

Évolutions comparées de la diversité de caractères quantitatifs, de marqueurs neutres et de gènes d'intérêt dans des (méta)populations expérimentales de blé et d'*Arabidopsis*

Claire LAVIGNE^{(1)*}, Isabelle GOLDRINGER⁽²⁾, Valérie LE CORRE⁽³⁾,
Emmanuelle PORCHER⁽¹⁾, Anne-Laure RAQUIN⁽²⁾,
Xavier REBOUD⁽³⁾, Thomas BATAILLON⁽⁴⁾, Jérôme ENJALBERT⁽⁵⁾

⁽¹⁾ INRA, Unité PSH, Domaine Saint Paul, Site Agroparc,
84914 Avignon Cedex 9, France

⁽²⁾ UMR de Génétique Végétale, INRA UPS INA P-G CNRS, Ferme du
Moulon, 91190 Gif-sur-Yvette, France

⁽³⁾ UMR 1210, INRA, Univ. Bourgogne, ENESAD, « Biologie et Gestion des
Adventices », BP 86510, 21065 Dijon Cedex, France

⁽⁴⁾ UMR 1097 « Diversité et Génomes des Plantes Cultivées »,
INRA Domaine de Melgueil, 34130 Mauguio, France

⁽⁵⁾ INRA, Pathologie végétale et Épidémiologie, 78850 Thiverval-Grignon, France

Abstract: Comparing the evolution of genetic diversity of quantitative traits, neutral markers and quantitative trait loci in experimental (meta)populations of wheat and *Arabidopsis*. We compare the influence of population structure on the genetic diversity of neutral markers, QTLs and traits encoded by these QTLs in predominantly selfing plant populations. To this end, we explored some of the processes governing the evolution of genetic diversity, notably those involved in spatial differentiation. We studied two groups of (meta)populations under experimental evolution. The first group contains seven populations of wheat that each evolved in a different environment during 10 generations. These seven populations belong to two subgroups (of three and four populations), within which initial genetic compositions were identical. The second group contains 12 metapopulations of *Arabidopsis thaliana* (20 populations each, three population sizes) that evolved in the greenhouse during eight generations under homogenous or heterogeneous selection. Evolution of a population depends on its effective population size (N_e); hence, we compared three methods for the estimation of N_e in experimental

* Correspondance et tirés à part : claire.lavigne@avignon.inra.fr

populations of wheat. When N_e is small, the method based on the observed evolution of allelic frequencies at neutral markers is almost as efficient as other methods based on likelihood estimations, and is much easier to implement. We also studied the distribution of changes in allelic frequencies at individual loci over generations to detect regions of the genome under selection in these populations. The method we developed accounts for the rate of false positives. Its application to one of the wheat populations did not reveal any significant selection, despite the observation of large changes in allelic frequency at two loci. In the experimental metapopulations of *A. thaliana*, we explored the impact of selection heterogeneity on the differentiation of microsatellite markers (estimated via F_{ST}) and quantitative traits (estimated via Q_{ST}). We show that Q_{ST} is positively correlated with selection heterogeneity among populations, the latter being estimated via a standard deviation of selection differentials. In addition, F_{ST} and Q_{ST} are strongly correlated, so that a comparison of these two indices of differentiation does not reveal the nonetheless efficient selection. This result suggests that such comparison between F_{ST} and Q_{ST} is a conservative test of detection of selection in selfing species. Finally, results from a simulation study comparing the differentiation at neutral markers, quantitative traits and loci coding for these traits (QTLs), suggest that QTLs might not be as useful as first thought to estimate directly the differentiation of quantitative traits in a metapopulation.

genetic resources/ dynamic management/ selection/ population structure/ effective population size

Résumé : Nos études visent à mieux comprendre comment le fractionnement d'une population en unités plus ou moins isolées peut affecter les relations entre les diversités génétiques des marqueurs neutres, des QTL et des caractères codés par ces QTL, dans des populations végétales autogames. Pour cela, nous nous sommes intéressés aux processus gouvernant l'évolution de ces diversités, et particulièrement leurs structurations spatiales. L'évolution des populations étant partiellement conditionnée par leurs effectifs efficaces, nous avons testé différentes méthodes d'estimation de ces effectifs dans des populations de blé. Nous avons ensuite cherché à détecter les zones du génome impliquées dans l'évolution de ces populations. Les marqueurs moléculaires neutres se sont révélés utiles pour ces deux objectifs. Nous avons ensuite étudié l'impact d'une sélection homogène ou hétérogène sur la structuration de la diversité de marqueurs neutres et de caractères quantitatifs. Ici aussi les marqueurs moléculaires se sont révélés utiles comme estimateurs de la structuration de la diversité génétique des caractères quantitatifs. Enfin par simulations, nous avons étudié la structuration aux locus codant un caractère soumis à sélection. Dans l'hypothèse d'un déterminisme génétique simple, nos modèles indiquent que le suivi de ces locus ne serait pas très informatif, contrairement à ce que nous pensions initialement.

