UNIVERSITE MONTPELLIER II

SCIENCES ET TECHNIQUES DU LANGUEDOC

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER II

Discipline : Biologie des Organismes et des Populations
Formation Doctorale : Biologie de l'Evolution et Ecologie
Ecole Doctorale : SIBAGHE

Fardeau de mutation, fardeau de dérive et fardeau de migration dans des populations fragmentées de plantes: approches théoriques

présentée et soutenue publiquement

par

Sébastien Lopez

Le 14 décembre 2007

Jury

Denis COUVET	Professeur MNHN Paris	Rapporteur
Bernard GODELLE	Professeur, Université de Montpellier II	Examinateur
Isabelle GOLDRINGER	Chargée de Recherche, INRA Gif-sur-Yvette	Membre Invité
Frédéric HOSPITAL	Directeur de Recherche, INRA Jouy-en-Josas	Rapporteur
Ophélie RONCE	Chargée de Recherche, CNRS Montpellier	Co-Directrice
François ROUSSET	Directeur de Recherche, CNRS Montpellier	Directeur
Denis ROZE	Chargé de Recherche, CNRS Roscoff	Examinateur

SOMMAIRE

INTRODUCTION GÉNÉRALE

Introduction

Le terme de fardeau génétique a été utilisé pour la première fois par Muller (1950) pour décrire comment la vigueur moyenne des individus dans une population décroît du fait de l'occurrence récurrente de mutations délétères spontanées. De façon plus générale, on peut voir ce fardeau comme le prix associé au maintien de la variabilité génétique, condition nécessaire de l'évolution des populations (Crow et Kimura 1970). Toutes les forces qui modèlent l'évolution de la diversité génétique au sein des populations affectent donc le fardeau. On parle ainsi de fardeau de mutation, de recombinaison, de ségrégation, de migration, ou de dérive. Si la connotation négative du terme de fardeau suggère clairement que l'on s'inquiète des conséquences démographiques putatives de ce déclin de la valeur sélective pour la persistance des populations, il ne faut pas oublier que l'autre face du fardeau génétique, la diversité, conditionne l'adaptation à un environnement changeant. Ces deux questions, celle de l'évolution de la valeur sélective et de la diversité génétique, sont donc au coeur des problématiques concernant l'adaptation et la persistence des populations dans le contexte des changements globaux.

Au cours de ma thèse, je me suis intéressé aux conséquences de la fragmentation de l'habitat pour l'évolution du fardeau génétique. Plus spécifiquement, j'ai modélisé les effets à long et court terme d'une réduction de la migration des graines et du pollen dans des populations de plantes à distribution discontinue. Je me suis particulièrement focalisé sur le cas de plantes exploitant des habitats hétérogènes, pour lesquelles la migration avait potentiellement des effets ambigus sur le fardeau génétique, en compromettant l'évolution de l'adaptation locale tout en s'opposant à l'effet de la dérive. Afin de replacer mes travaux dans un contexte plus large, je donnerai dans un premier temps quelques exemples de populations fragmentées de plantes, je ferai une brève revue de l'étendue actuelle des phénomènes de fragmentation des paysages, de leurs principales conséquences démographiques, de ce que l'on sait ou que l'on peut prédire à propos de l'évolution du fardeau génétique dans ces populations fragmentées. En particulier, je discuterai les choix méthodologiques des études théoriques ayant modélisé l'évolution du fardeau. Enfin, j'expliquerai comment les modèles développés pendant ma thèse prolongent les études théoriques précédentes.

I La fragmentation des populations

De nombreuses espèces de plantes présentent une aire de répartition discontinue, avec des sous-populations, généralement appelées dèmes, partiellement reliées entre elles par des flux migratoires. Parmi ces populations fragmentées, on peut faire la distinction entre celles pour qui ce type d'organisation spatiale est naturel et celles pour lesquelles il est d'origine anthropique. Cette distinction est importante dans la mesure où les populations naturellement fragmentées vivent dans ces conditions depuis de nombreuses générations. On peut alors supposer qu'elles sont adaptées à ce mode de fonctionnement, ce qui n'est probablement pas le cas pour les populations récemment fragmentées suite à l'action de l'homme.

1) Les populations naturellement fragmentées

La répartition spatiale des plantes dépend de nombreux facteurs. Une discontinuité spatiale de tels facteurs peut entraîner une structuration des populations en sous-populations occupant des poches d'habitat favorable et séparées par un ensemble de zones défavorables souvent appelé *matrice*. Dans la majorité des cas, les facteurs abiotiques conditionnent la niche fondamentale de l'espèce (Hutchinson 1957), c'est-à-dire l'ensemble des zones où l'espèce pourrait se perpétuer au cours du temps en l'absence d'autres espèces. Pour les plantes, les plus importants de ces facteurs abiotiques sont le climat et les caractéristiques du sol, qui ensemble conditionnent l'accès aux ressources. Une espèce inféodée aux mares possède une niche fondamentale naturellement discontinue et, par conséquent, des populations spatialement fragmentées. La Renoncule à noeuds fleuris (*Ranunculus nodiflorus*) vit dans un système de ce type, constitué de mares plus ou moins temporaires. On peut la rencontrer en forêt de Fontainebleau, où la présence de dalles de grès (les platières) empêche par endroits l'eau de s'écouler dans le sous-sol. La génétique de cette espèce a été étudiée pour comprendre le rôle de la migration entre ces mares temporaires (Noël et al 2006).

Les interactions de l'espèce avec son environnement ne se limitent cependant pas aux facteurs abiotiques: la présence d'autres êtres vivants peut également influencer la répartition spatiale d'une espèce. Compétition, parasitisme, prédation, mutualisme, peuvent ainsi empêcher une espèce d'être présente dans certaines parties de son "aire de répartition fondamentale" en la privant de ressources pourtant présentes. La compétition pour la lumière constitue par exemple

une des limitations les plus courantes. La dynamique forestière illustre très bien ce problème d'accès à la lumière: en forêt, l'apparition récurrente de chablis dus aux chutes d'arbres entraîne la formation de clairières, qui sont colonisées par des espèces friandes de lumière. Cependant, le milieu se referme progressivement au cours des successions végétales et les clairières disparaissent, tandis que d'autres se forment ailleurs, formant ainsi des populations discontinues pour les plantes inféodées à ce type d'habitats, comme l'espèce tropicale pionnière *Cecropia obtusifolia* (Alvarez-Buylla et Garcia-Barrios 1991). Contrairement aux espèces forestières évoquées ci-dessus, la répartition spatiale de la Centaurée de la Clape (*Centaurea corymbosa*), inféodée aux milieux ouverts, varie très peu au cours du temps, car son habitat favorable se situe au sommet de falaises calcaires, dont l'évolution est beaucoup plus lente que celle des chablis en forêt.

Enfin, indépendamment de l'environnement des populations, la répartition géographique d'une espèce peut également dépendre de la capacité de l'espèce à coloniser de nouvelles zones. On peut encore citer l'exemple de la Centaurée de la Clape, dont l'aire de répartition discontinue est aussi expliquée par sa très faible propension à s'installer dans de nouveaux sites (Colas et al 1997). Des raisons historiques peuvent également expliquer la fragmentation des populations à grande échelle. Par exemple, lors du dernier épisode glaciaire, les espèces d'arbres de climat tempéré ont été repoussées dans des zones-refuges discontinues (Bennett et al 1991, Cheddadi et al 2006). Outre ces phénomènes naturels, l'action de l'homme est également responsable de la fragmentation de nombreuses espèces en populations plus ou moins isolées.

2) La fragmentation d'origine anthropique

La destruction d'habitat constitue une des causes majeures de la fragmentation des populations. Si l'on considère une population continue, sa destruction progressive conduit à la formation de sous-populations isolées les unes des autres (Bogaert et al 2004) et donc à des populations fragmentées. L'ampleur locale des modifications des paysages dépend fortement de la valeur agricole des régions considérées (Baldi et al 2006, Madriñán et al 2007) ainsi que de la dynamique de colonisation par l'homme liée aux moyens de transport (par exemple les rivières, Madriñán et al 2007). La destruction des habitats est un processus extrêmement rapide: au Chili, les surfaces de forêts tempérée ont décru de 67% en 25 ans (Echeverria et al 2006), tandis que les forêts alluviales matures de l'est de la Colombie ont perdu 70% de leur surface en moins de

60 ans (Madriñán et al 2007). Au Mexique, la surface couverte par les forêts tropicales de montagne dans l'état du Chiapas a diminué de 50% entre les années 1975 et 2000. La vitesse de progression des fronts de déforestation est également très élevée (par exemple, 0.84 km/an pour Etter et al 2006). D'un point de vue spatial et temporel, la destruction des habitats a des conséquences complexes, qui peuvent dépendre de la proportion d'habitat détruit. Le nombre de fragments d'habitat persistant dans le paysage est une fonction non monotone de la proportion d'habitat détruit, avec un accroissement lorsque la destruction d'habitat est limitée, puis une diminution à partir d'un certain seuil de destruction (Echeverria et al 2006 par exemple). La destruction d'habitat entraîne une modification de la distribution de la taille des fragments. Ainsi, le nombre de fragments de petite taille augmente (Echeverria et al 2006), tandis que la taille des fragments se distribue de manière plus uniforme (Cayuela et al 2006), à cause du morcellement progressif des très grands fragments. Finalement, l'action de l'homme peut aussi accentuer les effets de la fragmentation des populations, en altérant les flux de gènes entre fragments. Par exemple, la construction et l'extension des réseaux de transport (ferroviaire ou routier), qui entraînent une destruction d'habitat en général assez réduite, peuvent quand même avoir des conséquences biologiques fortes, liées à une diminution des flux migratoires (voir par exemple une revue sur les amphibiens Cushman 2006). Chez les plantes, des changements d'abondance des vecteurs de dispersion du pollen et des graines, ou de leur comportement en réponse à des modifications de la matrice, ou encore l'évolution des traits favorisant la dispersion (structures de dispersion sur les graines, attractivité pour les pollinisateurs), sont susceptibles également d'affecter les flux de gènes entre fragments.

3) Fragmentation et migration

Pour résumer, la fragmentation d'une population en tant qu'état se caractérise par le fait que les individus se développent et se reproduisent au sein de groupes de taille limitée avec peu d'échanges entre des groupes différents. La fragmentation en tant que processus dynamique décrit des changements simultanés (ou non) dans la taille des fragments, leur nombre et leur isolement. Les conséquences de la fragmentation ont beaucoup été étudiées sous l'angle de celles de la taille réduite des populations locales. Dans ma thèse, je me suis plutôt focalisé sur les conséquences de l'isolement des populations et de l'intensité des flux migratoires, en cherchant à comprendre leurs interactions avec la taille des populations (voir Ouborg et al. 2006

pour une revue mettant en avant cette autre dimension de la génétique de la conservation des plantes).

Chez les plantes, la migration peut se produire par le biais des graines ou du pollen. Bien que certains événements puissent se produire sur des distances relativement grandes (voir par exemple l'étude empirique de Bacles et al (2006) sur le frêne), la migration des graines est un phénomène généralement limité dans l'espace. On suppose souvent que le pollen peut franchir de plus grandes distances que les graines, en raison notamment de sa petite taille (en particulier pour les espèces anémophiles). Cependant, des études récentes tendent à montrer que la distance des événements de dispersion par le pollen est elle aussi limitée (Koenig et Ashley 2003). Chez la Centaurée de la Clape par exemple, la distance de dispersion maximale des graines est de 168 centimètres (Colas et al 1997), tandis que seulement 20 % des événements de reproduction impliquent des partenaires éloignés de plus de 43 mètres (Hardy et al 2004). Les valeurs de F_{ST} entre paires de populations (calculés sur des marqueurs allozymes, Colas et al 1997) sont comprises entre 0.05 et 0.5, la plupart étant supérieures à 0.35 et correspondant à des populations très différenciées échangeant très rarement des gènes. De façon générale, les distances accrues entre fragments d'habitat peuvent donc limiter à la fois les flux de graines et de pollen et j'ai étudié les conséquences spécifiques à la perturbation de chaque mode de dispersion.

4) Modèles théoriques de populations fragmentées

Il existe deux principales classes de modèles décrivant les populations structurées génétiquement: les modèles spatialement continus et les modèles discrets, ces derniers étant utilisés pour étudier les populations fragmentées. Parmi les modèles discrets, deux types principaux sont très utilisés. Il s'agit des modèles d'isolement par la distance (comme les modèles en stepping-stone) et du modèle en îles de Wright (Wright 1969). Le modèle en îles de Wright considère des dèmes connectés entre eux par des flux migratoires symétriques et aléatoires entre toutes les paires de dèmes. Ce modèle est très utile car, parmi les modèles permettant d'étudier les conséquences d'une dispersion limitée, il est le moins compliqué à analyser. De plus, les prédictions de ce modèle se révèlent relativement robustes lorsque les taux de dispersion envisagés sont réduits, indépendamment de la distribution des distances de dispersion (Rousset 2004). Ce modèle est donc une alternative intéressante aux modèles plus

complexes d'isolement par la distance. C'est ce modèle que j'ai utilisé durant ma thèse, bien qu'il souffre de certaines faiblesses vis-à-vis de son réalisme biologique: l'hypothèse de migration aléatoire est une hypothèse très forte puisque, dans les populations naturelles, la migration se produit souvent à des distances limitées (voir plus haut) et en suivant des routes préférentielles (vents dominants pour les plantes anémochores, rivières pour les plantes hydrochores, voir par exemple Liu et al 2006). En outre, lorsque la dispersion fait appel à des animaux, on s'attend à ce que les événements de migration se produisent en nombre ou pas du tout, plutôt que de manière homogène dans le temps et l'espace.

II Conséquences de la fragmentation des populations

1) Conséquences démographiques directes

L'augmentation de la proportion d'habitat périphérique liée à un événement de fragmentation (voir par exemple Cayuela et al 2006) entraîne une augmentation de l'impact des effets de bord sur la démographie des espèces. De tels effets de bord peuvent se manifester par des changements concernant des phases cruciales du cycle de vie. Chez les espèces de sous-bois par exemple, la survie des juvéniles peut être négativement affectée par l'accroissement des températures et la diminution d'hygrométrie observés en lisière (Ohara et al 2006). En dehors de ces effets de bord, la fragmentation se manifeste en premier lieu par le fait que les individus se reproduisent au sein de populations de taille réduite. Par exemple, seulement un tiers des populations de *Globularia bisnagarica* étudiées par Honnay et ses collaborateurs (2007) comprennent plus de 250 individus. Dans ces conditions, les populations locales sont davantage exposées à la stochasticité démographique ou à d'autres phénomènes tels que les effets Allee. La méta-analyse d'Aguilar et al (2006) a montré que les effets négatifs de la fragmentation sur la reproduction des plantes semble s'expliquer principalement par une diminution de l'efficacité de la pollinisation, particulièrement importante pour les espèces auto-incompatibles. Un isolement accru des populations dans ce contexte a également des conséquences démographiques négatives, en empêchant les effets de rescousse démographique de la migration (Brown et Kodric-Brown 1977).

2) Changement des pressions de sélection

La fragmentation des populations peut modifier les pressions de sélection s'exerçant sur certains traits phénotypiques. Par exemple, la réduction des tailles de population associée à la fragmentation peut conduire à la perte d'allèles d'autoincompatibilité. La proportion de croisements compatibles au sein des populations est alors très souvent diminuée (voir par exemple Fischer et al 2003, Willi et al 2005 pour des études expérimentales sur les plantes *Cochlearia bavarica* et *Ranunculus reptans*), ce qui peut favoriser la sélection de variants autocompatibles (Reinartz et Les 1994). La fragmentation des habitats peut entraîner l'émergence de fortes pressions de sélection contre les comportements de dispersion si celle-ci devient plus coûteuse. Ceci est d'autant plus vrai pour les espèces dont la dispersion est passive, comme les plantes. Ce type de pressions de sélection pourrait expliquer par exemple les très faibles taux de colonisation observés chez la centaurée de la Clape (Colas et al 1997). Un dernier exemple concerne les structures florales. Dans l'étude de Hooftman et collaborateurs (2003), les individus de petites populations isolées de *Succisa pratensis* investissent davantage dans les structures reproductrices (plus d'inflorescences et plus de fleurs par inflorescence) que les individus vivant dans les grandes populations. Les auteurs expliquent ce résultat inattendu par la faible abondance de pollinisateurs dans les petites populations qui aurait sélectionné un investissement accru dans les structures attractives.

3) Evolution du fardeau génétique

a) Fardeau de mutation et fardeau de dérive

L'apparition récurrente de mutations spontanées ayant des effets délétères sur la valeur sélective, quel que soit l'environnement, a été identifiée très tôt comme la source principale de fardeau génétique. Le fardeau génétique est alors défini comme la diminution de valeur sélective moyenne dans la population d'intérêt par rapport à une population qui ne présenterait pas de mutations délétères. Plusieurs études théoriques récentes ont montré comment la fragmentation des populations pouvait affecter la dynamique de ce fardeau (e.g. Whitlock 2002, Glémin et al 2003, Roze et Rousset 2003, Roze et Rousset 2004, Theodorou et Couvet 2006a, Bouchy et al 2005). En particulier, les paramètres responsables du niveau de structuration des

populations fragmentées (nombre et taille des dèmes, patron et intensité de la migration) ont une forte influence sur la dynamique des mutations inconditionnellement délétères. Cependant, leur importance relative est variable. Ainsi, l'étude théorique de Theodorou et Couvet (2006a) a montré que la taille des dèmes et le taux de migration ont en général les plus forts effets. La plupart des études s'intéressant aux conséquences de la fragmentation des populations sur la dynamique des mutations inconditionnellement délétères étudient des populations à l'équilibre: les conséquences génétiques de la fragmentation des populations en tant qu'état sont donc assez bien connues, tandis que celles d'un épisode de fragmentation récent (fragmentation en tant que processus) sont pour la plupart inférées à partir des résultats obtenus sur les populations à l'équilibre (voir cependant Theodorou et Couvet 2006a,b et Bouchy et al 2005 pour des études théoriques suivant la dynamique du fardeau après un épisode de fragmentation).

On a vu précédemment que la fragmentation des populations a principalement deux conséquences directes: une diminution de la taille locale des populations et une diminution de la migration entre ces populations. La diminution de la migration entraîne un changement du rapport de force entre mutation, sélection et dérive. Le fardeau à l'équilibre dû a des mutations délétères codominantes augmente toujours quand la migration diminue du fait de l'accélération de la dérive. Pour des mutations récessives, une plus forte homozygotie liée à une diminution de la migration permet de contre-sélectionner plus efficacement les mutations. Il peut donc y avoir une diminution du fardeau à l'équilibre quand le taux de migration diminue, même si l'effet précédemment noté l'emporte toujours aux plus bas taux de migration. Lorsque la migration est très faible, on prédit donc que le fardeau génétique est en général augmenté par rapport à celui d'une grande population non structurée (Roze et Rousset 2004), quelle que soit la récessivité des mutations.

La taille efficace locale restreinte a également des conséquences distinctes sur la dynamique d'approche à l'équilibre des mutations selon qu'elles sont très récessives avec des effets forts à l'état homozygote, ou seulement partiellement récessives et d'effets faibles. Dans le premier cas, l'augmentation de l'homozygotie localement permet aux mutations récessives qui ségrégeaient à des fréquences fortes dans la population avant la fragmentation d'être démasquées. Cette augmentation du fardeau est seulement transitoire, car les mutations sont contre-sélectionnées au fur et à mesure qu'elles s'expriment (c'est le phénomène de purge génétique). Les mutations récessives d'effets forts sont susceptibles d'être purgées, tandis que les mutations d'effets plus faibles peuvent se fixer (voir Glémin 2003 pour des prédictions

théoriques, Paland et Schmid 2003 pour une étude expérimentale). Cependant, l'efficacité de la purge semble en pratique limitée notamment lorsque les populations sont de petite taille (voir Glémin 2003 ou Theodorou et Couvet 2006b pour des études théoriques et Byers et Waller 1999 pour une méta-analyse d'études expérimentales).

Le corollaire est que l'on s'attend à ce que les petites populations soient caractérisées par un fardeau de fixation plus grand mais une dépression de consanguinité plus faible (dûe aux mutations délétères récessives qui ségrègent dans la population) que les grandes populations (Bataillon et Kirkpatrick 2000). Ces prédictions théoriques sont en accord avec le résultat d'une étude expérimentale récente (Paland et Schmid 2003) qui donne une décomposition du fardeau génétique dû aux mutations récessives en composantes de dérive et de dépression de consanguinité. Cette étude montre que les populations plus petites que deux cents individus présentent un fardeau surtout dû à la fixation d'allèles délétères d'effets faibles (mais que les allèles d'effets très forts ont été purgés par la sélection) tandis que le fardeau génétique, dans les populations de plus grande taille, est davantage dû à des mutations non fixées responsables d'un certain niveau de dépression de consanguinité.

Le phénomène d'hétérosis, c'est-à-dire la plus grande valeur sélective des produits de croisements entre populations plutôt que intrapopulation, peut s'expliquer par le fait que certaines mutations sont masquées chez les hybrides lorsque différentes populations ont fixé des mutations délétères récessives à différents locus (pour des prédictions théoriques voir Glémin et al. 2003). La migration peut alors avoir un effet de rescousse génétique (voir l'étude sur *Silene alba*, Richards 2000, et Ingvarsson 2001, Tallmon et al. 2004 pour une discussion plus générale), d'autant plus fort que les populations sont petites et isolées (Willi & Fischer 2005).

b) Fardeau de migration

Les effets des mutations spontanées ne sont cependant pas toujours constants quel que soit l'environnement (e.g. Remold et Lenski 2004). L'existence de mutants montrant des effets sur la valeur sélective variables en fonction de l'environnement (par exemple délétères dans certains habitats, et neutres ou avantageux dans d'autres) est une condition requise pour voir émerger des patrons d'adaptation locale dans des paysages hétérogènes. On considère généralement qu'il existe un patron d'adaptation locale chez l'espèce étudiée lorsque, dans les différentes sous-populations, les individus issus de lignées locales réalisent de meilleures

performances que les individus en provenance d'autres zones (critère du "chez soi mieux que les autres"). De tels patrons peuvent être mis en évidence en comparant les performances d'individus issus de localités différentes dans chacune de ces localités d'origine (Kawecki et Ebert 2004). L'émergence de ces patrons lorsque la sélection naturelle varie dans l'espace nécessite également que les populations puissent être différenciées génétiquement, donc des flux de gènes restreints.

Des patrons d'adaptation locale ont été abondamment documentés chez les plantes. Ils peuvent par exemple être liés à une hétérogénéité climatique (Etterson et Shaw 2001, Rehfeldt et al 2002, Etterson 2004a et b, Savolainen et al 2004, Goldringer et al 2006) ou une hétérogénéité des sols (liée par exemple à une pollution par des métaux lourds: McNeilly 1968, Antonovics et Bradshaw 1970, MacNair 1987, Jimenez-Ambriz et al 2007). L'adaptation locale peut évoluer à des échelles spatiales très diverses: l'Impatience du Cap (*Impatiens capensis*) présente un patron d'adaptation locale entre des individus distants de quelques mètres seulement (Schmitt et Gamble 1990) tandis que, de manière moins surprenante, d'autres espèces montrent de tels patrons à l'échelle de continents entiers (Joshi et al 2001; pour des revues, voir Linhart et Grant 1996 ou Hufford et Mazer 2003).

Lorsque les flux migratoires augmentent, la composition génétique des populations s'homogénéise. De ce fait, on s'attend à ce qu'un patron d'adaptation locale ne puisse se maintenir que si les échanges entre populations sont limités et ce même si les pressions de sélection varient à travers l'espace. Lorsqu'un tel patron est présent, une augmentation de la migration est alors responsable d'une diminution de la valeur sélective des individus liée à une maladaptation moyenne croissante (voir par exemple Bolnick et Nosil 2007 pour une étude théorique et expérimentale sur le phasme *Timema cristinae)*. Cette maladaptation est de plus en plus couramment qualifiée de *fardeau de migration* (Hu et Li 2003, Hu 2006, Bolnick et Nosil 2007). Cependant, l'effet de la migration sur le niveau d'adaptation locale n'est pas toujours négatif. En effet, en même temps qu'elle éloigne les populations de l'optimum local par l'introduction d'individus maladaptés, la migration permet, dans une certaine messure, de convertir la variabilité génétique observée entre populations en variabilité génétique à l'échelle locale. Ainsi, tant que les populations restent différenciées, la migration peut agir localement comme une source de variabilité génétique (voir modèle de Phillips 1996), d'autant plus que l'environnement est spatialement hétérogène (Yeaman et Jarvis 2006). Cependant, à partir d'un certain seuil de migration, la différenciation entre les populations diminue (Lythgoe 1997). Si le

maintien d'une variabilité génétique peut être une source de maladaptation lorsque les populations sont bien adaptées aux conditions locales, celle-ci peut cependant se révéler avoir des effets positifs pour des populations génétiquement érodées et exposées à des conditions environnementales stressantes en bordure d'aire de répartition. Ainsi, des études théoriques ont montré que la migration peut favoriser l'évolution de niche, c'est-à-dire l'adaptation à de nouvelles conditions environnementales (Holt et al 2003) ou que de faibles taux de migration peuvent maximiser la valeur sélective moyenne sur l'ensemble de l'aire de répartition en contrecarrant l'effet délétère de la dérive dans les populations périphériques maladaptées (Alleaume-benharira et al 2006).

