breeding of *Crepidula fornicata* (L) in the River Blackwater, Essex. *Journal of the Marine Biological Association of the United Kingdom*, **30**, 49-71.
- Coe WR (1936) Sexual phases in *Crepidula*. *The Journal of Experimental Zoology*, **72**, 455-477.
- Coe WR (1938) Conditions influencing change of sex in mollusks of the genus *Crepidula*. *Journal of Experimental Zoology*, **77**, 401-424.
- Colautti RI, Grigorovich IA, MacIsaac HJ (2006) Propagule pressure: A null model for biological invasions. *Biological Invasions*, **8**, 1023-1037.
- Collin R (2001) The effects of mode of development on phylogeography and population structure of North Atlantic *Crepidula* (Gastropoda : Calyptraeidae). *Molecular Ecology*, **10**, 2249-2262.
- Collin R, McLellan M, Gruber K, Bailey-Jourdain C (2005) Effects of conspecific associations on size at sex change in three species of calyptraeid gastropods. *Marine Ecology-Progress Series*, **293**, 89-97.
- Conover DO, Clarke LM, Munch SB, Wagner GN (2006) Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology*, **69**, 21-47.
- Coum A (1979) La population de *Crepidula fornicata* (L, 1758) en rade de Brest - Ecologie et dynamique. *Thèse de doctorat. Université Bretagne Occidentale, France*.
- Cowen RK, Sponaugle S (2009) Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, **1**, 443-466.
- Coyer JA, Hoarau G, Pearson G, Mota C, Juterbock A, Alpermann T, John U, Olsen JL (2011) Genomic scans detect signatures of selection along a salinity gradient in populations of the intertidal seaweed *Fucus serratus* on a 12 km scale. *Marine Genomics*, **4**, 41-49.

## D

---

- Daguin-Thiébaud C, Le Cam S, Viard F (2009) Isolation of 11 microsatellite markers in *Crepidula convexa* (Gastropoda, Calyptraeidae) for parentage analyses. *Molecular Ecology Resources*, **9**, 917-920.
- Dame RF, Allen DM (1996) Between estuaries and the sea. *Journal of Experimental Marine Biology and Ecology*, **200**, 169-185.
- David P (1996) Structures génétiques et relation génotype-phénotype en milieu marin: le cas du bivalve *Spisula ovalis*. *Thèse de doctorat. Université de Montpellier 2, Montpellier*.
- David P (1998) Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity*, **80**, 531-537.
- David P, Delay B, Jarne P (1997) Heterozygosity and growth in the marine bivalve *Spisula ovalis*: testing alternative hypotheses. *Genetical Research*, **70**, 215-223.



- de Francisco AK, Galetti PM (2005) Genetic distance between broodstocks of the marine shrimp *Litopenaeus vannamei* (Decapoda, Penaeidae) by mtDNA analyses. *Genetics and Molecular Biology*, **28**, 258-261.
- de Montaudouin X, Sauriau PG (1999) The proliferating Gastropoda *Crepidula fornicata* may stimulate macrozoobenthic diversity. *Journal of the Marine Biological Association of the United Kingdom*, **79**, 1069-1077.
- de Montaudouin X, Labarraque D, Giraud K, Bachelet G (2001) Why does the introduced gastropod *Crepidula fornicata* fail to invade Arcachon Bay (France)? *Journal of the Marine Biological Association of the United Kingdom*, **81**, 97-104.
- DiBacco C, Levin LA (2000) Development and application of elemental fingerprinting to track the dispersal of marine invertebrate larvae. *Limnology and Oceanography*, **45**, 871-880.
- Domingues CP, Creer S, Taylor MI, Queiroga H, Carvalho GR (2011) Temporal genetic homogeneity among shore crab (*Carcinus maenas*) larval events supplied to an estuarine system on the Portuguese northwest coast. *Heredity*, **106**, 832-840.
- Dupont L (2004) Invasion des côtes françaises par le mollusque exotique *Crepidula fornicata*: contribution de la dispersion larvaire et du système de reproduction au succès de la colonisation. *Thèse de doctorat. Université Pierre et Marie Curie, Paris*.
- Dupont L, Bernas D, Viard F (2007a) Sex and genetic structure across age groups in populations of the European marine invasive mollusc, *Crepidula fornicata* L. (Gastropoda). *Biological Journal of the Linnean Society*, **90**, 365-374.
- Dupont L, Ellien C, Viard F (2007b) Limits to gene flow in the slipper limpet *Crepidula fornicata* as revealed by microsatellite data and a larval dispersal model. *Marine Ecology Progress Series*, **349**, 125-138.
- Dupont L, Jollivet D, Viard F (2003) High genetic diversity and ephemeral drift effects in a successful introduced mollusc (*Crepidula fornicata* : Gastropoda). *Marine Ecology-Progress Series*, **253**, 183-195.
- Dupont L, Richard J, Paulet YM, Thouzeau G, Viard F (2006) Gregariousness and protandry promote reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of larval paternity. *Molecular Ecology*, **15**, 3009-3021.
- Dupont L, Viard F (2003) Isolation and characterization of highly polymorphic microsatellite markers from the marine invasive species *Crepidula fornicata* (Gastropoda : Calyptraeidae). *Molecular Ecology Notes*, **3**, 498-500.

## E

---

- Edgell TC, Hollander J (2011) The evolutionary ecology of European green crab, *Carcinus maenas*, in North America. *Dordrecht*, **6**, 641-659.
- Ellien C, Thiébaud E, Dumas F, Salomon JC, Nival P (2004) A modelling study of the respective role of hydrodynamic processes and larval mortality on larval dispersal and recruitment of benthic invertebrates: example of *Pectinaria koreni* (Annelida : Polychaeta) in the Bay of Seine (English Channel). *Journal of Plankton Research*, **26**, 117-132.
- Elton CS (1958) The ecology of invasions by animals and plants. Methuen publ. University of Chicago Press, 196 pp.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113-4130.

## F

---

- Facon B, Genton BJ, Shykoff J, Jarne P, Estoup A, David P (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution*, **21**, 130-135.
- Faure MF, David P, Bonhomme F, Bierne N (2008) Genetic hitchhiking in a subdivided population of *Mytilus edulis*. *BMC Evolutionary Biology*, **8**, 164.
- Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic structure of Populations. *Genetics*, **174**, 875-891.

- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics*, **180**, 977-993.
- Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. *Molecular Ecology*, **11**, 1245-1262.
- Frontier S, Pichod-Viale D (1991) Ecosystèmes: structure, fonctionnement, évolution. *Masson, Paris*.

## G

---

- Gaines S, Roughgarden J (1985) Larval settlement rate - a leading determinant of structure in an ecological community of the marine intertidal zone. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 3707-3711.
- Galil BS (2000) A sea under siege – alien species in the Mediterranean. *Biological Invasions*, **2**, 177-186.
- Galil BS (2001) Exotics in the Mediterranean – Bioindicators of a sea change. *Biomare Newsletter*, **1**, 7-9.
- Galil B, Frogliani C, Noel P (2002). CIESM Atlas of exotic species in the Mediterranean. 2. Crustaceans. F. Briand (ed.), CIESM publ., Monaco: 1-192.
- Gilad Y, Rifkin SA, Pritchard JK (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends in Genetics*, **24**, 408-415.
- Goldstien SJ, Dupont L, Viard F, Hallas PJ, Nishikawa T, Schiel DR, Gemmell NJ, Bishop JDD (2011) Global phylogeography of the widely introduced North West Pacific ascidian *Styela clava*. *Plos One*, **6**, e16755.
- Gouilletquer P, Bachelet G, Sauriau PG, Noel P (2002) Open Atlantic coast of Europe - A century of introduced species into French waters. *Invasive Aquatic Species of Europe: Distribution, Impacts and Management*, 276-290.
- Gouilletquer P, Le Moine O (2002) Shellfish farming and coastal zone management (CZM) development in the Marennes-Oleron Bay and Charentais Sounds (Charente Maritime, France): A review of recent developments. *Aquaculture International*, **10**, 507-525.
- Grizel H, Heral M (1991) Introduction into France of the Japanese oyster (*Crassostrea gigas*). *Journal Du Conseil International pour l'exploration de la Mer*, **47**, 399-403.
- Grosholz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology & Evolution*, **17**, 22-27.

## H

---

- Hadfield MG, Carpizo-Ituarte EJ, del Carmen K, Nedved BT (2001) Metamorphic competence, a major adaptive convergence in marine invertebrate larvae. *American Zoologist*, **41**, 1123-1131.
- Hamon D (1996) Effets de la prolifération de la crépidule (*Crepidula fornicata*) sur la diversité de la macrofaune benthique en baie de Saint-Brieuc (Manche occidentale). *Colloque du Réseau "Diversité marine", Biodiversité en milieu dispersif, Boucher G et Feral JP édit., Muséum National Hist. nat. publ., Paris : 25p.*
- Hansen MM, Nielsen EE, Ruzzante DE, Bouza C, Mensberg KLD (2000) Genetic monitoring of supportive breeding in brown trout (*Salmo trutta* L.), using microsatellite DNA markers. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 2130-2139.
- Hare MP, Avise JC (1996) Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). *Evolution*, **50**, 2305-2315.
- Harms J (1999) The neozoan *Elminius modestus* Darwin (Crustacea, Cirripedia) : possible explanations for its successful invasion in European water. *Helgoland Marine Research*, **52**, 337-345.
- Hauser L, Adcock GJ, Smith PJ, Ramirez JHB, Carvalho GR (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11742-11747.
- Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, **9**, 333-362.
- Hedgecock D (1994a) Does variance in reproductive success limit effective population sizes of marine organisms? Pp. 122-134. In A. Beaumont, ed. "Genetics and evolution of aquatic organisms". Chapman and Hall, London.

- Hedgecock D (1994b) Temporal and spatial genetic structure of marine animal populations in the California current. *California Cooperative Oceanic Fisheries Investigations Reports*, **35**, 73-81.
- Hedgecock D, Barber PH, Edmands S (2007) Genetic approaches to measuring connectivity. *Oceanography*, **20**, 70-79.
- Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bulletin of Marine Science*, **87**, 971-1002.
- Hedrick P (2000) Genetics of populations - 2nd ed. *Jones and Bartlett, Boston (MA)* 737p.
- Henry JJ, Collin R, Perry KJ (2010) The slipper snail, *Crepidula*: An emerging lophotrochozoan model system. *Biological Bulletin*, **218**, 211-229.
- Hewitt CL, Campbell ML, Thresher RE, Martin RB, Boyd S, Cohen BF, Currie DR, Gomon MF, Keough MJ, Lewis JA, Lockett MM, Mays N, McArthur MA, O'Hara TD, Poore GCB, Ross DJ, Storey MJ, Watson JE, Wilson RS (2004) Introduced and cryptogenic species in Port Phillip Bay, Victoria, Australia. *Marine Biology*, **144**, 183-202.
- Hilbish TJ, Sasada K, Eyster LS, Pechenik JA (1999) Relationship between rates of swimming and growth in veliger larvae: genetic variance and covariance. *Journal of Experimental Marine Biology and Ecology*, **239**, 183-193.
- Hoagland KE (1978) Protandry and evolution of environmentally-mediated sex change - study of mollusca. *Malacologia*, **17**, 365-391.
- Hoagland KE (1985) Genetic relationships between one British and several north American populations of *Crepidula fornicata* based on allozymes studies (gastropoda, Calyptraeidae). *Journal of Molluscan Studies*, **51**, 177-182.
- Hoch JM, Cahill AE (2012) Variation in size at sex-change among natural populations of the protandrous hermaphrodite, *Crepidula fornicata* (Gastropoda, Calyptraeidae). *Marine Biology*, **159**, 897-905.
- Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, **21**, 797-807.
- Horta P, Oliveira EC (2000) Morphology and reproduction of *Anotrichium yagii* (Ceramiales, Rhodophyta) - a new invader seaweed in the American Atlantic? *Phycologia*, **39**, 390-394.
- Houde ED (1989) Comparative growth, mortality and energetics of marine fish larvae: temperature and implied latitudinal effects. *Fishery Bulletin*, **87**, 471-495.
- Hulme PE, Roy DB, Cunha T, T-B L (2008) A pan-european inventory of alien species: rationale, implementation and implications for managing biological invasions. In DAISIE, The Handbook of European Alien Species. Springer, Dordrecht.

## I

---

- Innan H, Kim Y (2004) Pattern of polymorphism after strong artificial selection in a domestication event. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10667-10672.

## J

---

- Jeffreys H (1961) The theory of probability (3rd ed.) *Oxford University Press*, 470pp,
- Johnson MS, Black R (1982) Chaotic genetic patchiness in an inter-tidal limpet, *Siphonaria* sp. *Marine Biology*, **70**, 157-164.
- Jones GP, Milicich MJ, Emslie MJ, Lunow C (1999) Self-recruitment in a coral reef fish population. *Nature*, **402**, 802-804.
- Jones GP, Planes S, Thorrold SR (2005) Coral reef fish larvae settle close to home. *Current Biology*, **15**, 1314-1318.
- Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, **10**, 551-555.

## K

---

- Kapuscinski ARD, Lannan JE (1986) A conceptual genetic fitness model for fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences*, **43**, 1606-1616.
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

- Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters*, **11**, 852-866.
- Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: A community perspective. *Ecology*, **84**, 2007-2020.
- Kinlan BP, Gaines SD, Lester SE (2005) Propagule dispersal and the scales of marine community process. *Diversity and Distributions*, **11**, 139-148.
- Koehn RK, Newell RIE, Immermann F (1980) Maintenance of an aminopeptidase allele frequency cline by natural selection. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, **77**, 5385-5389.
- Kollmann J, Banuelos MJ (2004) Latitudinal trends in growth and phenology of the invasive alien plant *Impatiens glandulifera* (Balsaminaceae). *Diversity and Distributions*, **10**, 377-385.
- Kozloff EN (1974) Keys to the marine invertebrates of Pudget Sound, The San Juan archipelago and adjacent regions. *Thesis University of Washington*, 226p

## L

---

- Laurent V, Planes S (2007) Effective population size estimation on *Sardina pilchardus* in the Bay of Biscay using a temporal genetic approach. *Biological Journal of the Linnean Society*, **90**, 591-602.
- Launey S (1998) Marqueurs microsatellites chez l'huître plate *Ostrea edulis* L.: caractérisation et applications à un programme de sélection pour une résistance au parasite *Bonamia ostreae* et à l'étude de populations naturelles. *Thèse de doctorat. Institut National Agronomique Paris Grignon, Paris*.
- Lasbleiz R (2003) Etude de la distribution spatio-temporelle des larves de mollusques en baie du Mont Saint-Michel. Cas des bivalves et de la crépidule. *Rapport de DES Sciences Naturelles. Université Pierre et Marie Curie, Paris*.
- Lawson-Handley LJ, Estoup A, Evans DM, Thomas CE, Lombaert E, Facon B, Aebi A, Roy HE (2011) Ecological genetics of invasive alien species. *Biocontrol*, **56**, 409-428.
- Le Cam S (2009) Grégarité, changement de sexe et polyandrie: modalités de la reproduction chez une espèce invasive *Crepidula fornicata*. *Thèse de doctorat. Université Pierre et Marie Curie, Paris*.
- Le Cam S, Pechenik JA, Cagnon M, Viard F (2009) Fast versus slow larval growth in an invasive marine mollusc: does paternity matter? *Journal of Heredity*, **100**, 455-464.
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, **17**, 386-391.
- Lee HJE, Boulding EG (2009) Spatial and temporal population genetic structure of four northeastern Pacific littorinid gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. *Molecular Ecology*, **18**, 2165-2184.
- Le Gall P (1980) Etude expérimentale de l'association en chaîne et son influence sur la croissance et la sexualité chez la crépidule *Crepidula fornicata*. *Thèse de Doctorat de l'Université de Caen*.
- Leger EA, Rice KJ (2003) Invasive California poppies (*Eschscholzia californica* Cham.) grow larger than native individuals under reduced competition. *Ecology Letters*, **6**, 257-264.
- Leinonen T, Cano JM, Makinen H, Merila J (2006) Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology*, **19**, 1803-1812.
- Lemer S, Rochel E, Planes S (2011) Correction method for null alleles in species with variable microsatellite flanking regions, a case study of the black-lipped pearl oyster *Pinctada margaritifera*. *Journal of Heredity*, **102**, 243-246.
- Leroy F (2011) Influence des conditions trophiques sur le développement larvaire de l'espèce invasive *Crepidula fornicata*. Conséquences sur sa capacité d'expansion. *Thèse de doctorat. Université Pierre et Marie Curie, Paris*.
- Levin LA (2006) Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology*, **46**, 282-297.
- Levin L, Bridges T (1995) Pattern and diversity in reproduction and development. *Pages 1-48 in L. McEdwards, editor. Ecology of marine invertebrate larvae. CRC Press, Inc., Florida*.
- Lewontin RC (1974) The problem of genetic diversity. *Harvey lectures*, **70** Series, 1-20.
- Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. *Genetics*, **74**, 175-195.