1. INTRODUCTION

La gestion et la conservation des espèces passent par une bonne connaissance des effectifs efficaces des populations et de leur variabilité génétique. La littérature suggère en effet qu'une faible diversité génétique va de pair avec un risque d'extinction plus élevé et une moindre capacité d'adaptation en réponse à une modification du milieu [4]. L'évolution de cette variabilité génétique sous l'effet de la dérive et de la sélection est en grande partie contrôlée par l'effectif efficace des populations [22]. Deux approches sont classiquement utilisées pour évaluer la variabilité au sein des espèces : l'utilisation de marqueurs enzymatiques ou moléculaires d'une part et la génétique quantitative d'autre part. L'utilisation de marqueurs est en général facile à mettre en œuvre et renseigne sur la structure des populations et les grands traits de leur histoire démographique. La seconde approche s'intéresse aux caractères quantitatifs (taille, architecture, rendement ou précocité...), c'est-à-dire au phénotype, cible de la sélection. Elle est la seule susceptible de renseigner directement sur le potentiel adaptatif à courte échéance des populations, mais reste peu utilisée car nécessitant des effectifs importants et des expérimentations longues [19]. De plus, la génétique quantitative donne accès aux composantes de la variance et non aux fréquences des allèles des gènes impliqués. Ceci peut conduire à des difficultés d'interprétation. Par exemple, une héritabilité faible pour un caractère pourrait être due aussi bien à des fréquences très faibles qu'à des fréquences fortes des allèles favorables, c'est-à-dire à une très mauvaise ou très bonne adaptation de la population à son environnement.

Les progrès réalisés dans la dissection génétique des caractères quantitatifs permettent d'envisager une troisième approche qu'on pourra nommer « génomique quantitative », fondée sur l'évaluation de la variabilité génétique aux QTL (Quantitative Trait Loci), gènes intervenant directement dans la variation des caractères quantitatifs. Cependant, des simulations fondées sur un modèle simple (caractère contrôlé par de nombreux gènes à effets faibles et additifs) ont montré une relative absence de corrélation entre variabilité des gènes codant un caractère et variabilité de ce caractère en métapopulation. Ce résultat, *a priori* surprenant, s'explique par le fait que la réponse à la sélection naturelle peut passer par la mise en place de forts

déséquilibres de liaison entre gènes, plutôt que par des changements de fréquences alléliques [8][11].

Dans le cadre de deux expériences de gestion dynamique, nous nous sommes intéressés à l'évolution de ces différents types de variabilités génétiques.

2. MATÉRIELS ET MÉTHODES

2.1. Matériel disponible

Les deux espèces sur lesquelles sont menées les expériences de gestion dynamique sont le blé tendre (*Triticum aestivum* L.) et l'arabette des dames (*Arabidopsis thaliana*).

2.1.1. Gestion dynamique de blé

L'UMR de Génétique végétale du Moulon conduit un programme de gestion dynamique de populations de blé depuis 1984 avec la participation de la DGER du ministère de l'Agriculture. Initialement, deux populations composites autogames (PA et PB) ont été distribuées dans un réseau d'une dizaine de sites couvrant l'aire de culture du blé en France. Chaque population est cultivée et récoltée dans un même site selon les mêmes conditions à chaque génération, certaines depuis 20 générations. 5 000 à 10 000 plantes sont cultivées par population, isolées des autres cultures de blé. La différenciation des populations au cours des générations a été étudiée préalablement à cette étude pour des caractères phénotypiques ([10], [13], [14], [6]) et pour des marqueurs neutres (RFLP, microsatellites : [3] et DEA A.-L. Raquin). Des phénomènes de sélection ont été mis en évidence, soit homogène sur l'ensemble des populations (ex : hauteur, résistance aux rouilles jaunes et brunes), soit hétérogène (ex : précocité de floraison, résistance à l'oïdium). Ainsi, même si la diversité est généralement maintenue sur l'ensemble du réseau, la sélection a un fort impact sur son organisation tant au niveau phénotypique que moléculaire.

2.1.2. Gestion dynamique d'*Arabidopsis*

Dix-huit métapopulations expérimentales d'*Arabidopsis thaliana* (allogamisée par un gène de stérilité mâle récessif nucléaire) sont gérées et évaluées depuis 10 générations à l'INRA de Dijon (UMR « Biologie et gestion des adventices ») et à l'université Paris-Sud (laboratoire ESE,

Orsay). Les métapopulations ont été constituées à partir d'une population initiale issue de croisements entre 15 écotypes. Chaque métapopulation contient 20 populations. À chaque génération, une population est fondée à partir de 10 % de graines issues de plantes mâles-stériles et de 88 % de graines issues de plantes hermaphrodites de la même population, et de 2 % de graines issues d'une autre population de la même métapopulation. Ceci assure un taux d'allofécondation de l'ordre de 10 % [17]. À Orsay, nous étudions l'influence de la taille des populations (10, 25, ou 100 individus) et du régime de sélection sur l'évolution des populations et de leur diversité génétique [9]. La sélection est soit homogène pour un cycle de vie court dans toutes les populations, soit hétérogène, la moitié seulement des populations étant sélectionnée pour ce caractère.