Si l'on commence à bien connaître le comportement des patrons de variabilité génétique et du fardeau de migration à l'équilibre en réponse à des variations de l'intensité des flux de gènes et de l'hétérogénéité de l'environnement, l'influence du mode de dispersion dans les populations fragmentées n'a en revanche pas du tout été étudiée pour un trait quantitatif (voir cependant Hu et Li 2001 ou Hu et Li 2003 pour des modèles monolocus), de même que la réponse dynamique des patrons d'adaptation locale à un événement de fragmentation. Dans ce dernier cas cependant, on peut s'attendre à ce que les conséquences d'un épisode récent de fragmentation soient globalement positives pour le niveau d'adaptation locale (à l'exception des cas où la dérive génétique serait très forte, comme expliqué ci-dessus). C'est en tout cas le résultat documenté par une étude récente (Goldringer et al 2006). À partir d'une souche homogène de blé tendre (*Triticum æstivum)*, plusieurs populations de grande taille ont été cultivées dans des environnements variables, sans que la migration soit possible entre les populations. Dans une certaine mesure, cette expérience mime l'effet de la fragmentation des habitats: une population génétiquement homogène, pas particulièrement adaptée à son environnement, est libérée de tout flux génétique entrant. Le résultat est très clair: en dix générations, les populations se sont différenciées en réponse à une pression de sélection sur les caractères liés à la résistance au froid hivernal et à la précocité des individus. En absence de contrôle avec migration entre localités, il est cependant difficile de savoir si la fragmentation (liée à une diminution des flux génétiques) a accéléré l'adaptation rapide aux pressions de sélection locale, ou pas.

c) Retour critique sur les modèles d'évolution du fardeau

Les études portant sur la relation entre le génotype et la valeur sélective des individus dans les populations structurées se répartissent en deux grands groupes: celles qui étudient cette relation sous l'hypothèse d'une sélection homogène dans l'espace et celles qui considèrent le cas d'une sélection hétérogène. En pratique, on s'aperçoit que le premier type d'études utilise très généralement un formalisme reliant directement le génotype à la valeur sélective des individus (voir encadré 1), tandis que la deuxième approche est le plus souvent basée sur des modèles de génétique quantitative, qui relient le génotype à la valeur sélective de manière indirecte par l'intermédiaire de la valeur d'un ou plusieurs traits phénotypiques (encadré 2). Si le choix de ce formalisme est en grande partie arbitraire, il n'est pas nécessairement sans implications sur les conclusions obtenues par ces modèles. J'ai au cours de ma thèse utilisé les deux types de formalisme, essentiellement pour des raisons de convenance de comparaison avec la littérature théorique précédente. On discutera en conclusion comment ces décisions de modélisation ont pu affecter les résultats obtenus. L'approche de génétique quantitative implique également une perspective différente sur le fardeau génétique dans le sens où l'attention est portée non plus seulement sur la valeur sélective mais sur la variabilité génétique pour des traits phénotypiques particuliers. Elle met donc en particulier l'accent sur l'évolution de la variabilité génétique comme une grandeur d'intérêt en soi, ce qui n'est pas vraiment le cas dans les approches qui relient directement la valeur sélective au génotype. Cet intérêt porté à des traits particuliers a également plus de sens dans des situations biologiques où on a des hypothèses précises à propos de la sélection sur ces traits en fonction du contexte écologique. Le fait que ces études théoriques se focalisent sur un petit nombre de caractères potentiellement sous sélection divergente implique néanmoins qu'elles peuvent décrire mal la dynamique du fardeau total. Finalement, la divergence entre ces deux formalismes a peut-être conduit au fait que fardeau de mutation et de dérive d'une part et fardeau de migration d'autre part ont jusqu'à présent été étudiés séparément alors que les données empiriques suggèrent de façon répétée que les deux phénomènes sont simultanément à l'oeuvre dans les populations naturelles.

Une grande variété d'approches théoriques sont utilisées pour étudier les conséquences génétiques de la fragmentation des populations. Cette diversité est dûe à deux raisons principales, la première liée à la façon de relier génotype et valeur sélective (encadrés 1 et 2), la seconde dépendant des méthodes de calcul employées et des hypothèses impliquées.

Encadré 1: *Relier directement le génotype à la valeur sélective*

Une première approche théorique utilisée est de considérer que les mutations ont un effet direct sur la valeur sélective des individus qui les portent. Pour exprimer cette relation directe, on suppose que les mutations sont associées à un coefficient de sélection *s* et un coefficient de dominance *h*. Lorsqu'un individu est homozygote pour une mutation à un locus donné, sa valeur sélective diminue d'un facteur *1-s*, tandis qu'elle diminue d'un facteur *1-hs* lorsque la mutation est présente à l'état hétérozygote (exemples d'études monolocus: Roze et Rousset 2003, Whitlock 2002, Theodorou et Couvet 2006a, Glémin et al 2003). Lorsqu'on s'intéresse à l'évolution de plusieurs locus, on néglige généralement les interactions entre eux en considérant un simple effet multiplicatif des effets de tous les locus sur la valeur sélective (exemples d'études multilocus: Couvet 2002, Bouchy et al 2005). De plus, les effets des mutations sont généralement considérés comme constants pour tous les locus. Afin d'introduire davantage de réalisme dans les modèles, certaines études prennent cependant en compte une variation de la valeur du coefficient de sélection entre les locus (par exemple, distribution des coefficients de sélection suivant une loi gamma: Theodorou et Couvet 2006b). Nous verrons dans le texte principal que les valeurs de *s* et de *h* ont une importance cruciale pour les résultats de ces modèles. Par conséquent, la calibration de ces paramètres est très importante.

Un sujet de débat en cours et pas du tout encore résolu est la proportion de mutations délétères parmi toutes les mutations apparaissant à chaque génération. Deux écoles se sont développées. D'un côté, les partisans de l'idée que la grande majorité des mutations sont délétères et qu'elles surviennent en grand nombre à chaque génération (voir par exemple Lynch et al 1999 pour une revue ou Schultz et al 1999 pour une étude sur *Arabidopsis thaliana*) et, de l'autre, ceux qui pensent que les mutations délétères ne sont pas plus fréquentes que les mutations avantageuses et qui, par conséquent, postulent des taux de mutation et des effets mutationnels moins élevés (pour un débat musclé, voir par exemple Shaw et al 2002, Keightley et Lynch 2003, Shaw et al 2003, Bataillon 2003). Ma conviction personnelle tendrait plutôt du côté des "modérés", ne serait-ce que pour la simple raison que, si la situation était vraiment si grave, nous ne serions sans doute plus là pour en parler (Lynch et al 1995, Higgins et Lynch 2001).

La relation entre le coefficient de sélection *s* et le coefficient de dominance *h* est également importante pour ces modèles théoriques. Basées sur l'observation que la plupart des mutations délétères sont plus ou moins récessives, deux théories ont été avancées (historique dans Charlesworth 1998). La théorie de Fisher (1928), reposant sur l'idée que les mutations sont initialement codominantes et deviennent progressivement

encadré 1 (suite)

récessives suite à la sélection de modificateurs de dominance et la théorie de Wright, basée sur l'idée que les réseaux métaboliques ont une marge de sécurité permettant une diminution importante de la quantité d'enzyme fonctionnel sans perte notable dans la vitesse de réaction (figure 1). Cette dernière théorie a reçu des appuis de plus en plus significatifs, si bien qu'elle est considérée comme la raison principale de la dominance des allèles sauvages sur les allèles mutés (voir cependant Bourguet 1999). Cette théorie explique la dominance pour des gènes codant des enzymes, mais il semblerait que la corrélation entre *s* et *h* s'observe pour une large gamme de types de gènes (Phadnis et Fry 2005). Par conséquent, les études théoriques considèrent souvent deux classes extrêmes de mutations: les mutations codominantes et d'effet faible ou les mutations récessives d'effet fort, dont les réponses à la fragmentation des populations sont très différentes (voir texte principal).

Enfin, il s'avère que les modèles de ce type sont peu utilisés pour étudier le phénomène d'adaptation locale (voir cependant Kawecki 2000 par exemple): ils sont en général utilisés pour modéliser la dynamique des mutations dont les effets ne varient pas dans l'espace.

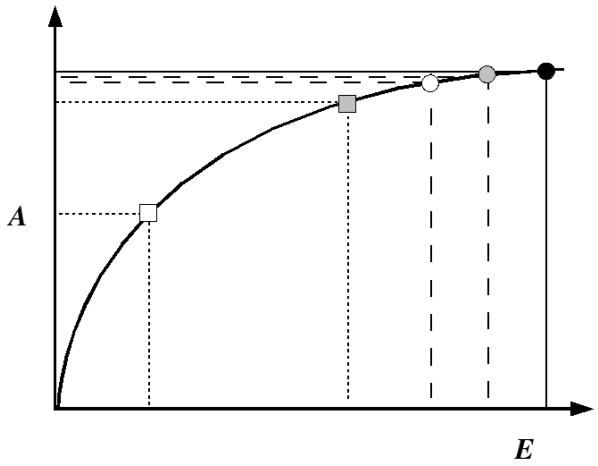


Figure 1: Relation entre quantité d'enzyme (*E*) et activité métabolique (*A*) d'une réaction enzymatique: la perte d'activité liée à une mutation d'effet faible à l'état hétérozygote (comparaison rond noir-rond gris) est proche de la moitié de celle dûe à cette même mutation à l'état homozygote (comparaison rond noir-rond blanc), tandis que la diminution d'activité dûe à une mutation d'effet fort à l'état hétérozygote (comparaison rond noir-carré gris) est beaucoup plus petite que celle dûe à la même mutation à l'état homozygote (comparaison rond noir-carré blanc).

Il est en général difficile d'obtenir des prédictions analytiques pour décrire l'effet des flux de gènes sur l'évolution de la variance génétique de traits polygéniques soumis à une sélection stabilisante, alors que la valeur moyenne du trait est relativement aisée à formuler lorsque la variance génétique est connue (Barton 1999). La variance génétique dépend de plusieurs composantes: la variance génique, qui mesure la variabilité présente si l'on considère la variation des fréquences des allèles pris isolément, la variance dûe au déséquilibre de liaison gamétique, causée par l'association préférentielle de certains allèles à différents locus entre eux, et la variance dûe aux associations entre allèles portés par des gamètes différents causées par les écarts à la panmixie (voir chapitre 1). Deux hypothèses extrêmes sont souvent utilisées, qui conditionnent les contributions relatives de chacune de ces composantes à la variance génétique totale. Le modèle infinitésimal suppose que la variance génique à l'équilibre de liaison est constante et que la plupart des changements de variance génétique sont dûs au déséquilibre de liaison (Barton 1999, Tufto 2000). Une autre approche consiste à négliger complètement la contribution du déséquilibre de liaison à la variance génétique totale (Barton 1999). La pertinence de ces deux approches en fonction des hypothèses biologiques est discutée dans le chapitre 1.

Pour le formalisme de type *h-s* (encadré 1), différentes méthodes de calcul sont utilisées, les plus courantes reposant sur l'utilisation d'équations de diffusion (Cherry et Wakeley 2003, Roze et Rousset 2003, Whitlock 2003). Ces méthodes effectuent des approximations ramenant les processus stochastiques discrets responsables de la dynamique des allèles délétères à des processus continus de diffusion. Le calcul de l'espérance et de la variance du changement des fréquences alléliques sur un temps très court permettent de connaître la densité de probabilité de la fréquence des allèles dans les populations à l'équilibre. Une autre méthode repose sur l'utilisation de matrices de transition (Couvet 2002, Bouchy et al 2005).

De manière générale, ces méthodes sont complexes d'un point de vue mathématique et reposent sur des hypothèses biologiques fortes, vis-à-vis de l'architecture génétique des traits quantitatifs par exemple. Dans de nombreux cas, le recours à des simulations est une alternative intéressante, pour tester les prédictions des modèles ou étudier des problèmes trop complexes pour être traités analytiquement. C'est ce type d'approche que j'ai utilisé dans ma thèse (simulations individu-centrées), bien qu'il implique d'autres problèmes méthodologiques. En premier lieu, de telles simulations ne permettent pas d'explorer l'espace des paramètres aussi efficacement que les modèles analytiques. De plus, les méthodes de simulations demandent

Encadré 2: *Relier le génotype à la valeur sélective par l'intermédiaire du phénotype*

De nombreux travaux étudient l'évolution des populations fragmentées par une approche de génétique quantitative, qui suppose qu'un grand nombre de locus codent la valeur d'un trait phénotypique de l'individu et que la sélection n'agit pas sur les allèles directement, mais sur le phénotype de l'individu. Le plus souvent, une variable aléatoire mimant les effets environnementaux sur le développement des individus est ajoutée à la valeur génétique du trait pour obtenir la valeur de phénotype de l'individu. Ce type de modèle est en général utilisé pour étudier les phénomènes d'adaptation. La valeur sélective dépend alors le plus souvent de la distance entre la valeur du phénotype et une valeur optimale, généralement sous la forme d'une fonction Gaussienne. Dans les études utilisant ce formalisme, les variables d'intérêt sont en général la moyenne de la valeur génotypique du trait ainsi que la variance génétique mesurée à l'intérieur des populations.

Souvent les interactions entre allèles (pour des paires d'allèles à un même locus ou entre locus) sont négligées car la valeur génétique du trait d'un individu est simplement calculée comme la somme des contributions génétiques de chacun des allèles qu'il porte. Cependant, des relations de dominance ou d'épistasie émergent de la relation non linéaire utilisée pour relier le phénotype à la valeur sélective. De récents travaux sur une extension (plusieurs traits) du modèle de Fisher (Martin et Lenormand 2006a,b, 2007) montrent que ce type de formalisme peut rendre compte de la distribution des effets des mutations observée pour de nombreuses espèces et dans différents environnements, et qu'il peut prédire la distribution des interactions épistatiques entre allèles à différents locus.

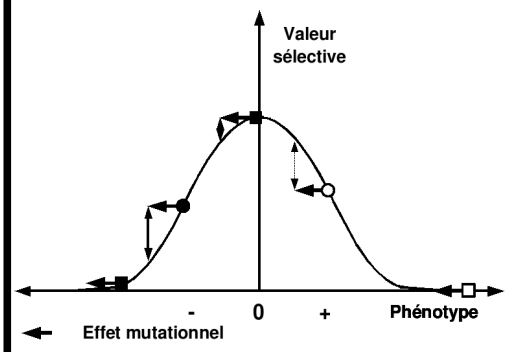


Figure 2: Variabilité des effets d'une mutation en fonction de la valeur de phénotype et de la relation phénotype-valeur sélective. L'optimum est ici arbitrairement fixé à 0. L'effet absolu de la mutation est représenté par la flèche en bas à gauche: elle diminue la valeur du phénotype. Si cette mutation se produit chez un individu dont le phénotype est bien supérieur à l'optimum, elle a tendance à rapprocher le phénotype de la valeur optimale (symboles vides): c'est une mutation favorable. Sinon, c'est une mutation défavorable (symboles pleins). L'effet sur la valeur sélective est variable: lorsque le phénotype de l'individu est très proche ou très éloigné de l'optimum, les effets sont faibles (carrés). Pour des phénotypes ni très proches, ni très éloignés de l'optimum, l'effet est fort (flèches verticales de grande taille pour les ronds). On voit aussi que plus un individu est proche de l'optimum, plus une mutation d'effet aléatoire a de chances de l'en éloigner et que lorsqu'il est maladapté, la probabilité d'une mutation favorable tend vers 0.5.

souvent beaucoup de temps. Cependant, cet inconvénient est maintenant beaucoup moins limitant du fait de la grande puissance de calcul des ordinateurs actuels.

III Objectifs de la thèse

L'impact génétique de la fragmentation des populations a été étudié par de nombreux travaux tant expérimentaux que théoriques. Ces différents travaux ont montré que le degré de fragmentation des populations (principalement à travers la taille des populations locales et les flux de gènes entre elles) peut avoir un impact contrasté selon que l'on considère des mutations sujettes à une sélection spatialement hétérogène ou, au contraire, des mutations contre-sélectionnées de la même manière à travers l'habitat. De manière très schématique, une augmentation de la migration entre dèmes se traduit généralement par une augmentation du fardeau de migration, tandis qu'elle entraîne une diminution du fardeau lié aux mutations inconditionnellement délétères. Cependant, les résultats présentés dans cette introduction générale ont tous été obtenus en considérant un seul type de pression de sélection à la fois: sélection homogène pour étudier les mutations inconditionnellement délétères et sélection spatialement hétérogène pour décrire le phénomène d'adaptation locale. On peut alors se demander dans quelle mesure les résultats obtenus en étudiant chaque type de sélection isolément seront retrouvés dans le cadre d'une approche conjointe.

Enfin, malgré le nombre considérable de travaux portant sur l'impact génétique de la fragmentation des populations, très peu est connu sur son influence à court terme, dans la mesure où la grande majorité des études théoriques portent sur des populations à l'équilibre. Or les activités humaines sont responsables d'événements de fragmentation récents affectant un très grand nombre d'espèces. Une partie de mon travail de thèse a donc consisté à étudier les conséquences génétiques à court terme de tels événements de fragmentation.

Plus précisément, nous avons essayé de combler trois manques. Dans un premier temps, nous avons voulu préciser le comportement des patrons d'adaptation locale en réponse à la migration: si l'influence de l'intensité du flux de gènes a été abondamment étudiée, la réponse de l'adaptation locale en fonction de la nature de ce flux de gènes n'est pas encore très bien connue. La question de base que nous nous sommes posée est la suivante: *la migration de deux grains de pollen a-t-elle les mêmes conséquences en termes d'adaptation locale que celle d'une graine ?* Pour apporter une réponse à cette question, deux approches théoriques ont été utilisées.

Ophélie Ronce a développé un modèle analytique de génétique quantitative prenant en compte la migration entre les populations, la sélection ainsi que la dérive, afin de décrire le fardeau génétique moyen et la divergence entre le populations, lorsque la variance génétique est connue. J'ai effectué des simulations individu-centrées pour tester les prédictions du modèle et étudier des questions plus détaillées que le modèle analytique ne permet pas d'explorer. Le résultat de ce travail est présenté sous la forme d'un article constituant le premier chapitre de ce document.

Les deux premières parties de cette introduction m'ont permis de montrer qu'il existe deux grandes classes d'études s'intéressant à la dynamique des mutations délétères dans les populations fragmentées: certaines études s'intéressent aux mutations dont les effets varient en fonction du milieu (étude de l'adaptation locale et du fardeau de migration), tandis que d'autres s'intéressent aux mutations dont les effets sont constants dans l'espace (étude des mutations inconditionnellement délétères et du fardeau de mutation associé). De surcroît, on a vu que la fragmentation des populations a des effets très variables selon l'approche utilisée. Par conséquent, la deuxième question à laquelle nous avons voulu apporter une réponse pourrait se formuler ainsi: *Comment la fragmentation des populations agit-elle sur des génomes complexes?* Pour étudier ce problème, j'ai effectué des simulations individu-centrées en utilisant des génomes présentant les deux types de mutations évoqués tout au long de cette introduction. Le résultat de ce travail est présenté sous la forme d'un manuscrit en préparation dans le deuxième chapitre. Ce travail porte sur deux points principaux. J'ai tout d'abord cherché à savoir dans quelle mesure les génomes complexes répondent à la fragmentation des habitats en tant qu'état, en étudiant l'état d'équilibre des populations. Cette étape m'a permis de mettre en évidence l'existence d'interactions entre les types de locus modélisés, ainsi que de décrire l'évolution du fardeau moyen des populations en fonction de l'intensité relative des pressions de sélection spatialement homogène et hétérogène et du degré de fragmentation des populations.

Le deuxième objectif de ce travail consistait à étudier la réponse à court terme d'un événement de fragmentation, mimé par une diminution des flux migratoires entre les populations. J'ai donc cherché à apporter une réponse à la question suivante: *Comment les génomes complexes étudiés précédemment répondent à un événement de fragmentation ?* afin également de mieux comprendre dans quelle mesure la fragmentation des habitats peut constituer une menace réelle pour les populations naturelles.

Chapitre 1

Hétérogénéité de l'environnement

et fardeau de migration chez les plantes:

rôle de la dispersion

par les graines et le pollen

Errata: Ce chapitre se présente sous forme d'épreuves (article à paraître dans *Journal of Evolutionary Biology*) non corrigées. Des erreurs sont à rectifier à la lecture:

- Annotations de la Figure 1:

  "*N*" au lieu de "*n*" au-dessus des panneaux

- Annotations de la Figure 2:

  "*N*" au lieu de "*n*" au-dessus des panneaux

- Annotations de la Figure 6:

  (i) lire "grey: part of the genotyptic variance contributed by gametic linkage disequilibrium" au lieu de "hatched: part of the genotypic variance contributed by gametic linkage disequilibrium"

  (ii) lire "hatched: part of the variance due to deviations from HWE" au lieu de "grey: part of the variance due to deviations from HWE"

- Légendes de la Figure 6:

  inverser $\hat{\delta}_1$ et $\hat{\delta}_2$

- Annotations de la Figure 7:

  lire "4 populations of 200 individuals" et non "32 populations of 25 individuals"

  lire "white, grey and hatched bars" au lieu de "white, hatched and grey bars"

  $\hat{\delta}_1$ au lieu de $\hat{\delta}_2{}^0$

- Légendes de la figure 7

  inverser $\hat{\delta}_1$ et $\hat{\delta}_2$

	J E B	1 4 4 2	B	Dispatch: 19.10.07	Journal: JEB	CE: Senthil
	Journal Name	Manuscript No.		Author Received:	No. of pages: 16	PE: Karthishwaran

doi:10.1111/j.1420-9101.2007.01442.x

Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes

S. LOPEZ,* F. ROUSSET,* F. H. SHAW*†, R. G. SHAW*‡ & O. RONCE*

*Institut des Sciences de l'Evolution (UM2-CNRS), Université Montpellier 2, Montpellier cedex 5, France
†Department of Mathematics, Hamline University, St Paul, MN, USA
‡Department of Evolution, Ecology and Behaviour, University of Minnesota, St Paul, MN, USA

Keywords:

dispersal;
local adaptation;
metapopulation;
pollen;
quantitative genetics;
seed.

Abstract

Evolution of local adaptation depends critically on the level of gene flow, which, in plants, can be due to either pollen or seed dispersal. Using analytical predictions and individual-centred simulations, we investigate the specific influence of seed and pollen dispersal on local adaptation in plant populations growing in patchy heterogeneous landscapes. We study the evolution of a polygenic trait subject to stabilizing selection within populations, but divergent selection between populations. Deviations from linkage equilibrium and Hardy–Weinberg equilibrium make different contributions to genotypic variance depending on the dispersal mode. Local genotypic variance, differentiation between populations and genetic load vary with the rate of gene flow but are similar for seed and pollen dispersal, unless the landscape is very heterogeneous. In this case, genetic load is higher in the case of pollen dispersal, which appears to be due to differences in the distribution of genotypic values before selection.

Introduction

Local adaptation is a well-studied phenomenon resulting from the action of spatially heterogeneous selection and characterized by a higher fitness of individuals in their native environment compared with immigrants (Kawecki & Ebert, 2004). Patterns of local adaptation have been particularly documented in plants, where genetic differentiation recurrently evolves in response to, for example, climatic gradients (Etterson & Shaw, 2001; Rehfeldt *et al.*, 2002; Etterson, 2004a,b; Savolainen *et al.*, 2004; Goldringer *et al.*, 2006) or spatially heterogeneous edaphic conditions (e.g. adaptation to soil contamination by heavy metals, McNeilly, 1968; Antonovics & Bradshaw, 1970; MacNair, 1987; Jimenez-Ambriz *et al.*, 2007). The evolution of local adaptation can sometimes be considered as a step toward speciation, in particular when it involves shifts in phenology as documented in *Mimulus guttatus* on mine tailings (McNeilly & Antonovics, 1968; Antonovics, 2006; Hall & Willis, 2006). Local adaptation can also be of crucial importance for conservation practices, because of the potentially harmful consequences of releasing maladapted individuals (Keller *et al.*, 2000; for a review, see McKay *et al.*, 2005).

Gene flow is generally thought to oppose the effect of divergent natural selection, and to increase the genetic load that depresses mean fitness in heterogeneous environments (Lenormand, 2002). There are, however, several dimensions to such migration load, which interact in a complex manner. First, gene flow may prevent adaptive divergence, measured by the difference in mean phenotype, of populations experiencing different selection pressures (for theoretical predictions in the case of polygenic traits, see Garcia-Ramos & Kirkpatrick, 1997; Tufto, 2000; Hendry *et al.*, 2001; for empirical evidence see, e.g. Hendry & Taylor, 2004). Second, gene flow may affect the evolution of genotypic variance within local populations (for theoretical predictions in the case of polygenic traits see in particular Goldstein & Holsinger, 1992; Phillips, 1996; Lythgoe, 1997; Barton, 1999; Tufto, 2000; for empirical evidence see Yeaman & Jarvis, 2006). Large genotypic variance for a phenotypic trait depresses mean fitness when selection on this trait is stabilizing

Correspondence: Ophélie Ronce, Institut des Sciences de l'Evolution (UM2-CNRS), Université Montpellier 2, Place Eugène Bataillon, CC 65, 34 095 Montpellier cedex 5, France.
Tel.: +33 4 67 14 32 50; fax: +33 4 67 14 36 22;
e-mail: ronce@isem.univ-montp2.fr

within local populations (Barton, 1986). Genotypic variance is, however, also the necessary fuel for adaptation by natural selection. Increased genotypic variance within populations may in turn facilitate adaptive divergence between them (e.g. Hendry *et al.*, 2001). The interactions between different components of the genetic load are especially complex in the case of polygenic characters as the evolution of the mean phenotype and the genotypic variance may in part be independent (e.g. see Spichtig & Kawecki, 2004). The net effect of gene flow on genetic load can thus be ambiguous. For instance, the prediction that gene flow limits species range (Kirkpatrick & Barton, 1997) does not hold when the effect of gene flow on genotypic variance is taken into account (Barton, 2001). Migration might replenish genetic variation eroded by drift and improve local adaptation in peripheral populations as predicted by recent simulation studies (see, e.g. Holt *et al.*, 2003; Alleaume-Benharira *et al.*, 2006). Higher migration rate also promotes local adaptation to antagonistic partners engaged in a coevolutionary arms race (for theoretical prediction see Gandon *et al.*, 1996; for empirical evidence see Morgan *et al.*, 2005).

The relative contributions of seeds and pollen to gene flow can vary greatly among species (for a review, see Ouborg *et al.*, 1999). Estimates of pollen to seed flow ratio can, for instance, range from 1.8 in *Eucalyptus nitens* (Moran, 1992) to 500 in *Quercus petraea* (Petit *et al.*, 1993) and vary also within species depending on the spatial scale considered (as in *Silene alba*, McCauley, 1997). For allogamous species, estimates of the ratio of pollen to seed flow are generally considered as high, both within and between populations (Latta, 2006, but see Bacles *et al.*, 2006). Few theoretical models exploring the effect of gene flow on genetic load and local adaptation explicitly deal with pollen and seed dispersal (Nagylaki, 1997; Hu & Li, 2001, 2002; Hu, 2005). The number of gene copies carried by a migrating seed is twice that carried by a migrating pollen grain. When allelic frequencies are heterogeneous through space, migration of zygotes through seeds results in local heterozygote deficiency, whereas migration of male gametes only, through pollen, generates, in contrast, a local excess of heterozygotes. When selection is spatially heterogeneous, immigrant genes are counter-selected in individuals combining resident and immigrant alleles with pollen dispersal, whereas selection acts more often on individuals carrying either two immigrant alleles or two resident alleles with seed dispersal. As a consequence, the response of migration load to these two modes of gene flow could be quite different. Considering a phenotypic trait under the control of a major locus, Hu & Li (2001) predicted that seed dispersal and pollen dispersal differently affect the shape of clines in genotypic variance along a sharp environmental transition. How such results generalize to polygenic traits in patchy landscapes is, however, unclear.

Here, we explore theoretically the specific consequences of gene flow through seed and pollen for the evolution of genetic load in a self-compatible plant inhabiting a patchy heterogeneous landscape, such as the heavy metal tolerant plant *Thlaspi caerulescens* growing in a network of polluted and nonpolluted sites (Jimenez-Ambriz *et al.*, 2007). We consider the evolution of a single quantitative character with polygenic variation, which is subject to both stabilizing selection within local populations and disruptive selection across populations in different habitats. We explore the consequences of gene flow for: (i) the maintenance of genetic variability within populations; (ii) the evolution of genetic divergence between populations in different environments; and (iii) the genetic load depressing the mean fitness within local populations, which integrates the effect of the former two parameters. Individual-based simulations are used to explore the evolution of within-population genotypic variance with seed and pollen dispersal. We derive analytical predictions for the divergence between populations and the genetic load assuming that the genotypic variance is known and the distribution of genotypic values is Gaussian within populations, following Tufto (2000) and Hendry *et al.* (2001). Our analytical model takes drift, selection and migration of both seed and pollen into account. We compare these predictions with the individual-based simulations and investigate how seed and pollen dispersal affect departures from the assumptions of the analytical model.