- Lima GM, Lutz RA (1990) The relationship of larval shell morphology to mode of development in marine prosobranch gastropods. *Journal of the Marine Biological Association of the United Kingdom*, **70**, 611-637.
- Limborg MT, Helyar SJ, de Bruyn M, Taylor MI, Nielsen EE, Ogden R, Carvalho GR, Bekkevold D, Consortium FPT (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Molecular Ecology*, **21**, 3686-3703.
- Lowe S, Browne M, Boudjelas S, de Poorder M (2004) 100 of the Worlds Worst Invasive Alien Species a selection from the global invasive species database ISSG and SSC of IUCN Auckland, New Zealand.
- Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, **19**, 3038-3051.
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981-994.
- Lutzen J (1999) *Styela clava* Herdman (Urochordata, Ascidiacea), a successful immigrant to North West Europe : ecology, propagation and chronology of spread. *Helgoland Marine Research*, **52**, 383-391.
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. *Sinauer*, 980 pp.

## M

---

- Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution*, **20**, 136-142.
- Mantel N (1967) Detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209.
- Martel C, Viard F, Bourguet D, Garcia-Meunier P (2004) Invasion of the marine gastropod *Ocenebrellus inornatus* on the French Atlantic coast. I- Scenario for the source of introduction. *Journal of Experimental Biology and Ecology*, **305**, 155-170.
- Marteil L (1963) La crépidule (*Crepidula fornicata*). *Science et Pêche*, **121**, 1-6.
- Mattersdorfer K, Koblmüller S, Sefc KM (2012) AFLP genome scans suggest divergent selection on colour patterning in allopatric colour morphs of a cichlid fish. *Molecular Ecology*, **21**, 3531-3544.
- McCormick MI, Molony BW (1995) Influence of water temperature during the larval stage on size, age and body condition of a tropical reef fish at settlement. *Marine Ecology-Progress Series*, **118**, 59-68.
- McDowell SCL (2002) Photosynthetic characteristics of invasive and noninvasive species of *Rubus* (Rosaceae). *American Journal of Botany*, **89**, 1431-1438.
- McGee BL, Targett NM (1989) Larval habitat selection in *Crepidula* (L) and its effect on adult distribution patterns. *Journal of Experimental Marine Biology and Ecology*, **131**, 195-214.
- McGlashan DJ, Ponniah M, Cassey P, Viard F (2008) Clarifying marine invasions with molecular markers: an illustration based on mtDNA from mistaken calyptreid gastropod identifications. *Biological Invasions*, **10**, 51-57.
- Mileikovski S (1971) Types of larval development in marine bottom invertebrates, their distribution and ecological significance - re-evaluation. *Marine Biology*, **10**, 193-213.
- Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, **6**, 485-492.

## N

---

- Natunewicz CC, Epifanio CE, Garvine RW (2001) Transport of crab larval patches in the coastal ocean. *Marine Ecology-Progress Series*, **222**, 143-154.
- Negrini R, D'Andrea M, Crepaldi P, Colli L, Nicoloso L, Guastella AM, Sechi T, Bordonaro S, Ajmone-Marsan P, Pilla F, Econogene C (2012) Effect of microsatellite outliers on the genetic structure of eight Italian goat breeds. *Small Ruminant Research*, **103**, 99-107.
- Nei M, Maruyama T (1975) Lewontin-krakauer test for neutral genes - comment. *Genetics*, **80**, 395-395.
- Nelson K, Soule M (1987) Genetical conservation of exploited fishes. In Ryman, N. et F. Utter (éds.). *Population genetics and fishery management. Washington Sea Grant program, Seattle*, pp. 345-368.,
- Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer C, Taranger G-L, Ogden R, Carvalho GR (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *Bmc Evolutionary Biology*, **9**, 276.

- Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375-402.
- Nunney L (1993) The influence of mating system and overlapping generations on effective population-size. *Evolution*, **47**, 1329-1341.

---

## O

---

- O'Connor MI, Bruno JF, Gaines SD, Halpern BS, Lester SE, Kinlan BP, Weiss JM (2007) Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 1266-1271.
- Occhipinti-Ambrogi A (2007) Global change and marine communities: Alien species and climate change. *Marine Pollution Bulletin*, **55**, 342-352.
- Orton JH (1912) An account of the natural history of the slipper-limpet (*Crepidula fornicata*), with some remarks on its occurrence on the oyster grounds of the Essex coast. *Journal of the Marine Biological Association of the United Kingdom*, **9**, 437-443.

---

## P

---

- Pagny J, Acou A, Ar Gall E, Blanchard M, Cabioc'h J, Canard A, Derrien-Courtel S, Feuteun E, Gentil F, Gouletquer P, Grall J, Gruet Y, Hamon D, Hily C, d'Hondt J-L, Le Duff M, Le Mao P, Le Roux A, Nézan E, Perrin B, Simon N, Stiger-Pouvreau V, Viard F, Ysnel F (2010) Les espèces marines invasives en Bretagne. GIP Bretagne Environnement publ., Rennes : 41 p.
- Palsboll PJ, Berube M, Allendorf FW (2007) Identification of management units using population genetic data. *Trends in Ecology & Evolution*, **22**, 11-16.
- Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428-3447.
- Palumbi SR (1994) Genetic-divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, **25**, 547-572.
- Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, **13**, S146-S158.
- Parker IM, Rodriguez J, Loik ME (2003) An evolutionary approach to understanding the biology of invasions: Local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology*, **17**, 59-72.
- Pawlik JR (1990) Natural and artificial induction of metamorphosis of *Phragmatopoma lapidosa californica* (polychaeta, Sabellariidae), with a critical look at the effects of bioactive compounds on marine invertebrate larvae. *Bulletin of Marine Science*, **46**, 512-536.
- Pechenik JA (1990) Delayed metamorphosis by larvae of benthic marine invertebrates: does it occur? Is there a price to pay? *Ophelia*, **32**, 63-94.
- Pechenik JA, Gee CC (1993) Onset of metamorphic competence in larvae of the gastropod *Crepidula fornicata* (L), judged by a natural and artificial cue. *Journal of Experimental Marine Biology and Ecology*, **167**, 59-72.
- Pechenik JA, Heyman WD (1987) Using KCL to determine size at competence for larvae of the marine gastropod *Crepidula fornicata* (L). *Journal of Experimental Marine Biology and Ecology*, **112**, 27-38.
- Pechenik JA, Levine SH (2007) Estimates of planktonic larval mortality using the marine gastropods *Crepidula fornicata* and *C. plana*. *Marine Ecology-Progress Series*, **344**, 107-118.
- Pechenik JA, Lima GM (1984) Relationship between growth, differentiation and length of larval life for individually reared larvae of the marine gastropod, *Crepidula fornicata*. *Biological Bulletin*, **166**, 537-549.
- Pelc RA, Warner RR, Gaines SD (2009) Geographical patterns of genetic structure in marine species with contrasting life histories. *Journal of Biogeography*, **36**, 1881-1890.
- Pepin P (1991) Effect of temperature and size on development, mortality, and survival rates of the pelagic early life-history stages of marine fish. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 503-518.

- Perez-Figueroa A, Garcia-Pereira MJ, Saura M, Rolan-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267-2276.
- Phillips PC (2005) Testing hypotheses regarding the genetics of adaptation. *Genetica*, **123**, 15-24.
- Pineda J, Hare JA, Sponaugle S (2007) Larval transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography*, **20**, 22-39.
- Poulsen NA, Nielsen EE, Schierup MH, Loeschcke V, Gronkjaer P (2006) Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). *Molecular Ecology*, **15**, 321-331.
- Prentis PJ, Wilson JRU, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Plant Science*, **13**, 288-294.
- Prentis PJ, Woolfit M, Thomas-Hall SR, Ortiz-Barrientos D, Pavasovic A, Lowe AJ, Schenk PM (2010) Massively parallel sequencing and analysis of expressed sequence tags in a successful invasive plant. *Annals of Botany*, **106**, 1009-1017.
- Price TD, Qvarnstrom A, Irwin DE (2003) The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 1433-1440.
- Prince JD, Sellers TL, Ford WB, Talbot SR (1987) Experimental-evidence for limited dispersal of haliotid larvae (genus *Haliotis*, mollusca, gastropoda). *Journal of experimental marine biology and ecology*, **106**, 243-263.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
- Proestou DA (2006) Isolation and characterization of microsatellite markers in the Atlantic slipper shell *Crepidula fornicata* for use in paternity analysis. *Molecular Ecology Notes*, **6**, 437-439.
- Proestou DA, Goldsmith MR, Twombly S (2008) Patterns of male reproductive success in *Crepidula fornicata* provide new insight for sex allocation and optimal sex change. *Biological Bulletin*, **214**, 192-200.
- Puebla O, Bermingham E, Guichard F (2009) Estimating dispersal from genetic isolation by distance in a coral reef fish (*Hypoplectrus puella*). *Ecology*, **90**, 3087-3098.
- Purcell JFH, Cowen RK, Hughes CR, Williams DA (2006) Weak genetic structure indicates strong dispersal limits: a tale of two coral reef fish. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 1483-1490.

## R

- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature*, **394**, 69-72.
- Richard J (2005) *Crepidula fornicata* : un modèle biologique pour l'étude du rôle de la variabilité des caractères phénotypique (reproduction, croissance et nutrition) sur les processus de colonisation en milieu marin. Thèse de Doctorat de l'Université de Bretagne Occidentale et l'Université de Genève.
- Richard J, Huet M, Thouzeau G, Paulet YM (2006) Reproduction of the invasive slipper limpet, *Crepidula fornicata*, in the Bay of Brest, France. *Marine Biology*, **149**, 789-801.
- Richards CL, Bosserdorf O, Muth NZ, Gurevitch J, Pigliucci M (2006) Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters*, **9**, 981-993.
- Richards CL, Pennings SC, Donovan LA (2005) Habitat range and phenotypic variation in salt marsh plants. *Plant Ecology*, **176**, 263-273.
- Richardson DM, Pysek P, Rejmanek M, Barbour MG, D. PF, C.J. W (2000) Naturalization and invasion of alien plants : concepts and definitions. *Diversity and Distributions*, **6**, 93-107.
- Rigal F (2009) Dynamique spatio-temporelle du nuage larvaire du gastéropode introduit *Crepidula fornicata* au sein d'une baie mégatidale, la baie de Morlaix (France). *Thèse de doctorat del'Université Pierre et Marie Curie, Paris*.
- Rigal F, Viard F, Ayata SD, Comtet T (2010) Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? *Biological Invasions*, **12**, 3171-3186.

- Rius M, Turon X, Ordonez V, Pascual M (2012) Tracking invasion histories in the sea: facing complex scenarios using multilocus data. *Plos One*, **7**, e35815.
- Robertson A (1975) Gene frequency distributions as a test of selective neutrality. *Genetics*, **81**, 775-785.
- Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 2453-2459.
- Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution*, **22**, 454-464.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219-1228.
- Rousset F (2001) Inferences from spatial population genetics. *Handbook of Statistical Genetics. D. Balding, M. Bishop & C. Cannings, Wiley: 239-269.*
- Rumrill SS (1990) Natural mortality of marine invertebrate larvae. *Ophelia*, **32**, 163-198.

## S

---

- Saenz-Agudelo P, Jones GP, Thorrold SR, Planes S (2011) Connectivity dominates larval replenishment in a coastal reef fish metapopulation. *Proceedings of the Royal Society B-Biological Sciences*, **278**, 2954-2961.
- Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305-332.
- Sax DF, Brown JH (2000) The paradox of invasion. *Global Ecology and Biogeography*, **9**, 363-371.
- Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, Grosberg RK, Hasting SA, Holt RD, Mayfield MM, O'Connor MI, Rice WR (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22**, 465-471.
- Scheltema R (1971) Larval dispersal as a means of genetic exchange between geographically separated populations of shallow-water benthic marine gastropods. *Biological Bulletin*, **140**, 284-322.
- Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology-Progress Series*, **436**, 291-305.
- Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin*, **216**, 373-385.
- Sibeaux A (2010) Etude d'une population de mollusques invasifs. Lien entre la dynamique de recrutement et la dynamique de changement de sexe dans la population de *Crepidula fornicata* établie en baie de Morlaix. *Rapport de stage de license, Université de Grenoble.*
- Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology-Progress Series*, **260**, 83-96.
- Simberloff D (2009) The role of propagule pressure in biological invasions. *Annual Review of Ecology Evolution and Systematics*, **40**, 81-102.
- Simon-Bouhet B, Garcia-Meunier P, Viard F (2006) Multiple introductions promote range expansion of the mollusc *Cyclope neritea* (Nassariidae) in France: evidence from mitochondrial sequence data. *Molecular Ecology*, **15**, 1699-1711.
- Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.
- Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, **47**, 264-279.
- Smaal AC, Kater BJ, Wijsman J (2009) Introduction, establishment and expansion of the Pacific oyster *Crassostrea gigas* in the Oosterschelde (SW Netherlands). *Helgoland Marine Research*, **63**, 75-83.
- Spitze K (1993) Population structure in *Daphnia obtusa* - quantitative genetic and allozymic variation. *Genetics*, **135**, 367-374.
- Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, **18**, 94-101.
- Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671-688.
- Sugg DW, Chesser RK (1994) Effective population sizes with multiple paternity. *Genetics*, **137**, 1147-1155.



- Sultan SE (2001) Phenotypic plasticity for fitness components in *Polygonum* species of contrasting ecological breadth. *Ecology*, **82**, 328-343.
- Szulkin M, Bierne N, David P (2010) Heterozygosity-Fitness Correlations: a time for reappraisal. *Evolution*, **64**, 1202-1217.

---

## T

---

- Taris N, Comtet T, Stolba R, Lasbleiz R, Pechenik JA, Viard F (2010) Experimental induction of larval metamorphosis by a naturally-produced halogenated compound (dibromomethane) in the invasive mollusc *Crepidula fornicata* (L.). *Journal of Experimental Marine Biology and Ecology*, **393**, 71-77.
- Thieltges DW, Strasser M, Reise K (2006) How bad are invaders in coastal waters? The case of the American slipper limpet *Crepidula fornicata* in western Europe. *Biological Invasions*, **8**, 1673-1680.
- Thiriot-Quievreux C, Scheltema RS (1982) Planktonic larvae of New England gastropods. V *Bittium alternatum*, *Triphora nigrocincta*, *Cerithiopsis emersoni*, *Lunatia heros* and *Crepidula plana*. *Malacologia*, **23**, 37-46.
- Thorson G (1950) Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews of the Cambridge Philosophical Society*, **25**, 1-45.
- Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 5948-5953.

---

## V

---

- Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623-3642.
- Vermeij GJ (1996) An agenda for invasion biology. *Biological Conservation*, **78**, 3-9.
- Viard F, Ellien C, Dupont L (2006) Dispersal ability and invasion success of *Crepidula fornicata* in a single gulf: insights from genetic markers and larval-dispersal model. *Helgoland Marine Research*, **60**, 144-152.
- Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics*, **158**, 1811-1823.
- Voisin M, Engel CR, Viard F (2005) Differential shuffling of native genetic diversity across introduced regions in a brown alga: Aquaculture vs. maritime traffic effects. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 5432-5437.

---

## W

---

- Wang T, Chen G, Zan Q, Wang C, Su Y-J (2012) AFLP genome scan to detect genetic structure and candidate loci under selection for local adaptation of the invasive weed *Mikania micrantha*. *Plos One*, **7**, e41310.
- Wang JL (2004) Sibship reconstruction from genetic data with typing errors. *Genetics*, **166**, 1963-1979.
- Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetics Research*, **89**, 135-153.
- Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, **11**, 141-145.
- Wang J (2012) Computationally efficient sibship and parentage assignment from multilocus marker data. *Genetics*, **191**, 183-94.
- Wang J, Santure AW (2009) Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, **181**, 1579-1594.
- Wang JL, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, **163**, 429-446.
- Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438-450.

- Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, **14**, 3335-3352.
- Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, **3**, 244-262.
- Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology-Progress Series*, **393**, 1-12.
- Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.
- Werner B (1955) Über die Anatomie, die Entwicklung und Biologie des Veligers und der Veliconcha von *Crepidula fornicata* L. (Gastropoda Prosobranchia). *Helgoländer Wissenschaftliche Meeresuntersuchungen*, **5**, 169-217.
- Whitlock MC (2008) Evolutionary inference from  $Q_{ST}$ . *Molecular Ecology*, **17**, 1885-1896.
- Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology*, **14**, 611-619.
- Williamson M, Fitter A (1996) The varying success of invaders. *Ecology*, **77**, 1661-1666.
- Willis AJ, Blossey B (1999) Benign environments do not explain the increased vigour of non-indigenous plants: a cross-continental transplant experiment. *Biocontrol Science and Technology*, **9**, 567-577.
- Willis AJ, Memmott J, Forrester RI (2000) Is there evidence for the post-invasion evolution of increased size among invasive plant species? *Ecology Letters*, **3**, 275-283.
- Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 0097-0159.
- Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354.

---

## Y

---

- Yu H, Li Q (2008) Exploiting EST databases for the development and characterization of EST-SSRs in the Pacific oyster (*Crassostrea gigas*). *Journal of Heredity*, **99**, 208-214.

---

## Z

---

- Zibrowius H (1991) Ongoing modification of the Mediterranean marine fauna and flora by the establishment of exotic species. *Mésogée*, 51 : 83-107.