2.2. Méthodes

2.2.1. Marquage moléculaire

Sur le blé, nous avons travaillé à partir de données de marquage obtenues dans différents projets. Dans trois populations PA et trois populations PB, 80 individus par population ont été génotypés aux générations 0 et 10 pour 29 marqueurs RFLP. Dans la population PA-Le Moulon, 200 individus ont été génotypés aux générations 1, 5 et 10 pour les mêmes marqueurs [3] [6]. Sur *Arabidopsis* nous avons utilisé cinq marqueurs microsatellites, situés chacun sur un chromosome. Nous avons génotypé 10 plantes dans 5 populations pour les métapopulations à 25 ou 100 individus par population et 7 plantes dans 6 populations pour les métapopulations à 10 individus par population, à la 8^e génération [15]. Nous avons estimé la structuration de la variabilité génétique pour ces marqueurs (F_{is} et F_{st}) avec le logiciel *Fstat*

2.2.2. Génétique quantitative et des populations

2.2.2.1. Blé

Les variations temporelles de fréquences alléliques aux 29 marqueurs RFLP entre les générations 0 et 10 pour 6 populations et entre les générations 1, 5 et 10 pour une population ont été utilisées pour estimer N_e , l'effectif efficace sur la période de temps considérée. Trois méthodes d'estimation, qui reposent sur l'hypothèse de populations fermées, ont été testées : (i) l'estimateur de la méthode des moments fondé sur la statistique F (cf équation (1), d'après [21]), (ii) un estimateur du maximum de vraisemblance calculé par simulations Monte Carlo

(programme MCLEEPS), [1], (iii) un estimateur du maximum de vraisemblance basé sur une approche par coalescence (programme TM3 <http://www.rubic.rdg.ac.uk/~mab/software.html>, [2]).

L'estimation de N_e par la méthode des moments repose sur la prise en compte des variations temporelles observées de fréquences alléliques à chaque locus l dans le cadre de la théorie des variances standardisées :

$$\hat{F}_{c,l} = \frac{1}{K_l} \sum_{k=1}^{K_l} \frac{(p_{ki,l} - p_{kf,l})^2}{(p_{ki,l} + p_{kf,l})/2 - p_{ki,l} p_{kf,l}} \quad (\text{selon [12]}), \text{ avec } p_{ki,l} \text{ (resp. } p_{fi,l}) \text{ la}$$

fréquence de l'allele k à la génération i (resp. f) au locus l , et K_l le nombre d'alleles au locus l . L'estimation multilocus F_c est obtenue par une moyenne pondérée des valeurs individuelles : $\hat{F}_c = \sum K_l \hat{F}_{c,l} / \sum K_l$ [20]. L'estimateur de N_e est alors :

$$\hat{N}_e = \frac{t}{2(\hat{F}_c - 1/2S_i - 1/2S_f)}$$

avec S_i (resp. S_f) le nombre d'individus échantillonnés à la génération i (resp. f), et t le nombre de générations entre i et f . Pour une population autogame, le nombre de gènes échantillonnés est S au lieu de $2S$ à chaque génération.

Ce calcul repose sur l'hypothèse que les fréquences alléliques évoluent à la même vitesse à tous les locus. Dans la réalité, certains locus, bien que neutres, peuvent varier plus ou moins vite que d'autres, car ils sont situés dans des zones du génome fortement soumises à sélection. Nous avons développé une méthode fondée sur des simulations d'échantillonnage multinomial des allèles à chaque génération à partir des fréquences dans l'échantillon initial, pour tester si les variations temporelles des fréquences des locus étudiés ($F_{c,l}$) sont homogènes ou si certains locus ont une variation extrême sortant de la distribution attendue (simulée) [5]. Cette méthode est utilisée pour identifier des locus qui seraient éventuellement indirectement soumis à sélection.

2.2.2.2. *Arabidopsis*

Dans chacune des 12 métapopulations d'Orsay, nous avons échantillonné 5 individus par population dans 6 populations (dont les populations utilisées pour les marqueurs microsatellites) aux générations 2 et 8 (soit 720 individus) pour l'analyse des caractères quantitatifs. Onze caractères ont été mesurés sur 4 individus de chacune des 720 familles formées par autofécondation des individus échantillonnés ainsi que sur les 15 lignées parentales. Nous avons

Évolutions comparées de diversités génétiques en métapopulations

estimé l'intensité de la sélection dans les métapopulations par des différentielles de sélection [7] en utilisant le nombre de siliques comme indice de la valeur sélective. La variance génétique pour les caractères quantitatifs a été estimée pour chaque population à partir de la variance inter-famille et du déficit en hétérozygotes. La structuration de la variabilité génétique quantitative a été estimée par le $Q_{ST} = (1 + F_{IS}) V_{BP} / [(1 + F_{IS}) V_{BP} + 2V_{WP}]$, où V_{BP} est la variance entre populations et V_{WP} la variance intra-population pour le caractère [22], [15].