Methods

General assumptions

Even though we conceive our model to be quite general, we find it useful to illustrate its relevance with some real case study. In southern France, the heavy metal tolerant plant *T. caerulescens* grows in a network of contaminated former mining sites and noncontaminated sites, only a few kilometres away from each other (Escarré *et al.*, 2000). Comparison of divergence for quantitative traits and molecular markers suggests strong divergent selection on several life-history traits in contaminated vs. noncontaminated sites, but weak stabilizing selection within each habitat type (Jimenez-Ambriz *et al.*, 2007). Gene flow between populations is weak but reveals no preferential exchanges between populations of the same habitat type (Dubois *et al.*, 2003; Jimenez-Ambriz *et al.*, 2007). The species is insect pollinated, self-compatible and seeds are dispersed ballistically at a short distance.

In our model, we consider a patchy population of a self-compatible hermaphroditic annual plant with no seed bank. There are n discrete patches of plants, each of size N. Migration of both seeds and pollen occurs between patches following an island model of migration. No selfing occurs prior to pollen dispersal. We consider

the effect of stabilizing selection on a single quantitative character. The landscape is heterogeneous, with different patches of habitat characterized by different optimal values for the phenotypic trait. Generations are discrete and nonoverlapping. The order of the events in the life cycle is: (1) selection acting via juvenile survival; (2) density regulation to a constant number of adults N in each population; (3) gametogenesis, pollen dispersal and syngamy; and (4) seed dispersal. We assume that survival probability in stage (1) is a function of genotype (see below) and that a very large number of seeds and pollen grains are produced at stages (3) and (4). A proportion m_p of pollen grains present after dispersal in a given local population originates from the $n-1$ other populations. Similarly, m_s is the seed migration rate. Mating occurs randomly after pollen dispersal and by considering a self-compatible species, we allow for random selfing to occur.

The genotypic value G of an individual for the quantitative character is determined by its diploid genotype at l loci. We assume that allelic effects on the quantitative character are additive, both within each locus (codominance) and across loci (no epistasis), i.e.

$$G = \sum_{i=1}^{l} (X_i^o + X_i^p),$$

where X_i^o (X_i^p) is the allelic value inherited through the ovule (pollen grain) at locus i. The phenotype Z of an individual is given by the sum of its genotypic value G and a random environmental effect distributed normally with mean 0 and variance σ_e^2. The survival probability of an individual with phenotype Z in population i is described by the Gaussian function:

$$W_i(Z) = \exp\left(\frac{-(Z-\theta_i)^2}{2\omega}\right)$$

with θ_i the optimal value of the quantitative character in population i and $1/\omega$ the intensity of stabilizing selection (here assumed to be the same in all populations). The expected survival probability of an individual with genotypic value G is

$$\tilde{W}_i(G) = \int \frac{1}{\sqrt{2\pi\sigma_e}} \exp\left(\frac{-(z-G)^2}{2\sigma_e^2}\right) W_i(z)\mathrm{d}z,$$

which is also:

$$\tilde{W}_i(G) = \sqrt{\frac{\omega}{\sigma_s^2}} \exp\left(\frac{-(G-\theta_i)^2}{2\sigma_s^2}\right), \quad \text{with } \sigma_s^2 = \sigma_e^2 + \omega \quad (1)$$

We define $\bar{G}_i$ as the mean genotypic value in population i, $\bar{\bar{G}}$ the mean genotypic value in the metapopulation, V_{gi} is the variance of genotypic values in population i, $\bar{V}_g$ is the mean within-population variance of the metapopulation, V_y is the variance of mean genotypic values among populations in the metapopulation. Similarly, $\bar{W}_i$ is the mean fitness in population i and $\bar{\bar{W}}$ the mean fitness at the scale of the metapopulation. These are random variables, which fluctuate through time because of drift. We are interested in the expected values of these random variables. We follow the expected mean genotypic variance within populations σ_g^2, the expected variance in genotypic values between populations σ_y^2 and the expected mean genetic load λ in the metapopulation:

$$\sigma_g^2 = E[\bar{V}_g] = E\left[\sum_{i=1}^{n} \frac{V_{gi}}{n}\right] \quad (2)$$

$$\sigma_y^2 = E[V_y] = E\left[\sum_{i=1}^{n} \frac{(\bar{G}_i - \bar{\bar{G}})^2}{n}\right] \quad (3)$$

$$\lambda = 1 - E[\bar{\bar{W}}] = 1 - E\left[\sum_{i=1}^{n} \frac{\bar{W}_i}{n}\right] \quad (4)$$

Our measure of load describes how the mean fitness at the scale of the metapopulation deviates from its maximal value of 1. Without loss of generality, we can rescale the phenotype so that:

$$E[\bar{\bar{G}}] = E\left[\sum_{i=1}^{n} \frac{\bar{G}_i}{n}\right] = 0 \quad (5)$$

To better understand the effect of pollen and seed migration on the evolution of the genotypic variance, we further decompose the variance of genotypic values in different terms (see, e.g. Bulmer, 1989 for a similar decomposition). Using the expression for genotypic value G, the expected within-population genotypic variance is:

$$\begin{aligned}\sigma_g^2 &= \frac{1}{n}\sum_{i=1}^{n} E_i\left[(G-\bar{G}_i)^2\right] \\ &= \frac{1}{n}\sum_{i=1}^{n} E_i\left[\left(\sum_{j=1}^{l} X_j^o + X_j^p - E_i[X_j^o + X_j^p]\right)^2\right],\end{aligned} \quad (6)$$

where E_i denotes the expectation in population i. This sum can be rearranged as:

$$\sigma_g^2 = \sigma_0^2 + \delta_1 + \delta_2 \quad (7)$$

with

$$\sigma_0^2 = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{l} E_i[(X_k^p)^2] + E_i[(X_k^o)^2] - 2E_i[\tilde{X}_k]^2$$

where

$$\tilde{X}_k = \frac{X_k^p + X_k^o}{2}, \quad (8)$$

$$\delta_1 = \frac{1}{n}\sum_{i=1}^{n} 2\sum_{k=1}^{l}\sum_{j>k}^{l}\left(E_i[X_k^p X_j^p] + E_i[X_k^o X_j^o] - 2E_i[\tilde{X}_k]E_i[\tilde{X}_j]\right), \quad (9)$$

and

$$\delta_2 = \frac{1}{n}\sum_{i=1}^{n} 2\sum_{k=1}^{l}\sum_{j=1}^{l}(E_i[X_k^{\mathrm{p}}X_j^{\mathrm{o}}] - E_i[\tilde{X}_k]E_i[\tilde{X}_j]). \qquad (10)$$

The term σ_0^2 is the expected genic variance at Hardy–Weinberg equilibrium (HWE) and linkage equilibrium, δ_1 is the part of the variance contributed by gametic linkage disequilibrium at HWE and δ_2 is the part of the variance due to deviations of genotypic frequencies with respect to HWE. Note that δ_1 and δ_2 are covariance terms and can be negative. Expressions in eqns 8–10 are slightly different from that in Bulmer (1989) to take into account potentially different allelic frequencies in male and female gametes. In particular, with pollen dispersal in heterogeneous environments, δ_2 is nonzero even if reproduction is panmictic because of the different allelic frequencies in pollen and ovules.

Analytical model

We assume that genotypic values are distributed normally after seed migration and before selection. This is unlikely to hold if both migration and selection are strong. Comparison of the present predictions to simulations will allow measuring the effect of non-Gaussian distribution of phenotypes. We further assume that, before selection, the within-population genotypic variance V_{gi} varies little around its expected value and does not vary much across populations, so that $V_{gi} \approx \sigma_g^2$. Finally, we assume that all loci recombine freely (no physical linkage). Equations for changes of: (i) the mean phenotype $\bar{G}_i$ in each local population; (ii) the expected genotypic variance σ_g^2 within populations; and (iii) the genotypic variance between populations σ_y^2, along the life cycle, are given in Appendix A. Solving for the equilibrium value of σ_g^2, however, requires computing the effect of selection, drift and migration on complex components of the genotypic variance (see eqns A12–A14 and 7–10), which is out of the scope of the present paper. Instead, we will consider the value of σ_g^2 as a parameter and derive predictions for the divergence between populations σ_y^2 and the mean genetic load λ when the within-population genotypic variance is known (as in Hendry *et al.*, 2001, or Garcia-Ramos & Kirkpatrick, 1997). Evolution of the mean variance in genotypic values within populations σ_g^2 and its components (eqns 7–10) will be studied through simulations.

Simulations

In line with the analytical model, we ran individual-centred simulations. Throughout, we use a value of $\sigma_e^2 = 1$, so that all phenotypic measures are standardized by the environmental standard deviation. We simulate highly or slightly fragmented metapopulations: $n = 32$ and $N = 25$ or $n = 4$ and $N = 200$, respectively, so that the metapopulation total size is constant and equal to 800 adults despite different degrees of fragmentation. We do not investigate cases where the local population size is very large. For the sake of simplicity, we assume that there are only two types of habitat with optimal phenotypes θ_1 and θ_2, respectively (e.g. polluted vs. non-polluted), and an equal number of patches of each type. Note that such symmetry facilitates the maintenance of genetic diversity. Habitat heterogeneity is measured by $\Delta\theta = \theta_1 - \theta_2$. We study three levels of landscape heterogeneity corresponding to homogeneous, moderately heterogeneous and highly heterogeneous landscapes with $\Delta\theta$ values of 0, 1 and 3 respectively. When both seeds and pollen disperse, preliminary simulations showed that results were always intermediate between those obtained for pure pollen grains and pure seed dispersal. To illustrate the effect of the dispersal mode, we therefore show simulation results for these two extreme cases only (pure pollen dispersal $m_s = 0$ and pure seed dispersal $m_p = 0$).

In most of the simulations presented here, we assume that the trait value is determined by 10 freely recombining loci. Simulations were also run for two, five and 30 loci with little qualitative or quantitative differences (results not shown). To illustrate the effect of physical linkage, we also present results with 10 linked loci, with a recombination probability of 0.01 between adjacent loci. The number of alleles segregating simultaneously at each locus is not limited. Mutation creates a new allele with an allelic effect obtained as the sum of the parental allelic effect and of a normally distributed deviation with mean 0 and variance α^2. Mutations occur independently at each locus on each gamete with probability μ. According to estimates from mutation–accumulation studies (e.g. Shaw *et al.*, 2002), we used two values for the diploid genomic mutation rate ($U = 2l\mu = 10^{-2}$ or 10^{-1}), which represents the expectation for the number of new mutations of a diploid zygote. We standardized values of α^2 so the total variance introduced by mutation $\sigma_m^2 = U\alpha^2$ is constant and $\sigma_m^2 = 10^{-3}\sigma_e^2$ as suggested by the literature (see, e.g. Bürger *et al.*, 1989). Thus, we simulate either rare mutations with large effects on the phenotype, or more frequent mutations with smaller effects. We use a rather high intensity of selection, $\omega = 1$ (which is, however, compatible with some empirical estimations, see discussion in Johnson & Barton, 2005). Ovule production is Poisson distributed with expectation $F = 12$ per individual. All ovules are fertilized and F is sufficiently large to ensure that the number of surviving juveniles in the simulations is always greater than the carrying capacity and that the number of adult plants per patch is constant.

We used the batch method for Markov chains (Hastings, 1970) to obtain estimates and confidence intervals for variables of interest (see complete description in Appendix B). We checked the validity of our simulation programme by comparing its results with analytical predictions and previously published simulation results. We ran simulations with a single isolated population and verified that, when using the same parameter values as

in Bürger *et al.* (1989), the estimated variance in our simulations was in close agreement with that observed in their simulations (see their Table 1). We also compared estimates of the within-population genotypic variance with predictions of the 'Stochastic house-of-card approximation' (SHOC, Bürger *et al.*, 1989). The estimated genotypic variance in our simulations was, in general, close to the SHOC approximation. Yet, when both population size and mutation rate were large ($N = 200$, $U \geq 10^{-1}$) and the number of loci was not very large ($l < 50$), the estimated genotypic variance was overesti- 3 pp., *sqq*). Here, we simply compute $\hat{\sigma}_y^2/\hat{\sigma}_g^2$ as a standardmated by the SHOC approximation. We also ran simulations with a neutral quantitative character and compared the mean local genotypic variance at equilibrium, the variance in genotypic values between populations and the time to reach this equilibrium to analytical approximations derived by Lande (1992), and to more exact analytical results derived in the case of pollen and seed dispersal. We obtained a very good agreement between our simulation results and the analytical predictions in the neutral case (data not shown).

Results

Divergence between populations when the genotypic variance is known

Analytical predictions

Let us first assume that the within-population genotypic variance at equilibrium ($\hat{\sigma}_g^2$) is known. Then, solving eqn A19 and recursions (eqns A4–A18) shows that the expected value of the mean genotypic value in population i at equilibrium is:

$$\hat{E}[\bar{G}_i] = \theta_i(1-\tilde{m}_t)\frac{\hat{\sigma}_g^2}{\tilde{m}_t\sigma_s^2+\hat{\sigma}_g^2}. \qquad (11)$$

where

$$\tilde{m}_t = \tilde{m}_s + \frac{\tilde{m}_p}{2}(1-\tilde{m}_s),\ \ \tilde{m}_p = \frac{n}{n-1}m_p \text{ and } \tilde{m}_s = \frac{n}{n-1}m_s. \qquad (12)$$

The term $\tilde{m}_t$ in eqn 12 can be interpreted as the net migration rate, which takes into account gene copies introduced by immigrant seeds and those introduced by immigrant pollen having fertilized nondispersed seeds. The migration rate of pollen must be twice that of seeds to achieve the same level of gene flow. Pollen and seed dispersal rates are furthermore rescaled as functions of the number of demes (see eqn 12). Equation (11) leads to the same prediction about divergence in mean phenotype as eqn 8 in Hendry *et al.* (2001). In particular, the expected mean phenotype in population i depends on the net migration rate, but not on how pollen and seeds contribute to such migration rate. The mean phenotype is closer to the local optimum if the net migration rate is low, the intensity of selection high (σ_s^2 is smaller) and the within-population genotypic variance high.

By substituting into eqn A6 and solving recursions (eqns A5–A18), we obtain the between-population variance at equilibrium as a function of the within-population genotypic variance. Differentiation between populations for quantitative traits is classically measured through statistics such as Q_{ST} (e.g. Bonnin *et al.*, 1996), which are functions of the ratio of between- and within-population variances $\hat{\sigma}_y^2/\hat{\sigma}_g^2$. The expression for Q_{ST} also depends on F_{IS} (Bonnin *et al.*, 1996, Appendix), which itself depends on the selfing rate and dispersal rates (Rousset, 2004, 132 ized measure of between-population differentiation:

$$\frac{\hat{\sigma}_y^2}{\hat{\sigma}_g^2} = \frac{(1-\tilde{m}_t)^2}{\tilde{m}_t\sigma_s^2+\hat{\sigma}_g^2}\left(\frac{\sigma_s^2(\sigma_s^2+\hat{\sigma}_g^2)}{\hat{\sigma}_g^2+(2-\tilde{m}_t)\sigma_s^2}\frac{n-1}{nN}+\frac{\hat{\sigma}_g^2}{\tilde{m}_t\sigma_s^2+\hat{\sigma}_g^2}\sum_{i=1}^{n}\frac{\theta_i^2}{n}\right). \qquad (13)$$

Equation (13) predicts that differentiation between populations increases with increasing variance in optimal phenotypes $\sum_{i=1}^{n}\theta_i^2/n$ and decreasing patch size N (see also Fig. 1). Differentiation is predicted to depend only on the net migration rate, independent of the allocation between seed and pollen dispersal. It declines with increasing migration. Increasing the genotypic variance within populations $\hat{\sigma}_g^2$ has nonmonotone effects on the differentiation between populations. When the migration rate is very low, differentiation decreases with increasing genotypic variance within populations (see Fig. 1). Conversely, differentiation increases with increasing genotypic variance within populations when the migration rate is large (Fig. 1). When selection is weak (σ_s^2 is very large), we can verify that differentiation between populations tends towards that expected under neutrality of the phenotypic trait, i.e.

$$\frac{\hat{\sigma}_y^2}{\hat{\sigma}_g^2} \to \frac{(n-1)(1-\tilde{m}_t)^2}{nN\tilde{m}_t(2-\tilde{m}_t)} \quad \text{when } \sigma_s^2 \to +\infty, \qquad (14)$$

which is consistent with earlier results (Rousset, 2004, equation 8.16).

In the presence of disruptive selection between populations $\sum_{i=1}^{n}\theta_i^2/n \neq 0$ but stabilizing selection within populations (σ_s^2 small), differentiation between populations for selected traits as predicted by eqn 13 is larger than for neutral traits (eqn 14) when the migration rate is large. When the migration rate is low, however, differentiation between populations for selected traits may be lower than that expected under neutrality of the trait, despite the divergent selection.

Comparison with simulations

To check on the accuracy of our predictions, we used the estimated genotypic variance measured in the simulations $\hat{\sigma}_g^2$ to compute the differentiation between populations using eqn 13 and compared these predictions to the estimated differentiation between populations from the simulations. Our aim here is not to assess, through simulations, the validity of the predicted relationship

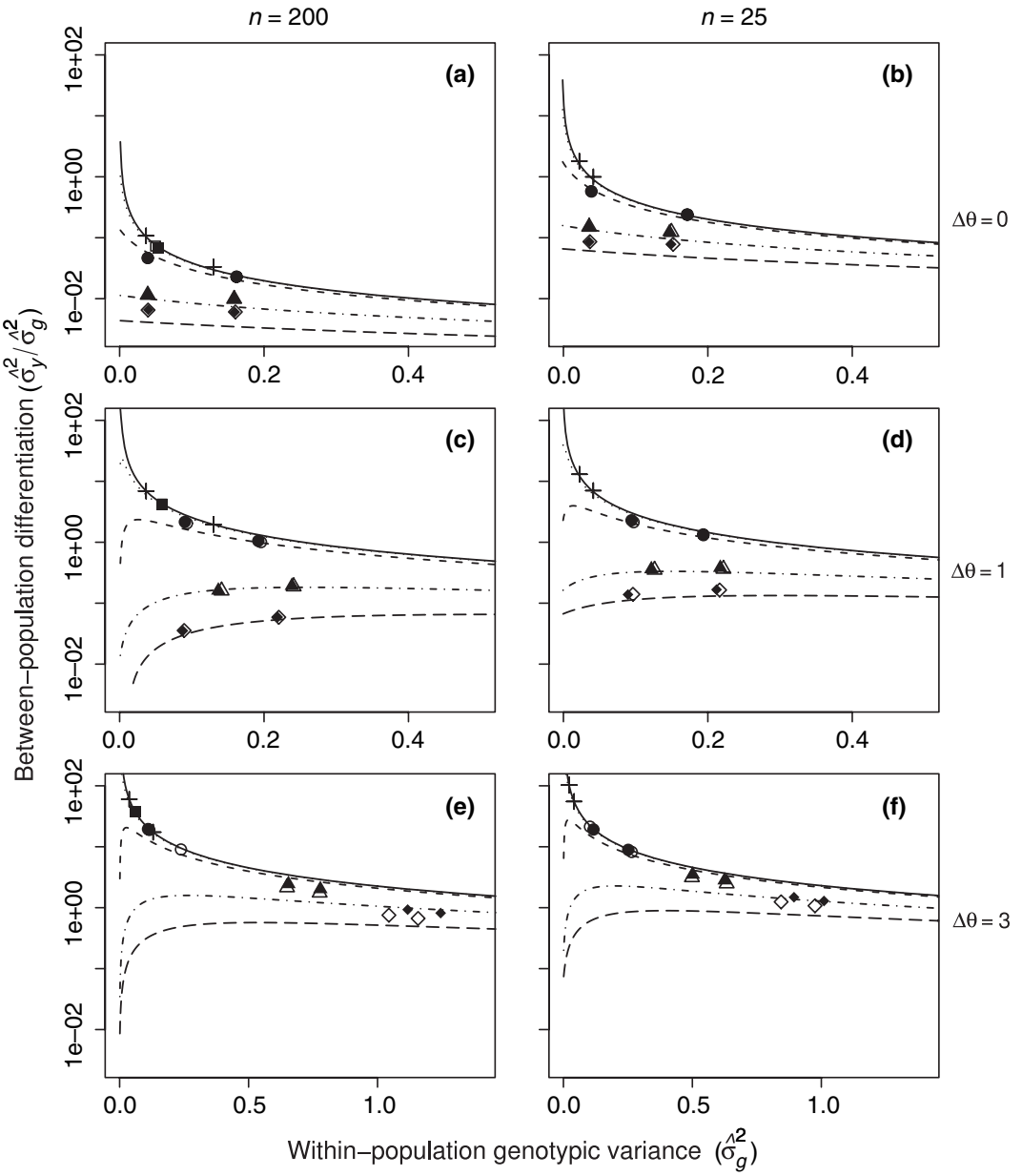


Fig. 1 Between-population differentiation measured by the ratio of between- to within-population genotypic variance $\hat{\sigma}_y^2/\hat{\sigma}_g^2$ as a function of mean within-population genotypic variance $\hat{\sigma}_g^2$ in different landscapes. Lines are analytical predictions from eqn 13, for different values of the net migration rate. Open (filled) symbols correspond to simulation results with pure pollen (seed) dispersal, for two values of the genomic mutation rate ($U = 10^{-2}$, 10^{-1}) and five values of the net migration rate; crosses and continuous lines: $\tilde{m}_t = 0$; squares and dotted lines: $\tilde{m}_t, = 10^{-3}n/(n-1)$; circles and short-dashed lines: $\tilde{m}_t = 10^{-2}n/(n-1)$); triangles and dot-dashed lines: $\tilde{m}_t = 10^{-1}n/(n-1)$; diamonds and long-dashed lines: $\tilde{m}_t = 2 \times 10^{-1}n/(n-1)$. Left panels (a,c,e) show moderately fragmented landscapes of $n = 4$ populations of $N = 200$ individuals. Right panels (b,d,f) show highly fragmented landscapes of $n = 32$ populations of $N = 25$ individuals. Habitat heterogeneity increases from top to bottom: (a,b) $\Delta\theta = 0$, (c,d) $\Delta\theta = 1$, (e,f) $\Delta\theta = 3$. Free recombination between loci. Confidence intervals are smaller than size of symbols. Confidence intervals on the *x*-axis range from 10^{-4} to 9×10^{-3}. Some simulation points are missing (20 over 120 parameters sets) due to lack of convergence to a stable equilibrium (see *Methods*). Missing points correspond to the following parameter sets: ($N = 25$, n=32, $\tilde{m}_t = 10^{-3} \times 32/31$, for all values of the seed dispersal rate, mutation rate and habitat heterogeneity) ($N = 200$, $\tilde{m}_t = 10^{-3} \times 4/3$, U =1, for all values of the seed dispersal rate and habitat heterogeneity) ($N = 200$, $\tilde{m}_t = 10^{-3} \times 4/3$, $U = 10^{-2}$, $\Delta\theta = 3$ only for pure pollen dispersal) ($N = 200$, $\tilde{m}_t = 10^{-2} \times 4/3$, $U = 10^{-1}$, $\Delta\theta = 3$ only for pure seed dispersal).

between $\hat{\sigma}_y^2/\hat{\sigma}_g^2$ and $\hat{\sigma}_g^2$ for all values of the latter parameter. In particular, only a limited range of values for $\hat{\sigma}_g^2$ are biologically relevant for a given landscape structure and genetic architecture. Instead, we compare the deviation between analytical predictions and simulation results for a discrete number of parameter sets. This comparison helps us to evaluate the effect of deviations from a Gaussian distribution of genotypic values. Figure 1 shows that eqn 13 performs best when habitat heterogeneity is moderate and local population size is large. A very good quantitative agreement is then found for both pure pollen dispersal and pure seed dispersal. In general, eqn 13, however, underestimates differentiation between populations when migration is large, especially when habitat heterogeneity is large (Fig. 1). Whereas between-population differentiation observed in the simulations is very similar for pure pollen and pure seed dispersal when habitat heterogeneity is moderate, observed differentiation is higher with seed dispersal in strongly heterogeneous landscapes. It appears that seed dispersal causes stronger departures from model predictions in strongly heterogeneous landscapes than pollen dispersal does. Quantitative differences between the two modes of dispersal remain, however, small.

Genetic load when the genotypic variance is known

Analytical predictions

Using eqn 11 and assuming that the mean phenotype in population i varies little around its expected value, the expected mean fitness in population i can be approximated by:

$$E[\bar{W}_i] \approx \sqrt{\frac{\omega}{\sigma_s^2 + \hat{\sigma}_g^2}} \exp\left(\frac{-\theta_i^2 \tilde{m}_t^2 (\sigma_s^2 + \hat{\sigma}_g^2)}{2(\tilde{m}_t \sigma_s^2 + \hat{\sigma}_g^2)^2}\right). \qquad (15)$$

Note that such approximation should be less accurate when local population size is small and the mean phenotype fluctuates widely around its expected value because of drift. Equation (15) predicts that, for a given within-population genotypic variance $\hat{\sigma}_g^2$, the local genetic load depends only on the net migration rate $\tilde{m}_t$ and not on how pollen and seed dispersal contribute to total gene flow. In particular, eqn 15 predicts that the genetic load always increases with increasing net migration rate $\tilde{m}_t$ when the genotypic variance is held constant (see also Fig. 2). Equation (15) also suggests that the genetic load decreases with increasing genotypic variance when such variance is low, but increases again with

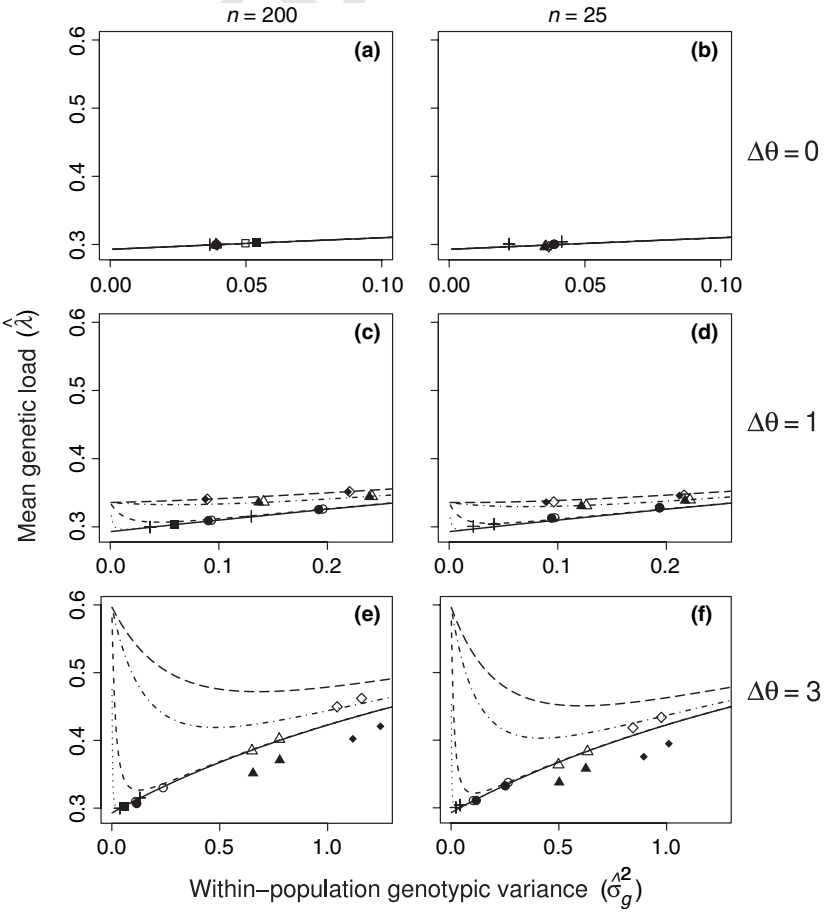


Fig. 2 Mean genetic load as a function of mean within-population genotypic variance $\hat{\sigma}_g^2$ in different landscapes. Lines are analytical predictions from eqn 15, for different values of the net migration rate. Open (filled) dots correspond to simulation results with pure pollen (seed) dispersal for two values of the mutation rate ($U = 10^{-2}$, 10^{-1}) and five values of the migration rate (same values of the migration rate and associated symbols as in Fig. 1). Left panels (a,c,e) show moderately fragmented landscapes of $n = 4$ populations of $N = 200$ individuals. Right panels (b,d,f) show highly fragmented landscapes of $n = 32$ populations of $N = 25$ individuals. Habitat heterogeneity increases from top to bottom: (a,b) $\Delta\theta = 0$, (c,d) $\Delta\theta = 1$, (e,f) $\Delta\theta = 3$ Free recombination between loci. Confidence intervals are smaller than size of symbols. Confidence intervals on the x-axis range from 10^{-4} to 9×10^{-3}. Missing data points are as in Fig. 1.

higher level of variance as illustrated in Fig. 2. Such nonmonotone patterns of variation of the genetic load with increasing genotypic variance are predicted for intermediate dispersal rates in heterogeneous landscapes. The genetic load is predicted to increase with increasing habitat heterogeneity.