---

## **Annexes**

---



## **Liste des annexes :**

### **Annexes relatives à la Première Partie**

#### **Article 1 : Identification *in silico* et optimisation de 12 locus EST-SSRs chez l'espèce invasive *Crepidula fornicata***

- **Annexe 1** : Ségrégation mendélienne des 12 locus EST-SSRs nouvellement développés

#### **Article 2 : Comparaison des patrons de polymorphismes génomiques entre aire d'origine et aires d'introduction chez le mollusque marin invasif *Crepidula fornicata***

- **Annexe 2** : Développement des marqueurs AFLPs
- **Annexe 3** : Poster présenté à la conférence Jacques Monod « Développements théoriques et empiriques en génomique évolutive »
- **Annexe 4** : Résultats d'analyses d'assignation statistiques réalisées avec le logiciel STRUCTURE
- **Annexe 5** : Résultats détaillés de l'identification d'outliers pour les 22 populations
- **Annexe 6** : Résultats détaillés de l'identification d'outliers entre paires de populations

### **Annexes relatives à la Deuxième Partie**

- **Annexe 7** : Structures de taille de la population benthique échantillonnée mensuellement au Château du Taureau
- **Annexe 8** : Caractéristiques des composantes modales décrivant la structure de taille de la population benthique échantillonnée mensuellement au Château du Taureau

#### **Article 3 : Des effets de dérive génétique accompagnent les épisodes annuels de recrutement des juvéniles chez le gastéropode invasif *Crepidula fornicata***

- **Annexe 9** : Données d'apparentement des juvéniles échantillonnés sur neuf années successives (2002 à 2010) et sur sept mois consécutifs (Septembre 2010 à Mars 2011)
- **Annexe 10** : Diversité génétique des juvéniles échantillonnés sur neuf années successives (2002 à 2010) et sur sept mois consécutifs (Septembre 2010 à Mars 2011)
- **Annexe 11** : Isolement par le temps et test associé pour les échantillons de juvéniles échantillonnés sur neuf années successives (2002 à 2010) et sur sept mois consécutifs (Septembre 2010 à Mars 2011)

#### **Article 4 : Des différences génétiques et la présence d'individus apparentés dans le nuage larvaire sont en accord avec l'hypothèse de loterie familiale chez le gastéropode *Crepidula fornicata***

- **Annexe 12** : Nombre de paires d'individus apparentés en fonction du nombre de runs réalisés
- **Annexe 13** : Distributions des coefficients d'apparentement de 100 plein-frères, 100 demi-frères et 100 individus non apparentés

#### **Article 5 : Des déséquilibres d'identité au sein du pool larvaire mais pas de corrélation hétérozygotie-fitness détectés chez le gastéropode marin *Crepidula fornicata***

- **Annexe 14** : Vérification des hypothèses d'application du modèle linéaire et transformation des données

**ANNEXE 1 : Ségrégation mendélienne des 12 locus EST-SSRS caractérisés dans l'article 1.**

Locus	Parent		génotypes attendus des descendants	No de PCRs qui n'ont pas marché	No de descendants qui ont été génotypes	ratio observé	ratio attendu	chi-2	p-value
	Mère	Père							
DA1YL23	A1A2	A3A4	A1A3:A1A4:A2A3:A2A4	0	25	6:7:4:8	1:1:1:1	1.40	0.70
DA1YN22	B1B2	B1B3	B1B1:B1B3:B2B1:B2B3	6	19	5:5:3:6	1:1:1:1	1.20	0.75
CL83	C1C2	C3C3	C1C3:C2C3	6	19	12:7	1:1	1.32	0.36
BI13YE17	D1D2	D3D3	D1D3:D2D3	0	25	14:11	1:1	0.36	0.69
DA8YN14	E1E2	E3E4	E1E3:E1E4:E2E3:E2E4	0	25	7:6:4:8	1:1:1:1	1.40	0.70
CL322	F1F1	F1F2	F1F1:F1F2	0	25	11:14	1:1	0.60	0.68
DA1YA23	G1G1	G2G2	G1G2	0	25	25	1	0	1
DA5YM24	H1H2	H2H3	H1H2:H1H3:H2H2:H2H3	0	25	11:4:3:7	1:1:1:1	6.20	0.10
CL191	I1I2	I1I2	I1I1:I1I2:I2I2	0	25	6:15:4	1:2:1	1.32	0.52
CL270	J1J1	J1J2	J1J1:J1J2	0	25	16:9	1:1	1.96	0.22
CL291	K1K1	K1K1	K1K1	0	25	25	1*		
DA4YO03	L1L2	L3L4	L1L3:L1L4:L2L3:L2L4	0	25	9:8:4:4	1:1:1:1	3.32	0.34

\* locus non informatif : les deux parents sont homozygotes pour différents allèles, la ségrégation mendélienne ne peut donc être testée

**ANNEXE 2 : Conditions de digestion, ligation et d'amplification de marqueurs AFLPs chez *Crepidula fornicata*.** Le développement des protocoles de génotypages des marqueurs AFLPs chez *C. fornicata* a été réalisé par C. Daguin-Thiébaud, ingénieur au laboratoire. Parmi 62 combinaisons d'amorces sélectives testées sur 16 individus, quatre combinaisons ont été sélectionnées, produisant un nombre de marqueurs en grande quantité et présentant de très bons taux de répétabilité. Le protocole utilisé est le suivant :

A- 1<sup>ère</sup> étape : réaction de digestion avec MseI et EcoRI

DIGESTION double	V(μL) pour 1 réaction	[C] finale ou quantité
ADN génomique (5-150ng)	5	
H2O milliQ	10.25	
NEBuffer 4 10X	2	1x
BSA 10x	2	1x
MseI 10u/μl (NEB)	0.5	5U
EcoRI 20u/μl (NEB)	0.25	5U
<b>Volume réactionnel (μL)</b>	<b>20</b>	

Programme de digestion : 3h à 37°C

B- 2<sup>ème</sup> étape : réaction de ligation

Ligation seule	V(μL) pour 1 réaction	[C] finale ou quantité
Mix de digestion	20	
Ligase buffer 10X	3	1x
ECO ADAPT* 5μM	0.5	0.08μM
MSE ADAPT* 50μM	0.5	0.8μM
T4 DNA ligase 20u/μL (NEB)	0.25	5U
H2O milliQ	5.75	
<b>Volume réactionnel (μL)</b>	<b>30</b>	

Séquences des adaptateurs (d'après Vos *et al.* 1995) :

EcoAdapt1 : CTCGTAGACTGCGTACC  
 EcoAdapt2 : AATTGGTACGCAGTCTAC  
 MseAdapt1 : GACGATGAGTCCTGAG  
 MseAdapt2 : TACTCAGGACTCAT

Programme de ligation : 16h à 16°C

C- 3<sup>ème</sup> étape : réaction de pré-amplification

Deux combinaisons d'amorces présélectives ont été utilisées : Eco+A/Mse+A ou Eco+C/Mse+C. :

Séquence des amorces Eco+1 et Mse+1 (d'après Vos *et al.* 1995):

Eco+1 : GACTGCGTACCAATTC- C ou A  
 Mse+1 : GATGAGTCCTGAGTAA-A ou C

<b>PCR Pré-sélective - Pré-amplification -</b>	<b>V(μL) pour 1 réaction</b>	<b>[C] finale ou quantité</b>
ADN (digestion-ligation dilué 1/10)	<b>5</b>	
Taq Buffer 10x	2.5	1x
MgCl <sub>2</sub> 25mM	1.5	1.5mM
dNTP 2,5mM	2	0.2mM
Eco+1 10μM	<b>1.5</b>	0.6μM
Mse+1 10μM	<b>1.5</b>	0.6μM
BSA (1mg/ml)	<b>2.5</b>	0.1μM
ABgene Thermoprime Pol 5U/μl	0.1	0.5U
H <sub>2</sub> O	8.4	
<b>Volume réactionnel (μL)</b>	<b>25</b>	

Programme de pré-amplification :

dénaturation initiale	72°	6 min	
	94°	2 min	
dénaturation	94°	30s	répété pendant 20 cycles
hybridation	56°	1min	
élongation	72°	1min	
élongation finale	72°	5mn	

A ce stade, est réalisé un gel d'agarose 1.5% afin de vérifier que l'ADN soit bien digéré. Si ce n'est pas le cas, on recommence les trois premières étapes avec 5μL d'ADN dilué au 1/10<sup>ème</sup>.

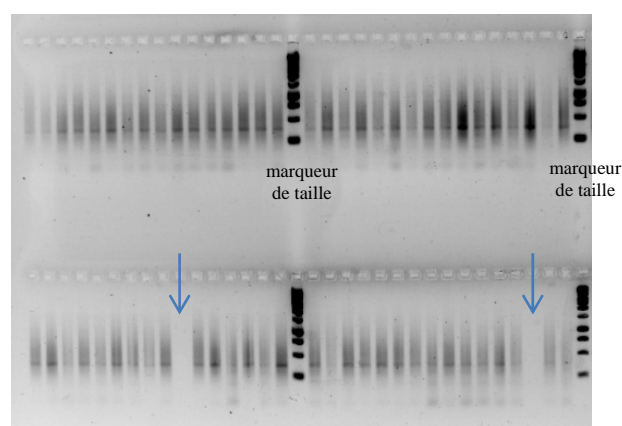


Fig. 1 Gel 1.5% de la pré-amplification Eco+C/Mse+C sur 48 individus. Deux individus ne présentent pas de smears (indiqués par la flèche bleue) : l'ADN de ces deux individus n'a pas été bien digéré.



D- 4<sup>ème</sup> étape : PCR Sélectives

Suite à la Pré amplification Eco+A/Mse+A, deux combinaisons d'amorces sont utilisées : -ACG (amorce marquée en NED)/ - AGC et -AAC (amorce marquée en PET)/ - ACC.

De même, deux combinaisons d'amorces sont utilisées suite à la Pré-amplification Eco+C/Mse+C : -CTG (amorce marquée en FAM)/-CAG et -CTC (amorce marquée en VIC)/ - CTC

Séquence des amorces Eco+3 et Mse+3 (d'après Vos *et al.* 1995):

Eco+3            GACTGCGTACCAATTC-ACG ou AAC ou CTG ou CTC  
Mse+3            GATGAGTCCTGAGTAA-AGC ou ACC ou CAG ou CTC

PCR sélective	V(µL) pour 1 réaction	[C] finale ou quantité
ADN (préamp diluée 1/10)	5	
Taq Buffer 10x	2	1x
MgCl2 25mM	1.2	1.5µM
dNTP 2,5mM	1.2	0.15µM
Eco+3 FLUO 10µM	0.4	0.2µM
Mse+3 10µM	0.4	0.2µM
Thermoprime Pol 5U/µl	0.12	0.6U
BSA 1mg/ml	2	0.1µM
H2O	7.68	
<b>Volume réactionnel (µL)</b>	<b>20</b>	

Programme des PCR-sélectives :

	dénaturation initiale	94°	5min	
Touch-down	dénaturation	94°	30s	-0.7°/cycle sur 12 cycles
	hybridation	65°	30s	
	élongation	72°	60s	
	dénaturation	94°	30s	25 cycles
	hybridation	56°	30s	
	élongation	72°	60s	
	élongation finale	72°	20min	

**ANNEXE 3. Poster relatif à l'article 2** présenté à la conférence Jacques Monod « Développements théoriques et empiriques en génomique évolutive » (Avril 2012)

# Did selective processes occur following introduction of the mollusc *Crepidula fornicata*? - Insights from a genome-scan study -

Florentine Riquet, Claire Daguin-Thiébaud, Frédérique Viard

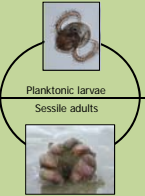
Team DivCo: Lab. Adaptation & Diversité en Milieu Marin (UMR 7144 CNRS UPMC), Station Biologique de Roscoff, France  
Contact: riquet@sb-roscoff.fr

- Natural selection has been invoked as an important **evolutionary driver** for the success of **invasive species**
- Genome scan** approach can be powerful to **detect selective events** between native and introduced populations
- Evidences are however still scarce notably because of the difficulty in disentangling the different evolutionary forces
- Several attributes of marine invertebrates make them useful case studies:**  
large genetic diversity on which selection could act  
large population size, reducing effect of genetic drift

This is exemplified with the slipper limpet *Crepidula fornicata*, an emblematic marine invader:

## BOX 1. *Crepidula fornicata*:

- A **free-swimming larval stage** (up to 7 weeks)
- A **stacking behavior** (males and females piling-up on each other)
- High fecundity** >10 000 larvae per female
- Large environmental tolerance**



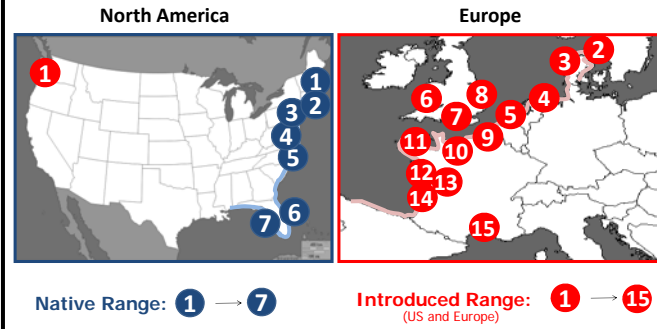
- Long history of invasion**
- High propagule pressure** leading to large introduced populations
- Large genetic polymorphism** in the populations studied up to now (mainly in the European introduced range)

## Objectives:

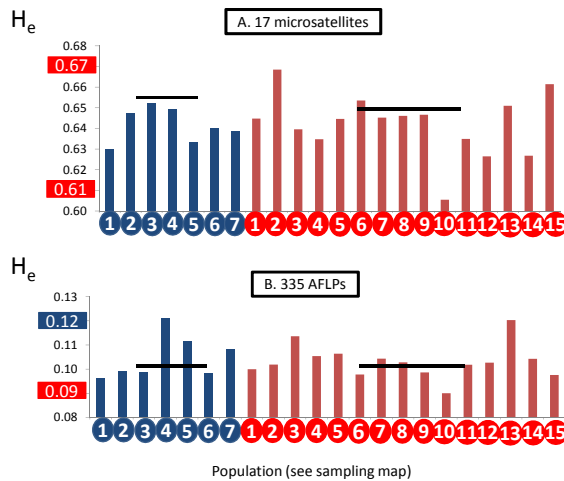
- A detailed comparison of the genetic diversity in the introduced and native ranges
- Looking for outlier loci (AFLP and microsatellites) to assess the likelihood of selective events following introduction

## Sampling design and genotyping

- 22 localities (7 in the native range, 15 in the introduced range)
- c.a. **32 individuals** in each sample
- 17 microsatellites**<sup>(1)</sup> and **335 AFLPs**<sup>(2,3)</sup>



## Native and invasive populations display similar levels of genetic diversity



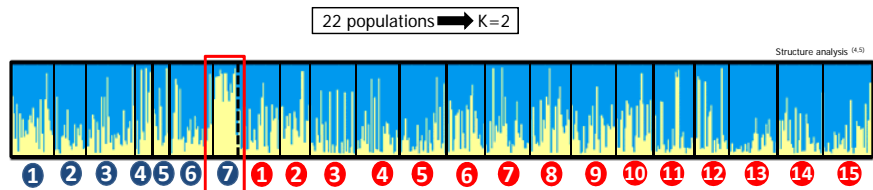
- A large genetic diversity on which selection may act
- A singular population **7** which did not contribute to the introduction in Europe nor in Western USA

## New insights about the routes of invasion

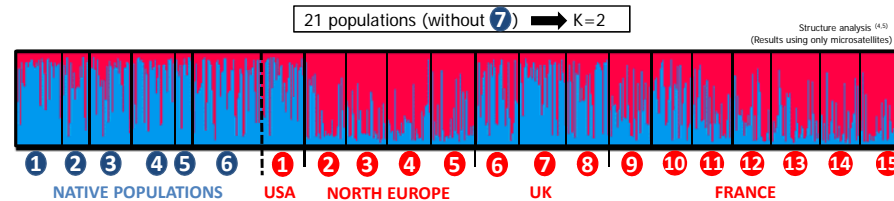
- A genetic structure more important within the **native range** than between **native** and **introduced** populations

$$F_{ST}(\text{native pop}) = 0.026^{***} > F_{ST}(\text{native vs. European pop}) = 0.004^{***}$$

- Source regions are hidden by the population **7**, which is highly genetically differentiated:



- An introduction history in accordance to historical observations:



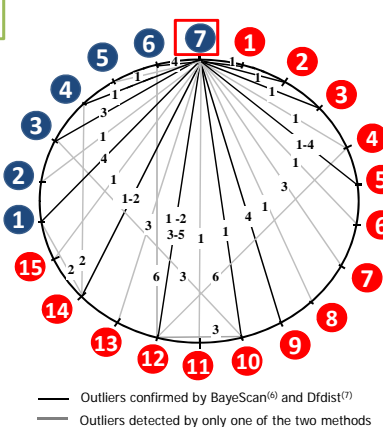
- First introductions in UK,
- Then from UK to northern Europe (gene surfing?)
- Introduction in East-USA, a replicate of the introduction in UK
- Massive introductions in France (with oyster-farming)

## Footprints of selective events? Yes but not where they were first expected ...

Native range: 3 outliers (1, 3, 4); all point out the most southern part of the distribution range (pop **7**)

Selective events in the native range at a well-known biogeographic barrier: coupling of endogenous and exogenous barriers trapped in an environmental boundary<sup>(8)</sup>

Outlier loci in pairwise comparisons (352 loci examined)

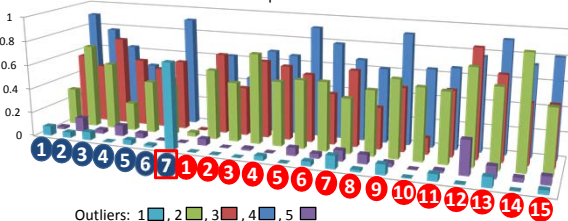


Native vs. Introduced ranges: 5 outliers (1, 2, 3, 4, 5); all involving the West Florida population (pop **7**), which is not a source of the introduction