2.2.3. Simulations informatiques

Les simulations ont été effectuées avec des versions modifiées du programme *Metapop*, initialement développé au laboratoire ESE [11]. Ce programme simule l'évolution de métapopulations dont les populations sont spatialisées. Les individus ne sont décrits que par leur génotype mais ne sont pas spatialement explicites. Nous nous sommes intéressés (1) aux relations entre la diversité génétique de marqueurs neutres, de QTL et des caractères codés par ces QTL et (2) et à la structuration spatiale de ces diversités génétiques pour différents scénarios évolutifs (migration forte ou faible, sélection homogène ou hétérogène, régime de reproduction autogame ou allogame) dans des métapopulations à l'équilibre [11]. Par ailleurs, nous avons simulé l'évolution des métapopulations d'*A. thaliana* afin d'interpréter les patrons d'évolution observés (voir les résultats ci-après) et en particulier de comprendre le rôle des forts déséquilibres de liaison existant dans les populations F_2 initiales [15]. Dans les deux cas, le caractère quantitatif était codé par un nombre variable de QTL physiquement indépendants et à effet additif sur le caractère et la sélection était plus ou moins hétérogène [11], [15].

3. RÉSULTATS

3.1. Approche méthodologique de la différenciation des (méta)populations expérimentales

3.1.1. Différenciation temporelle

Les N_e estimés sur les 6 populations de blé entre les générations 0-10 et 1-5-10 varient entre 37 et 216. Les estimations obtenues avec les trois méthodes sont très cohérentes entre elles et permettent d'identifier trois populations de $N_e < 60$ et trois autres de $N_e > 100$. C'est principalement pour les N_e supérieurs à 150 que les deux méthodes par maximum de vraisemblance [1],

[2] s'écartent le plus de l'estimateur fondé sur le F-statistique [21] (- 30 %) et que ces méthodes présentent un gain de précision. Notons que le classement des populations selon leur N_e est conservé d'une méthode à l'autre. Par ailleurs, nous avons montré que la précision sur N_e calculé à partir du F-statistique peut-être améliorée en utilisant la distribution des F_c obtenue par simulation d'échantillonnage multinomial sur plusieurs générations, plutôt qu'en utilisant l'approximation classique du Chi2 [5]. Nous avons ensuite testé l'homogénéité des variations de fréquences alléliques individuelles sur l'ensemble des locus. La méthode que nous avons développée tient compte des tests multiples (plusieurs locus sont testés simultanément) et permet de contrôler le taux de faux positifs parmi les locus identifiés comme significatifs [18]. Dans la population PA-Le Moulon, l'application de cette méthode aux F_c individuels des 29 marqueurs RFLP entre les générations 1-10 ne permet pas d'identifier de façon certaine des locus dont les variations seraient plus rapides que l'évolution moyenne du génome, même si certaines semblent plus élevées (fig. 1a). En effet, si l'on retient les locus Fba242-C et Fba280-C (dont les p -values sont les plus faibles, fig. 1b) comme « candidats sous sélection », nous avons un taux de faux positifs d'environ 40 % parmi ces deux locus [5]. L'utilisation de cette méthode de détection de sélection est en cours pour les autres populations, en particulier au voisinage de gènes *a priori* sélectionnés (p. ex. gène de nanisme *Rht1*, thèse A.-L. Raquin).

3.1.2. Différenciation spatiale

Sur *Arabidopsis*, le calcul des différentiels de sélection montre que tous les caractères mesurés (sauf la hauteur du premier fruit) étaient bien sous sélection directionnelle dans les métapopulations sélectionnées pour la précocité. Dans les métapopulations sous sélection hétérogène, six caractères étaient effectivement sous sélection hétérogène, les autres caractères étant, en fait, soumis à une sélection uniforme (fig. 2). Dans chaque métapopulation, nous avons quantifié le niveau d'hétérogénéité de la sélection pour chaque caractère en calculant l'écart-type entre son différentiel de sélection moyen dans les trois populations sous sélection pour la précocité et son différentiel moyen dans les trois autres populations. Le niveau d'hétérogénéité dans les métapopulations en sélection homogène a été estimé en échantillonnant deux groupes de trois populations au hasard [16].