Comparison with simulations

As previously with between-population differentiation, Fig. 2 shows that eqn 15 performs best when habitat heterogeneity is moderate and local population size is large. Genetic load observed in the simulations is in particular very similar for pure pollen and pure seed dispersal when habitat heterogeneity is moderate. Genetic load in small isolated populations is underestimated by eqn 15. Deviations from analytical predictions are, however, small in moderately heterogeneous landscapes even for small populations. When habitat heterogeneity is large, for both small and large local population size, eqn 15 greatly overestimates the genetic load and more strikingly so in the case of pure seed dispersal than for pure pollen dispersal (Fig. 2). In the range of parameters used in the simulations, genetic variance evolves to levels such that genetic load always increases with the local genetic variance (Fig. 2). Even with unrealistically low mutation rates (results not shown), intermediate migration rates in heterogeneous landscapes help maintain relatively large amounts of within-population variance in our model, so that situations predicted by eqn 15, where genetic load decreases when genetic variance increases were never observed.

The observed distribution of phenotypic values within a population before selection (Fig. 3) strongly deviates from a Gaussian distribution, for both pure seed and pollen dispersal when habitat heterogeneity is high. In particular with pure seed dispersal in strongly heterogeneous landscapes, the distribution of phenotypes is characterized by a fat tail, corresponding to immigrants from a different habitat, which are clearly distinct from the local resident population.

Figure 4 compares the mean genetic load at the scale of the metapopulation for pure pollen and pure seed dispersal as a function of habitat heterogeneity. As habitat heterogeneity increases, the genetic load caused by pollen migration is always higher than with pure seed dispersal. Contrary to our analytical prediction, above some level of heterogeneity, the genetic load eventually decreases as habitat heterogeneity increases in the case of seed dispersal while it reaches a plateau for pollen dispersal.

Effect of pollen and seed dispersal on the genotypic variance

Analytical predictions

Our analytical model does not allow us to solve for the within-population genotypic variance at equilibrium. Equations A11–A18, however, give some insights about

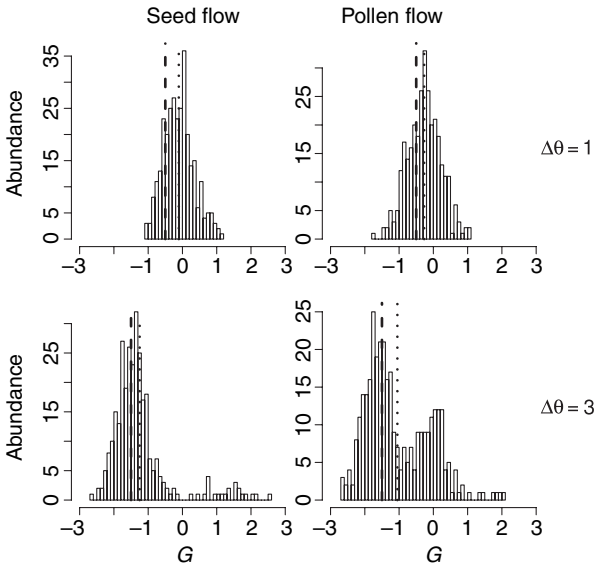


Fig. 3 Examples of the within-population distribution of genotypic values observed in the simulations as a function of landscape heterogeneity and dispersal mode. The data correspond to a single census after 495 000 generations of simulations, within one population after reproduction and before selection, in a slightly fragmented landscape of four populations of 200 individuals. Top panels: slightly heterogeneous landscape ($\Delta\theta = 1$), bottom panels: highly heterogeneous landscape ($\Delta\theta = 3$); left (right) panels: pure seed (pollen) dispersal. Dashed lines represent the local optimal phenotype value and the dotted lines the local mean phenotype at the census time. Genomic mutation rate U =0.1; net migration rate $\tilde{m}_t = 0.1 \times 4/3$. Free recombination between loci.

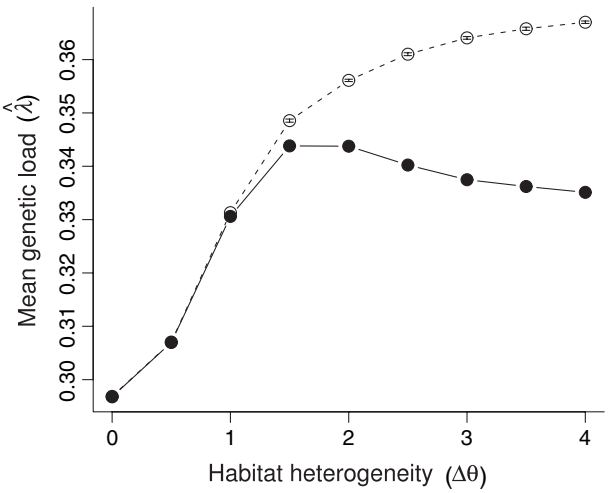


Fig. 4 Genetic load as a function of landscape heterogeneity and dispersal mode in a highly fragmented landscape of 32 populations of 25 individuals. Open (filled) circles correspond to pure pollen (seed) dispersal. Genomic mutation rate $U = 0.1$. Free recombination between loci. Net migration rate $\tilde{m}_t = 0.1 times 32/31$.

the effect of pollen and seed dispersal on the evolution of the within-population genotypic variance. Consider the case of pure pollen dispersal ($\tilde{m}_s = 0$). Then eqn A18 reads:

$$\sigma_g^{2(4)} = \frac{\sigma_0^{2(2)}}{2} - \frac{\rho^{(2)}}{2} + \frac{\sigma_g^{2(2)}}{2} + \frac{\tilde{m}_t}{1 - m_t}\hat{\sigma}_y, \qquad (16)$$

where $\sigma_g^{2(2)}$ and $\sigma_g^{2(4)}$ are the within-population genotypic variance, respectively, in adults before dispersal and in juveniles after dispersal, $\hat{\sigma}_y^2$ is the between-population genotypic variance at equilibrium as given by eqn 13, and $\sigma_0^{2(2)} - \rho^{(2)}$ is a complex term which varies with the genic variance and deviations from HWE in adults after selection (see eqn A13). Conversely, with pure seed dispersal ($\tilde{m}_p = 0$), eqn A18 becomes:

$$\sigma_g^{2(4)} = \frac{\sigma_0^{2(2)}}{2} - \frac{\rho^{(2)}}{2} + \frac{\sigma_g^{2(2)}}{2} + \frac{\tilde{m}_t(2 - \tilde{m}_t)}{(1 - m_t)^2}\sigma. \qquad (17)$$

Seed dispersal is therefore predicted to introduce more genotypic variance within populations than pollen dispersal for the same net migration rate. This is because a large part of the total genotypic variance is contained within hybrid individuals, rather than between individuals, after pollen dispersal and syngamy.

Equations (16–17) further suggest that the variance due to migration does not increase monotonically with dispersal, as $\hat{\sigma}_y^2$ itself decreases with increasing migration (see previous section about divergence between populations). Yet, fully understanding how pollen and seed dispersal affect the evolution of the genotypic variance requires describing their effects on each of the different components of the genotypic variance (in particular computing the term $\sigma_0^{2(2)} - \rho^{(2)}$), which our analytical model does not explore. We therefore resort to simulations to investigate this question.

Simulations results

Figure 5 compares the within-population genotypic variance and its components ($\hat{\sigma}_0^2$, $\hat{\delta}_1$ and $\hat{\delta}_2$) as a function of landscape heterogeneity $\Delta\theta$, in the cases of pure pollen or pure seed dispersal. In the case of pure pollen dispersal, the amount of genic variance $\hat{\sigma}_0^2$ represents the major component of genotypic variance (Fig. 5) and increases regularly with landscape heterogeneity. The two other components are negative and comparatively weak in magnitude. The component of variance due to deviations from panmixia $\hat{\delta}_2$ decreases as habitat heterogeneity increases (Fig. 5). With pure seed dispersal, we observe a different pattern. Genic variance $\hat{\sigma}_0^2$ is always smaller with pure seed dispersal than with pure pollen dispersal. It is maximal for intermediate habitat heterogeneity. Contrary to the case of pollen dispersal, in highly heterogeneous landscapes, the other components of variance ($\hat{\delta}_1$ and $\hat{\delta}_2$) are positive and grow large as habitat heterogeneity increases (Fig. 5).

Therefore, in moderately heterogeneous landscapes, higher total genotypic variance with pollen dispersal results from a higher genic variance. On the contrary, in highly heterogeneous landscapes, seed dispersal generates strong positive linkage disequilibrium and heterozygote deficiency, inflating the genotypic variance above that expected with pollen dispersal. Note that, even though the components of genotypic variance are widely different between the two dispersal modes, their combined effects lead to very similar level of total variance. For both modes of dispersal, total genotypic variance increases with landscape heterogeneity.

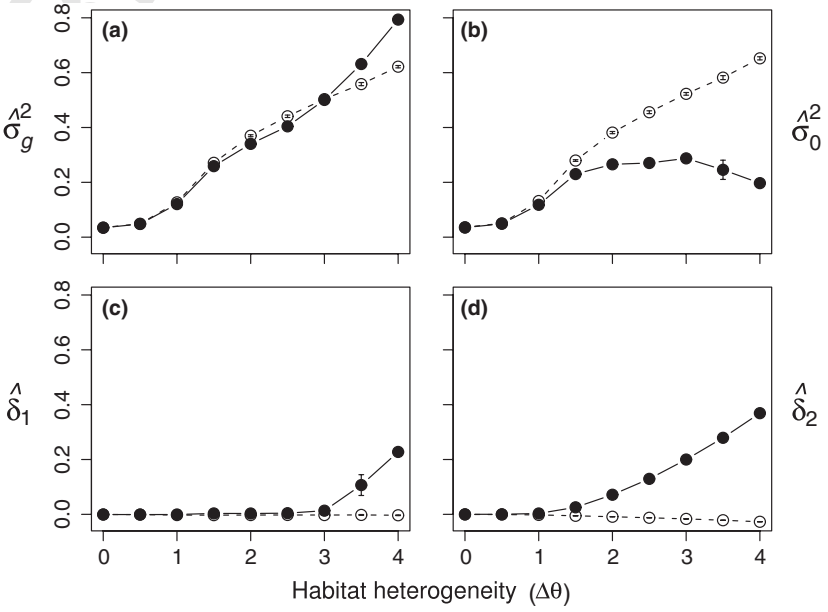


Fig. 5 Within-population genotypic variance and its components as a function of landscape heterogeneity and dispersal mode in a highly fragmented landscape of 32 populations of 25 individuals. Open (filled) symbols correspond to pure pollen (pure seed) dispersal. (a) Total within-population genotypic variance ($\hat{\sigma}_g^2$), (b) mean genic variance at HWE ($\hat{\sigma}_0^2$), (c) mean part of the genotypic variance contributed by gametic linkage disequilibrium at HWE ($\hat{\delta}_1$), (d) mean part of the variance due to deviations of genotypic frequencies with respect to HWE ($\hat{\delta}_2$). See eqns 7–10 for definitions. Genomic mutation rate $U = 0.01$. Free recombination between loci. Net migration rate $\tilde{m}_t = 0.1 \times 32/31$. Confidence intervals are smaller than plotting symbols.

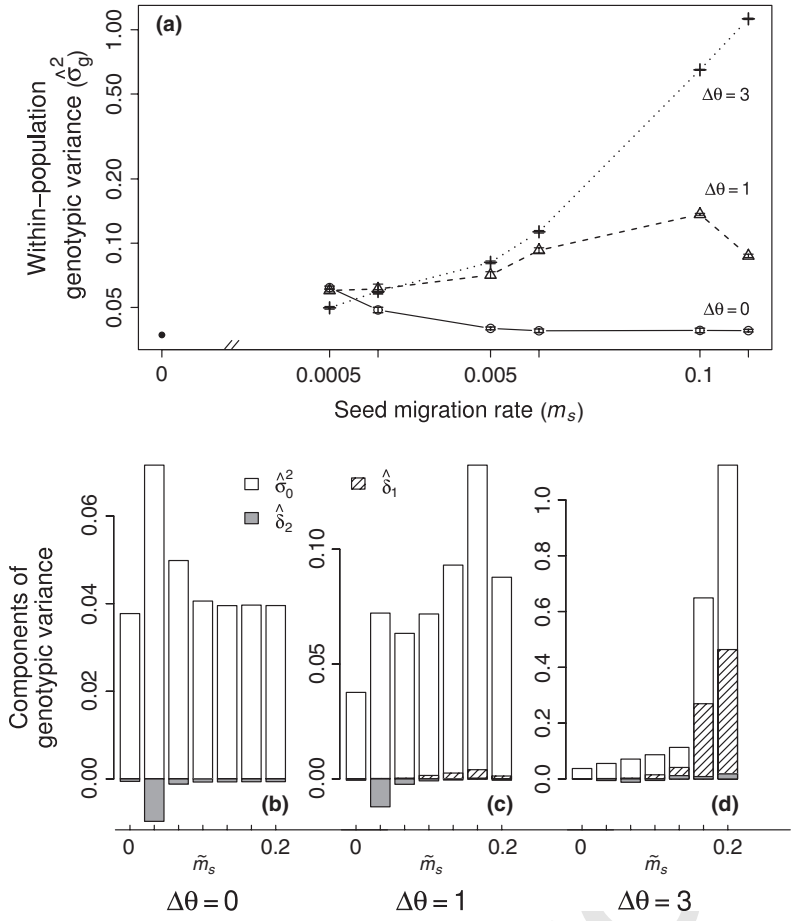


Fig. 6 Within-population genotypic variance and its components as a function of seed migration rate m_s in a highly fragmented landscape of 32 populations of 25 individuals. (a) Top panel shows mean within-population total genotypic variance $\hat{\sigma}_g^2$ for different levels of landscape heterogeneity; continuous line and open circles: homogeneous landscape ($\Delta\theta = 0$); dashed line and open triangles: slightly heterogeneous landscape ($\Delta\theta = 1$); dotted line and crosses: highly heterogeneous landscape ($\Delta\theta = 3$); the point on the left shows the level of within-population genotypic variance without migration; (b,c,d) bottom panels show within-population genotypic variance components as a function of the net seed migration rate; white: genic variance at HWE ($\hat{\sigma}_0^2$); hatched: part of the genotypic variance contributed by gametic linkage disequilibrium at HWE ($\hat{\delta}_1$); grey: part of the variance due to deviations from HWE ($\hat{\delta}_2$). See eqns 7–10 for definitions. Landscape heterogeneity increases from left to right. Pure seed dispersal: $m_p = 0$. Genomic mutation rate $U = 0.01$. Free recombination between loci.

Increasing the migration rate has the same qualitative effect on within-population genotypic variance with pure seed dispersal and pure pollen dispersal (results not shown). In homogeneous landscapes, low migration rates raise the genotypic variance much above that expected in isolated populations (Fig. 6a). Yet, a critical threshold in migration is soon reached, such that genotypic variance decreases with higher migration rate and converges again towards levels expected in isolated populations (Fig. 6). As habitat heterogeneity increases, migration increases the within-population genotypic variance more strongly and the critical migration threshold, above which the genotypic variance starts to decrease, is much higher (Fig. 6). The evolution of the genic variance $\hat{\sigma}_0^2$ as a function of the migration rate follows identical patterns and contributes the major part of the genotypic variance (Fig. 6). The component of variance due to gametic linkage disequilibrium $\hat{\delta}_1$ contributes little to the total genotypic variance: it is negative in isolated populations, reaches a minimum for low migration rates and becomes positive only in highly heterogeneous landscapes with high rates of seed dispersal (Fig. 6). The contribution of deviations from HWE genotypic frequencies $\hat{\delta}_2$ to the total genotypic variance increases with seed migration rate in highly heterogeneous landscapes (Fig. 6d).

Effect of physical linkage on genetic load and genotypic variance

We also investigated the effects of physical linkage between loci involved in local adaptation, using simulations. The effect of physical linkage on the genetic load and genotypic variance depends on the mutation rate and is noticeable only when habitat heterogeneity is not too high (see Fig. 7). Physical linkage affects the evolution of genotypic variance and genetic load, but similarly for seed and pollen dispersal and conclusions about differences between the two modes of dispersal are unaffected (results not shown). When the mutation rate is high, both the genetic load and genotypic variance are higher when the loci are freely recombining than when they are physically linked. The reverse is true when the mutation rate is low. Figure 7 shows that this pattern is due to the fact that: (i) linkage increases the genic variance $\hat{\sigma}_0^2$, but decreases the component of genotypic variance due to gametic linkage disequilibrium $\hat{\delta}_1$ (linkage disequilibrium, which is negative, increases in abso-

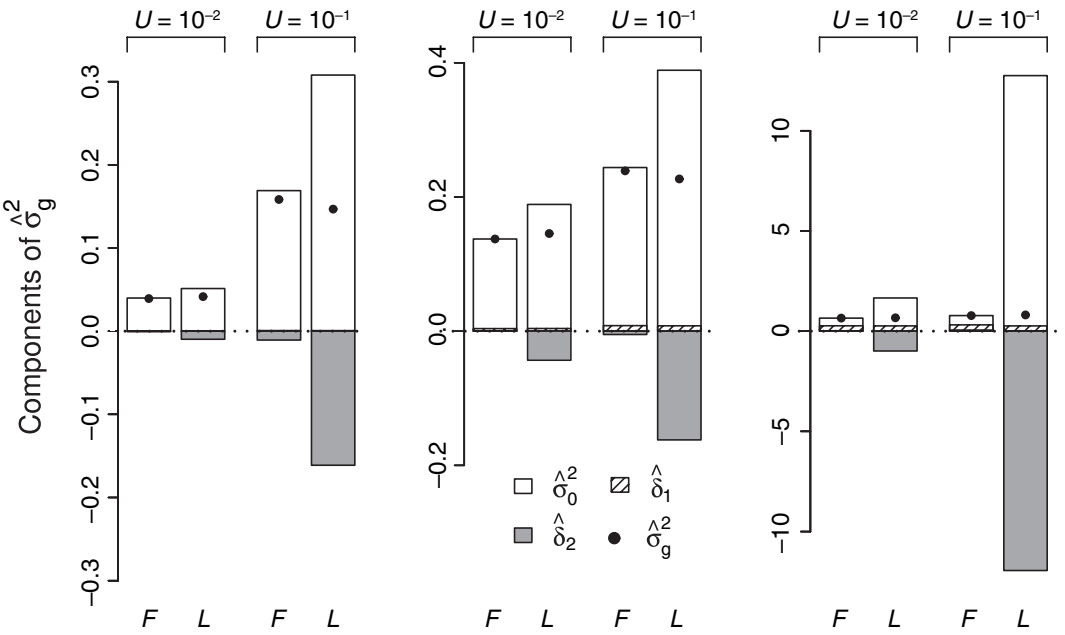


Fig. 7 Partition of the within-population genotypic variance as a function of landscape heterogeneity, mutation and recombination rates in a highly fragmented landscape of 32 populations of 25 individuals. F = free recombination between loci. L = all loci are linked on the same chromosome with a recombination probability of 0.01 between adjacent loci. In each panel, the two first bars correspond to a genomic mutation rate $U = 10^{-2}$, and $U = 10^{-1}$, respectively, for the latter two bars. White, hatched and grey bars, respectively, correspond to the mean genic variance at HWE $\hat{\sigma}^2_0$, the mean part of the genotypic variance contributed by gametic linkage disequilibrium at HWE ($\hat{\delta}^0_2$) and the mean part of the variance due to deviations of genotypic frequencies with respect to HWE ($\hat{\delta}_2$). Dots indicate the mean within-population genotypic variance $\hat{\sigma}^2_g$. Habitat heterogeneity measured by the difference in optimal phenotypic values in the two habitats increases from left to right: $\Delta\theta = 0$, $\Delta\theta = 1$ and $\Delta\theta = 3$.Migration occurs through seeds only, with $\tilde{m}_s = 0.1 \times 4/3$. Confidence intervals are plotted for $\hat{\sigma}^2_g$ only and are smaller than symbols.

lute value); (ii) the proportional contribution of $\hat{\delta}_1$ to the total genotypic variance is higher when the mutation rate is higher. Therefore, the negative effect of linkage on the evolution of the genotypic variance dominates when the mutation rate is high, and smaller within-population genotypic variance then causes less genetic load.

Discussion

Using analytical predictions and individual-centred simulations, we studied the impact of dispersal on the within-population genotypic variance, population differentiation and mean genetic load and investigated how the dispersal mode (through seeds or pollen grains) or the level of landscape heterogeneity affect the results.

Effect of seed vs. pollen dispersal on genetic load and between-population differentiation

Our analytical model predicts that, for a given net migration rate, the dispersal mode should affect neither the mean genetic load nor the differentiation between populations, provided that similar levels of within-population genotypic variance evolve for various rates of pollen and seed dispersal. For one locus, previous models have also found that spatial changes in allelic frequencies in plant populations under selection could be described as in animal models when pollen and seed dispersal parameters are aggregated in some appropriate combination (Nagylaki, 1997; Hu & Li, 2001). Our simulations confirm in part such predictions, as the effect of the dispersal mode on either the within-population variance, the between-population differentiation or the genetic load, was very small for most of the range of parameters explored. In highly heterogeneous landscapes, however, the genetic load was greater and the differentiation between populations was smaller with pollen dispersal than with seed dispersal. Selection thus better opposes the effect of gene flow when seeds disperse rather than pollen. This increased efficacy of selection is not simply due to differences in genotypic variance, as our analytical model based on a Gaussian distribution of phenotypes largely underestimates the difference in genetic load between seed and pollen dispersal (Fig. 4). Instead, differences in distribution of genotypic values before selection could be involved. This is reminiscent of results obtained by Ronce & Kirkpatrick (2001), who found that the deleterious effects of dispersal on maladaptation depended on the

relative order of migration and recombination affecting the distribution of genotypic values before selection. In highly heterogeneous landscapes, individuals originating from migrant seeds from the wrong habitat indeed represent a separate class of phenotypes (Fig. 3). For extreme values of habitat heterogeneity, most immigrants may be so distant from the local optimum that they are eliminated immediately after migration. Under the assumptions of our simulations, the genetic load is then expected to converge toward the immigration rate from the wrong habitat, i.e. to $\tilde{m}_t/2$, in the case of pure seed dispersal (Fig. 4). In the case of pollen dispersal, badly adapted immigrant alleles are counter-selected only as heterozygotes together with locally fit alleles. At values of landscape heterogeneity for which most immigrant seeds fail to recruit, many of the heterozygous individuals derived from immigrant pollen survive, which introduces maladapted genes within local populations and increases the genetic load. Effective migration rate is thus higher with pollen than with seed dispersal. The second consequence of pollen dispersal is that twice the number of selective deaths is necessary to eliminate maladapted alleles from the population just after migration than with seed dispersal. With pollen dispersal, the limit value of the genetic load when habitat heterogeneity increases is thus twice that expected with seed dispersal. Consequences for the joint evolution of seed and pollen dispersal (see Ravigné *et al.*, 2006) in presence of local adaptation loci (see Billiard & Lenormand, 2005) deserves further exploration. Our model does not consider selfing in excess that expected under random mating in a self-compatible species. Higher selfing rate would decrease the amount of gene flow achieved by pollen dispersal. Higher homozygosity due to inbred mating might also result in the distribution of genotypic values deviating more strongly from a Gaussian distribution in heterogeneous landscapes. We can therefore conjecture that increasing the selfing rate would have comparable qualitative effects on genetic load and genetic variance as increasing the contribution of seeds migration to gene flow.

Effect of seed vs. pollen dispersal on genotypic variance

The overall effect of the dispersal mode on within-population genotypic variance is relatively small in the range of parameters explored in this study. Dispersal mode has, however, strong opposite effects on different components of the genotypic variance. Genic variance is higher in the case of pollen dispersal compared with seed dispersal because of the higher effective immigration rate with pollen in heterogeneous landscapes and the more efficient removal of maladapted alleles with seed dispersal (see above). Conversely, heterozygote deficiency generated by seed dispersal inflates the genotypic variance considerably in heterogeneous landscapes. Greater differentiation between populations with seed dispersal also results in greater contribution of gametic disequilibrium to genotypic variance. Whereas only weak contributions to genotypic variance arise from gametic linkage disequilibrium (δ_1) and deviation of genotypic frequencies from HWE (δ_2) in the case of pure pollen dispersal, $\delta_1 + \delta_2$ can reach up to 80% of the total variance in very heterogeneous landscapes for pure seed dispersal. Such components of variance appear to be quickly converted in genic variance with pollen dispersal. These different effects of the dispersal mode on different components of the genotypic variance, however, largely compensate each other, leading to values of local genotypic variance very similar for pure seed or pure pollen dispersal. The robustness of this surprising outcome of our simulations remains to be confirmed by further analytical studies. Hu & Li (2001) showed that, for a trait under the control of a major locus, the ratio of pollen to seed dispersal affected the shape and position of clines in additive and dominance variance along some sharp environmental transition, as well as spatial patterns in deviation of genotypic frequencies with respect to HWE. The mechanisms underlying the specific effects of pollen and seed dispersal were not discussed explicitly. Our model considers only additive effects of alleles on the phenotype expression. Effects of seed and pollen dispersal on the expression of genotypic variance might be different in the presence of dominance interactions between alleles.