No evidence for selection following introduction

- What may explain the success of this species?
- Introduction of pre-adapted genotypes
  - Important role played by plasticity
  - Traces of selection too weak to be detected

Allelic frequencies of the 5 outliers



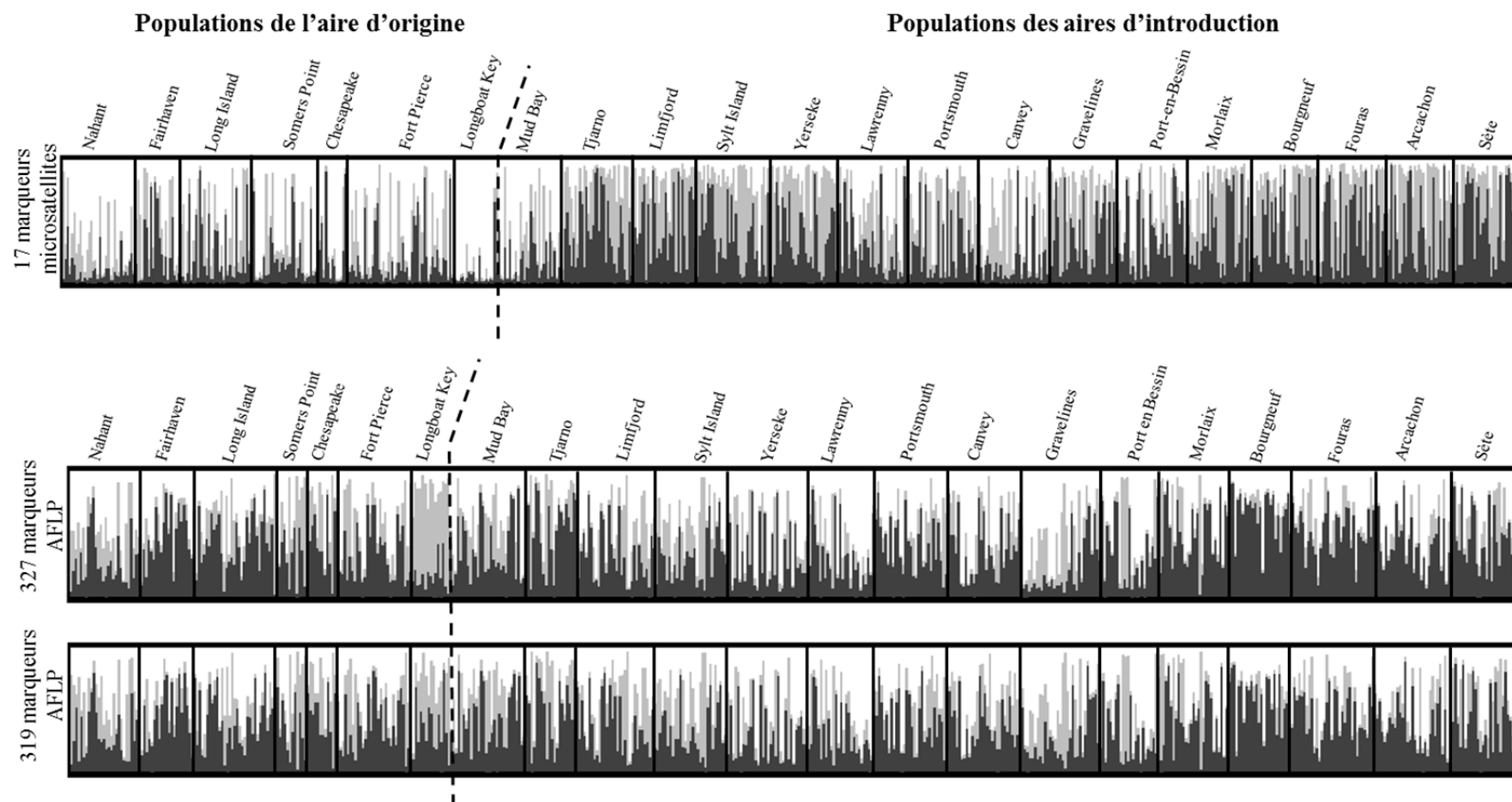
- ### Perspectives
- To better investigate the changes in phenotypic traits and develop quantitative genetic approach
  - To compare results of genome-scan of several introduced marine species to determine the generality of our findings

References:  
<sup>(1)</sup> Riquet et al. (2011) In silico mining and characterization of 12 EST-SSRs for the invasive slipper limpet *Crepidula fornicata*. *Marine Genomics*.  
<sup>(2)</sup> Vos et al. (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*.  
<sup>(3)</sup> Whitlock et al. (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*.  
<sup>(4)</sup> Pritchard et al. (2000) Inference of population structure using multilocus genotype data. *Genetics*.  
<sup>(5)</sup> Javanm et al. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*.  
<sup>(6)</sup> Foll et al. (2008) An approximate Bayesian computation approach to overcome biases that arise when using amplified fragment length polymorphism markers to study population structure. *Genetics*.  
<sup>(7)</sup> Beaumont and Nichols (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London Series B: Biological Sciences*.  
<sup>(8)</sup> Bierne et al. (2011) The coupling hypothesis: why genome-scan may fail to map local adaptation genes. *Molecular Ecology*.

Acknowledgments:  
 We are grateful to the many people who helped us for the sampling in Northern America and Europe. We are also grateful to R. Whitlock for providing helpful advice in AFLP analysis. This work is part of the PhD of FR who benefits from a grant by the CNRS and Région Bretagne with additional help to attend this meeting supported by the GIS Europe Mer. This work was supported by the Agence Nationale pour la Recherche (ANR - project '14-Els').

**ANNEXE 4. Résultats d'analyses d'assignation statistiques réalisées avec le logiciel Structure** (Pritchard *et al.* 2000) à partir des données obtenues pour les 22 populations étudiées dans l'article 2 (7 populations de l'aire d'origine et 15 des aires d'introduction).

Les résultats présentés ici correspondent au cas où le nombre de groupes identifiés par STRUCTURE est  $K=3$ . Les couleurs correspondent à chacun des groupes identifiés. La distribution des 3 couleurs le long de chaque ligne horizontale représente la probabilité d'appartenance de chaque individu (représenté par une ligne horizontale) à chacun des trois groupes. Les populations sont séparées par une ligne noire continue et les populations de l'aire d'origine des populations des aires d'introduction par une ligne en pointillée.



**ANNEXE 5. Résultats détaillés de l'identification d'outliers** par les logiciels Dfdist (A-) et BayeScan (B-) pour les 22 populations analysées dans l'article 2. Les numéros des outliers correspondent aux numéros des outliers dans l'article 2. Un locus est identifié comme outlier (C-) lorsqu'il est détecté à un niveau de confiance de 99% par Dfdist (indiqué « oui » dans A-) et lorsque la probabilité postérieure (p) est supérieure à 0,91 par BayeScan (B-). Concernant les résultats obtenus par le logiciel BayeScan, la « Posterior Odds » (PO, *i.e.* le ratio des probabilités postérieures du modèle incluant et n'incluant pas la sélection), la probabilité postérieure ainsi que l'interprétation de ces résultats selon l'échelle de Jeffreys (1961) sont données. Les marqueurs sont triés par valeurs de probabilité postérieure (BayeScan) décroissantes. Les lignes sont grisées lorsque les marqueurs sont détectés comme outliers de manière « substantielle » ou « valant la peine d'être mentionnés » (interprétation selon l'échelle de Jeffreys). Un « x » indique que les marqueurs ne sont pas identifiés comme outliers. Un « x » (-C) indique un marqueur identifié uniquement par une des deux méthodes.

marqueurs	A- Identification par Dfdist	B- Résultats BayeScan			C- OUTLIERS
		PO	p	Force de la détection	
<b>1</b>	oui	infini	1.00	décisive	<b>1</b>
<b>2</b>	oui	infini	1.00	décisive	<b>2</b>
<b>3</b>	oui	infini	1.00	décisive	<b>3</b>
<b>4</b>	oui	infini	1.00	décisive	<b>4</b>
7	x	infini	1.00	décisive	<b>x</b>
<b>5</b>	oui	713.18	1.00	décisive	<b>5</b>
<b>6</b>	oui	141.84	0.99	décisive	<b>6</b>
11	x	30.24	0.97	forte	<b>x</b>
12	x	17.86	0.95	forte	<b>x</b>
autre locus	x	8.34	0.89	substantielle	x
autre locus	x	6.31	0.86	substantielle	x
autre locus	x	5.03	0.83	substantielle	x
autre locus	x	4.36	0.81	substantielle	x
autre locus	x	3.75	0.79	substantielle	x
autre locus	x	2.71	0.73	vaut la peine d'être mentionné	x
autre locus	x	2.64	0.73	vaut la peine d'être mentionné	x
autre locus	x	2.48	0.71	vaut la peine d'être mentionné	x
autre locus	x	2.32	0.70	vaut la peine d'être mentionné	x
autre locus	x	2.29	0.70	vaut la peine d'être mentionné	x
autre locus	x	2.22	0.69	vaut la peine d'être mentionné	x
autre locus	x	1.60	0.61	vaut la peine d'être mentionné	x
autre locus	x	1.51	0.60	vaut la peine d'être mentionné	x
autre locus	x	1.37	0.58	vaut la peine d'être mentionné	x
9	oui	x	x	x	<b>x</b>
10	oui	x	x	x	<b>x</b>
8	x	x	x	x	x

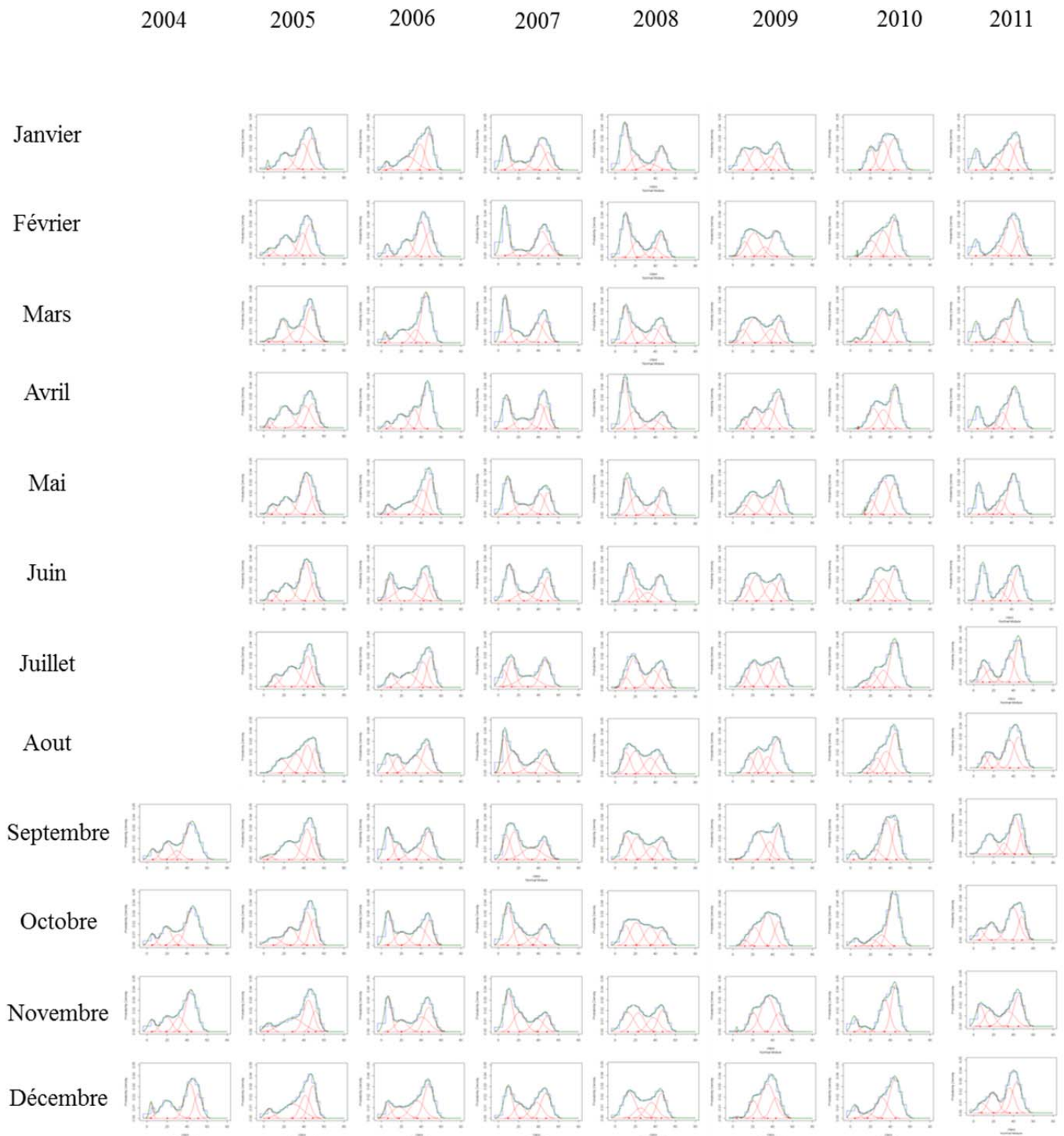
**ANNEXE 6. Résultats détaillés de l'identification d'outliers** par les logiciels Dfdist (A-) et BayeScan (B-) pour les 231 comparaisons de populations deux à deux menées dans l'article 2. Les numéros des outliers correspondent aux numéros des outliers dans l'article 2. Un locus est identifié comme outlier (symbolisé par « 1 » dans C-) lorsqu'il est détecté à un niveau de confiance de 99% par Dfdist (indiqué par « 1 » dans A-) et lorsque la probabilité postérieure (p) est supérieure à 0,91 par BayeScan (symbolisé par « 1 » dans B-). Concernant les résultats obtenus par le logiciel BayeScan, la « Posterior Odds » (PO, *i.e.* le ratio des probabilités postérieures du modèle incluant et n'incluant pas la sélection), la probabilité postérieure ainsi que l'interprétation de ces résultats selon l'échelle de Jeffreys (1961) sont données. Les marqueurs sont triés par valeurs de probabilité postérieure (BayeScan) décroissantes. Les lignes sont grisées lorsque les marqueurs sont détectés comme outliers de manière « substantielle » ou « valant la peine d'être mentionnés » (interprétation selon l'échelle de Jeffreys). Un « x » indique que les marqueurs ne sont pas identifiés comme outliers, un « x » (-C) indique un marqueur identifié uniquement par une des deux méthodes et un « 1 » (-C) indique que les marqueurs sont identifiés à un seuil de 99% par Dfdist et méritent d'être mentionnés par BayeScan.

Paire de populations comparées	N marqueurs AFLPs utilisés (polymorphes)	Outliers									B- Détails des résultats de BayeScan		
		1	2	3	4	5	6	7	8	Autre locus	PO	probabilité postérieure	Force de la détection
		A B C	A B C	A B C	A B C	A B C	A B C	A B C	A B C	A B C	A B		
all populations	327	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	0 1 x	0 1 x	2 15	détails en annexe 5		
Mud Bay	Longboat Key	210	1 1 1		1 0 x						85.2	0.99	très forte
Arcachon	Longboat Key	225	1 1 1	1 1 1		1 0 x	1 0 x	1 0 x			52.2	0.98	
Tjarno	Longboat Key	207	1 1 1	1 0 x	1 0 x						46.2	0.98	
Long Island	Longboat Key	213	1 0 x		1 1 1						39.0	0.97	
Limfjord	Longboat Key	211	1 1 1	1 0 x	1 0 x				1 0 x		35.0	0.97	
Port-en-Bessin	LongBoatKey	194	1 1 1				1 0 x				21.9	0.96	forte
Somers Point	Longboat Key	197	1 1 1		1 0 x						20.4	0.95	
Fort Pierce	Longboat Key	210	1 0 x	1 0 x	1 0 x	1 1 1	1 0 x		1 0 x		16.6	0.94	
Yerseke	Longboat Key	212	1 1 1	1 0 x	1 0 x	1 1 1	1 0 x		1 0 x		15.3	0.94	
Gravelines	Longboat Key	226	1 0 x	1 0 x	1 0 x	1 1 1	1 0 x		1 0 x		10.9	0.92	
Nahant	Longboat Key	197	1 0 x		1 0 x	1 1 1	1 0 x				10.3	0.91	

Bourgneuf	<b>Longboat Key</b>	233	1 1 <b>1</b>	1 1 <b>1</b>	1 1 <b>1</b>	1 0 x			1 1 <b>1</b>			10.2	0.91	
Canvey Island	<b>Longboat Key</b>	218	1 1 <b>1</b>	1 0 x				1 0 x	1 0 x			9.1	0.90	substantielle
Port-en-Bessin	Bourgneuf	215			1 1 <b>1</b>					1 0 x		6.1	0.86	
Fouras	<b>Longboat Key</b>	229	1 0 x		1 1 <b>1</b>	1 0 x	1 0 x	1 0 x				5.3	0.84	
Portsmouth	<b>Longboat Key</b>	235	1 0 x		1 1 <b>1</b>				1 0 x			4.4	0.81	
Port-en-Bessin	<b>Long Island</b>	200			1 1 <b>1</b>							3.3	0.77	
Sète	<b>Longboat Key</b>	219	1 1 <b>1</b>	1 0 x	1 0 x	1 0 x	1 0 x	1 0 x	1 0 x			2.9	0.74	vaut la peine d'être mentionné
<b>Chesapeake</b>	<b>Longboat Key</b>	189	1 1 <b>1</b>		1 0 x		1 0 x					2.7	0.73	
Sylt Island	<b>Longboat Key</b>	216	1 1 <b>1</b>		0 1 x				1 0 x			2.5	0.72	
Sylt Island	Bourgneuf	227								1 1 <b>1</b>		2.2	0.69	
Lawrenny	<b>Longboat Key</b>	213	1 1 <b>1</b>	1 0 x		1 0 x						1.9	0.65	
<b>Fairhaven</b>	<b>Longboat Key</b>	208	1 1 <b>1</b>				1 0 x		1 0 x			1.5	0.60	
Arcachon	<b>Somers Point</b>	219		1 1 <b>1</b>								1.4	0.59	
Morlaix	<b>Longboat Key</b>	211	1 1 <b>1</b>	1 0 x	1 0 x							1.4	0.58	
Bourgneuf	<b>Fort Pierce</b>	230								1 1 <b>1</b>	1 0	1.1	0.53	
Arcachon	<b>Nahant</b>	219		1 1 <b>1</b>								1.1	0.53	
Canvey Island	Bourgneuf	234			1 1 x							0.6	0.36	aucune
Port-en-Bessin	Fouras	213			1 1 x						1 0	0.6	0.36	
Limfjord	<b>Fort Pierce</b>	211									2 1	0.4	0.29	
Arcachon	<b>Fort Pierce</b>	221		1 1 x								0.4	0.28	
Gravelines	<b>Fairhaven</b>	217									1 1	0.3	0.25	
Gravelines	Port en Bessin	204									2 1	0.3	0.25	
Portsmouth	<b>Somers Point</b>	222									1 1	0.3	0.15	
Portsmouth	Sweden	226									1 1	0.3	0.20	
Gravelines	Bourgneuf	237								1 1 x		0.2	0.19	
Bourgneuf	<b>Nahant</b>	228		1 0 x						0 1 x		0.2	0.18	
Limfjord	<b>Somers Point</b>	205									1 1	0.2	0.00	
Portsmouth	Port en Bessin	236									1 1	0.2	0.17	
Tjarno	Bourgneuf	222			1 0 x							x	x	

**ANNEXE 7.** Evolution de la structure de taille de la population benthique échantillonnée mensuellement au Château du Taureau. Données de Sabrina Le Cam.