Évolutions comparées de diversités génétiques en métapopulations

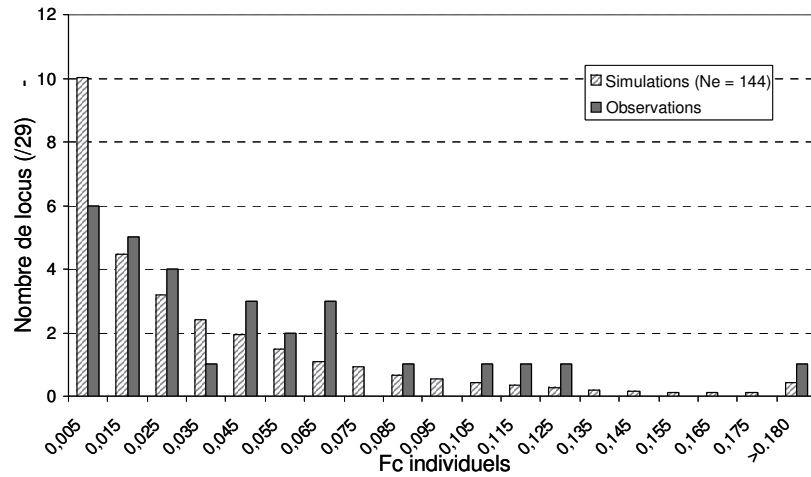


Figure 1 : a) Distributions observée et simulée des F_c individuels (F_c = variance temporelle standardisée des fréquences alléliques à chaque locus) dans la population PA-Moulon : $N_e = 144$, évolution entre les générations 1 et 10. La distribution simulée résulte du cumul des distributions individuelles pour chacun des 29 locus. Elle est présentée à titre indicatif.

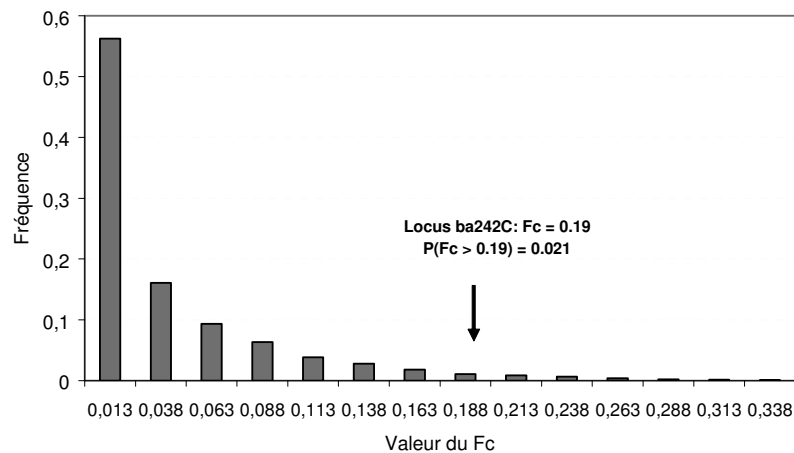


Figure 1 : b) Distribution des F_c individuels estimés à partir de simulations d'une population représentant PA-Moulon entre les générations 1 et 10 ($N_e = 144$, $S_1 = 168$, $S_{10} = 214$, fréquences alléliques initiales au locus Fba242C = {0,19, 0,81}).

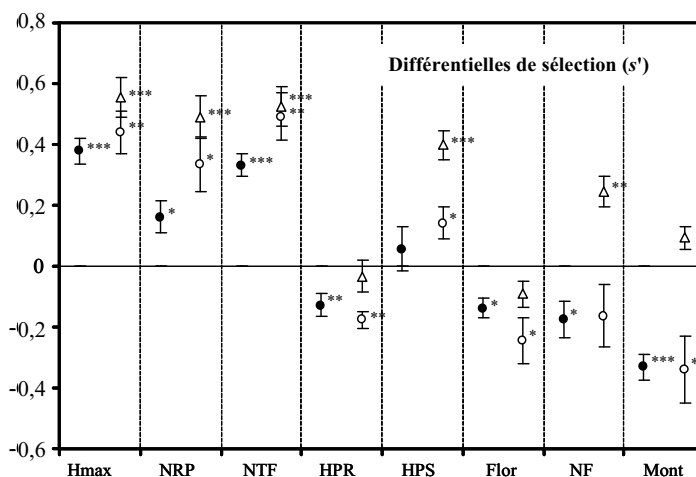


Figure 2 : Différentiels de sélection standardisés à la génération 2. Les valeurs moyennes, calculées sur les 6 métapopulations suivant un même régime de sélection, sont figurées pour les métapopulations en sélection homogène (cercles noirs) et les métapopulations en sélection hétérogène (symboles blancs), pour lesquelles on a distingué les populations sous sélection pour un cycle de vie court (cercles blancs) de celles sans sélection artificielle (triangles blancs). Les barres d'erreur indiquent les erreurs standard sur les 6 métapopulations et les étoiles indiquent le niveau de signification des moyennes (test t): *** : $P < 0,001$; ** : $0,001 < P < 0,01$; * : $0,01 < P < 0,05$. Hmax = Hauteur maximale; NRP = Nombre de ramifications primaires; NTF = Nombre de têtes de floraison; HPR = Hauteur de la première ramification; HPS = Hauteur de la première silique; Flor = Date de floraison; NF = Nombre de feuilles de la rosette; Mont = Date de montaison) [16].