Effect of habitat heterogeneity

The global influence of the level of habitat heterogeneity on genetic load or within-population genotypic variance was much higher than the effect of the dispersal mode. Local genotypic variance increases with landscape heterogeneity in our simulations, in agreement with Yeaman & Jarvis' (2006) experimental work on populations of lodgepole pine. We also showed, both analytically and with our simulation model, that an increase in habitat heterogeneity always led to higher levels of population differentiation. This contrasts with predictions of theoretical models taking into account feedback between population dynamics and local adaptation, which have shown that beyond some critical level of habitat heterogeneity, differentiation between populations collapsed (Kirkpatrick & Barton, 1997; Ronce & Kirkpatrick, 2001; Holt *et al.*, 2003). Strong assumptions about symmetry in our simulations prevent us to witness phenomena of gene swamping from a particular habitat type. We found that the genetic load can increase and then decrease with increasing habitat heterogeneity, as observed with seed dispersal. Genetic load decreases in highly heterogeneous landscapes when selection against migrants is so strong that maladapted alleles do not introgress into the local population, which is not predicted by analytical models based on a Gaussian distribution of phenotypic values.

Effect of migration on differentiation between populations

In agreement with other studies (see, e.g. Hendry *et al.*, 2001 for theoretical predictions, and Hendry & Taylor, 2004 for empirical results), we found a negative relationship between the level of gene flow and differentiation between populations. Differentiation in our model results from both genetic drift and divergent selection in different habitats. It is opposed by gene flow and stabilizing selection within habitats. Interestingly, our model predicts that differentiation between populations for a selected character could then be larger or smaller than for a neutral character depending on the migration rate, a result also observed in simulations by Le Corre & Kremer (2003). The implications of these findings for inferences about divergent selection based on Q_{ST}–F_{ST} comparisons (e.g. Porcher *et al.*, 2006, for more references see Latta, 2006) deserve further exploration.

Effect of migration on the maintenance of genetic variation within populations

Whereas stabilizing selection tends to reduce variability within populations, levels of quantitative variation found in natural populations are usually high (for a general review see Johnson & Barton, 2005). A putative mechanism for the maintenance of variation invokes the effect of gene flow between genetically differentiated populations. Using a model of stabilizing selection, Phillips (1996) showed that migration can serve as a source of variation in the same way as mutation does, even in a homogeneous environment. This is due to genetic redundancy and local drift, which maintain different genotypes with similar phenotypes in different localities (see Goldstein & Holsinger, 1992). Mixing then increases the local genotypic variance but, as no strong evolutionary force opposes the homogenizing effect of migration, a critical migration rate is soon reached so that differentiation between populations vanishes (see also Lythgoe, 1997).

In homogeneous landscapes, we indeed showed that within-population genotypic variance reaches a maximum for very low dispersal rates, in good agreement with these previous results. In heterogeneous landscapes, however, selection opposes the effect of gene flow on differentiation more strongly. The critical migration threshold increases dramatically with habitat heterogeneity, which is in agreement with results obtained by Tufto (2000), see his Fig. 2b). Note, however, that demographical asymmetries in heterogeneous landscapes could much decrease such critical migration rate. Our simulations showed that genotypic variance before selection could be multiplied by a factor of 30 due to the effect of migration in highly heterogeneous landscapes. This suggests that migration could be a major force maintaining high levels of genotypic variance despite stabilizing selection, as suggested by the data collected by Yeaman & Jarvis (2006).

Effect of migration on the genetic load

The overall influence of migration on the genetic load is difficult to predict *a priori*. Ignoring the effect of migration on genotypic variance, our analytical model predicts that the genetic load increases with increasing dispersal. We also found that increasing gene flow has nonmonotone effects on the level of within-population genotypic variance, which, in turn, has a nonmonotone effect on the genetic load. Overall, simulations suggest that the genetic load generally increases with increasing dispersal rates. Genetic load was sometimes observed to decline with increasing migration in homogeneous landscapes, but it was largely explained by the decrease in genotypic variance when migration was above a critical threshold. In heterogeneous landscapes, we observed no instance when increasing migration improved the mean fitness in small populations, as found in species margins by previous simulation studies (Holt *et al.*, 2003; Alleaume-Benharira *et al.*, 2006). This suggests that demographic asymmetries included in the latter studies, but not in the present model, are critical to the observation of such rescue effects of dispersal. Finally, genetic load has no demographic consequences under the assumptions of our model. An important perspective of this work would be to incorporate feedbacks between population dynamics and genetics to predict the effect of seed and pollen migration for the persistence of populations in heterogeneous landscapes.

Effect of physical linkage between loci

We investigated whether physical linkage can alter the general results obtained for unlinked loci. We show that linkage does not affect the general conclusions concerning the influence of the dispersal mode. For limited habitat heterogeneity, it, however, influences the level of within-population genotypic variance and hence the genetic load. This effect is due to contrasting effects of linkage on different components of the genotypic variance, leading to an increase or a decrease in the local variance depending on the relative contribution of linkage disequilibrium to total within-population genotypic variance. This phenomenon deserves more detailed investigations.

Implications for analytical models

Analytical predictions about the effect of gene flow on the evolution of genotypic variance for polygenic characters under stabilizing selection are difficult to obtain and generally rely on strong assumptions (see Barton, 1999). Two extreme models are often considered. The infinitesimal model assumes that the local

genic variance at linkage equilibrium is constant and that most changes in genotypic variance are due to changes in linkage disequilibrium (for applications in the case of dispersal in heterogeneous landscapes see, e.g. Barton, 1999; Tufto, 2000). Conversely, some other models neglect the contribution of linkage disequilibrium to the genotypic variance altogether (see also Barton, 1999). In the range of parameters explored in our simulations, the latter approach appears more appropriate as genic variance contributed to a large part of the total genotypic variance and responded strongly to habitat heterogeneity, dispersal intensity and mode of dispersal. Gametic linkage disequilibrium, however, makes significant contribution to genotypic variance in highly heterogeneous landscapes with seed dispersal. Finally, our model also suggests that deviation of genotypic frequencies from HWE in the case of pollen and seed dispersal could significantly alter the evolution of the genotypic variance, which is little studied by available analytical approaches (see, however, Hu & Li, 2001).

In agreement with Tufto (2000), we showed that, except for high levels of landscape heterogeneity and dispersal intensity, a model assuming a Gaussian distribution of genotypic values before selection produces reasonably accurate predictions for the mean genetic load and between populations differentiation when the local variance is known. In highly heterogeneous landscapes, violation of this assumption has more important consequences: analytical models based on the Gaussian assumption largely overestimate the genetic load and underestimate differentiation between populations.

Implications for empirical studies

Altogether, we showed that the dispersal mode has a rather small influence on local adaptation, unless the landscape is very heterogeneous. It means that, in general, species with different dispersal modes are unlikely to show consistent strong differences in migration load. In the rare endangered *Centaurea corymbosa* and *Brassica insularis*, gene flow among extant populations is extremely rare (Hurtrez-Bousses, 1996; Fréville *et al.*, 2001; Wilson & Rannala, 2003) and there is evidence that loss of genetic diversity in some populations may compromise their persistence (Glémin *et al.*, 2005). Human-mediated gene flow might be considered as a management practice to improve the long-term viability of such species. In the absence of evidence for strong disruptive selection among populations (Petit *et al.*, 2001), our model suggests that the choice of transferring pollen or seeds should depend on other considerations than their putative effects on local adaptation.

Situations with strongly divergent selection between sites and relatively large gene flow are more favourable to observing large differences in migration load depending on dispersal mode. In this case, gene flow mediated by pollen is expected to be more harmful than seed flow, with regard to local adaptation. Our model then predicts that anemophilous plants with pollen dispersal over large spatial scales (like many tree species) might suffer from higher migration load than plants with reduced pollen dispersal but higher seed dispersal (e.g. autogamous plants). Some introgressions between cultivated plants and their wild relatives present a combination of high gene flow and strongly divergent selection on domestication traits (Ellstrand, 2003). Gene flow from cultivated rice *Oryza sativa* has, for instance, been shown to result in decreased fitness in wild endemic rice populations of *Oryza rufipogon* (Ellstrand, 2003). Gene flow between cultivated and wild species is, in general, mediated by pollen flow, as crop species have been selected for reduced seed dispersal during domestication. Our model predicts that pollen flow from abundant crops may cause more genetic pollution than if seeds escape from cultivated fields, even if seeds and pollen disperse at the same distance. Plant communities found on heavily contaminated soils are also candidate case studies with strong divergent selection and potentially large gene flow. Most species in such communities are indeed also found on normal soils, sometimes at a short distance from contaminated sites (Escarré *et al.*, 2000), and have developed ecotypes specialized on contrasting edaphic conditions, as documented, for instance, in *T. caerulescens* (Jimenez-Ambriz *et al.*, 2007). Studying the distribution of dispersal syndromes in such communities, as well as contrasting the pattern of differentiation for traits under divergent selection to the differentiation for nuclear and cytoplasmic neutral molecular markers (using Q_{ST} vs. F_{ST} comparisons, for instance) may help to assess the relative importance of pollen and seed migration in constraining local adaptation. Finally, our model predicts that restoration of genetic diversity through artificial gene flow in endangered species showing strong patterns of local adaptation is more likely to succeed if pollen flow is increased, rather than seed migration, due to the more difficult introgression in the latter case and to the faster breakage of linkage disequilibrium between locally maladapted alleles and beneficial migrant alleles in the former (Papa *et al.*, 2005).

Acknowledgments

Roland Vergilino helped much with the simulations at an early stage of this project. We thank Isabelle Olivieri, Joëlle Ronfort, Yannis Michalakis and Mark Kirkpatrick for advice, discussions and comments on the manucript. Financial support was provided by: (1) French Ministry of Research, through the programme ACI 'Jeune chercheur', allocated to F.R., and through the ACI 'Ecologie Quantitative', allocated to Isabelle Olivieri and Dominique Jolly; (2) CNRS through the programme 'Impact des Biotechnologies dans les agroécosystèmes' allocated to Isabelle Olivieri; (3) the European Union Fifth

Framework through the programme 'Plant dispersal' (headed by Ben Vosman); (4) the Agence Nationale pour la Recherche through the programme 'Activités humaines, dynamique et gestion de la biodiversité en milieu méditerranéen' headed by John Thompson (A-BIME, contract ANR-05-BDIV-014); and (5) the John Simon Guggenheim Foundation through support to Ruth Shaw. This is publication ISEM 2007XXX of the Institut des Sciences de l'Evolution.

References

Alleaume-Benharira, M., Pen, I.R. & Ronce, O. 2006. Geographical patterns of adaptation within a species range: interactions between drift and gene flow. *J. Evol. Biol.* **19**: 203–215.

Antonovics, J. 2006. Evolution in closely adjacent plant populations. 10. Long-term persistence of prereproductive isolation at a mine boundary. *Heredity* **97**: 33–37.

Antonovics, J. & Bradshaw, A.D. 1970. Evolution in closely adjacent plant populations. 8. Clinal patterns at a mine boundary. *Heredity* **25**: 349–362.

Bacles, C.F.E., Lowe, A.J. & Ennos, R.A. 2006. Effective seed dispersal across a fragmented landscape. *Science* **311**: 628–628.

Barton, N.H. 1986. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet. Res.* **47**: 209–216.

Barton, N.H. 1999. Clines in polygenic traits. *Genet. Res.* **74**: 223–236.

Barton, N.H. 2001. Adaptation at the edge of a species' range. In: *Integrating Ecology and Evolution in a Spatial Context*, pp. 365–392. Blackwell Science, New York.

Billiard, S. & Lenormand, T. 2005. Evolution of migration under kin selection and local adaptation. *Evolution* **59**: 13–23.

Bonnin, I., Prosperi, J.M. & Olivieri, I. 1996. Genetic markers and quantitative genetic variation in *Medicago truncutula* (Leguminosae): a comparative analysis of population structure. *Genetics* **143**: 1795–1805.

Bulmer, M.G. 1980. *The mathematical theory of quantitative genetics*. Clarendon Press, Oxford.

Bulmer, M.G. 1989. Maintenance of genetic variability by mutation-selection balance: a child's guide through the jungle. *Genome* **31**: 761–767.

Bürger, R., Wagner, G.P. & Stettinger, F. 1989. How much heritable variation can be maintained in finite populations by mutation selection balance. *Evolution* **43**: 1748–1766.

Dubois, S., Cheptou, P.O., Petit, C., Meerts, P., Poncelet, M., Vekemans, X., Lefebvre, C. & Escarre, J. 2003. Genetic structure and mating systems of metallicolous and nonmetallicolous populations of *Thlaspi caerulescens*. *New Phytol.* **157**: 633–641.

Ellstrand, N.C. 2003. *Dangerous liaisons? When cultivated plants mate with their wild relatives*. The Johns Hopkins University Press, Baltimore, MD and London.

Escarré, J., Lefebvre, C., Gruber, W., Leblanc, M., Lepart, J., Riviere, Y. & Delay, B. 2000. Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. *New Phytol.* **145**: 429–437.

Etterson, J.R. 2004a. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the great plains. *Evolution* **58**: 1446–1458.

Etterson, J.R. 2004b. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. II. Genetic architecture of three populations reciprocally planted along an environmental gradient in the great plains. *Evolution* **58**: 1459–1471.

Etterson, J.R. & Shaw, R.G. 2001. Constraint to adaptive evolution in response to global warming. *Science* **294**: 151–154.

Fréville, H., Justy, F. & Olivieri, I. 2001. Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). *Mol. Ecol.* **10**: 879–889.

Gandon, S., Capowiez, Y., Dubois, Y., Michalakis, Y. & Olivieri, I. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proc. R. Soc. Lond. B Biol.* **263**: 1003–1009.

Garcia-Ramos, G. & Kirkpatrick, M. 1997. Genetic models of adaptation and gene flow in peripheral populations. *Evolution* **51**: 21–28.

Glémin, S., Gaude, T., Guillemin, M.L., Lourmas, M., Olivieri, I. & Mignot, A. 2005. Balancing selection in the wild: testing population genetics theory of self-incompatibility in the rare species *Brassica insularis*. *Genetics* **171**: 279–289.

Goldringer, I., Prouin, C., Rousset, M., Galic, N. & Bonnin, I. 2006. Rapid differentiation of experimental populations of wheat for heading time in response to local climatic conditions. *Ann. Bot.* **98**: 805–817.

Goldstein, D.B. & Holsinger, K.E. 1992. Maintenance of polygenic variation in spatially structured populations – roles for local mating and genetic redundancy. *Evolution* **46**: 412–429.

Hall, M.C. & Willis, J.H. 2006. Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. *Evolution* **60**: 2466–2477.

Hastings, W.K. 1970. Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**: 97–109.

Hendry, A.P. & Taylor, E.B. 2004. How much of the variation in adaptive divergence can be explained by gene flow? – An evaluation using lake-stream stickleback pairs. *Evolution* **58**: 2319–2331.

Hendry, A.P., Day, T. & Taylor, E.B. 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution* **55**: 459–466.

Holt, R.D., Gomulkiewicz, R. & Barfield, M. 2003. The phenomenology of niche evolution via quantitative traits in a black-hole sink. *Proc. R. Soc. Lond. B Biol.* **270**: 215–224.

Hu, X.S. 2005. Tension versus ecological zones in a two-locus system. *Theor. Popul. Biol.* **68**: 119–131.

Hu, X.S. & Li, B. 2001. Assessment of the ratio of pollen to seed flow in a cline for genetic variation in a quantitative trait. *Heredity* **87**: 400–409.

Hu, X.S. & Li, B. 2002. Seed and pollen flow and cline discordance among genes with different modes of inheritance. *Heredity* **88**: 212–217.

Hurtrez-Bousses, S. 1996. Genetic differentiation among natural populations of the rare corsican endemic *Brassica insularis* Moris: implications for conservation guidelines. *Biol. Conserv.* **76**: 25–30.

Jimenez-Ambriz, G., Petit, C., Bourrie, I., Dubois, S., Olivieri, I. & Ronce, O. 2007. Life history variation in the heavy metal tolerant plant *Thlaspi caerulescens* growing in a network of contaminated and noncontaminated sites in southern France: role of gene flow, selection and phenotypic plasticity. *New Phytol.* **173**: 199–215.

Johnson, T. & Barton, N. 2005. Theoretical models of selection and mutation on quantitative traits. *Proc. R. Soc. Lond. B Biol.* **360**: 1411–1425.

Kawecki, T.J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecol. Lett.* **7**: 1225–1241.

Keller, M., Kollmann, J. & Edwards, P.J. 2000. Genetic introgression from distant provenances reduces fitness in local weed populations. *J. Appl. Ecol.* **37**: 647–659.

Kirkpatrick, M. & Barton, N.H. 1997. Evolution of a species range. *Am. Nat.* **150**: 1–23.

Lande, R. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution* **46**: 381–389.

Latta, R.G. 2006. Integrating patterns across multiple genetic markers to infer spatial processes. *Landsc. Ecol.* **21**: 809–820.

Le Corre, V. & Kremer, A. 2003. Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* **164**: 1205–1219.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**: 183–189.

Lythgoe, K.A. 1997. Consequences of gene flow in spatially structured populations. *Genet. Res.* **69**: 49–60.

MacNair, M.R. 1987. Heavy-metal tolerance in plants: a model evolutionary system. *Trends Ecol. Evol.* **2**: 354–359.

McCauley, D.E. 1997. The relative contributions of seed and pollen movement to the local genetic structure of *Silene alba*. *J. Hered.* **88**: 257–263.

McKay, J.K., Christian, C.E., Harrison, S. & Rice, K.J. 2005. ''How local is local?'' – a review of practical and conceptual issues in the genetics of restoration. *Restor. Ecol.* **13**: 432–440.

McNeilly, T. 1968. Evolution in closely adjacent plant populations. 3. *Agrostis tenuis* on a small copper mine. *Heredity* **23**: 99–108.

McNeilly, T. & Antonovics, J. 1968. Evolution in closely adjacent plant populations. 4. Barriers to gene flow. *Heredity* **23**: 205–218.

Moran, G.F. 1992. Patterns of genetic diversity in Australian tree species. *New For.* **6**: 49–66.

Morgan, A.D., Gandon, S. & Buckling, A. 2005. The effect of migration on local adaptation in a coevolving host–parasite system. *Nature* **437**: 253–256.

Nagylaki, T. 1997. The diffusion model for migration and selection in a plant population. *J. Math. Biol.* **35**: 409–431.

Ouborg, N.J., Piquot, Y. & Van Groenendael, J.M. 1999. Population genetics, molecular markers and the study of dispersal in plants. *J. Ecol.* **87**: 551–568.

Papa, R., Acosta, J., Delgado-Salinas, A. & Gepts, P. 2005. A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica. *Theor. Appl. Genet.* **111**: 1147–1158.

Petit, R.J., Kremer, A. & Wagner, D.B. 1993. Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theor. Appl. Genet.* **87**: 122–128.

Petit, C., Fréville, H., Mignot, A., Colas, B., Riba, M., Imbert, E., Hurtrez-Bousses, S., Virevaire, M. & Olivieri, I. 2001. Gene flow and local adaptation in two endemic plant species. *Biol. Conserv.* **100**: 21–34.

Phillips, P.C. 1996. Maintenance of polygenic variation via a migration-selection balance under uniform selection. *Evolution* **50**: 1334–1339.

Porcher, E., Giraud, T. & Lavigne, C. 2006. Genetic differentiation of neutral markers and quantitative traits in predominantly selfing metapopulations: confronting theory and experiments with *Arabidopsis thaliana*. *Genet. Res.* **87**: 1–12.

Ravigné, V., Olivieri, I., Gonzalez-Martinez, S.C. & Rousset, F. 2006. Selective interactions between short-distance pollen and seed dispersal in self-compatible species. *Evolution* **60**: 2257–2271.

Rehfeldt, G.E., Tchebakova, N.M., Parfenova, Y.I., Wykoff, W.R., Kuzmina, N.A. & Milyutin, L.I. 2002. Intraspecific responses to climate in *Pinus sylvestris*. *Glob. Chang. Biol.* **8**: 912–929.

Ronce, O. & Kirkpatrick, M. 2001. When sources become sinks: migrational meltdown in heterogeneous habitats. *Evolution* **55**: 1520–1531.

Rousset, F. 2004. *Genetic structure and selection in subdivided populations*. Princeton University Press, Princeton, NJ.

Savolainen, O., Bokma, F., Garcia-Gil, R., Komulainen, P. & Repo, T. 2004. Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of *Pinus sylvestris* to climatic changes. *For. Ecol. Manage.* **197**: 79–89.

Shaw, F.H., Geyer, C.J. & Shaw, R.G. 2002. A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* **3**: 453–463.

Spichtig, M. & Kawecki, T.J. 2004. The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *Am. Nat.* **164**: 70–84.

Tufto, J. 2000. Quantitative genetic models for the balance between migration and stabilizing selection. *Genet. Res.* **76**: 285–293.

Wilson, G.A. & Rannala, B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**: 1177–1191.

Yeaman, S. & Jarvis, A. 2006. Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proc. R. Soc. Lond. B Biol.* **273**: 1587–1593.

Supplementary material

The following supplementary material is available for this article:

This material is available as part of the online article from:

XXXXXXXX

http://www.blackwell-synergy.com/doi/abs/10.1111/j.1420-9101.2007.01442.x

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Received 16 July 2007; revised 3 September 2007; accepted 6 September 2007

APPENDIX A: **Changes in mean phenotype, genotypic variance within and between populations along the life cycle**

We assume that genotypic values are distributed normally after seed migration and before selection in each population with variance $V_{gi} \approx \sigma_g^2$. Throughout we use superscripts (1) to (4) to designate variable of interest measured at different stages of the life cycle (*e.g.* $V_{gi}^{(1)}$ is the genotypic variance in population *i* sampled just after selection and before density regulation). We are interested in computing variable of interest at equilibrium between drift, selection and migration. Once such equilibrium has been reached the expected mean genotypic value at the scale of the metapopulation does not change along the life cycle and equation (5) of the main text holds at each stage.

Under such assumptions, after selection, we have:

$$\bar{W}_i = \int \tilde{W}_i(g) \frac{1}{\sqrt{2\pi V_{gi}}} \exp\left(\frac{-\left(g - \bar{G}_i\right)^2}{2V_{gi}} \right) dg = \sqrt{\frac{\omega^2}{\sigma_s^2 + V_{gi}}} \exp\left(\frac{-\left(\bar{G}_i - \theta_i\right)^2}{2\left(\sigma_s^2 + V_{gi}\right)} \right) \qquad \text{(A1)}$$

and

$$\bar{G}_i^{(1)} = \int g \frac{\tilde{W}_i(g)}{\bar{W}_i} \frac{1}{\sqrt{2\pi V_{gi}}} \exp\left(\frac{-\left(g - \bar{G}_i\right)^2}{2V_{gi}} \right) dg = \frac{\sigma_s^2}{\sigma_s^2 + V_{gi}} \bar{G}_i + \frac{V_{gi}}{\sigma_s^2 + V_{gi}} \theta_i \,. \qquad \text{(A2)}$$

Similarly, we obtain the genetic variance after selection as:

$$V_{gi}^{(1)} = \frac{\sigma_s^2}{\sigma_s^2 + V_{gi}} V_{gi} \qquad \text{(A3)}$$

(for similar results see *e.g.* Bulmer 1980 page 151, equation 9.23).

Thus $E\left[\bar{G}_i^{(1)}\right] = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_g^2} E\left[\bar{G}_i\right] + \frac{\sigma_g^2}{\sigma_s^2 + \sigma_g^2} \theta_i$, (A4)

$$\sigma_g^{2(1)} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_g^2}\sigma_g^2 \,, \tag{A5}$$

and $$\sigma_y^{2(1)} = \left(\frac{\sigma_s^2}{\sigma_s^2 + \sigma_g^2}\right)^2 \left(\sigma_y^2 + \left(\frac{\sigma_g^2}{\sigma_s^2}\right)^2 \sum_{i=1}^{n} \frac{\theta_i^2}{n} + 2\left(\frac{\sigma_g^2}{\sigma_s^2}\right) \sum_{i=1}^{n} \frac{E\left[\bar{G}_i\right]\theta_i}{n}\right). \tag{A6}$$

During density regulation, N adults are chosen at random from the surviving juveniles in each population to produce the next generation. The mean and variances of genotypic values among adults within and between populations then become:

$$E\left[\bar{G}_i^{(2)}\right] = E\left[\bar{G}_i^{(1)}\right], \tag{A7}$$

$$\sigma_g^{2(2)} = \left(1 - \frac{1}{N}\right)\sigma_g^{2(1)} \,, \tag{A8}$$

$$\text{and } \sigma_y^{2(2)} = \sigma_y^{2(1)} + \frac{n-1}{nN}\sigma_g^{2(1)} \,. \tag{A9}$$

We assume that all loci recombine freely. After gametogenesis, pollen dispersal and syngamy, we have:

$$\bar{G}_i^{(3)} = \left(1 - m_p\right)\bar{G}_i^{(2)} + m_p\left(\frac{\bar{G}_i^{(2)}}{2} + \frac{1}{n-1}\sum_{j=1, j\neq i}^{n} \frac{\bar{G}_j^{(2)}}{2}\right), \tag{A8}$$

which can be rearranged as:

$$\bar{G}_i^{(3)} = \left(1 - \tilde{m}_p\right)\bar{G}_i^{(2)} + \tilde{m}_p \frac{\bar{G}_i^{(2)} + \bar{G}}{2} \text{ where } \tilde{m}_p = \frac{n}{n-1} m_p \,. \tag{A9}$$

$$\text{Thus } E\left[\bar{G}_i^{(3)}\right] = \left(1 - \frac{\tilde{m}_p}{2}\right)E\left[\bar{G}_i^{(2)}\right], \tag{A10}$$

$$\text{and } \sigma_y^{2(3)} = \left(1 - \frac{\tilde{m}_p}{2}\right)^2 \sigma_y^{2(2)} \,. \tag{A11}$$

To obtain the within-population genetic variance among zygotes after pollen dispersal, we first compute the mean variance in genotypic values among pollen

grains produced within each population σ_γ^2. Maternal and paternal genes are recombined in the newly produced gametes, which affects associations between genes, present at the same locus, or at different loci. We then have:

$$\sigma_\gamma^2 = \frac{\sigma_0^{2(2)}}{2} + \frac{\delta_1^{(2)}}{4} + \frac{\delta_2^{(2)}}{4} - \frac{\rho^{(2)}}{4} = \frac{\sigma_0^{2(2)}}{4} - \frac{\rho^{(2)}}{4} + \frac{\sigma_g^{2(2)}}{4}, \tag{A12}$$

where σ_0^2 is the expected genic variance at Hardy-Weinberg equilibrium (HWE) and linkage equilibrium, δ_1 is the part of the variance contributed by gametic linkage disequilibrium at HWE, δ_2 is the part of the variance due to deviations of genotypic frequencies with respect to HWE in the parental population (see main text for complete expressions), and the last term is:

$$\rho = \frac{2}{n}\sum_{i=1}^{n}\sum_{k=1}^{l} E_i\left[\left(X_k^p\right)\left(X_k^o\right)\right] - E_i\left[\tilde{X}_k\right]^2. \tag{A13}$$

Note that at HWE, $\rho = 0$. Negative correlation between allelic values of paternal and maternal copies at the same locus makes this term larger, while positive correlation makes it smaller than the genic variance at HWE. The same expression holds for variance among ovules. After pollen dispersal and syngamy with non-dispersed ovules, we have:

$$\sigma_g^{2(3)} = 2\sigma_\gamma^2 + \frac{\tilde{m}_p}{2}\left(1 - \frac{\tilde{m}_p}{2}\right)\sigma_y^{2(2)} = \frac{\sigma_0^{2(2)}}{2} - \frac{\rho^{(2)}}{2} + \frac{\sigma_g^{2(2)}}{2} + \frac{\tilde{m}_p}{2}\left(1 - \frac{\tilde{m}_p}{2}\right)\sigma_y^{2(2)}. \tag{A14}$$

At HWE, (A14) leads to similar equations for change in genetic variance to equations (4-6) in Tufto (2000).