Les graphiques indiquent la distribution des fréquences de taille basées sur la longueur droite et représenté par les histogrammes en bleu, ainsi que les résultats des décompositions modales associées : les courbes rouges représentent chacune des composantes gaussiennes et la courbe verte est la résultante polymodale de l'ensemble des gaussiennes. Les classes de taille (4mm) sont indiquées en abscisse et les fréquences relatives en ordonnées.





**ANNEXE 8.** Caractéristiques des composantes modales décrivant la structure des histogrammes de fréquences de tailles dans la population de crépidule du Château du Taureau, pour chacun des échantillons disponibles de septembre 2004 à décembre 2011. Données de Sabrina Le Cam.

N est la taille de l'échantillon,  $\chi^2$ , ddl et P sont les valeurs, degré de liberté et probabilité associée au test de  $\chi^2$  d'ajustement de la distribution théorique à la distribution observée. Quatre composantes gaussiennes ont été identifiées au maximum et les valeurs de la proportion, de la moyenne et de l'écart-type de chaque composante sont précisées.

Mois	N		Mode 1	Mode 2	Mode 3	Mode 4	Test
sept-04	314	Proportion	0.09	0.21	0.10	0.60	ddl = 9
		Moyenne (mm)	5.65	19.98	30.06	44.88	2.1896
		Ecart-type (mm)	3.62	5.00	5.00	6.66	P = 0.9881
oct-04	400	Proportion	0.09	0.22	0.13	0.57	ddl = 9
		Moyenne (mm)	4.95	19.00	31.06	45.87	2.2387
		Ecart-type (mm)	3.22	5.00	5.00	6.14	P = 0.9871
nov-04	287	Proportion	0.10	0.18	0.17	0.54	ddl = 7
		Moyenne (mm)	4.66	19.50	32.05	44.06	1.0088
		Ecart-type (mm)	3.49	5.00	5.00	5.56	P = 0.9947
déc-04	213	Proportion	0.06	0.30	0.41	0.23	ddl = 8
		Moyenne (mm)	3.98	20.18	42.66	51.03	1.9159
		Ecart-type (mm)	1.70	7.00	5.00	4.00	P = 0.9835
janv-05	316	Proportion	0.02	0.24	0.37	0.37	ddl = 8
		Moyenne (mm)	3.84	20.77	38.89	47.81	2.2034
		Ecart-type (mm)	1.05	6.00	5.99	5.00	P = 0.9741
févr-05	255	Proportion	0.06	0.32	0.24	0.37	ddl = 6
		Moyenne (mm)	5.28	21.89	38.60	46.17	0.665
		Ecart-type (mm)	3.94	6.59	4.49	5.00	P = 0.9952
mars-05	395	Proportion	0.02	0.25	0.32	0.41	ddl = 8
		Moyenne (mm)	4.81	19.06	36.83	47.01	1.3656
		Ecart-type (mm)	2.58	4.86	8.69	5.00	P = 0.9947
avr-05	411	Proportion	0.05	0.36	0.31	0.28	ddl = 7
		Moyenne (mm)	5.69	21.05	41.22	48.38	0.7235
		Ecart-type (mm)	2.60	6.88	6.02	5.00	P = 0.9981
mai-05	322	Proportion	0.08	0.26	0.49	0.17	ddl = 6
		Moyenne (mm)	7.95	22.43	41.42	49.84	0.8097
		Ecart-type (mm)	3.36	6.19	5.31	4.00	P = 0.9918
juin-05	385	Proportion	NA	NA	NA	NA	NA
		Moyenne (mm)	NA	NA	NA	NA	NA
		Ecart-type (mm)	NA	NA	NA	NA	NA
juil-05	382	Proportion	0.09	0.38	0.33	0.20	ddl = 6
		Moyenne (mm)	11.07	27.61	43.58	49.19	0.7725
		Ecart-type (mm)	3.68	7.75	4.55	4.00	P = 0.9928
août-05	305	Proportion	0.20	0.27	0.33	0.20	ddl = 7
		Moyenne (mm)	17.38	32.13	43.56	51.24	1.3198
		Ecart-type (mm)	6.02	6.00	5.00	3.40	P = 0.9879
sept-05	314	Proportion	0.04	0.39	0.36	0.21	ddl = 8
		Moyenne (mm)	3.67	25.14	42.93	49.53	0.844
		Ecart-type (mm)	4.00	8.76	5.00	4.00	P = 0.9991
oct-05	303	Proportion	0.12	0.25	0.39	0.24	ddl = 7
		Moyenne (mm)	9.34	26.80	43.01	49.77	1.124
		Ecart-type (mm)	6.46	5.86	5.00	4.00	P = 0.9926

Mois	N		Mode 1	Mode 2	Mode 3	Mode 4	Test
nov-05	307	Proportion	0.07	0.35	0.37	0.20	ddl = 10
		Moyenne (mm)	4.03	28.31	44.48	51.54	1.6075
		Ecart-type (mm)	4.00	11.05	5.00	4.00	P = 0.9986
déc-05	312	Proportion	0.08	0.34	0.27	0.31	ddl = 8
		Moyenne (mm)	5.04	28.52	41.44	48.82	1.5174
		Ecart-type (mm)	3.73	9.92	5.00	4.00	P = 0.9924
janv-06	308	Proportion	0.06	0.28	0.31	0.36	ddl = 7
		Moyenne (mm)	4.84	26.48	38.80	48.28	1.1171
		Ecart-type (mm)	2.93	8.75	5.00	4.00	P = 0.9927
févr-06	239	Proportion	0.09	0.22	0.41	0.28	ddl = 8
		Moyenne (mm)	5.76	24.90	40.13	48.11	1.8162
		Ecart-type (mm)	3.07	5.32	5.00	5.00	P = 0.9861
mars-06	300	Proportion	0.05	0.22	0.16	0.57	ddl = 7
		Moyenne (mm)	3.83	20.72	34.76	45.12	1.0312
		Ecart-type (mm)	1.93	7.19	5.00	5.00	P = 0.9943
avr-06	300	Proportion	0.05	0.19	0.19	0.56	ddl = 8
		Moyenne (mm)	5.25	19.63	33.07	46.31	1.606
		Ecart-type (mm)	2.93	5.46	4.00	5.00	P = 0.9908
mai-06	277	Proportion	0.07	0.29	0.30	0.35	ddl = 8
		Moyenne (mm)	6.76	27.52	41.99	49.55	1.547
		Ecart-type (mm)	3.06	9.83	5.00	4.00	P = 0.9919
juin-06	309	Proportion	0.21	0.30	0.34	0.16	ddl = 7
		Moyenne (mm)	8.69	23.16	42.31	49.35	1.0393
		Ecart-type (mm)	3.63	9.51	5.00	4.00	P = 0.9942
juil-06	304	Proportion	0.14	0.27	0.30	0.29	ddl = 7
		Moyenne (mm)	9.47	26.91	41.70	49.50	0.9226
		Ecart-type (mm)	4.55	7.75	5.00	4.00	P = 0.996
août-06	322	Proportion	0.20	0.15	0.31	0.33	ddl = 8
		Moyenne (mm)	5.88	16.03	34.57	46.38	1.0174
		Ecart-type (mm)	4.41	4.00	7.82	5.00	P = 0.9981
sept-06	336	Proportion	0.30	0.15	0.21	0.34	ddl = 8
		Moyenne (mm)	6.57	17.20	35.32	47.36	0.8573
		Ecart-type (mm)	3.97	4.00	7.76	5.00	P = 0.999
oct-06	305	Proportion	0.29	0.19	0.29	0.23	ddl = 7
		Moyenne (mm)	6.74	20.25	38.77	48.05	0.7696
		Ecart-type (mm)	3.65	6.00	6.75	4.00	P = 0.9977
nov-06	308	Proportion	0.28	0.17	0.26	0.29	ddl = 8
		Moyenne (mm)	6.61	20.02	40.71	47.64	0.5638
		Ecart-type (mm)	3.44	6.00	7.47	5.00	P = 0.9998
déc-06	299	Proportion	0.11	0.19	0.29	0.41	ddl = 9
		Moyenne (mm)	6.53	17.51	35.93	47.65	2.2344
		Ecart-type (mm)	3.40	7.00	7.00	5.00	P = 0.9872
janv-07	302	Proportion	0.29	0.15	0.36	0.21	ddl = 8
		Moyenne (mm)	6.56	21.02	41.07	49.42	0.4755
		Ecart-type (mm)	3.58	7.00	5.84	5.00	P = 0.9999
févr-07	306	Proportion	0.42	0.09	0.35	0.14	ddl = 8
		Moyenne (mm)	6.43	21.48	43.45	49.74	0.7948
		Ecart-type (mm)	3.53	6.00	6.18	5.00	P = 0.9992
mars-07	300	Proportion	0.36	0.17	0.27	0.21	ddl = 7
		Moyenne (mm)	6.60	18.20	39.71	47.06	0.3732
		Ecart-type (mm)	3.32	6.00	6.04	4.00	P = 0.9998

Mois	N		Mode 1	Mode 2	Mode 3	Mode 4	Test
avr-07	302	Proportion	0.32	0.18	0.28	0.21	ddl = 7
		Moyenne (mm)	7.88	24.23	42.04	47.67	0.6787
		Ecart-type (mm)	4.09	7.00	4.87	4.00	P = 0.9985
mai-07	301	Proportion	0.37	0.18	0.24	0.21	ddl = 7
		Moyenne (mm)	9.03	24.05	41.36	49.50	1.1203
		Ecart-type (mm)	4.19	7.00	4.97	4.00	P = 0.9926
juin-07	308	Proportion	0.40	0.14	0.23	0.22	ddl = 7
		Moyenne (mm)	10,88	25,36	42,53	49,53	0.8648
		Ecart-type (mm)	4.74	7.00	5.26	4.00	P = 0.9967
juil-07	438	Proportion	0.16	0.26	0.26	0.32	ddl = 9
		Moyenne (mm)	4.94	13.09	29.50	46.95	2.5695
		Ecart-type (mm)	4.00	4.00	9.64	5.00	P = 0.979
août-07	320	Proportion	0.26	0.30	0.21	0.22	ddl = 9
		Moyenne (mm)	5.93	14.59	36.94	47.64	1.869
		Ecart-type (mm)	3.00	6.00	7.68	5.00	P = 0.9934
sept-07	426	Proportion	0.22	0.33	0.27	0.18	ddl = 9
		Moyenne (mm)	8.16	15.81	33.22	46.32	1.2667
		Ecart-type (mm)	3.98	5.00	9.57	4.00	P = 0.9985
oct-07	529	Proportion	0.41	0.25	0.10	0.25	ddl = 7
		Moyenne (mm)	9.27	20.06	34.46	46.48	1.0059
		Ecart-type (mm)	4.38	6.00	5.19	5.00	P = 0.9947
nov-07	530	Proportion	0.40	0.27	0.18	0.15	ddl = 7
		Moyenne (mm)	9.90	20.78	39.27	48.43	0.5506
		Ecart-type (mm)	4.33	6.00	5.84	4.00	P = 0.9992
déc-07	303	Proportion	0.31	0.17	0.22	0.30	ddl = 8
		Moyenne (mm)	9.69	21.08	38.09	47.57	1.4977
		Ecart-type (mm)	4.06	5.00	5.95	5.00	P = 0.9927
janv-08	418	Proportion	0.47	0.17	0.11	0.25	ddl = 7
		Moyenne (mm)	9.35	20.50	35.00	46.51	0.9885
		Ecart-type (mm)	4.26	5.00	7.49	4.79	P = 0.995
fevr-08	307	Proportion	0.44	0.18	0.14	0.25	ddl = 7
		Moyenne (mm)	10.18	21.40	38.00	46.25	0.4231
		Ecart-type (mm)	4.22	5.00	5.75	4.50	P = 0.9997
mars-08	356	Proportion	0.37	0.22	0.21	0.20	ddl = 7
		Moyenne (mm)	10.04	21.42	39.00	46.99	0.7607
		Ecart-type (mm)	4.31	6.00	6.20	4.70	P = 0.9978
avr-08	449	Proportion	0.45	0.26	0.14	0.15	ddl = 6
		Moyenne (mm)	9.31	19.02	38.20	47.91	0.8736
		Ecart-type (mm)	3.86	6.00	6.56	4.58	P = 0.99
mai-08	387	Proportion	0.33	0.27	0.16	0.24	ddl = 7
		Moyenne (mm)	11.13	20.45	38.00	48.11	0.8253
		Ecart-type (mm)	3.86	6.00	5.81	4.07	P = 0.9972
juin-08	329	Proportion	0.38	0.16	0.13	0.33	ddl = 8
		Moyenne (mm)	13.82	22.00	32.31	45.24	0.6449
		Ecart-type (mm)	4.61	5.00	6.00	5.24	P = 0.9997
juil-08	302	Proportion	0.10	0.44	0.25	0.21	ddl = 10
		Moyenne (mm)	10.49	19.00	38.00	47.30	1.8972
		Ecart-type (mm)	4.00	6.00	6.00	4.27	P = 0.9971
aout 08	349	Proportion	0.22	0.32	0.23	0.22	ddl = 9
		Moyenne (mm)	10.70	20.00	35.00	44.76	0.7942
		Ecart-type (mm)	4.40	6.00	6.00	5.21	P = 0.9998