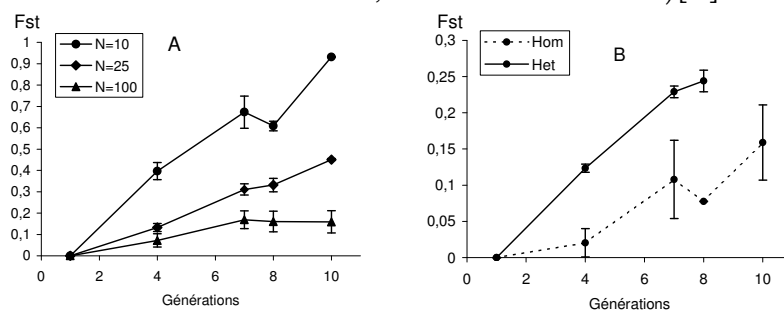


Figure 3 A et B : Différenciation observée pour les marqueurs et les caractères quantitatifs. Évolution de la différenciation pour les marqueurs neutres pendant 10 générations, en fonction de la taille des populations (A) et du régime de sélection dans les métapopulations avec 100 plantes par population (B). Les valeurs aux générations 4, 7 et 10 sont estimées à partir des marqueurs enzymatiques, celles de la génération 8 à partir des marqueurs microsatellites. Les données enzymatiques manquent à la génération 10 pour les métapopulations sous sélection hétérogène. Pour cette génération, les moyennes par taille de population sont donc calculées à partir des seules métapopulations en sélection homogène.

Évolutions comparées de diversités génétiques en métapopulations

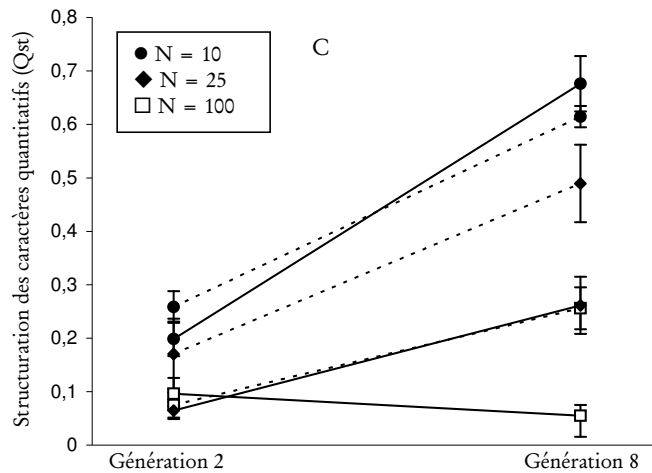


Figure 3 C : Différenciation observée pour les marqueurs et les caractères quantitatifs. Évolution de la différenciation moyenne pour les caractères quantitatifs, sous sélection homogène (traits continus) et hétérogène (pointillés), pour les trois tailles de population. Les barres d'erreur correspondent aux erreurs standard, calculées sur l'ensemble des caractères et les deux métapopulations d'un même traitement. L'évolution est quasi identique pour les métapopulations à 25 individus sous sélection homogène et celles à 100 individus sous sélection hétérogène (courbes superposées).

Dans les populations de grande taille ($N = 100$), l'hétérogénéité de la sélection a une influence directe sur la structuration des caractères quantitatifs : on observe une corrélation significative positive entre les valeurs du Q_{st} à la génération 8 et l'hétérogénéité de la sélection. Pour les autres tailles de population, la dérive semble prédominer et les valeurs des Q_{st} ne sont plus corrélées à celle de l'hétérogénéité de la sélection [16].

Connaissant l'impact de la sélection sur la structuration des caractères quantitatifs, nous avons ensuite comparé la structuration des caractères quantitatifs (Q_{st}) à celle des marqueurs neutres (F_{st}). Plusieurs approches théoriques prédisent que ces deux indices de structuration sont identiques dans le cas d'un caractère neutre [22]. Pour un caractère soumis à sélection, on attend $Q_{st} < F_{st}$ si la sélection est uniforme et $Q_{st} > F_{st}$ si elle est hétérogène. La comparaison du F_{st} au Q_{st} est donc susceptible de renseigner sur l'hétérogénéité de la sélection agissant sur le caractère, et notre approche expérimentale permet de tester la validité de cette méthode de détection de la sélection couramment employée. Les évolutions des indices de différenciation sont présentés sur la figure 3. Contrairement aux prédictions, nous n'avons pas pu mettre en évidence de différence significative entre F_{st} et Q_{st} , y compris dans les métapopulations de

grande taille [15], bien que la structuration des caractères soit sous le contrôle de la sélection.