After seed migration, we similarly have:

$$\bar{G}_i^{(4)} = \left(1 - \tilde{m}_s\right)\bar{G}_i^{(3)} + \tilde{m}_s\bar{\bar{G}}^{(3)} \text{ where } \tilde{m}_s = \frac{n}{n-1}m_s, \tag{A15}$$

$$E\left[\bar{G}_i^{(4)}\right] = \left(1 - \tilde{m}_s\right)E\left[\bar{G}_i^{(3)}\right], \tag{A16}$$

$$\sigma_y^{2(4)} = \left(1-\tilde{m}_s\right)^2 \sigma_y^{2(3)} \tag{A17}$$

and $\sigma_g^{2(4)} = \sigma_g^{2(3)} + \tilde{m}_s \left(2-\tilde{m}_s\right) \sigma_y^{2(3)}$. (A18)

At equilibrium, $E\left[\overline{G}_i^{(4)}\right] = E\left[\overline{G}_i\right]$, $\sigma_g^{2(4)} = \sigma_g^2$, and $\sigma_y^{2(4)} = \sigma_y^2$. (A19)

References:

Bulmer, M.G. 1980. The mathematical theory of quantitative genetics. Clarendon Press, Oxford.

Tufto, J. 2000. Quantitative genetic models for the balance between migration and stabilizing selection. *Genet. Res*. **76**: 285-293.

APPENDIX B: **Method for the analysis of simulations results**

The simulations are initiated by randomly assigning alleles to individuals. Allelic values are drawn in some arbitrary distribution and then standardized to ensure that, within each population, the initial expected mean trait value fits the local optimum and the initial expected genotypic variance is close to a typical value of $0.5\sigma_e^2$. For each set of parameters, we ran a single long simulation for 500 000 generations. The first 50 000 generations were discarded to eliminate transient dynamics from consideration. The remaining temporal sequence was analyzed in order to obtain estimates and confidence intervals for variables of interest using the batch method for Markov chains (Hastings 1970) as follows. Variables of interest were recorded before selection every 10 generations. Preliminary visual inspections showed that the mean within-population genotypic variance $\overline{V}_g$ had the slowest rate of evolution, with the strongest level of temporal autocorrelation. We therefore used this variable to define the optimal batch length. The temporal sequence was split into batches of t generations. $\overline{V}_g$ was averaged within each batch to provide separate estimations of σ_g^2 . We used a minimum number of ten independent batches to build our estimate. The suitable batch length corresponds to the batch length t such that the correlation of means between adjacent batches was minimized and not significantly different from 0 (correlations are tested using the Pearson method with R stats library, with a type I error of 5%). Means in adjacent batches can then be considered as approximately independent. An estimate for σ_g^2 and its confidence interval were then computed from the means for the different batches. The same batch length is then used to estimate components of the within-population genotypic

variance σ_0^2, δ_1 and δ_2, the divergence between populations σ_y^2 and the mean genetic load λ. We checked the robustness of the method by running two independent replicates for many different parameters sets and obtained estimates that did not differ significantly (results not shown). For some parameter sets, characterized in particular by a low but non null migration rate, this method however performed poorly due to a very long time for reaching an apparent equilibrium (e.g. in some simulations, the equilibrium had not been reached after 3.10^6 generations). In such cases, we were unable to compute reliable estimates of variables of interest with a meaningful confidence interval and did not include the corresponding results in our figures.

Reference:

Hastings, W. K. 1970. Monte-Carlo sampling methods using markov chains and their applications. *Biometrika*. **57**: 97-109.

CHAPITRE 2

RÉPONSE À COURT TERME ET À LONG TERME

DU FARDEAU GÉNÉTIQUE À LA FRAGMENTATION

DANS LES PAYSAGES HÉTÉROGÈNES

Manuscrit en préparation

Short-term and long-term response of genetic load to fragmentation of heterogeneous landscapes

Auteurs: S. Lopez, F. Rousset et O. Ronce

Adresse: Institut des Sciences de l'Evolution de Montpellier (UM2-CNRS)

Université Montpellier , Montpellier cedex 5, France

Résumé: La modélisation de génomes complexes comprenant des locus soumis à une sélection spatialement homogène et d'autres à une sélection spatialement hétérogène permet d'étudier l'évolution conjointe du fardeau de mutation, du fardeau de dérive et du fardeau de migration, en réponse à la fragmentation des populations. La migration a des effets antagonistes sur la dynamique des mutations aux deux types de locus étudiés. L'intensité de la migration minimisant le fardeau génétique total est compatible avec la règle classique "*One-Migrant-Per-Generation*". Nous montrons que les dynamiques des fardeaux liés aux pressions de sélection homogène ou hétérogène dans l'espace ne sont pas indépendantes.

L'étude de la réponse transitoire des composantes du fardeau à une diminution des flux génétiques entre les populations montre que les petites populations subissent une augmentation du fardeau génétique global, tandis que la réponse des grandes populations dépend de l'intensité de la sélection spatialement hétérogène. La dynamique temporelle du fardeau dépend également des paramètres mutationnels.

Short-term and long-term response of genetic load to fragmentation of heterogeneous landscapes

INTRODUCTION

Human societies have been modifying natural landscapes for a very long time, in particular through their agricultural practices (e.g. Lentfer & Torrence 2007). More recently, human activities such as intensive agriculture or the improvement of communication networks have led to increasing levels of landscapes fragmentation (Zhang & Guindon 2005), which has become a threat for wildlife. As discussed in Fahrig (2003), the expression *landscape fragmentation* has been increasingly used to describe habitat loss. Habitat fragmentation is however a complex phenomenon, which is generally characterized by an increase in the number of patches that can support a population. Patch isolation also increases while a decrease in mean patch size is observed in addition to habitat loss (Fahrig 2003). Human-induced landscape alterations can be very fast in comparison to evolutionary processes. For instance, Cayuela et al (2006) showed that montane forest cover in the Mexican Chiapas has been reduced by 50% in 25 years while the mean size of forest fragments was divided by 7, patch number was multiplied by 3 and isolation between fragments (as measured by mean proximity index) increased 70 fold.

Among the numerous consequences of habitat fragmentation, theoretical studies have investigated its long-term genetic consequences on fitness (e.g. Whitlock 2002, Glémin et al 2003, Roze & Rousset 2003, Roze & Rousset 2004, Theodorou & Couvet 2006a, Bouchy et al 2005). These models consider mutations with constant deleterious effects and variable levels of recessivity and predict how mean fitness at equilibrium varies with population fragmentation (but see Bouchy et al 2005 and Theodorou & Couvet 2006b for predictions about the short-term consequences of fragmentation). The frequency of deleterious mutations at equilibrium and the

associated genetic load are greatly sensitive to several parameters that determine the level of structuration of the populations (i.e. demes size, demes number and level of migration between them). The local deme size appears to have the most important effect (Theodorou & Couvet 2006a). Higher levels of population structuration generally induce a higher homozygosity at a local scale, allowing the purging of recessive deleterious alleles. However, it also leads to lower levels of local genetic variance, which in turn decreases the efficacy of selection (Roze & Rousset 2003). Increase in local genetic drift can lead to the accumulation of mutations of small effects, as showed in Higgins & Lynch's (2001) simulation study. When structuration is very important, the mean equilibrium load is generally expected to be larger than in undivided populations (Roze & Rousset 2004).

A lot of experimental studies investigated the consequences of habitat fragmentation on plant populations. Direct demographic consequences of fragmentation have been amply documented, such as increased side effects on recruitmernt (Ohara et al 2006), or decreasing reproductive performance in fragments due to lack of pollinators or compatible pollen (see a review in Aguilar et al 2006). Fewer empirical studies have investigated the longer term genetic consequences of fragmentation. In many plant populations, fragmentation is too recent to have already left a strong signature in patterns of molecular diversity (e.g. Galeuchet et al 2005, but see Honnay et al 2007, Willi et al 2007, Leimu et al 2006). Effect of fragmentation on gene flow, through seeds or pollen, can be complex (Young et al 1996; White et al 2002; Mix et al 2006, Ouborg et al 2006). The effects of fragmentation on genetic load has been studied by examining the performance of plants grown in a common environment (for reviews see e.g. Young et al 1996 or Honnay & Jacquemyn 2006). Most of such studies focused on the impact of reduced population size, comparing the performance of plants from large and small populations. Reduced levels of local genetic diversity in small populations were in general found to be

associated with higher genetic load and poorer fitness (e.g. Fischer & Mathies 1998, Fischer et al 2000, Hooftman et al 2003, Fischer et al 2003, see review in Leimu et al 2006). In agreement with theoretical predictions (Bataillon and Kirkpatrick 2000), Paland and Schmid (2003) showed that the architecture of the genetic load contributed by recessive mutations in fragmented populations depends critically on the size of the demes. When demes are small, most of the genetic load is due to drift load contributed by fixed alleles of small effects, whereas demes larger than 200 individuals rather suffer from inbreeding depression due to segregating mutations with larger effects. Empirical evidence suggests that effect of purging in small populations is very weak (Byers & Waller 1999), in agreement with theoretical predictions (Glémin 2003). A small number of empirical studies have investigated how increasing isolation interacts with reduced population size in molding the genetic load in fragmented populations (see Ouborg et al 2006 for a review). Even small values of the migration rate seem to be able to enhance the viablity of fragmented populations through a genetic rescue effect (Richards 2000, Ingvarsson 2001, Tallmon et al 2004), in particular for low deme size (Willi & Fischer 2005).

Empirical studies have however shown that the consequences of increasing isolation due to landscape fragmentation may not be always negative, as suggested by numerous evidence for outbreeding depression accompanying gene flow (Montalvo and Ellstrand 2001, Waser et al 2000, see review in Tallmon et al 2004). Local adaptation in particular is one potential cause of outbreeding depression (Lynch 1991, Edmands and Timmerman 2003). Local adaptation describes the better performance of local genotypes than foreign genotypes in their their site of origin (Kawecki and Ebert 2004). It implies that mutations have different effects in different environments (Lenormand 2002). Local adaptation patterns in plants have been documented in many situations, at a broad range of spatial scales (e.g. Schmitt & Gamble 1990, Joshi et al 2001; for reviews, see Linhart & Grant 1996 or Hufford & Mazer 2003). Both theoretical

(Garcia-Ramos & Kirkpatrick 1997, Tufto 2000, Hendry et al 2001) and empirical (e.g. Hendry & Taylor 2004) studies showed that migration can impair adaptive divergence and increase the genetic load (Lopez et al 2007). In a recent experimental study, Vergeer et al (2004) found that both the size and the distance of the population of origin of migrants affected the efficiency of the genetic rescue effect. Even if the introduction of foreign genotypes can have immediate beneficial effects in a population, outbreeding depression can occur in later generations (Keller et al 2000). Thus, migration can have contrasted fitness consequences and the reality of the genetic rescue effect is still uncertain when local adaptation patterns are present (Tallmon et al 2004). The «One Migrant Per Generation» (OMPG) rule states that between one and ten immigrants each generation can prevent the increase of the local genetic load through drift or inbreeding, while still allowing for population divergence leading to patterns of local adaptation (Mills & Allendorf 1996, but see Bouchy et al 2005 for different recommendations when the total number of populations is small).

Overall, it seems that landscape fragmentation can have contrasted consequences for the fitness of local populations. Previous theoretical studies however fail to encapsulate this complexity by (i) studying separately the phenomena of heterosis and local adaptation, which are deeply entangled in empirical studies of fragmentation, (ii) by focusing exclusively on long-term effects while most empirical studies document effects in recently fragmented landscapes. Here, we extend previous theoretical work in two directions, using individual-centered simulations: (i) we consider the effect of fragmentation on complex genomes by investigating how unconditionnally deleterious and antagonistic mutations interact and differentially contribute to the genetic load in a fragmented population and (ii) we investigate both the equilibrium patterns of the genetic load and its transient dynamics in the generations following the onset of fragmentation. This allows us to contrast the short- and long-term consequences of

habitat fragmentation. We also put stronger emphasis on the effect of migration on the genetic load, by simulating landscapes with different connectivity and a fragmentation event characterized by a progressive reduction of the migration rate. We show that the consequences of habitat fragmentation are in general negative for metapopulations composed of very small demes even when many loci are under heterogeneous selection. However, metapopulations made of larger demes can benefit from increased isolation coming with fragmentation when divergent selection across the landscape affects a fraction of loci in the genome, both in the short and in the long terms.

METHODS

Landscape structure and life cycle

We rely on individual-centered multilocus simulations of a patchy population. There are n discrete patches of individuals, each of size N, exchanging juveniles at rate m according to an island model. Both deme size and migration rate are varied. We simulate highly or slightly spatially fragmented populations, with {n=32; N=25} and {n=4; N=200} respectively, so that total population size remains constant. We also change the level of landscape fragmentation by using several values for the migration rate ($m \in 0, 10^{-3}, 10^{-2}, 10^{-1}, 2.10^{-1}$). For most simulations migration is allowed only through juveniles, so that we can compare our results to analytical predictions (Roze & Rousset 2004). Generations are discrete and non- overlapping. The order of the events in the life-cycle is: 1) selection acting via juvenile survival, 2) density regulation to a constant number of adults N, 3) gametogenesis and local syngamy, 4) juvenile migration. Juvenile production is Poisson distributed with expectation large enough so that deme size remains constant after selection, whatever the mean juvenile survival is. Our model

could therefore apply to a hermaphroditic annual plant species without a seed bank and with seed dispersal but no pollen dispersal between patches, although it could also apply to some animal species with strict juvenile dispersal. In some simulations we consider the specific effect of pollen dispersal instead of seed dispersal. For these simulations, the pollen dispersal rate was adjusted so that it achieved the same level of gene flow as seeds, ie. $m_P = 2m_S$.

Genetic architecture

Demes are equally distributed among two different habitats, and we distinguish two types of diallelic loci. For type A loci, each allele is deleterious in one habitat and advantageous in the other, exhibiting antagonistic effects. For type D loci, mutational effects are habitat-independent, the deleterious allele being equally so in both habitats. The dominance coefficient of the deleterious allele on the advantageous one is h for both types of loci. The selection coefficient is s for any deleterious effect in the homozygous state. Juvenile survival of an individual in habitat x (W_x) is the product of the contribution of all loci across the genome. Thus, juvenile surviving probability is computed as $W_x = W^D W_x^A$, with $W^D = (1-hs)^{n_{01}^D} (1-s)^{n_{11}^D}$ and $W_x^A = (1-hs)^{n_{01}^A} (1-s)^{n_{xx}^A}$,where n_{01}^D and n_{11}^D are the number of type D loci in heterozygous and homozygous state, respectively, n_{01}^A is the number of type A loci in heterozygous state for the locally deleterious allele and n_{xx}^A the number of type A loci homozygous for this same deleterious allele (i.e. $n_{xx}^A = n_{00}^A$ in habitat 1 and $n_{xx}^A = n_{11}^A$ in habitat 0). Mean genetic load L is computed across the whole population as $L = 1 - \bar{W}$, where $\bar{W}$ is the mean of W_x taken over all juvenile present in each habitat after reproduction and dispersal and before selection. Similarly, $\bar{W}_D$ and $\bar{W}_A$ are the mean of the fitness components W^D and W_x^A . We also compute the mean load per locus of each

type, defined as:

$l_D = 1 - \sqrt[n_D]{\bar{W}_D}$ and $l_A = 1 - \sqrt[n_A]{\bar{W}_A}$, where n_D and n_A are the number of type D and type A loci, respectively, in the genome.

We studied genomes of 1000 selected loci differing with respect to the proportion p_A of type A loci ($p_A \in$ 0%, 5%, 10%, 20%, 50%). All genomes are structured as 10 pairs of chromosomes, each of them containing 100 selected loci with a recombination probability of 0.01 between two adjacent loci. In most simulations, types D and A loci are randomly located across the genome. A genetic architecture such that local adaptation loci are grouped together in the same area of the genome can be favored by selection (for a theoretical study about chromosomic inversions, see Kirkpatrick & Barton 2006). In some simulations, we investigated the effect of aggregation of type A loci. Genomes with aggregated loci where created as follows. The entire genome is first considered as a single vector. The proportion p_A of type A loci determines the probability that the first locus in the genome is of type A. While all type A loci have not been placed in the genome, the *i*-th locus is of type A with probability

$p = \frac{N_{A\,remaining}}{N_{tot\,remaining}} + \rho X (1 - \frac{N_{A\,remaining}}{N_{tot\,remaining}})$, with X=0 if (*i-1*)-th locus is of type D and 1 otherwise. $N_{A\ remaining}$ is the number of type A loci that have not been placed yet and $N_{tot\ remaining}$ is the total number of loci that have not been assigned, and ρ is the expected correlation between the identity of two adjacent loci and is set to a value of 0.9. When $N_{tot\ remaining}$ =0, p=0. When all 1000 positions have been assigned an identity, chromosomes are created simply by separating loci 100 by 100, so that there are ten chromosomes of 100 loci, as for random genomes. We created several hundreds of aggregated loci using this method. For each of these genomes, the real correlation between two adjacent loci differed slightly from ρ. The simulation were run

with the genome with observed ρ closest to 0.9. Mutations and reverse mutations occur independently at each locus on each gamete with probability μ and v respectively. When a new mutation toward allelic state 1 occurs at a given gene copy (with probability μ), the mutated gene receives allele 1, irrespective of its previous value. Thus, the realized number of changes in allelic state depends on μ and v, but also on allelic frequency at each locus. Throughout we assume $\mu=v=5.10^{-4}$ (for estimates of genomic mutation rates, see Shaw et al 2002). New mutations generally exhibit a negative correlation between s and h (e.g. Phadnis & Fry 2005); hence we considered that mutations are either codominant and moderately deleterious (h=0.5 and s=0.01) or recessive and very deleterious (h=0 and s=0.1). These two types also represent extreme cases that are often distinguished for their contrasted effect on genetic load and inbreeding depression (i.e. very deleterious and recessive mutations or mildly deleterious and codominant mutations).

Simulation protocols

Two types of simulations have been run, depending on the question addressed:

- Structured populations at equilibrium

We first ran simulations with constant demographic parameters in order to clarify the response of the genetic load to population structuration for complex genomes and investigate whether some interactions can appear between different types of loci. For each set of parameters, an initial population is created. We run two replicates for each set of parameters, differing in the initial conditions. In the first replicate, individuals are initially free of deleterious mutations at both types of loci (thus exhibit no genetic load), while the second replicate is run with initial frequencies of deleterious mutations of at least 1.25-fold that observed after 3000 generations in the first replicate. Each replicate lasts for 8000 generations.

We record L, l_D, l_A and alleles frequencies every ten generations. We thus obtain two temporal sequences for these variables, the first starting from below the putative equilibrium allele frequencies and the second from above. We consider that an equilibrium has been reached when for both types of loci the mean allele frequencies in each replicate have crossed each other at least once. The subsequent generations of the two replicates are then pooled and confidence intervals are obtained using the batch method for Markov chains (Hastings 1970), with batch length determined as in Lopez et al (2007).

- Fragmentation scenario

We also ran simulations mimicking a landscape fragmentation event through a progressive decrease in the migration between demes. We studied the response of the genetic load to a fragmentation event in the specific case where the initial migration rate among demes is high, i.e. m=20%. We first ran a long simulation following the equilibrium-simulation protocol and, after 3000 generations, started recording population state each 500 generations (we had previously checked that 3000 generations were sufficient to reach an equilibrium for the sets of parameters used with this simulation protocol, and that 500 generations between two records were also sufficient to prevent pseudoreplication). We thus obtained 30 equilibrium metapopulations used as replicates to simulate a fragmentation event. For each of these replicates, a new simulation was launched. After 50 generations, the migration rate was reduced by 1% of its initial value each generation, thus reaching 0 after 100 generations of fragmentation. The transient dynamics was further followed for 200 generations. The 30 replicates were used to compute means and confidence intervals for the different variables monitored, as a function of time since the onset of fragmentation.

RESULTS

Equilibrium simulations

Figure 1 presents the equilibrium genetic load as a function of the migration rate, the proportion of loci with unconditionally deleterious mutations (type D loci) or antagonistic effects (type A loci), landscape structure (i.e. demes number and size) and types of mutations (i.e. codominant and slightly deleterious or recessive and strongly deleterious mutations). It also shows analytical predictions from Roze & Rousset 2004 which can be compared to our results in absence of locus with antagonistic effects (see below). Note that, in absence of migration, type D and type A loci are equivalent. Thus, the proportion of the different types of loci in the genome has no influence on the genetic load when m=0.

Influence of patch size

The comparison between left and right panels in Figure 1 shows that the genetic load is higher when demes are smaller, irrespective of genome organisation or type of mutations. The difference between large and small populations is greater when the migration rate is low, because of the higher sensitivity of small populations to local genetic drift.

Influence of mutations type

The comparison between top and bottom panels in Figure 1 shows that the genetic load is higher when mutations are codominant and slightly deleterious rather than recessive and highly deleterious. The selection coefficient has little influence on genetic load in the range (s=10^{-2} or s=10^{-1}) that we consider (Roze and Rousset 2004, see their Figures 4 and 9). Increasing the dominance coefficient increases the genetic load, for the well-known reason that only one selective death is sufficient to eliminate two copies of a recessive deleterious allele, where two deaths are necessary for codominant mutations.

Influence of migration rate in absence of type A loci

In absence of loci with habitat-specific effects (type A loci), the genetic load is a decreasing function of migration for the two types of mutations we used. The effect of migration is more striking in small demes. Our results are in good qualitative agreement with analytical predictions from Roze & Rousset (2004) for extreme values of the dominance coefficient *h*. It should be noted that the same authors predict non-monotone relationships between genetic load and migration for intermediate levels of the dominance coefficient *h* (Figure 4 in Roze & Rousset 2004). The quantitative agreement between their analytical predictions and our simulation results is good for large migration rates but very bad when the migration rate is lower than the selection coefficient, as predicted by Roze & Rousset (2003).

Influence of the presence of type A loci

When a fraction of loci are under heterogeneous selection (type A loci), the genetic load is increased compared to the case where they are absent (compare stripped to empty bars in Figure 1). The effect of type A loci on the genetic load generally increases with migration. The genetic load can then vary non-monotonically with migration because, while decreasing the load associated with type D loci (as seen above), higher population-mixing introduces maladapted alleles at type A loci, which enhances the load associated to these loci. For low migration rates ($m_s \leq 10^{-2}$), the genetic load at equilibrium is then generally a decreasing function of migration, while for higher migration, a further increase in migration rate can lead to an increase of the genetic load, particularly when deme size is high and for recessive and highly deleterious mutations. The resulting optimal number of migrants is about 2, except for very small demes when mutations are codominant and slightly deleterious, where it could reach a value greater than 5 because of the very high level of local genetic drift in this case.

Interactions between type A loci

As seen above, the equilibrium genetic load is increased when a fraction of loci is subject to spatially heterogeneous selection. However, the genetic load is not a monotone function of the proportion of type A loci in the genome. Figure 1 clearly shows that in most cases the mean genetic load increases with the proportion of type A loci when it is small and then decreases when it reaches some threshold value. To better understand the behaviour of types D and A loci, Figure 2 shows the load per locus of each type. The load per type A locus generally decreases with increasing proportion of type A loci in the genome (Figure 2). This happens because migration and divergent selection among habitats generate strong positive linkage disequilibrium among type A loci. As a result, selection eliminates badly adapted migrant alleles in sets, which allows more efficient purging and smaller genetic load at each locus. For high proportions of type A loci in particular, introgression can be prevented because most immigrants have very poor fitness and die just after migration. Then, a further increase in the proportion of type A loci leads to a decrease in their individual contribution since their global contribution cannot increase further. The comparison between top and bottom panels in Figure 2 shows that the threshold proportion of type A loci, above which their individual contribution to the genetic load significantly decreases, increases with the deleterious allele selection coefficient s.

Effect of migration on type A loci

Even though the part of the genetic load associated with A loci generally increases with migration, as expected, our simulations reveal more complex patterns when populations are small (figure 2). In particular, there might exist an intermediate migration rate that minimizes the load per A locus, despite the heterogeneous selection. Such optimal migration rate increases with the fraction of A loci in the genome. In large populations, however, the load per A locus

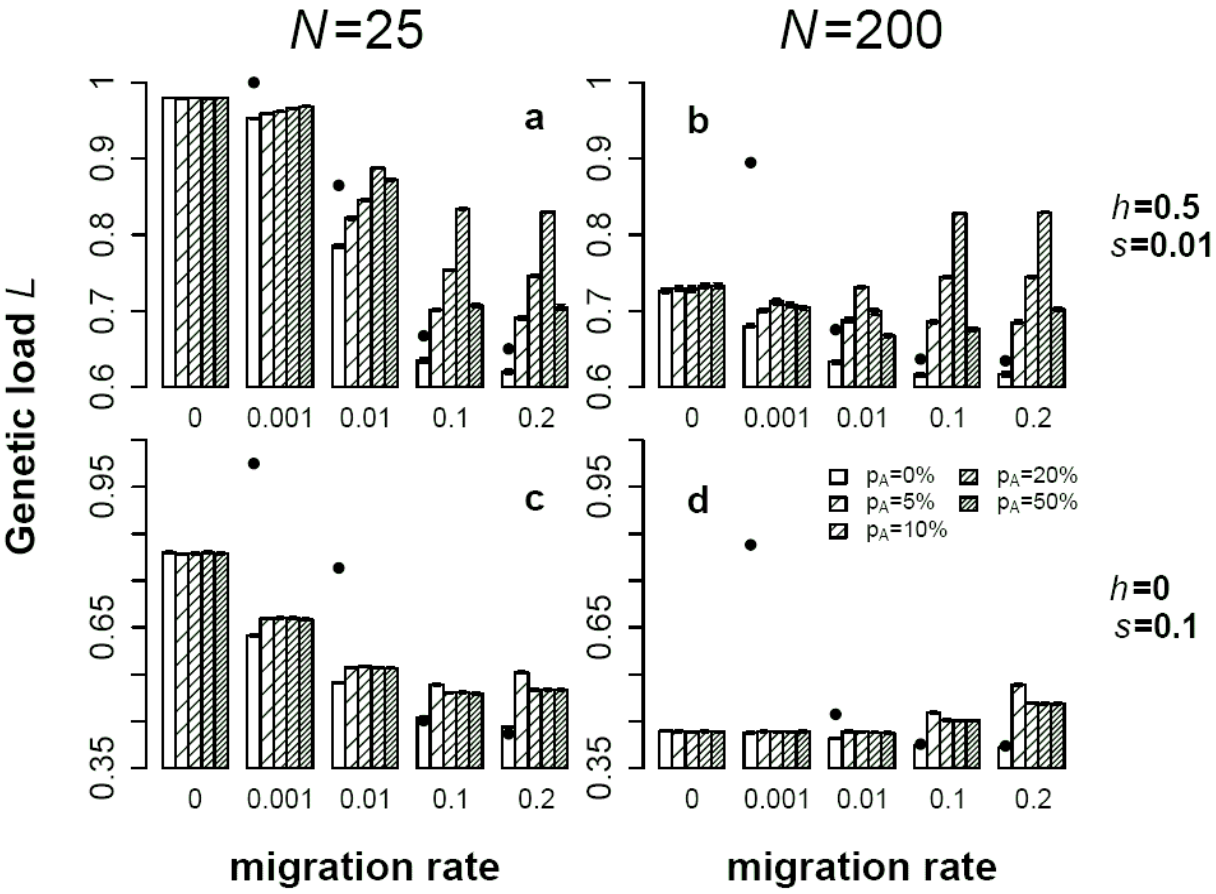


Figure 1: Mean genetic load as a function of migration rate *m*, metapopulation structure (i.e. number and size of the demes), mutational parameters and genome composition. Each group of bars corresponds to a value of *m*, with *m* increasing from left to right and taking the values 0, 0.1%, 1%, 10% and 20%. Within each group of bars, the proportion p_A of type A loci used in the simulations increases from left to right, taking the values 0, 5%, 10%, 20% and 50%. Top panels show the genetic load for codominant and slightly deleterious mutations (s=0.01 and h=0.5), bottom panels show the genetic load for recessive and highly deleterious mutatons (s=0.1 and h=0), left panels show the genetic load for a metapopulation composed of small demes (N=25 and n=32) and right panels show the genetic load for a metapopulation made of large demes (N=200 and n=4). Dots indicate analytical approximation (Roze & Rousset 2004) in the case where p_A=0.