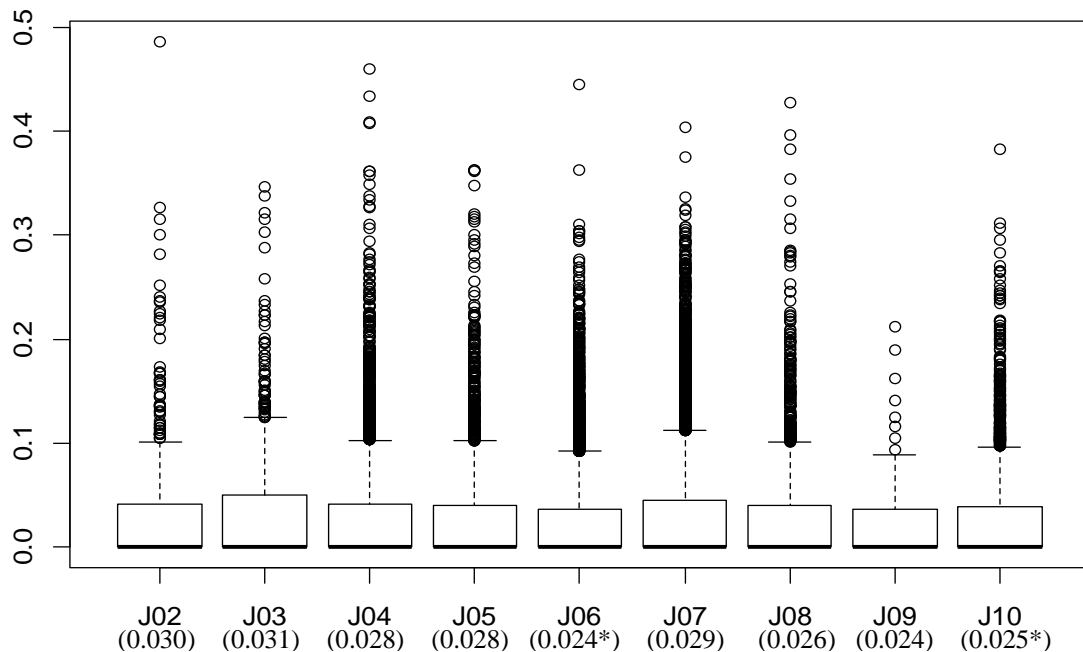
Mois	N		Mode 1	Mode 2	Mode 3	Mode 4	Test
sept-08	401	Proportion	0.27	0.32	0.19	0.22	ddl = 9
		Moyenne (mm)	10.80	22.00	37.00	47.59	1.0683
		Ecart-type (mm)	4.89	6.00	6.00	4.70	P = 0.9993
oct-08	346	Proportion	0.23	0.33	0.26	0.18	ddl = 9
		Moyenne (mm)	10.02	21.00	35.00	47.61	0.669
		Ecart-type (mm)	4.92	6.00	6.00	4.61	P = 0.9999
nov-08	433	Proportion	0.29	0.27	0.20	0.24	ddl = 9
		Moyenne (mm)	12.47	23.00	38.00	47.21	0.8248
		Ecart-type (mm)	6.33	6.00	6.00	4.66	P = 0.9997
déc-08	451	Proportion	0.38	0.14	0.22	0.26	ddl = 8
		Moyenne (mm)	13.68	25.03	36.00	45.98	0.566
		Ecart-type (mm)	6.24	6.00	6.00	4.45	P = 0.9998
janv-09	459	Proportion	0.26	0.30	0.19	0.25	ddl = 8
		Moyenne (mm)	10.28	23.61	38.00	46.08	0.8625
		Ecart-type (mm)	5.48	6.00	6.00	4.84	P = 0.999
févr-09	417	Proportion	0.16	0.38	0.11	0.35	ddl = 9
		Moyenne (mm)	12.00	21.36	32.00	44.38	1.9166
		Ecart-type (mm)	4.50	6.72	5.00	5.86	P = 0.9927
mars-09	446	Proportion	0.15	0.43	0.20	0.23	ddl = 7
		Moyenne (mm)	10.37	21.85	39.00	47.80	0.9495
		Ecart-type (mm)	4.72	7.38	6.00	4.38	P = 0.9956
avr-09	296	Proportion	0.07	0.25	0.29	0.40	ddl = 8
		Moyenne (mm)	10.77	21.99	37.00	47.04	2.1897
		Ecart-type (mm)	2.95	5.02	6.00	5.00	P = 0.9746
mai-09	403	Proportion	0.13	0.29	0.27	0.31	ddl = 8
		Moyenne (mm)	11.34	21.27	37.00	47.41	0.628
		Ecart-type (mm)	5.59	6.00	6.00	4.32	P = 0.9997
juin-09	524	Proportion	0.18	0.37	0.26	0.20	ddl = 8
		Moyenne (mm)	13.90	23.60	39.00	48.33	0.6455
		Ecart-type (mm)	4.48	6.00	6.00	3.92	P = 0.9997
juil-09	334	Proportion	0.09	0.30	0.33	0.28	ddl = 8
		Moyenne (mm)	11.78	20.99	35.00	46.46	1.7472
		Ecart-type (mm)	3.42	5.00	6.00	4.70	P = 0.9878
août-09	268	Proportion	0.16	0.21	0.19	0.44	ddl = 10
		Moyenne (mm)	16.49	25.49	35.00	44.86	1.7414
		Ecart-type (mm)	4.00	4.00	5.00	5.52	P = 0.998
sept-09	412	Proportion	0.01	0.45	0.21	0.33	ddl = 7
		Moyenne (mm)	3.06	24.99	37.00	46.35	0.8364
		Ecart-type (mm)	4.76	7.07	5.00	4.25	P = 0.9971
oct-09	275	Proportion	0.05	0.23	0.46	0.26	ddl = 7
		Moyenne (mm)	12.09	21.69	35.00	46.34	0.6548
		Ecart-type (mm)	3.39	5.25	6.00	4.51	P = 0.9987
nov-09	218	Proportion	0.01	0.20	0.58	0.21	ddl = 7
		Moyenne (mm)	3.95	20.58	35.00	46.08	0.2932
		Ecart-type (mm)	0.75	4.47	7.00	4.77	P = 0.9999
déc-09	217	Proportion	0.01	0.18	0.57	0.24	ddl = 7
		Moyenne (mm)	4.00	20.67	35.00	43.35	1.0858
		Ecart-type (mm)	2.25	4.40	7.00	4.94	P = 0.9933
janv-10	253	Proportion	0.00	0.22	0.34	0.44	ddl = 10
		Moyenne (mm)	9.55	20.03	33.00	44.00	2.0438
		Ecart-type (mm)	1.22	4.00	5.04	6.00	P = 0.996

Mois	N		Mode 1	Mode 2	Mode 3	Mode 4	Test
févr-10	303	Proportion	0.00	0.22	0.37	0.41	ddl = 8
		Moyenne (mm)	6.66	20.65	33.00	44.63	2.7996
		Ecart-type (mm)	0.20	5.76	6.00	4.62	P = 0.9463
mars-10	245	Proportion	0.03	0.20	0.45	0.32	ddl = 7
		Moyenne (mm)	5.42	21.23	33.00	46.67	0.7369
		Ecart-type (mm)	2.62	5.63	6.00	4.50	P = 0.9998
avr-10	282	Proportion	0.00	0.28	0.27	0.45	ddl = 7
		Moyenne (mm)	7.50	22.96	33.00	45.24	0.3483
		Ecart-type (mm)	0.63	5.75	6.00	4.58	P = 0.9998
mai-10	238	Proportion	0.00	0.13	0.49	0.37	ddl = 7
		Moyenne (mm)	14.30	21.20	33.00	44.21	1.1978
		Ecart-type (mm)	0.24	3.83	6.00	5.15	P = 0.991
juin-10	320	Proportion	0.00	0.30	0.31	0.39	ddl = 8
		Moyenne (mm)	7.61	24.40	33.00	45.03	0.9565
		Ecart-type (mm)	0.65	6.24	6.00	5.24	P = 0.9985
juil-10	237	Proportion	0.05	0.14	0.24	0.57	ddl = 8
		Moyenne (mm)	12.76	23.70	33.00	44.88	0.8417
		Ecart-type (mm)	3.39	4.46	6.00	5.24	P = 0.9991
août-10	222	Proportion	0.05	0.17	0.31	0.46	ddl = 6
		Moyenne (mm)	16.73	26.79	35.94	44.77	0.3354
		Ecart-type (mm)	3.40	4.70	6.00	4.80	P = 0.9993
sept-10	209	Proportion	0.07	0.12	0.47	0.34	ddl = 6
		Moyenne (mm)	3.18	24.84	36.26	46.02	1.4421
		Ecart-type (mm)	3.12	5.00	5.00	3.90	P = 0.9632
oct-10	249	Proportion	0.08	0.08	0.13	0.71	ddl = 7
		Moyenne (mm)	4.96	21.71	31.08	43.76	0.5594
		Ecart-type (mm)	4.26	5.00	5.00	5.28	P = 0.9992
nov-10	278	Proportion	0.11	0.05	0.38	0.45	ddl = 6
		Moyenne (mm)	3.73	14.53	34.31	45.07	1.2586
		Ecart-type (mm)	3.11	4.00	5.00	4.15	P = 0.9739
déc-10	281	Proportion	0.11	0.06	0.29	0.53	ddl = 8
		Moyenne (mm)	4.13	17.03	31.58	44.99	1.8992
		Ecart-type (mm)	3.601	5.000	6.000	5.619	P = 0.9839
janv-11	340	Proportion	0.19	0.15	0.38	0.29	ddl = 7
		Moyenne (mm)	4.26	23.65	36.98	46.27	0.5567
		Ecart-type (mm)	3.61	5.00	6.00	4.38	P = 0.9992
févr-11	274	Proportion	0.14	0.15	0.53	0.19	ddl = 7
		Moyenne (mm)	4.01	26.66	39.45	47.65	0.9691
		Ecart-type (mm)	3.43	6.00	6.00	4.03	P = 0.9953
mars-11	306	Proportion	0.17	0.06	0.26	0.51	ddl = 8
		Moyenne (mm)	4.33	22.00	32.17	45.87	1.0633
		Ecart-type (mm)	3.45	6.00	5.00	5.03	P = 0.9978
avr-11	323	Proportion	0.21	0.04	0.16	0.59	ddl = 9
		Moyenne (mm)	5.73	22.00	30.69	43.16	1.6965
		Ecart-type (mm)	4.01	4.00	4.00	5.89	P = 0.9954
mai-11	371	Proportion	0.26	0.06	0.14	0.54	ddl = 7
		Moyenne (mm)	7.13	21.39	30.00	42.08	1.179
		Ecart-type (mm)	3.44	4.00	4.00	5.62	P = 0.9914
juin-11	367	Proportion	0.35	0.10	0.20	0.36	ddl = 7
		Moyenne (mm)	11.13	28.96	38.38	46.85	1.2318
		Ecart-type (mm)	3.79	4.00	4.00	4.63	P = 0.9902

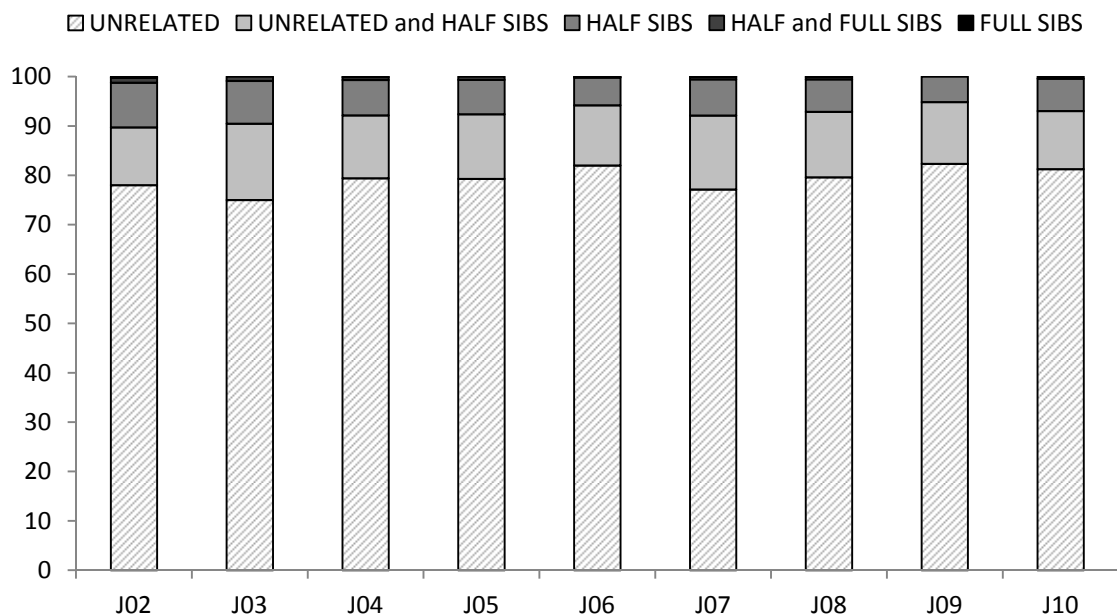
Mois	N		Mode 1	Mode 2	Mode 3	Mode 4	Test
juil-11	317	Proportion	0.15	0.14	0.30	0.40	ddl = 8
		Moyenne (mm)	9.00	16.00	36.07	45.68	2.0083
		Ecart-type (mm)	3.90	5.00	5.00	4.06	P = 0.9808
août-11	301	Proportion	0.09	0.14	0.40	0.38	ddl = 8
		Moyenne (mm)	11.00	18.29	36.37	45.10	1.6145
		Ecart-type (mm)	3.05	4.00	6.00	5.18	P = 0.9906
sept-11	377	Proportion	0.30	0.11	0.44	0.15	ddl = 7
		Moyenne (mm)	16.05	30.65	41.59	48.57	1.841
		Ecart-type (mm)	6.23	4.00	5.00	3.00	P = 0.9681
oct-11	349	Proportion	0.11	0.22	0.47	0.21	ddl = 7
		Moyenne (mm)	6.10	18.17	39.65	48.56	0.6832
		Ecart-type (mm)	4.00	5.00	6.00	3.70	P = 0.9985
nov-11	369	Proportion	0.20	0.19	0.26	0.35	ddl = 7
		Moyenne (mm)	6.40	16.09	34.54	45.11	0.6861
		Ecart-type (mm)	4.00	5.00	7.16	4.74	P = 0.9984
déc-11	303	Proportion	0.11	0.24	0.28	0.36	ddl = 8
		Moyenne (mm)	8.00	19.68	36.20	43.45	1.6994
		Ecart-type (mm)	4.89	5.00	4.75	5.00	P = 0.9889

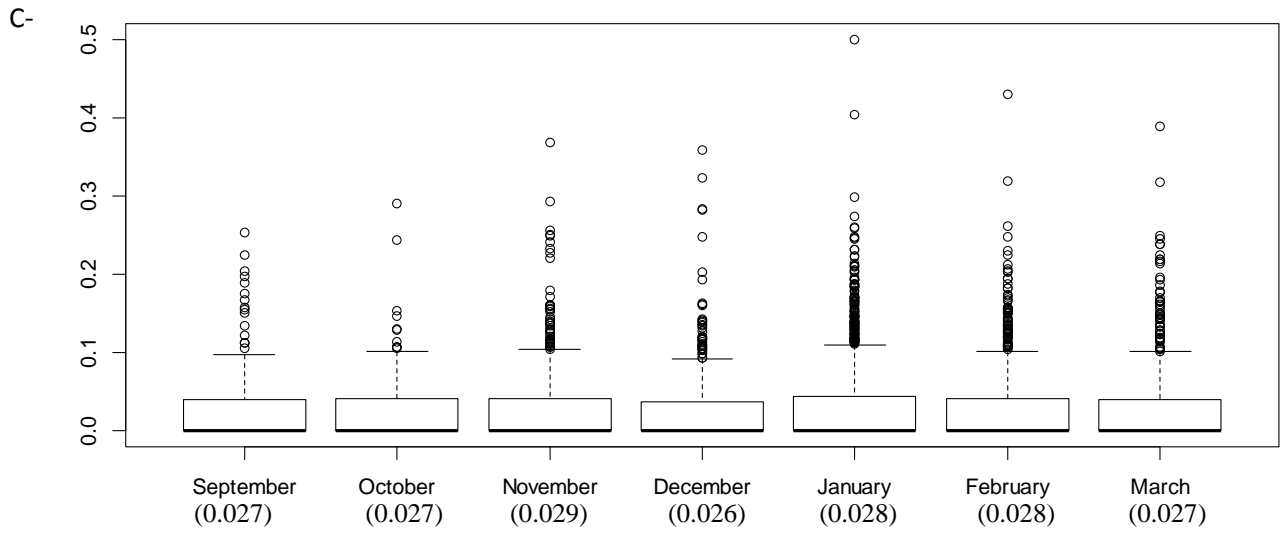
**ANNEXE 9.** Données d'apparement relatives aux échantillons de juvéniles étudiés dans l'article 4. En complément des analyses de taille efficace réalisées par la méthode de Wang (2009) dans le logiciel COLONY, nous avons estimé les coefficients d'apparement par la méthode de Wang (2007) implémentée dans le logiciel Coancestry (Wang 2011) en utilisant les échantillonnages de juvéniles sur neuf années, de 2002 et 2010 (A-, B-) et sur sept mois, de Septembre 2010 à Mars 2011 (C-, D-). Pour chacun des échantillons étudiés, les coefficients d'apparement entre chaque paire d'individus sont estimés et la distribution de ces coefficients au sein de chaque échantillon est donnée dans les figures A- et C-. L'apparement moyen de chaque échantillon est donné entre parenthèse et les échantillons significativement différents des autres échantillons sont signalés par \*. Deux individus sont identifiés sans ambiguïté comme plein-frères lorsqu'ils ont un coefficient d'apparement supérieur à 0,40 ; deux individus sont identifiés sans ambiguïté comme demi-frères lorsqu'ils ont un coefficient d'apparement variant entre 0,1 et 0,25 ; enfin, deux individus sont identifiés sans ambiguïté comme non-apparentés lorsqu'ils ont un coefficient d'apparement inférieur à 0,05 (cf. article 5 pour l'établissement de ces seuils obtenus par simulations). Les figures B- et D- présentent, par échantillon, le pourcentage de paires d'individus identifiées sans ambiguïté comme plein-frères, demi-frères et non-apparentés ainsi que le pourcentage de paires d'individus étant non-apparentés ou demi-frères et demi ou plein-frères.