3.2. Approche par modélisation de la différenciation : comparaison caractères quantitatifs, QTL et marqueurs neutres

Les simulations indiquent que dans des populations allogames à l'équilibre, la structuration des marqueurs neutres est généralement peu influencée par le régime de sélection. En revanche, cette structuration est sensible à la force de la sélection chez les autogames : plus la sélection est intense, plus les marqueurs neutres sont différenciés, que la sélection soit homogène ou hétérogène. Nous retrouvons ce même résultat dès les premières générations dans les simulations mimant les métapopulations d'*A. thaliana* en l'absence de déséquilibres de liaison initiaux entre les locus neutres et les QTL. En revanche, quand l'ensemble du génome est soumis à de forts déséquilibres de liaisons initiaux, les F_{st} des marqueurs neutres sont contrôlés par l'hétérogénéité de la sélection, et non par son intensité, avec en particulier des valeurs plus faibles en sélection uniforme que dans le cas neutre. Dans les deux cas, il existe une corrélation significative entre F_{st} aux QTL, aux marqueurs neutres et valeurs du Q_{st} dans les populations à l'équilibre. Cependant, sous sélection diversifiante, la différenciation des QTL est intermédiaire entre le F_{st} et le Q_{st} alors qu'elle est la plus élevée lorsque la sélection est uniforme (si la migration n'est pas trop forte).

Globalement, les résultats à l'équilibre indiquent que la comparaison du Q_{st} et du F_{st} devrait généralement permettre de détecter une sélection hétérogène lorsque la migration efficace entre population est suffisante et en l'absence de déséquilibres de liaison initiaux. En revanche, ce test de détection de la sélection ne semble pas fonctionner dans le cas d'espèces autogames soumises à de forts déséquilibres de liaison sur l'ensemble du génome, au moins lorsque l'évolution porte sur peu de générations.

4. DISCUSSION

Les processus évolutifs majoritairement à l'œuvre dans les populations dépendent fortement des effectifs efficaces des populations. Pour notre jeu de données, les estimations obtenues avec trois méthodes d'estimation d'effectifs efficaces [1], [2], [21] sont très cohérentes entre elles. Les deux méthodes reposant sur un estimateur par maximum de vraisemblance sont

nettement plus lourdes à mettre en œuvre et ne sont pas beaucoup plus précises pour des N_e inférieurs à 150, il semble donc que les gestionnaires de populations puissent se contenter dans un premier temps de l'estimateur fondé sur la statistique F [21] en calculant éventuellement l'intervalle de confiance à partir des simulations d'échantillonnage [5]. Pour les valeurs estimées élevées, une des deux autres méthodes pourra alors être mise en œuvre. Par ailleurs, nous notons à nouveau ici que les N_e génétiques estimés pour les populations de blé étudiées, bien inférieurs aux effectifs démographiques que l'on peut calculer à partir du nombre de plantes ($\approx 2\ 600$), ne peuvent s'expliquer que par une sélection intense ([6], travaux en cours) suggérant que les phénomènes de sélection jouent un rôle important dans l'évolution de la diversité des populations conservées *in situ* ou en gestion dynamique [5]. Les méthodes proposées permettent d'identifier des locus potentiellement sous sélection parmi des marqueurs neutres. Les gestionnaires peuvent ainsi obtenir une estimation plus fiable de l'effectif efficace et de la structuration spatiale en les écartant et, de plus, ils ont ainsi accès à des zones du génome potentiellement impliquées dans l'adaptation d'une population à son milieu.

L'utilisation des marqueurs semble également possible pour estimer la structuration de la diversité génétique de caractères phénotypiques soumis à sélection chez les espèces autogames. En effet, les indices de structuration F_{st} et Q_{st} sont fortement corrélés à la fois dans nos simulations et dans les métapopulations expérimentales. Les simulations indiquent que ceci est en grande partie dû à l'impact de la sélection sur l'évolution de la diversité de ces marqueurs, pourtant neutres. Cet impact s'explique ici par des déséquilibres de liaisons initiaux assez élevés, déséquilibres qui sont maintenus au cours des générations du fait du faible taux d'allogamie [15]. De ce fait, le test de détection de la sélection par la comparaison des F_{st} et des Q_{st} n'est pas efficace dans nos métapopulations. De manière générale, des simulations préliminaires suggèrent que ce test serait donc conservatif dans des populations autogames qui maintiennent de forts déséquilibres de liaison, c'est-à-dire qu'il y a une forte chance de ne pas y détecter une sélection hétérogène.