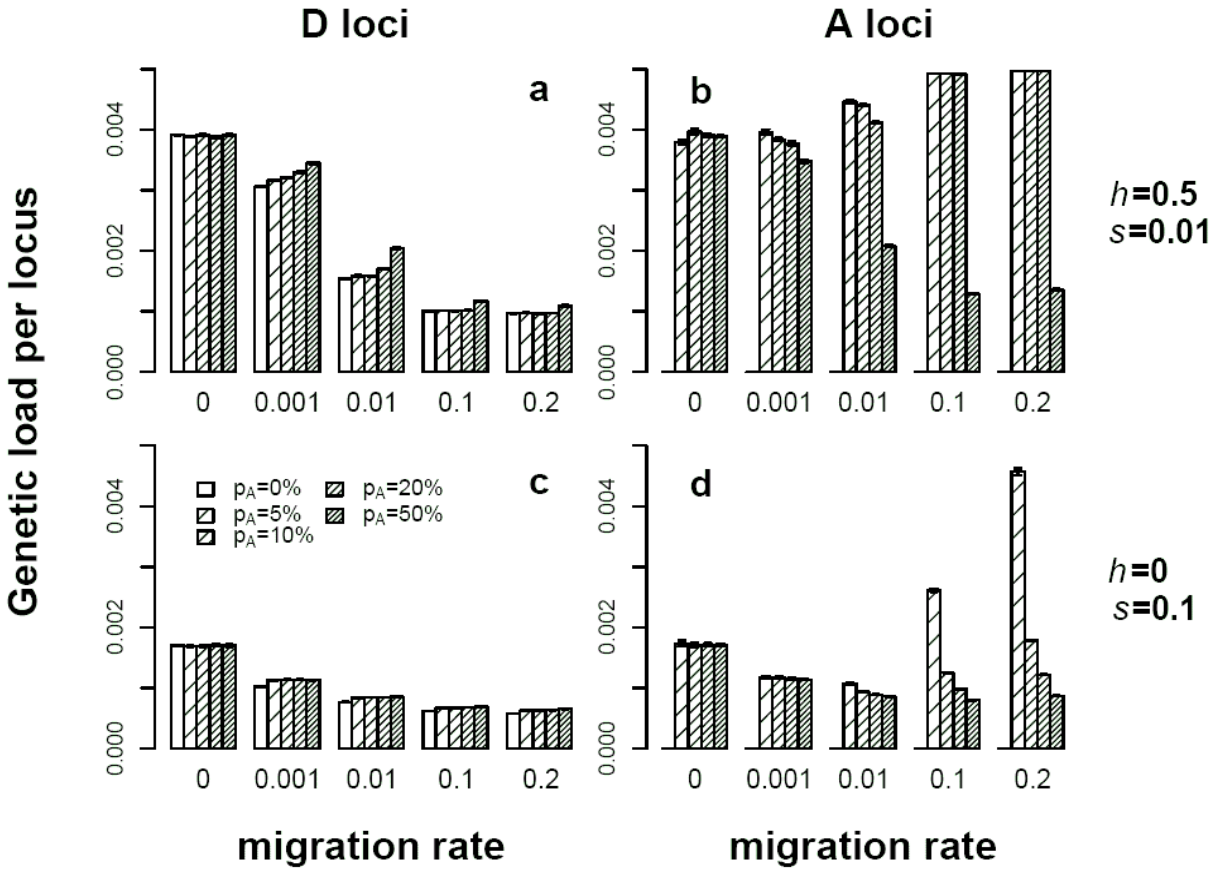


Figure 2: Load per type D (left panels) and type A (right panels) locus. Simulation parameters and symbols are the same as in figure 1, N=25; n=32.

rarely decreases when the migration rate increases (see supplementary figure).

Interactions between type D and type A loci

Figure 2 shows that the contribution of each type D locus to the genetic load is a slightly increasing function of the proportion of type A loci, although this increase is rather small. The presence of type A loci increases differentiation between populations. Increasing the proportion of type A loci indeed results in a decrease of the effective migration rate among demes (i.e. the proportion of immigrants that succesfully establish in the population). Restricted effective gene flow between demes then affects the dynamics of the load at type D loci as seen above.

Influence of distance between type A loci

Figure 3 compares simulation results with loci randomly located or grouped according to their type. The effect of increased linkage between type A loci is to decrease the global load for intermadiate migration rates. The load per type D locus slightly increases and the load per type A locus decreases for the same migration rates. The effect of an increase in physical linkage between type A loci is thus similar to that of an increase of the proportion of type A loci. For intermediate dispersal rates, linkage disequilibrium between type A loci is high. When loci are grouped together, recombination needs more time to break this strong linkage disequilibrium. This delayed breakdown of linkage disequilibrium leads to more efficient purging, hence a lower load per type A locus when they are linked together and a lower effective migration rate leading to an increase of the load per type D locus.

Influence of the dispersing stage

We also investigated the effect of pollen migration on the mean equilibrium load and its components for the case of small demes and recessive and highly deleterious mutations (Figure 4). When mutations are recessive and highly deleterious, the genetic load is increased when migration is due to pollen grains, compared to the case of seed migration, and this difference

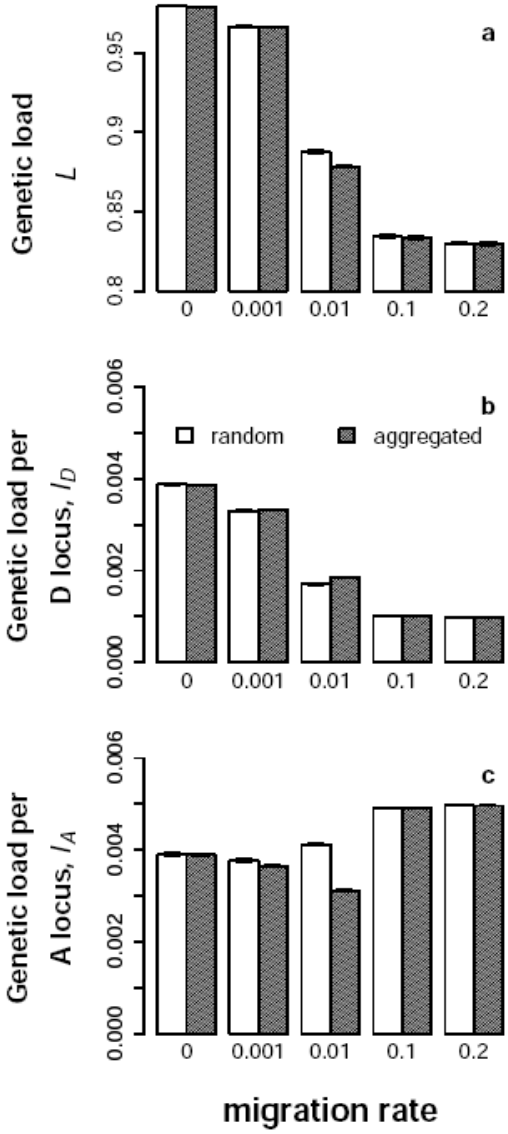


Figure 3: Effect of physical linkage between type A loci as a function of the migration rate *m* in a highly structured population (*N*=25 and *n*=32) and for codominant and slightly deleterious mutations. Top, middle and bottom panels show the mean genetic load *L*, the load per type D locus l_D and the load per type A locus l_A, respectively. Within each panel, each group of bars corresponds to a value of *m*, with *m* increasing from left to right and taking the values 0%, 0.1%, 1%, 10% and 20%. All genomes contain 800 type D loci and 200 type A loci and differ only with respect to the location of type D and A loci (see Methods). White bars correspond to genomes with randomly located loci. Black bars show the results for genomes with type A loci preferentially grouped.

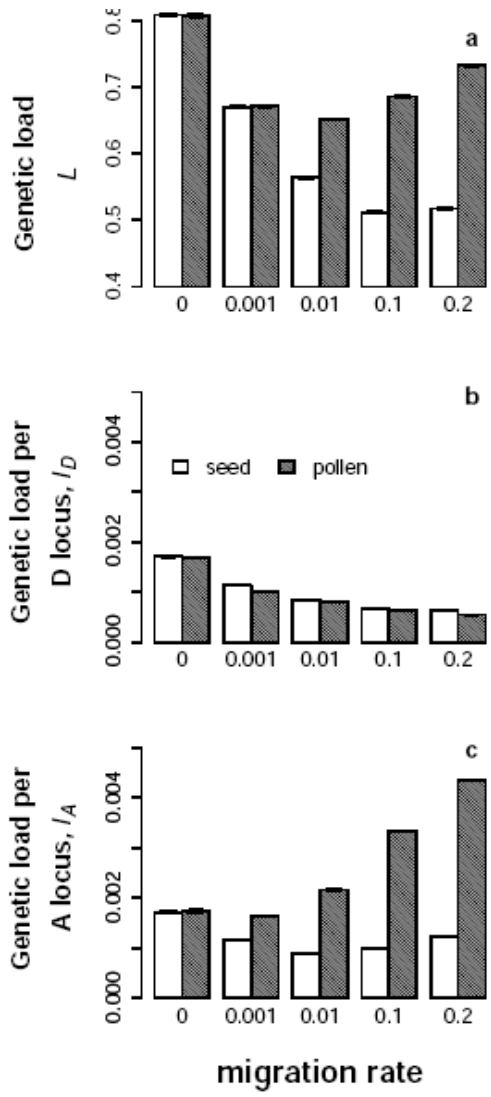


Figure 4: Effect of the dispersing stage as a function of the migration rate in a slightly structured population (N=200 and n=4). Top, middle and bottom panels show the mean genetic load L, the load per type D locus l_D and the load per type A locus l_A, respectively. Within each panel, each group of bars corresponds to a value of m, with m increasing from left to right and taking the values 0%, 0.1%, 1%, 10% and 20%. All genomes contain 800 type D loci and 200 type A loci.White bars indicate migration through seeds (at a rate m) and black bars indicate migration through pollen grains only (at a rate $2m$).

increases with migration. Figure 4 shows that, the load associated with type D loci then decreases when pollen migrates rather than seeds, while the contrary is true for type A loci. Differences between pollen and seed migration are however small when mutations are codominant *(data not shown*). This is because genes of the wrong habitat can introgress more easily when dispersed in haploid gametes. Indeed, in this case, recessive deleterious mutations are masked in the heterozygous progeny arising from migrant pollen. This increases their contribution to genetic load in later generations. However, this also increases the proportion of immigrants that survive selection, relative to the seed dispersal case, which leads to a better response of type D loci to the selection, hence a lower load at these loci.

Impact of a fragmentation event

Unconditionally deleterious mutations

In absence of type A loci and when mutations are codominant and slightly deleterious, the genetic load increases with time since the onset of fragmentation, because of the progressive accumulation of mutations due to the decreased local effective population size (Figure 5). This increase is faster when the migration rate becomes very small (Figure 5 after thesecond dashed bar). This accumulation of mutations is faster in small demes because they are subject to higher levels of genetic drift (compare lines in top-left panel). For this type of mutations, genetic load is thus mostly due to the fixation of deleterious alleles (drift load). When mutations are recessive and highly deleterious, the response to the fragmentation event is qualitatively different both during and after fragmentation. In the medium term ($t \leqslant 100$ generations after the onset of fragmentation), the increase in genetic load is faster than in the case of codominant and slightly deleterious mutations (compare bottom and top left panels). In the very long term ($t \geqslant 150$ generations) and for large populations, the evolution of the total genetic load is non

monotone, which is not the case for codominant and slightly deleterious mutations. In the first phase of the fragmentation event, matings involve increasingly related individuals, which causes inbreeding depression because recessive deleterious mutations become unmasked. This leads to a fast increase of the genetic load, before it decreases owing to purging effects: once they are expressed in homozygous state, recessive deleterious mutations become easier to eliminate by selection. However, the genetic load is higher at the end of the simulations than before fragmentation due to the fixation of deleterious mutations. In small populations, homozygous individuals appear more rapidly than in large demes, with two consequences: (i) the initial increase of the genetic load is faster in small populations (compare continuous and dotted lines for bottom left panel) and (ii) selection cannot eliminate mutations fast enough causing only a very limited purge. Thus, in small populations, inbreeding depression is rapidly converted into drift load contributed by deleterious mutations that became fixed. Genetic load thus remains far higher that in large demes and keeps increasing because of the intensity of genetic drift, despite the large coefficient of selection used.

Presence of type A loci

In presence of type A loci, genetic load is higher before fragmentation than when they are absent (compare left and right panels in Figure 5). In addition, the relative increase in post-fragmentation genetic load is reduced. Genetic load can even decrease through time after fragmentation in larger populations, both in the short term (recessive and strongly deleterious mutations, see panel d in Figure 5) and in the long term (codominant and slightly deleterious mutations, panel b). As a result, the load can be lower after fragmentation than before for larger deme size, while it is not the case for very small populations. Load due to recessive and strongly deleterious mutations can therefore exhibit complex patterns of variation through time in large

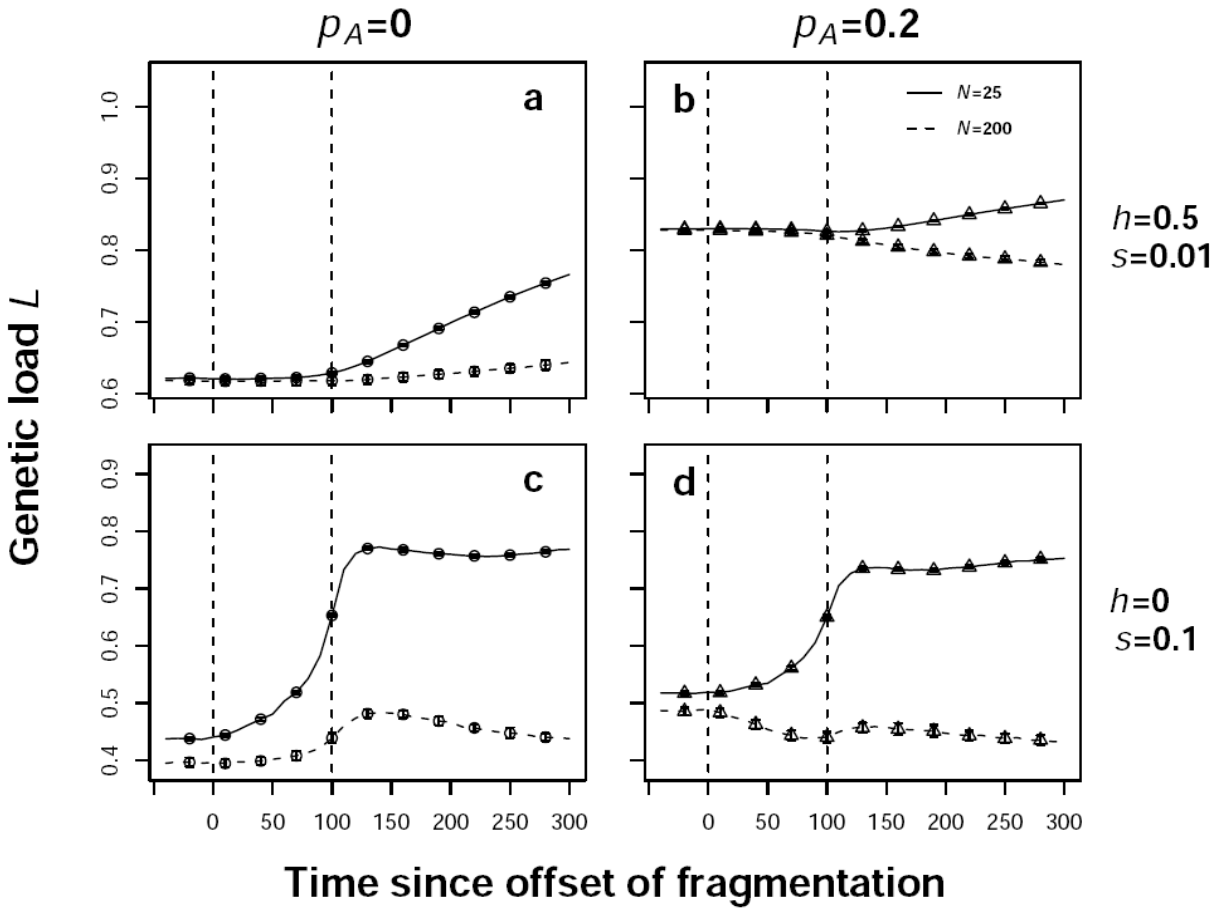


Figure 5: Effect of fragmentation on the global genetic load as a function of time since the onset of the fragmentation event, genome composition, mutational parameters and demes size. Fragmentation starts after 50 generations and consists in a progressive decrease of the migration rate by 1% of its initial value each generation. Fragmentation is thus complete after 150 generations. Left (a and c) and right (b and d) panels show genetic load response for p_A=0% and p_A=20%, respectively. Top (a and b) and bottom (c and d) panels show genetic load response for codominant and slightly deleterious mutations (s=0.01 and h=0.5) and recessive and highly deleterious mutations (s=0.1 and h=0), respectively. Continuous lines: small demes (N=25 and n=32); dotted lines: large demes (N=200 and n=4). Left and right vertical dashed lines show the generations where the migration rate starts to decrease (t=0) and reaches 0 (t=100), respectively.

populations (panel d): decreasing, then increasing in the longer term ($t \geq 150$ generations) and decreasing again in the very long-term ($t \geq 150$ generations), due to the different timing of purging at both type D and type A loci.

Figure 6 shows the dynamic of the load per one type D and one type A loci in response of the fragmentation event, for a highly fragmented population. It shows that codominant and slightly deleterious mutations progressively accumulate at type D loci for both very small and larger populations (panel a), while maladapted alleles are progressively purged at type A loci (panel b). When deme size is very small, purging of maladapted type A loci alleles is very slow and incomplete and the accumulation of mutations at type D loci is faster (because of the influence of genetic drift). Global genetic load therefore progressively increases for such populations. When deme size is larger, purging of maladapted type A loci alleles is fast, which explains the progressive decrease of the global genetic load in this case.

When mutations are recessive and highly deleterious, purging of maladapted alleles at type A loci is very weak when deme size is small: although the load per one type A locus starts to decrease when fragmentation begins, migration eventually becomes so small that genetic drift locally counteracts selection, leading to the accumulation of maladapted alleles. This and the weakness of purging at type D loci explains how global genetic load remains very high after fragmentation in small populations. For higher deme size, purging of maladapted alleles at type A loci does happen, although it is incomplete. Thus, the load per one type A locus decreases until generation 100 after the onset of the fragmentation event. In the same time, the load per one type D locus slightly increases because of increased of homozygosity (panel c in Figure 6). The load for one type D locus keeps increasing until $t \geq 150$ generations. Then it reaches a maximum, and starts to decrease because of purging. The very long time needed for purging to occur explains the highly non-monotone response of the global genetic load in the case of

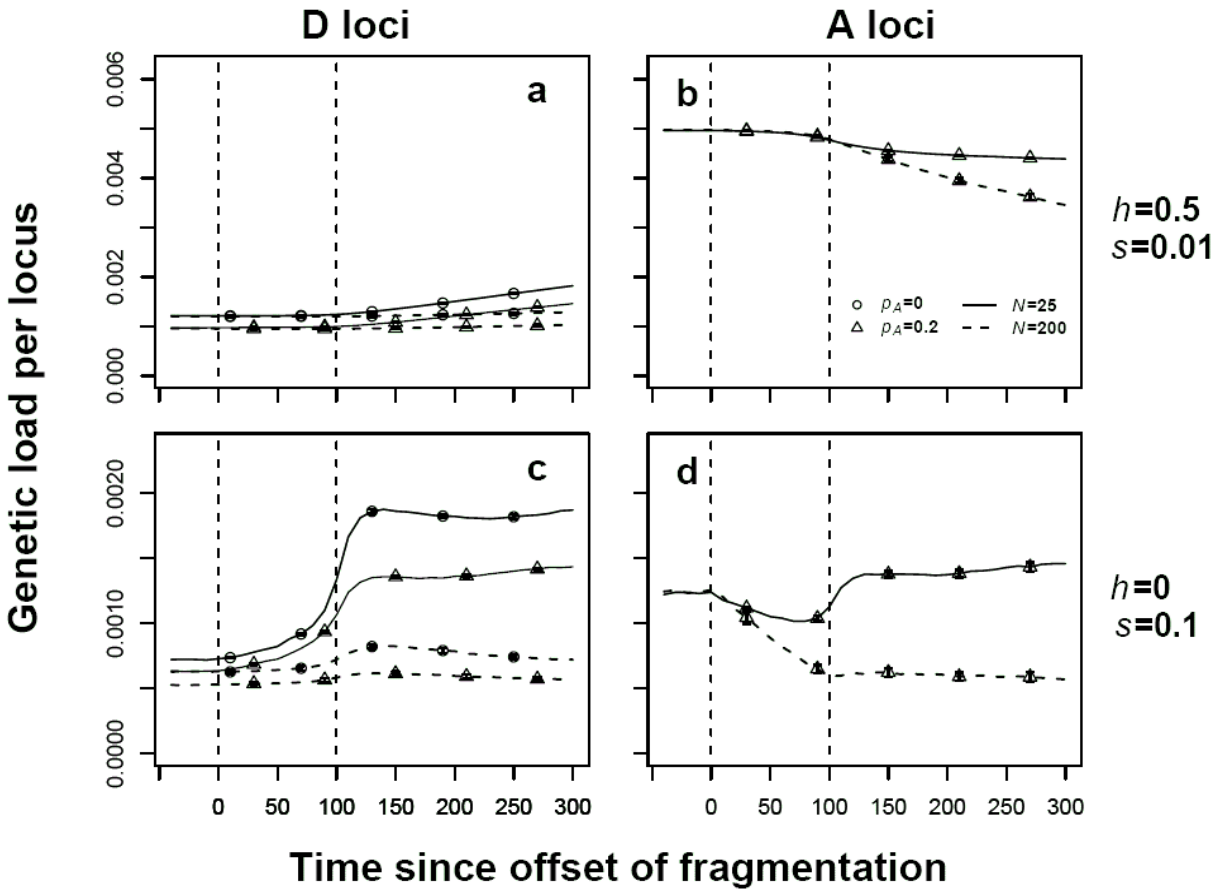


Figure 6: Effect of a fragmentation event on the contribution of one locus of type D (left panels, a and c) and one locus of type A (right panels, b and d) to the genetic load as a function of time since the onset of the fragmentation event, metapopulation structure, mutational parameters and genome composition. Continuous lines: small demes (N=25 and n=32); dotted lines: large demes (N=200 and n=4); circles: p_A=0%; triangles: p_A=20%. Left and right vertical dashed lines as in figure 5.

higher deme size and recessive and highly deleterious mutations.

Whether type A loci are present or not, it is important to notice that the long term genetic load is far higher after fragmentation when local populations are very small. When deme size is not very small and type A loci are present, the global load can be smaller after than before fragmentation.

Interactions between types of loci

Figures 6 shows that the presence of type A loci in the genome has some influence on the load associated with one type D locus, when mutations are recessive and highly deleterious (compare lines with circles to lines with triangles in panel c). As seen previously, the presence of type A loci in the genome induces an increase in the contribution of one type D locus to the genetic load before fragmentation. However, the presence of type A loci in the genome also slows down the increase of the load contributed by each type D locus in response to the fragmentation event. In the long term ($t \geq 100$ generations), the contribution of each type D locus to the genetic load is then smaller when there are type A loci in the genome. This is due to the fact that, before fragmentation, the reduced immigration rate due to the presence of type A loci results in both higher genetic load (due to increased homozygosity) but also lower frequency of deleterious alleles at type D loci (due to increased purging, *data not shown*). When habitat fragmentation occurs, there are therefore less mutations expressed through an increase of homozygosity, hence a lower load associated to one type D locus in the long term.

DISCUSSION

Throughout this work, we investigated the genetic consequences of habitat fragmentation for structured populations. By incorporating local adaptation in our model, we were able to take into account both positive and negative effects of gene flow between

populations, leading to a theoretical approach closer to empirical observations. Our results for populations at equilibrium show that the genetic load is generally increased when local adaptation is accounted for. This increase arises because of two reasons. First, gene flow between different habitats induce a migration load because of the maladaptation of immigrants. Second, the two types of loci we simulate can interact so that the presence of spatially heterogeneous selection pressures induce a slight decrease in the effective gene flow between populations. This in turn increases the genetic load at loci that are not subject to such heterogeneous pressures, but such effect is weak. Because of the widespread occurrence of local adaptation patterns in plants populations, global genetic load in structured populations is thus likely to be higher than previous theoretical studies predicted.

We also showed that the genetic load is maximised when a limited number of loci is under heterogeneous selection. This brings to the conclusion that local adaptation must be taken into account for conservation programs to be relevant. Indeed, many programs failed because they did not take enough care of local adaptation (Storfer 1999). In addition, we showed that migration through pollen grains can increase the equilibrium genetic load compared to the case where seeds migrate. Our results concerning the influence of the migrating stage are reminiscent of those obtained by Lopez et al (2007).