A-

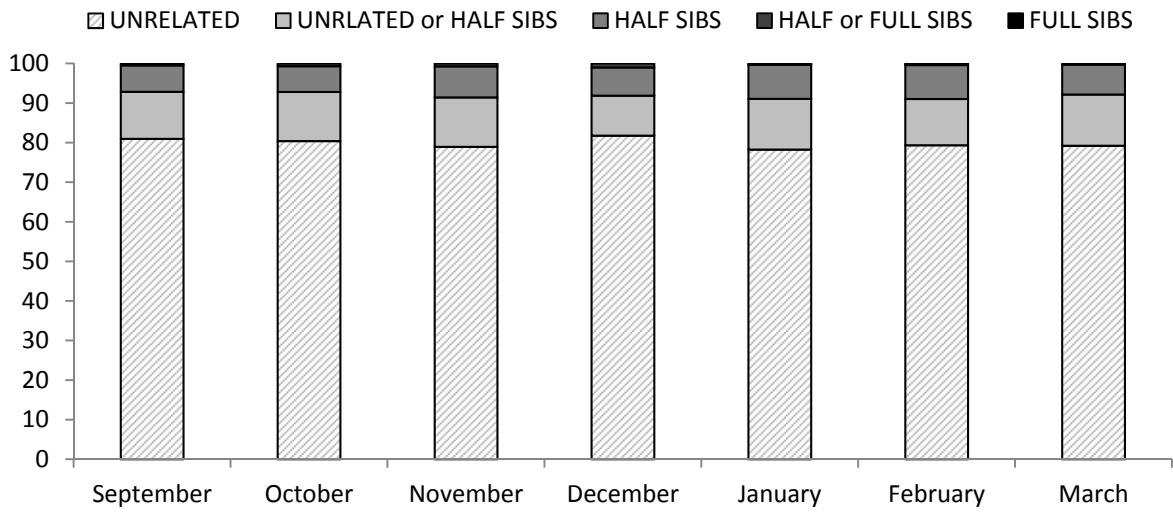


B-



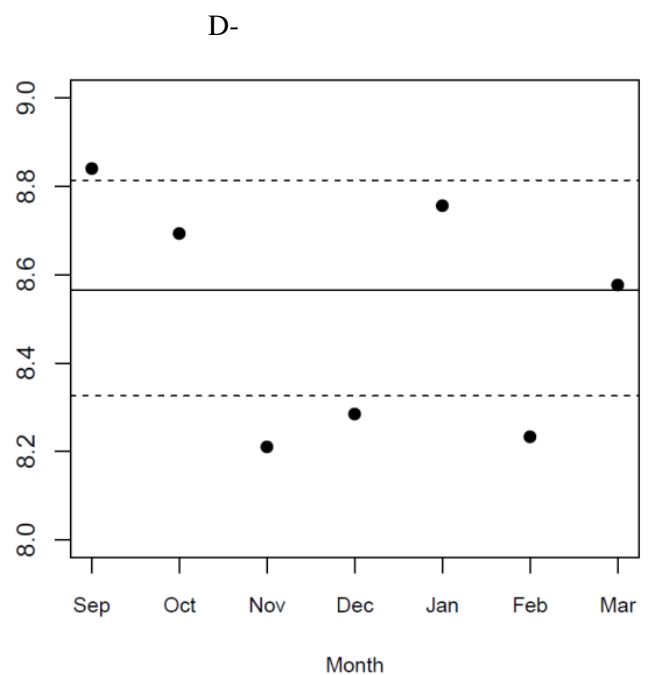
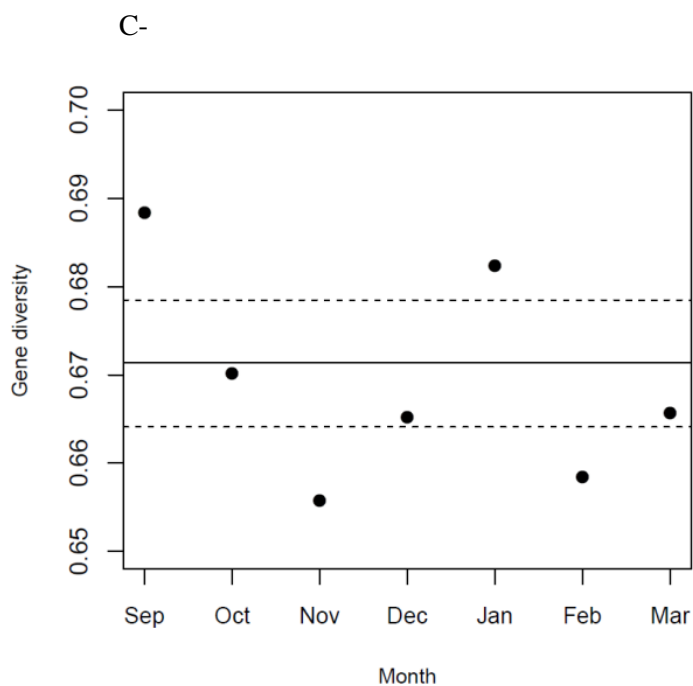
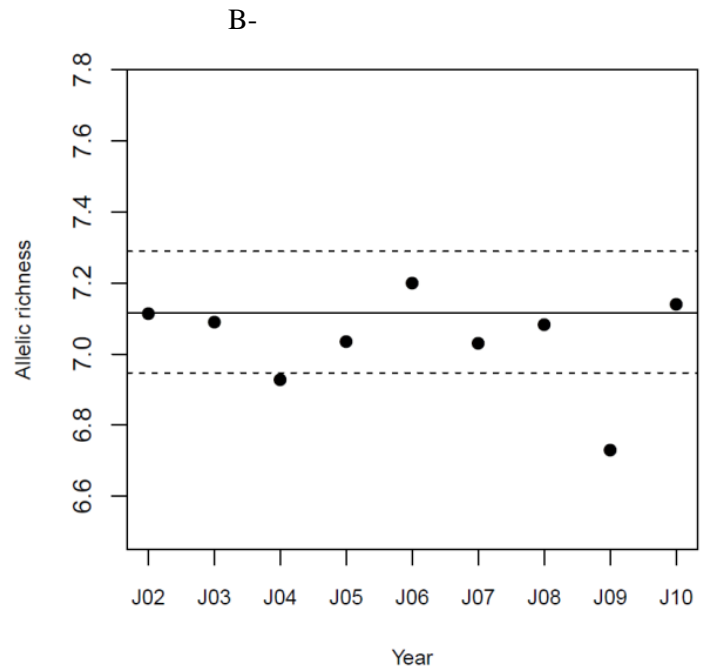
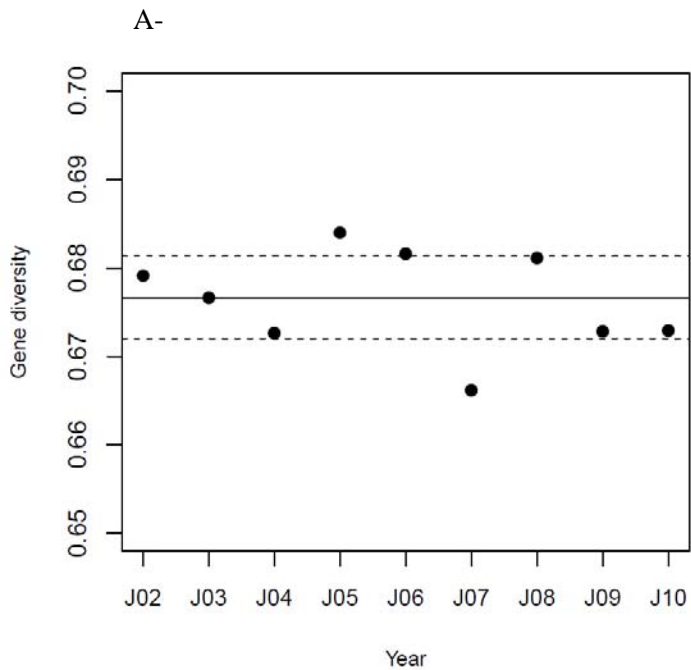


D-



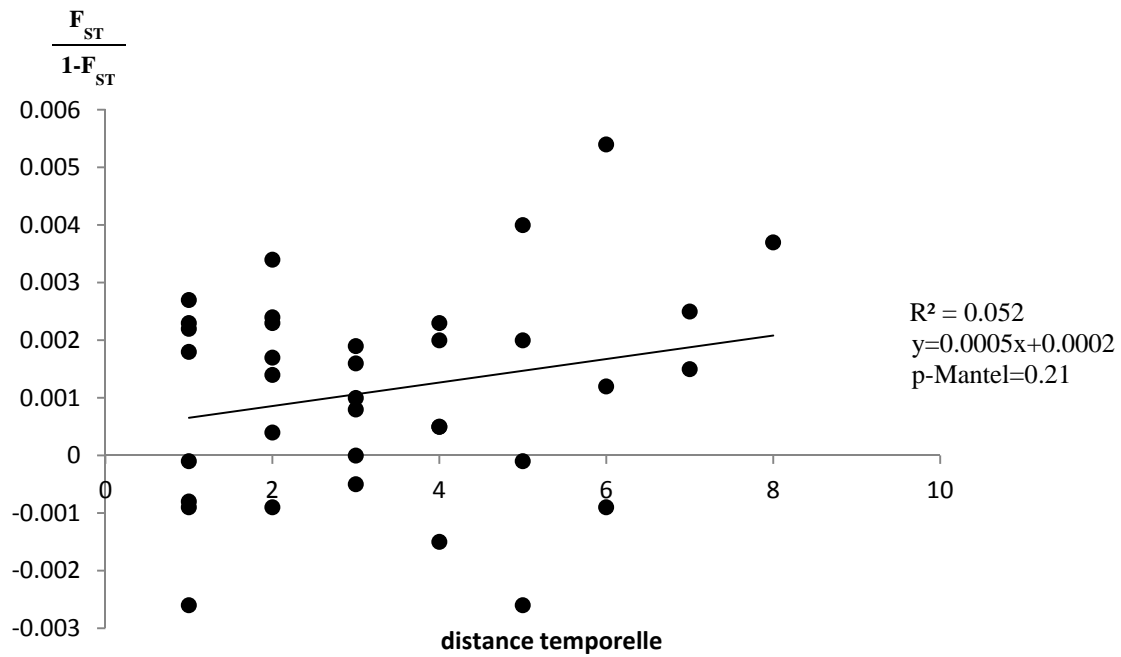


**ANNEXE 10.** Distributions de deux indices de diversité génétique (hétérozygotie attendue (A-, C-) et richesse allélique (B-, D-)) estimés pour les juvéniles échantillonnés sur neuf années successives (de 2002 à 2010, A-, B-) et sur sept mois successifs (de Septembre 2010 à Mars 2011, C-, D-, **cf. article 4**). La moyenne pour chaque indice de diversité par échantillonnage est représentée par la ligne continue. A partir de la moyenne, 1000 valeurs ont été simulées et l'intervalle de confiance a été calculé. Il est représenté par les lignes en pointillées. Les indices situés en-dessous ou au-dessus de l'intervalle de confiance sont alors plus faibles ou plus forts qu'attendus. Les simulations ont été réalisées avec Thomas Broquet.

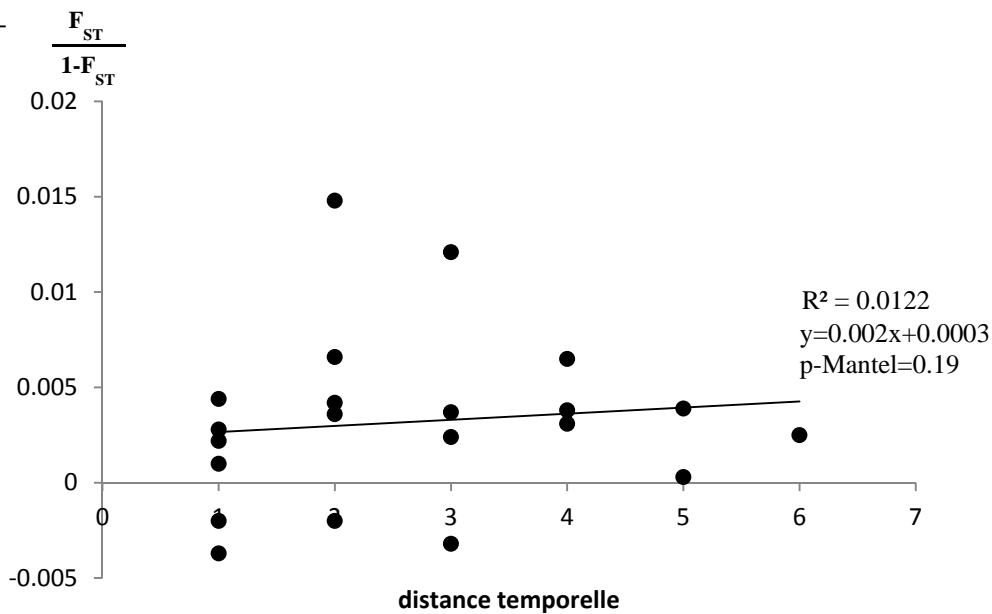


**ANNEXE 11.** Isolement par le temps et test de Mantel associé pour les échantillons de juvéniles de 2002 à 2010 (A-) et de Septembre 2010 à Mars 2011 (B-, cf. article 4). Les régressions linéaires et tests de Mantel n'ont pas montré de corrélation significative entre les distances génétiques et temporelles que ce soit entre juvéniles échantillonnés de 2002 à 2009 et de Septembre 2010 à Mars 2011. La ligne continue représente la droite de régression linéaire.

A-



B-

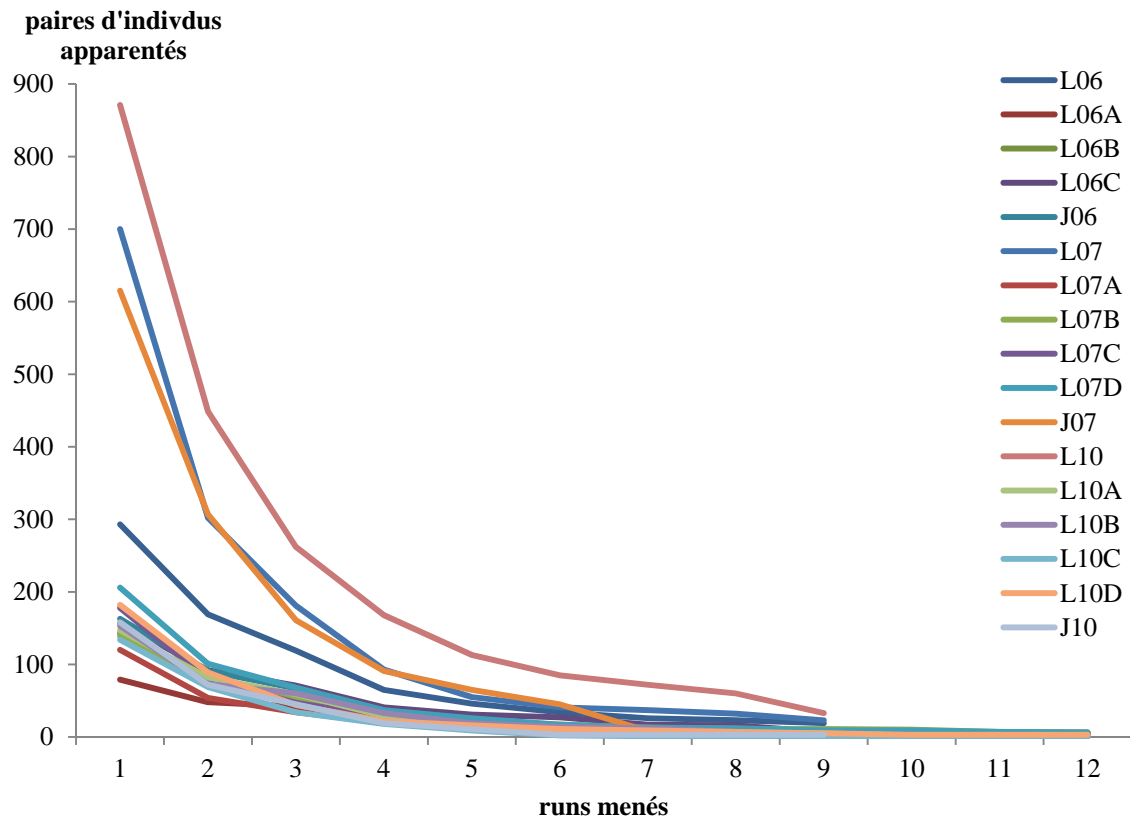


**ANNEXE 12.** Nombres de paires d'individus apparentés identifiées en fonction du nombre de runs réalisés avec le logiciel COLONY (Wang 2004 ; Jones & Wang 2010).

Les noms d'échantillon indiqués dans la légende correspondent à ceux présentés dans le tableau 1 de l'article 4.

Pour ces analyses, une erreur de génotypage de  $10^{-3}$  est appliquée à chaque locus. Les fréquences alléliques sont estimées à partir de l'ensemble des individus adultes échantillonnés dans la baie de Morlaix (2006, 2007 et 2010, cf. matériel et méthode et annexes de l'article 4). Les inférences sont réalisées selon le modèle "full-likelihood" de Wang & Santure (2009), préconisé par les auteurs (voir également Wang 2012).

Ce graphique montre qu'un faible nombre de paires d'individus reconnus comme étant apparentés sont retrouvés entre différents runs. Pour notre étude nous n'avons gardé que les paires qui étaient retrouvées de façon répétées dans 12 runs pour les échantillons larvaires et dans 9 runs pour les juvéniles.