Enfin, le suivi de la structuration des populations directement aux QTL ne semble pas pouvoir se substituer à l'étude directe de la structuration des caractères et semble même parfois induire en erreur quant à cette structuration. Ce comportement surprenant est vraisemblablement dû au fait que la différenciation aux QTL

dépendrait plus de l'intensité de la sélection que de son hétérogénéité alors que le Q_{st} dépend à la fois de l'intensité et de l'hétérogénéité de la sélection. Cependant, les avancées récentes en physiologie et génomique montrent que l'architecture génétique des caractères est autrement plus complexe que ne le supposent les modèles classiques utilisés ici. Ceux-ci pécheraient avant tout en sous-estimant de manière systématique la part de la variation génétique dépendant d'interactions entre loci. Disposer de modèles plus réalistes semble une étape décisive.

REMERCIEMENTS

Nous remercions M. LEFRANC, C. SAUNOIS, T. GIRAUD, T. BOUTIN, V. PELLETIER, L. SAUNOIS et G. FÉLIX pour leur aide précieuse, P.-H. GOUYON et B. GODELLE pour les idées à l'origine des métapopulations.

RÉFÉRENCES

- [1] Anderson E.C., Williamson E.G., Thompson E.A., Monte Carlo evaluation of the likelihood for $N(e)$ from temporally spaced samples, *Genetics* 156 (2000) 2109-2118.
- [2] Berthier P., Beaumont M.A., Cornuet J.M., Luikart G., Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach, *Genetics* 160 (2002) 741-751.
- [3] Enjalbert J., Goldringer I., Paillard S., Brabant P., Molecular markers to study genetic drift and selection in wheat populations, *J Exp Bot* 50 (1999) 283-290.
- [4] Frankel O.H., Soulé M.E., Conservation and evolution, Cambridge University Press, 1981.
- [5] Goldringer I., Bataillon T., On the distribution of temporal variations in allelic frequency: consequences for the estimation of effective population size and the detection of loci undergoing selection, *Genetics* 168 (2004) 563-568.
- [6] Goldringer I., Enjalbert J., Raquin A.-L., Brabant P., Strong selection in wheat populations during ten generations of dynamic management, *Genet. Sel. Evol.* 33 (2001) 441-463.
- [7] Lande R., Arnold S.J., The measurement of selection on correlated characters, *Evolution* 37 (1983) 1210-1226.
- [8] Latta R.G., Gene flow, adaptive divergence and comparative population structure among loci, *New Phytol.* 161 (2003) 51-58.

Évolutions comparées de diversités génétiques en métapopulations

- [9] Lavigne C., Reboud X., Lefranc M., Porcher E., Roux F., Olivieri I., Godelle B., Evolution of genetic diversity in metapopulations: *Arabidopsis thaliana* as an experimental model, *Genet. Sel. Evol.* 33 (2001) S399-423.
- [10] Le Boulc'h V., David J.L., Brabant P., de Vallavieille-Pope C., Dynamic conservation of variability: responses of wheat populations to different selective pressures including powdery mildew, *Genet. Sel. Evol.* 26 (1994) 221s-240s.
- [11] Le Corre V., Kremer A., Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection, *Genetics* 164 (2003) 1205-1219.
- [12] Nei M., Tajima F., Genetic drift and effective population size, *Genetics* 98 (1981) 625-640.
- [13] Paillard S., Goldringer I., Enjalbert J., Doussinault G., de Vallavieille-Pope C., Brabant P., Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. I- Is specific seedling resistance selected?, *Theor. Appl. Genet.* 101 (2000a) 449-456.
- [14] Paillard S., Goldringer I., Enjalbert J., Trottet M., David J., de Vallavieille-Pope C., Brabant P., Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. II- Adult plant resistance, *Theor. Appl. Genet.* 101 (2000b) 457-462.
- [15] Porcher E., Evolution de la diversité génétique neutre et sélectionnée en métapopulation, Thèse de l'université Paris-XI (2002).
- [16] Porcher E., Giraud T., Goldringer I., Lavigne C., Experimental demonstration of a causal relationship between heterogeneity of selection and genetic differentiation at quantitative traits, *Evolution* 58 (2004) 1434-1445.
- [17] Porcher E., Gouyon P-H., Lavigne C., Dynamic management of genetic resources: maintenance of outcrossing in experimental metapopulations of a predominantly inbreeding species, *Conservation Genetics* 5 (2004) 259-269.
- [18] Storey J.D., Tibshirani R., Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* 100 (2003) 9440-9445.
- [19] Storfer A., Quantitative genetics: a promising approach for the assessment of genetic variation in endangered species, *Trends Ecol. Evol.* 11 (1996) 343-348.
- [20] Tajima F., Nei M., Note on genetic drift and effective population size, *Genetics* 106 (106) 569-574.
- [21] Waples R.S., A generalized approach for estimating effective population size from temporal changes in allele frequency, *Genetics* 121 (1989) 379-391.
- [22] Wright S., Evolution and the genetics of populations, vol. 2. The theory of gene frequencies, The University of Chicago Press. Chicago, 1969.