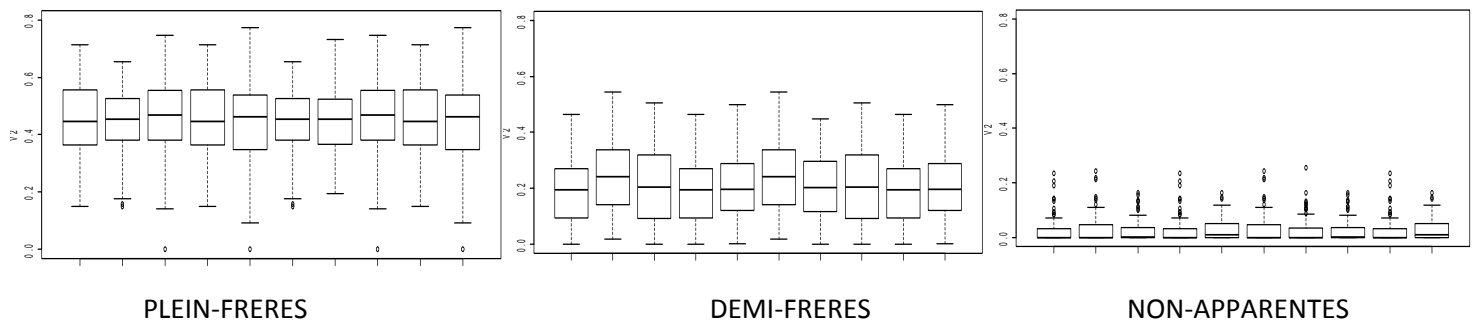


**ANNEXE 13.** Résultats des simulations effectuées avec le logiciel COANCESTRY (Wang 2011) afin de déterminer la distribution des coefficients d'apparentement chez des pleins-frères, demi-frères et individus non apparentés dans la population étudiée dans l'**article 4**.

Les figures résument le résultat de dix simulations permettant de calculer les coefficients d'apparentement de 100 plein-frères, 100 demi-frères et 100 individus non apparentés chacune.

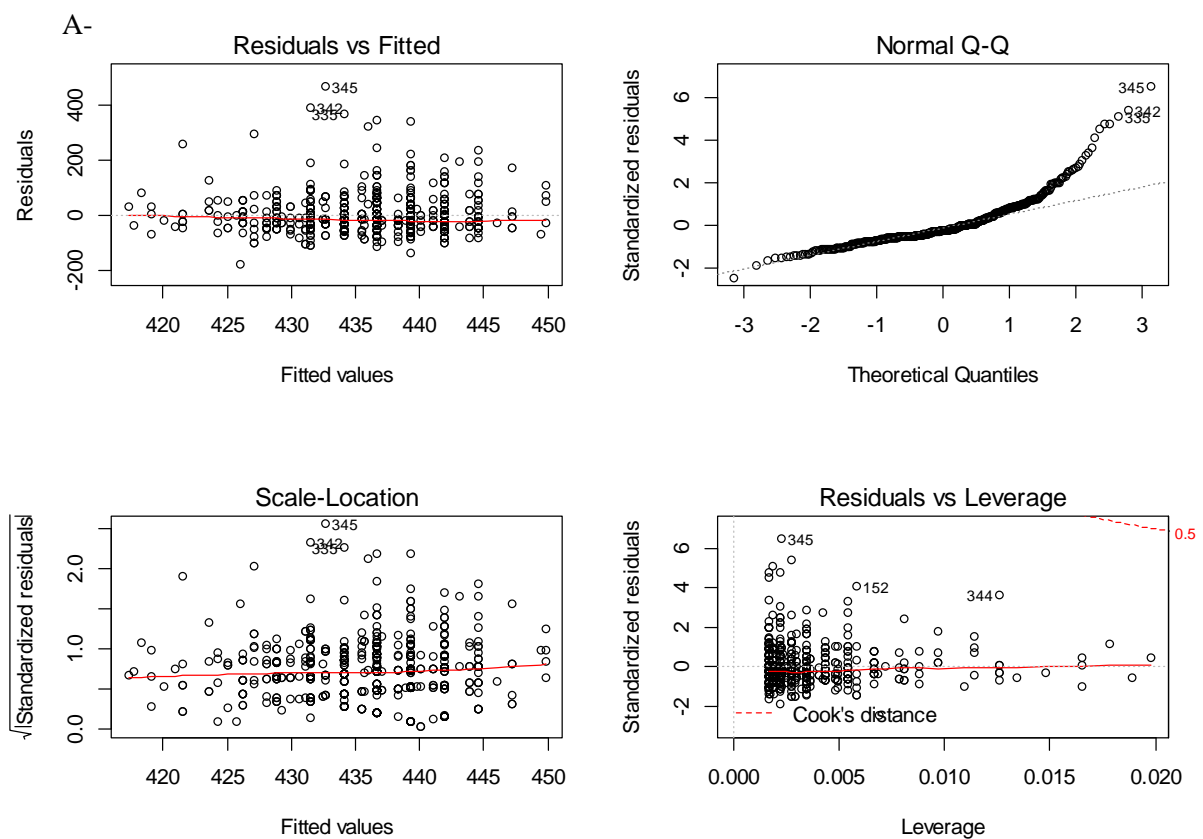
Pour les simulations, une erreur de génotypage de  $10^{-3}$  est appliquée à chaque locus et les fréquences alléliques sont estimées à partir l'ensemble des individus adultes échantillonnés dans la baie de Morlaix (2006, 2007 et 2010, cf. **article 4**).

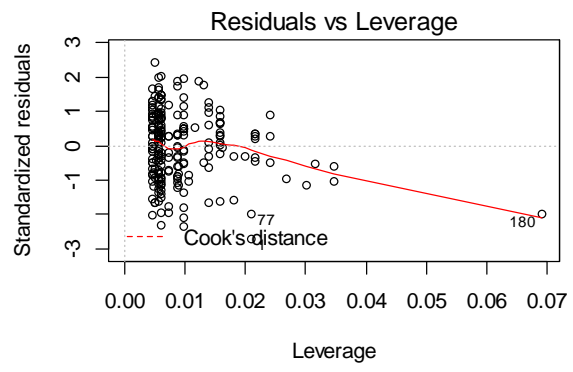
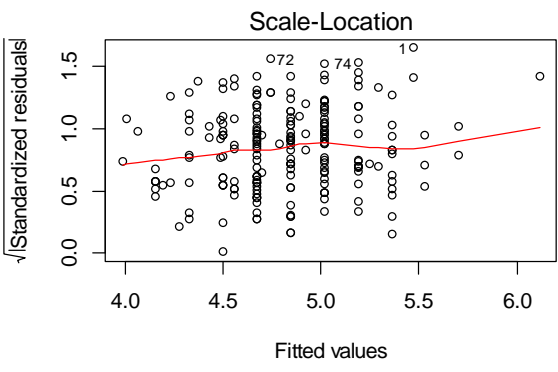
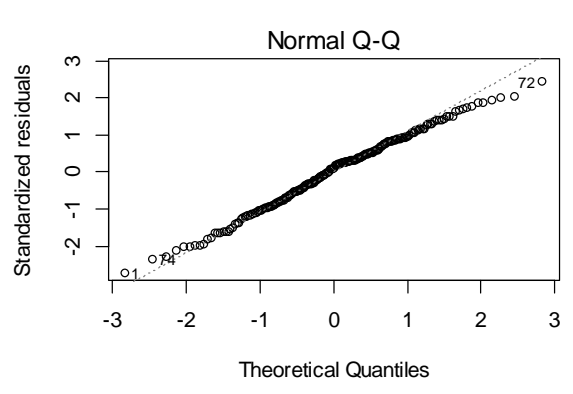
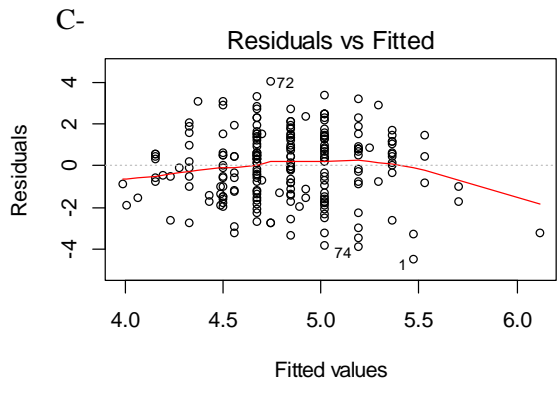
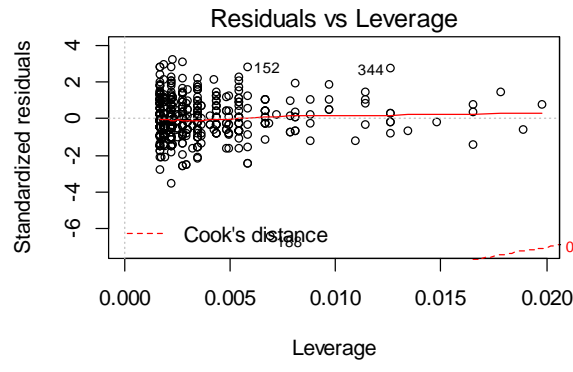
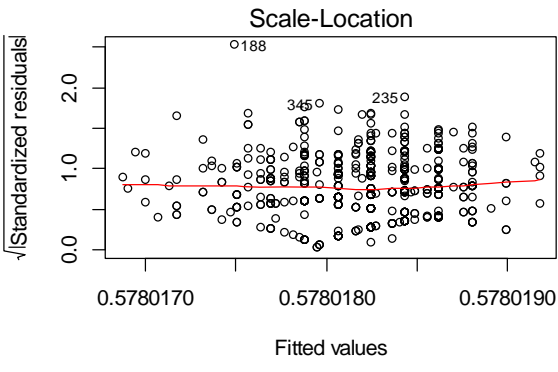
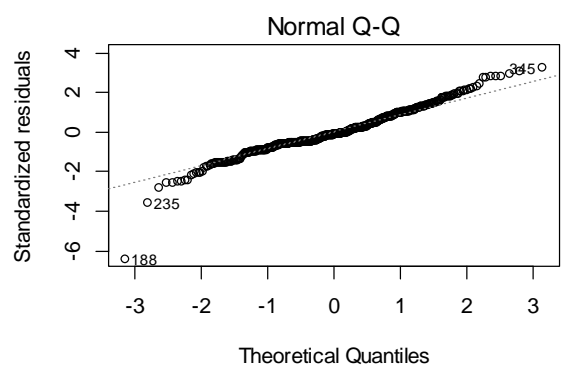
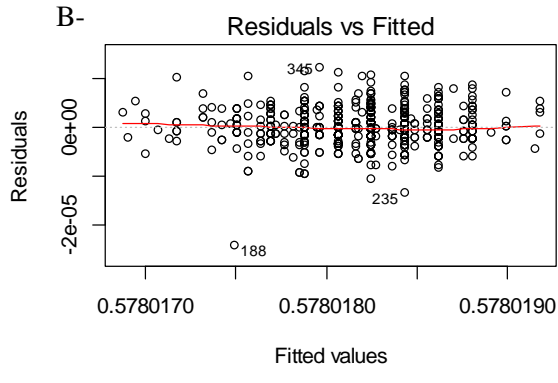
Ces simulations indiquent que les distributions des coefficients d'apparentement sont significativement différents entre les trois catégories (plein-frères, demi-frères et non-apparentés). Au sein de chacune de ces trois catégories, aucune différence significative des distributions des coefficients d'apparentement n'a été identifiée entre simulations.



**ANNEXE 14.** Modèles linéaires utilisés pour l'analyse des échantillons larvaires et juvéniles de l'article 5. Les analyses ont été réalisées avec T. Broquet, sous le logiciel R. Les hypothèses d'application du modèle linéaire sont vérifiées (indépendance des échantillons, normalité des résidus, homogénéité des variances) et si elles ne le sont pas, les données sont transformées avec la fonction boxcox de la library MASS. Le 1<sup>er</sup> graphe représente la distribution des résidus non standardisés, le 2<sup>ème</sup> (en haut à droite) la normalité des résidus (si les points sont proches d'une droite, cela signifie que les observations pourront être modélisées par un modèle normal), le 3<sup>ème</sup> graphe (en bas à gauche) la distribution des résidus standardisés et enfin, le 4<sup>ème</sup> graphe, le levier, *i.e.* si un individu a plus de poids. Le modèle linéaire peut ensuite être testé.

- A- Hypothèses non vérifiées pour les échantillons larvaires
- B- Hypothèses vérifiées pour les échantillons larvaires, après transformation des données
- C- Hypothèses vérifiées pour les échantillons juvéniles





## Résumé de la thèse

Les espèces marines à cycle bento-pélagique présentent des caractéristiques démographiques originales (*i.e.* forte fécondité, grande taille de populations et une phase larvaire dispersive) suggérant des effets de sélection et de dérive génétique limités. Néanmoins, dans le cas des espèces invasives, des changements évolutifs liés aux pressions de sélection dans le nouvel environnement sont attendus. Par ailleurs, il existe des exemples d'organismes marins chez lesquels la taille efficace est réduite par rapport à la taille réelle des populations. Cette thèse a analysé l'importance de ces deux forces évolutives, sélection et dérive génétique, chez le gastéropode *Crepidula fornicata*. Originaire de la côte est de l'Amérique du Nord et introduit en Europe à plusieurs reprises depuis la fin du 19<sup>ème</sup> siècle, il est emblématique des processus d'introduction biologique en milieu marin.

La première partie de cette thèse visait à tester l'hypothèse de processus de sélection ayant agi postérieurement à l'introduction par l'homme de *C. fornicata* en dehors de son aire d'origine. Une approche par scan génomique (344 AFLPs et microsatellites) a été menée sur 22 populations des aires d'origine et d'introduction. Une forte diversité génétique dans les deux types d'aires et une faible structure génétique entre elles indiquent une forte pression en propagules et de multiples introductions de différentes sources. Aucune signature de sélection n'a été identifiée entre aires native et introduites suggérant que les différentiels de sélection, s'ils existent, sont trop faibles pour laisser une trace perceptible par des scans génomiques. En revanche, huit locus « outliers », séparant les populations le long des côtes de la Floride, ont été identifiés, indiquant la probable évolution sur le long terme d'adaptations locales dans l'aire d'origine de la crépidule.

Dans une seconde partie, les effets, à une échelle locale, de la dispersion naturelle et de la dérive génétique ont été analysés afin de tester l'hypothèse de « loterie familiale » (*i.e.* un nombre limité de familles assure le renouvellement complet de la population) chez *Crepidula fornicata*. A partir d'échantillonnages répétés de différents stades du cycle de vie (adultes, larves et juvéniles) sur plusieurs années au sein d'une baie (Morlaix, Bretagne Nord), des analyses ont été menées avec 16 locus microsatellites. Les variations de composition génétique des pools larvaires et des juvéniles suggèrent une variance du succès reproducteurs des adultes. En outre, de façon inattendue, des larves et juvéniles fortement apparentés (plein- ou demi-frères) sont observés. Enfin, les tailles efficaces estimées à partir de l'analyse des juvéniles (échantillonnés durant 9 ans) sont inférieures à celle de la population *in situ*. Ces résultats sont en faveur de l'hypothèse de loterie familiale chez cette espèce.

Ces deux parties de thèse concourent à montrer que sélection et dérive génétique peuvent être, dans certaines conditions, des forces évolutives importantes jouant sur la dynamique éco-évolutive des populations de *Crepidula fornicata*.

Mots- clés : Larves, connectivité, succès reproducteur, adaptation, invasions biologiques, mollusques

## Abstract

Benthic-pelagic marine species display unique demographic characteristics (*i.e.* high fecundity, large population size and a larval phase ensuring dispersion) that suggest limited effects of selection and genetic drift. Evolutionary changes due to new selective pressures are however expected in the introduced range of invasive species. Besides, many marine species exhibit an effective size reduced relative to the census population size. This work aimed to analyze the importance of these two evolutionary forces, namely selection and genetic drift, in the marine gastropod *Crepidula fornicata*. This mollusk, native from the East coast of North America and introduced in Europe several times since the late 19th century, is emblematic of biological invasion processes in marine environment.

First, the hypothesis of selective events following introduction of *C. fornicata* outside its native range was tested. A genome scan approach (344 AFLP and microsatellites markers) was conducted on 22 populations sampled in native and introduced ranges of *C. fornicata*. A large genetic diversity in both ranges and a low genetic structure between them pointed out the occurrence of high propagule pressure and recurrent introductions from several sources. No evidence of selective processes following introduction was shown suggesting that selection strength, if existing, is too weak to leave a noticeable footprint when using genome scan approach. However, eight outlier loci, responsible for a major population divergence along the coasts of Florida, were identified: they are likely to trace back long-term adaptation processes that occurred in the native range of the slipper limpet.

Then, effects of natural dispersal and genetic drift at a local scale were jointly analyzed. The primary objective was here to test for the familial lottery hypothesis (*i.e.* a small number of families ensure the renewal of the population) in *C. fornicata*. Based on repeated samplings of different life cycle stages (adults, larvae, juveniles) over three years within a bay (Morlaix, North Brittany), genetic analyzes were conducted with 16 microsatellite loci. Variations in genetic composition of the larval pool and of juveniles suggested variance in reproductive success among individuals. In addition, related (full and half sibs) larvae and juveniles were observed, a result unexpected in this long dispersive mollusk. Finally, effective size, estimated from temporal analyses of juvenile samples was shown to be smaller than the census size. All these results make plausible the scenario of sweepstake events and familial lottery processes in the study species.

Altogether, these results show that selection and genetic drift may be, under some specific conditions, important drivers of the eco-evolutionary population dynamics of *Crepidula fornicata*.

Keywords: Larvae, connectivity, reproductive success, adaptation, biological invasion, mollusc