The One Migrant Per Generation rule has been used as a tool in conservation programs for a long time (Storfer 1999). Whereas increasing migration between populations lead to lower genetic load when selection is spatially homogeneous, this is no longer true when a fraction of loci is under spatially heterogeneous selection. In this case, there is an intermediate migration rate that minimizes the genetic load. The optimal number of migrants is poorly sensitive to both the deme size, the strength of divergent selection (i.e. the proportion of type A loci) or the type of mutations considered (i.e. codominant and slightly deleterious or recessive and highly

deleterious mutations). The optimal value of *Nm* is about 2, except in the case of very small demes and codominant and slightly deleterious mutations where it could be greater than 5 because of the very high level of local genetic drift in this case (Figure 1). Thus, the suggestion of Mills & Allendorf (1996) that the optimal number of migrants shoud range between 1 and 10 seems to be robust, even when many loci are under divergent selection.

Note that the antogonistic consequences of migration on genetic load at the scale of the whole genome can also be observed at the scale of individual loci under heterogeneous selection. In small populations, the specific genetic load per locus for those loci contributing to local adaptation is minimized for some intermediate value of the migration rate, in agreement with some previous theoretical predictions (Alleaume-Benharira et al 2006) but not others (Lopez et al 2007). The reason why such optimal migration rate unexpectedly increases with the fraction of the genome under heterogeneous selection deserves further exploration.

We also investigated the short-term genetic consequences of habitat fragmentation by simulating a fragmentation event characterized by a fast decrease of the migration rate among demes. We show that such a decrease of the migration rate has large and negative consequences when deme size is very low, because the balance between selection and local genetic drift rapidly shifts to the latter, leading to the accumulation and fixation of deleterious mutations. However, for higher deme size, a decrease in the migration rate can lead to a decrease of the global genetic load when many loci are under heterogeneous selection. We showed that purging of unconditionally deleterious mutations, when it occurs, has a very limited beneficial influence on mean fitness, confirming numerous empirical studies on plants (for a review, see Byers & Waller 1999) or a recent theoretical study (Theodorou & Couvet 2006b, see their Figure 1). Our results also show that the genetic response to a fragmentation event is not instantaneous. For instance, when mutations are codominant and slightly deleterious, a significant increase of the

genetic load occurs at least 75 generations after the beginning of the fragmentation event, even more when deme size is not very small. It means that conservation measures could be implemented in fragmented landscapes before the genetic consequences of fragmentation have become a real concern.

We used two values of deme size to study the influence of the level of population structuration on the genetic load. Although both values we used could appear to be rather small for natural populations, distinct results are obtained for each of them. In agreement with both experimental (Paland & Schmid 2003 on fragmented populations of the plant *Gentianella germanica*) or theoretical (Bataillon & Kirkpatrick 2000) studies, we showed that very small populations exposed to a cessation of migration rapidly develop a high genetic drift load due to the fixation of deleterious mutation, whereas demes as large as 200 individuals exhibit a very small drift load but suffer higher inbreeding depression.

Even if the introduction of two types of loci is a step toward more realistic theory, our model suffers from several limitations, which deserve further investigations. In our simulations, we did not allow for demogenetic retroactions: individuals had a virtually infinite fecundity, leading to sometimes unrealistic values of genetic load. Our model does not explore the consequences of such high genetic loads for the demography, while it has been shown that this can bring populations into an extinction vortex (see e.g. Higgins & Lynch 2001). Models incorporating feedbacks between demography and migration load in heterogeneous landscapes have also lead to more complex patterns of evolution of local adaptation (Ronce and Kirkpatrick 2000, Holt et al 2003). It would be interesting to investigate how demogenetic retroactions would affect the interactions between loci under homogeneous and divergent selection.

We chose to study two types of mutations representing arbitrary extreme cases (either completely recessive with strong effects or codominant with weak effects). In reality both types

of mutation shape the genetic load simultaneously. Incorporating a more continuous distribution for the dominance and selection coefficients of mutations would be a great improvement of our approach toward more realistic genetic assumptions. There is however little empirical evidence about whether such distribution differs between loci under heterogeneous selection and those under homogeneous selection across space. We also totally neglected epistasis. However, coadapted genes can play an important role in outbreeding depression (Tallmon et al 2004). Thus, improving the genetic model, by incorporating compensatory mutations (see e.g. Poon & Otto 2000) or some form of epistasis for instance, should rise interesting results.

Implications for conservation strategies

Although our results must be considered carefully as concerns their potential implications for wildlife management, we can make several conclusions. Populations composed of very small demes suffer large and rapidly increasing genetic load and decreased genetic variation when migration between demes vanishes, because of increased local genetic drift. The outcome is the same when local adaptation is taken into account. Thus, when populations of interest are very small, it is particularly important to improve their fitness through an increase of migration between such demes. Indeed, such a migration pattern is able to enhance genetic diversity, as empirically evidenced by Willi & Fischer (2005). Moreover, even if fitness could be in some cases better improved when local material is introduced in the populations (e.g. Vergeer et al 2004), our results show that this is not a large concern for populations composed of very small demes. Indeed, very small (N=25) populations suffer very high local genetic drift so that they are not adapted to their local conditions anymore (plain curves in Figure 6 right panels, but see Willi et al 2007). As a consequence, local adaptation is not the highest priority for efficient conservation strategies in such extreme cases. By contrast, when the populations are not very small (N=200), local adaptation should be taken into account in conservation programs. Indeed,

in this case, a fragmentation event reduces the genetic load at loci under spatially heterogeneous selection, while it slightly increases that of loci which are not subject to such selection pressures. Thus, the introduction of migrants from the same habitat type would not impair local adaptation while still enhancing the genetic variation, hence leading to a higher efficacy of selection.

ACKNOWLEDGEMENTS

Roland Vergilino helped much with the simulations at an early stage of this project. We thank Isabelle Olivieri, Joëlle Ronfort, Denis Roze, Sylvain Glémin, Yannis Michalakis and Thomas Menormand for advice, discussion and comments on this project. Financial support was provided by: (1) french Ministry of Research, through the programme ACI 'Jeune chercheur', allocated to F.R. And through the ACI 'Ecologie Quantitative', allocated to Isabelle Olivieri and Dominique Jolly; (2) CNRS through the programme 'Impact des Biotechnologies dans les agroécosystèmes' allocated to Isabelle Olivieri; (3) the european Union fifth Framework through the programme 'Plant Dispersal' (headed by Ben Vosman); (4) the Agence Nationale pour la Recherche through the programme 'Activités humaines, dynamique et gestion de la biodiversité en milieu méditerranéen' headed by John Thompson (A-BIME, contract ANR-05-BDIV-014). This is publication ISEM XXXX xxx ofthe Institut des Sciences de l'Evolution.

REFERENCES

Aguilar, R., L. Ashworth, L. Galetto, and M. A. Aizen. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. Ecology Letters **9**:968-980.

Bataillon, T., M. Kirkpatrick. 2000. Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Genetical Research **75**:75-81.

Bouchy, P., K. Theodorou, and C. Couvet. 2005. Metapopulation viability: influence of migration. Conservation Genetics **6**:75-85.

Byers, D. L. , and D. M. Waller. 1999. Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annual Review Of Ecology And Systematics **30**:479-513 .

Cayuela, L., J. M. Rey-Benayas, and C. Echeverria. 2006. Clearance and fragmentation of tropical montane forests in the Highlands of Chiapas, Mexico (1975-2000). Forest Ecology and Management **226**:208-218.

Edmands, S., and C. C. Timmerman. 2003. Modeling factors affecting the severity of outbreeding depression. Conservation Biology **17**:883-892 .

Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology Evolution and Systematics **34**:487-515.

Fischer, M., M. Hock, and M. Paschke. 2003. Low genetic variation reduces cross-compatibility and offspring fitness in populations of a narrow endemic plant with a self-incompatibility system. Conservation Genetics **4**:325-336.

Fischer, M., and D. Matthies. 1998. Effects of population size on performance in the rare plant *Gentianella germanica*. Journal of Ecology **86**:195-204.

Fischer, M., M. van Kleunen, and B. Schmid. 2000. Genetic Allee effects on performance, plasticity and developmental stability in a clonal plant. Ecology Letters **3**:530-539.

Galeuchet, D. J., C. Perret, C., and M. Fischer. 2005. Microsatellite variation and structure of 28 populations of the common wetland plant, *Lychnis flos-cuculi* L., in a fragmented landscape.

Molecular Ecology **14**:991-1000.

Garcia-Ramos, G., and M. Kirkpatrick. 1997. Genetic models of adaptation and gene flow in peripheral populations. Evolution **51**:21-28.

Glémin S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. Evolution **57**:2678-2687.

Glémin, S., J. Ronfort, and T. Bataillon. 2003. Patterns of inbreeding depression and architecture of the load in subdivided populations. Genetics **165**:2193-2212.

Goldringer, I., J. Enjalbert, A. L. Raquin, and P. Brabant. 2001. Strong selection in wheat populations during ten generations of dynamic management. Genetics Selection Evolution **33**:S441-S463 .

Hastings, W. K. 1970. Monte-Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika **57**:97-109.

Hendry, A. P., T. Day, and E. B. Taylor. 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: A theoretical framework for empirical tests. Evolution **55**:459-466.

Hendry, A. P., and E. B. Taylor. 2004. How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake-stream stickleback pairs. Evolution **58**:2319-2331.

Higgins, K., and M. Lynch. 2001. Metapopulation extinction caused by mutation accumulation. Proceedings of the National Academy of Sciences of the United States of America **98**:2928-2933.

Holt, R. D., R. Gomulkiewicz, and M. Barfield. 2003. The phenomology of niche evolution via quantitive traits in a 'black-hole' sink. Proceedings of the Royal Society of London Series B-

Biological Sciences **270**:215-224.

Honnay, O., and H. Jacquemyn. 2007. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. Conservation Biology **21**:823-831.

Honnay, O., D. Adriaens, E. Coart, H. Jacquemyn, and I. Roldan-Ruiz. 2007. Genetic diversity within and between remnant populations of the endangered calcareous grassland plant *Globularia bisnagarica* L. Conservation Genetics **8**:293-303.

Hooftman, D. A. P., M. van Kleunen, and M. Diemer. 2003. Effects of habitat fragmentation on the fitness of two common wetland species, *Carex davalliana* and *Succisa pratensis*. Oecologia **134**:350-359.

Hufford K. M., and S. Mazer. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. Trends in Ecology and Evolution **18**:147-155.

Ingvarsson P. K. 2001. Restoration of genetic variation lost-the genetic rescue hypothesis. Trends in Ecology and Evolution **16**:62-63.

Joshi J., Schmid B., Caldeira M. C., Dimitrakopoulos, J. Good, R. Harris, E. Hector, K. Huss-Danell, A. Jumpponen, A. Minns, C. P. H. Mulder, J. S. Pereira, A. Prinz, M. Scherer-Lorenzen, A. -S. D. Siamantziouras, A. C. Terry, A. Y. Troumbis, and J. H. Lawton. 2001. Local adaptation enhances performance of common plant species. Ecology Letters **4**:536-544.

Kawecki J. T., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters **7**:1225-1241.

Keller, M., J. Kollmann, and P. J. Edwards. 2000. Genetic introgression from distant provenances reduces fitness in local weed populations. Journal of Applied Ecology **37**:647-659.

Kirkpatrick M., and N. Barton. 2006 Chromosome inversions, local adaptation and speciation. Genetics **173**:419-434.

Leimu, R., P. Mutikainen, J. Koricheva, and M. Fischer. 2006. How general are positive relationships between plant population size, fitness and genetic variation? Journal of Ecology **94**:942-952.

Lenormand T. 2002. Gene flow and the limits to natural selection. Trends in Ecology and Evolution **17**:183-189.

Lentfer, C., and R. Torrence. 2007. Holocene volcanic activity, vegetation succession, and ancient human land use: Unraveling the interactions on Garua Island, Papua New Guinea. Review of Palaeobotany and Palynology **143**:83-105.

Linhart Y. B., and M. C. Grant. 1996. Evolutionary significance of local genetic differentiation in plants. Annual Review of Ecology and Systematics **27**:237-277.

Lopez S., F. Rousset, F. H. Shaw, R. G. Shaw, and O. Ronce. Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes. Journal of Evolutionary Biology *In press*.

Lynch, M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution **45**:622-629.

Mills, L. S., and F. W. Allendorf. 1996. The one-migrant-per-generation rule in conservation and management. Conservation Biology **10**:1509-1518.

Mix, C., F. X. Picó, J. M. van Groenendael, and N. J. Ouborg. 2006. Inbreeding and soil conditions affect dispersal and components of performance of two plant species in fragmented landscapes. Basic and Applied Ecology **7**:59-69.

Montalvo, A. M., and N. C. Ellstrand. 2001. Nonlocal transplantation and outbreeding depression in the subshrub *Lotus scoparius* (Fabaceae). American Journal Of Botany **88**:258-269.

Ohara, M., H. Tomimatsu, T. Takada, and S. Kawano. 2006. Importance of life history studies for conservation of fragmented populations: A case study of the understory herb, *Trillium camschatcense*. Plant Species Biology **21**:1-12.

Ouborg, N. J., P. Vergeer, and C. Mix. 2006. The rough edges of the conservation genetics paradigm for plants. Journal of Ecology **94**: 1233-1248.

Paland S., and B. Schmid. 2003. Population size and the nature of genetic load in *Gentianella germanica*. Evolution **57**:2242-2251.

Phadnis, N., and J. D. Fry. 2005. Widespread correlations between dominance and homozygous effects of mutations: Implications for theories of dominance. Genetics **171**:385-392.

Poon A., and S. P. Otto. 2000 Compensating for our load of mutations: freezing the meltdown of small populations. Evolution**54**:1467-1479.

Richards C. M. 2000. Inbreeding depression and genetic rescue in a plant metapopulation. The American Naturalist **155**:383-394.

Ronce O., and M. Kikrpatrick. 2000. When sources becomes sinks: migrational meltdown in heterogeneous habitats; Evolution **55**:1520-1531.

Roze, D., and F. Rousset. 2003. Selection and drift in subdivided populations: A straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. Genetics **165**:2153-2166.

Roze, D., and F. Rousset. 2004. Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. Genetics **167**:1001-1015.

Schmitt J., and S. E. Gamble. 1990. The effect of ditance from the parental site on offspring performance and inbreeding depression in *Impatiens capensis*: a test of the local adaptation hypothesis. Evolution **44**:2022-2030.

Shaw, F. H., C. J. Geyer, and R. G. Shaw. 2002 . A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. Evolution **56**:453-463.

Storfer, A. 1999. Gene flow and endangered species translocations: a topic revisited. Biological Conservation **87**:173-180.

Tallmon D. A., G. Luikart, and R. S. Waples. 2004. The alluring simplicity and complex reality of genetic rescue. Trends in Ecology and Evolution **19**: 489-496.

Theodorou, K., and D. Couvet. 2006a. Genetic load in subdivided populations: interactions between the migration rate, the size and the number of subpopulations. Heredity **96**:69-78.

Theodorou, K., and D. Couvet. 2006b. On the expected relationship between inbreeding, fitness, and extinction. Genetics Selection Evolution **38**:371-387.

Tufto, J. 2000. Quantitative genetic models for the balance between migration and stabilizing selection. Genetical Research **76**:285-293.

Vergeer P., E. Sonderen, and N. J. Ouborg. 2004. Introduction strategies put to the test: local adaptation versus heterosis. Conservation biology **18**:812-821.

Waser, N. M., M. V. Price, and R. G. Shaw. Outbreeding depression varies among cohorts of *Ipomopsis aggregata* planted in nature. Evolution **54**:485-491.

White, G. M., D. H. Boshier, and W. Powell. 2002. Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. Proceedings of the National Academy of Sciences of the United States of America **99**:2038-2042.

Whitlock, M. C. 2002. Selection, load and inbreeding depression in a large metapopullation. Genetics **160**:1191-1202.

Willi Y., and M. Fischer. 2005. Genetic rescue in interconnected populations of small and large

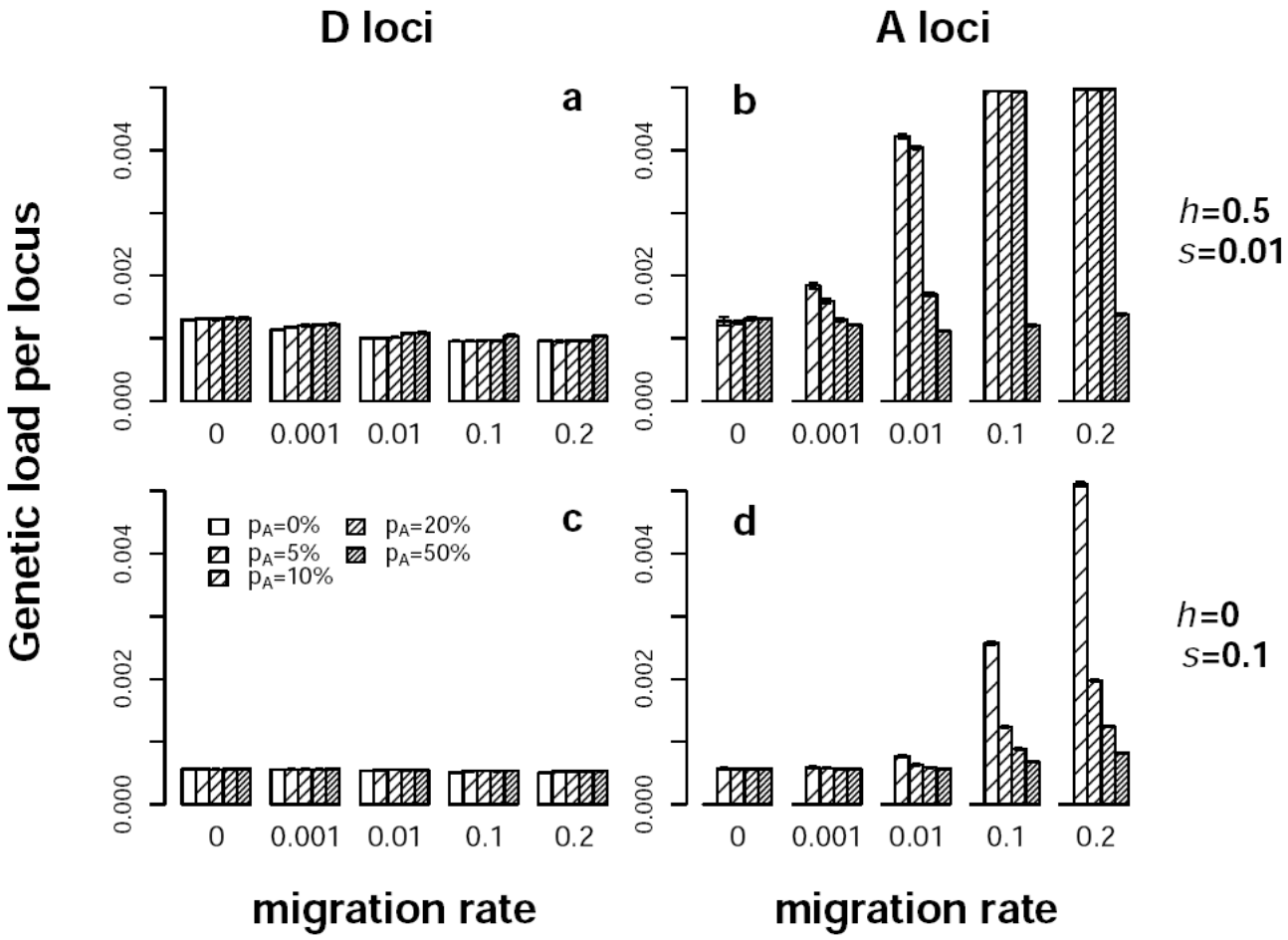
size of the self-incompatible *Ranunculus reptans*. Heredity **95**:437-443.

Willi, Y., J. Van Buskirk, B. Schmid, and M. Fischer. 2007. Genetic isolation of fragmented populations is exacerbated by drift and selection. Journal of Evolutionary Biology **20**:534-542.

Young, A., T. Boyle, and T. Brown. 1996. The population genetic consequences of habitat fragmentation for plants. Trends in Ecology & Evolution **11**:413-418.

Zhang, Y., and B. Guindon. 2005. Landscape analysis of human impacts on forest fragmentation in the Great Lakes region. Canadian Journal of Remote Sensing **31**:153-166.

Supplementary material:



Load per type D (left panels) and type A (right panels) locus for a slightly fragmented population (N=200; n=4). Each group of bars corresponds to a value of m, with m increasing from left to right and taking the values 0, 0.1%, 1%, 10% and 20%. Within each group of bars, the proportion p_A of type A loci used in the simulations increases from left to right, taking the values 0, 5%, 10%, 20% and 50%. Top panels show the genetic load for codominant and slightly deleterious mutations (s=0.01 and h=0.5), bottom panels show the genetic load for recessive and highly deleterious mutations (s=0.1 and h=0).

CONCLUSION ET PERSPECTIVES

Les résultats théoriques obtenus au cours de cette thèse éclairent comment la fragmentation des paysages (en tant qu'état ou en tant que processus) affecte, d'une part, l'évolution de la diversité génétique sélectionnée et sa distribution spatiale et, d'autre part, l'évolution du fardeau génétique. Plus particulièrement, je me suis attaché à comprendre comment l'existence de pressions de sélection divergentes affecte les conséquences de la fragmentation dans des paysages hétérogènes. La fragmentation a été abordée essentiellement par l'étude des conséquences de variations de l'intensité de la migration et du mode de migration chez les plantes. Dans ce contexte, mes résultats montrent que conserver la diversité génétique locale, à l'échelle de l'aire de répartition de l'espèce, ou encore minimiser le fardeau génétique qui compromet la persistance à court terme des populations, correspondent à des objectifs de gestion partiellement conflictuels. Je vais tout d'abord rappeler les conclusions principales de mon travail concernant la diversité génétique sélectionnée et les différentes causes du fardeau génétique. Ceci mènera à une réflexion plus méthodologique et critique sur les choix de modélisation de ces processus et des conséquences de ces choix. Enfin, je discuterai des implications pratiques potentielles des résultats obtenus pour la biologie de la conservation. Je terminerai par les perspectives de ce travail et les questions qui restent en suspens à son issue.

I Fragmentation et maintien de la diversité génétique sélectionnée

Cette question a été abordée essentiellement dans le chapitre 1. Une réduction de la migration conduit toujours à plus de divergence génétique entre populations. Les conséquences d'une réduction des flux de gènes pour la variabilité génétique à l'intérieur des populations locales sont plus complexes. Or c'est ce dernier type de variabilité qui va contraindre la réponse évolutive de ces populations si elles sont confrontées à des changements de l'environnement. Nos résultats confirment les prédictions d'études théoriques précédentes montrant qu'il existe un taux intermédiaire optimal de migration qui maximise la variance génétique locale. Ce taux est très faible quand la sélection favorise le même phénotype partout dans le paysage, mais il augmente significativement quand la sélection est divergente entre sites. Dans un paysage hétérogène, la migration apparaît alors comme la force principale expliquant le maintien d'une forte diversité génétique locale. Ceci remet en cause les prédictions de certains modèles (Lythgoe 1997) qui avaient sous-estimé l'effet de la migration sur la variance génétique sélectionnée en présence de sélection à la fois stabilisante localement et divergente entre

localités.

Notre comparaison des effets spécifiques de la migration du pollen et des graines a montré que, si l'intensité des flux géniques (indépendamment du mode de dispersion) était le facteur principal déterminant le niveau de variance génétique totale, la nature de cette variance (variance génique, contribution du déséquilibre de liaison et des écarts à la panmixie) quant à elle, variait fortement en fonction de la contribution de la migration des graines et du pollen. Nous avons de même observé que la liaison physique entre locus affectait la contribution des différentes composantes de la variance génétique mais peu le niveau de variance totale. L'explication de cette relative constance du niveau de variance totale mériterait des explorations plus approfondies.

II Fragmentation et fardeau génétique

Les résulats obtenus montrent que l'influence de la migration sur le fardeau dépend de la façon dont la sélection varie dans l'espace. Si la direction de la sélection diverge entre localités, une migration réduite a globalement des conséquences positives sur la valeur sélective moyenne des individus, ce qui confirme les résultats d'autres travaux (voir Lenormand 2002), que la sélection agisse de façon stabilisante sur un trait phénotypique avec des optimums différents (chapitre 1), ou qu'elle agisse de façon directionnelle sur des allèles différents selon les sites (chapitre 2). En présence d'optimums phénotypiques variables dans l'espace, des flux génétiques moins importants permettent une plus grande divergence entre les populations et, dans une certaine gamme de paramètres, une variance génétique locale moins élevée, ce qui se traduit par un fardeau moins élevé en présence de sélection stabilisante (chapitre 1). De même, le fardeau associé aux locus soumis à une sélection directionnelle divergente entre sites (chapitre 2) augmente en général avec la migration. Ce résultat est cependant à tempérer (voir ci-dessous). Nous avons par ailleurs montré que l'adaptation locale est plus facilement compromise par la migration du pollen que de celle des graines (chapitre 1 et chapitre 2), encore que cette différence entre migration des graines et du pollen ne soit observée que dans des situations bien spécifiques: pour un trait quantitatif, elle n'apparaît que pour des niveaux très élevés d'hétérogénéité du paysage (chapitre 1), tandis que l'effet spécifique de la migration du pollen pour le fardeau génétique est visible surtout pour des mutations d'effet fort et très récessives dans le chapitre 2.

Au contraire, quand la sélection est homogène à travers l'espace, une diminution de la migration conduit en général à une augmentation du fardeau génétique. C'est le cas quand l'optimum phénotypique ne diffère pas entre localités et que les populations sont très petites (chapitre 1) et que des niveaux mêmes modérés de migration conduisent rapidement à une réduction de la variance génétique, alors synonyme de réduction du fardeau en présence de sélection stabilisante. C'est le cas aussi pour des mutations inconditionnellement délétères soumises à une sélection directionnelle (chapitre 2): pour les deux classes de mutations que nous avons étudiées (codominantes et faiblement délétères ou récessives et fortement délétères), la valeur sélective moyenne augmente toujours avec la migration, en accord avec certains résultats expérimentaux (Willi et Fischer 2005, Paschke et al 2005) ou théoriques (Roze et Rousset 2004).

Enfin, nous avons étudié la réponse génétique des populations à divers niveaux de migration lorsque les deux types de pression de sélection évoqués ci-dessus sont présents et agissent sur différents locus. L'effet de la migration sur le fardeau génétique en présence de génomes composites varie alors de façon complexe en fonction de la taille des populations ou du nombre de locus sous sélection hétérogène (chapitre 2). Le fardeau génétique total augmente généralement quand la migration diminue et que la taille des populations locales est très petite. Il peut diminuer, augmenter ou varier de façon non monotone quand les populations sont grandes. Cette étude a permis aussi de mettre en évidence des interactions entre les locus soumis à une sélection homogène ou hétérogène dans l'espace. Ainsi, le fardeau dû aux mutations inconditionnellement délétères augmente lorsque d'autres locus sont sous sélection spatialement hétérogène, du fait de l'influence de ces derniers sur le taux de migration efficace. L'augmentation du fardeau dû aux mutations inconditionnellement délétères est cependant très limitée. On peut d'ailleurs se demander pourquoi, dans la mesure où les différents types de locus sont physiquement liés dans nos simulations. Une analyse fine, à l'échelle des locus et non plus du génome dans son ensemble, permettra peut-être de comprendre comment ce fardeau supplémentaire se répartit au sein du génome en fonction de l'organisation des différents types de locus. Si le type des pressions de sélection (homogènes ou héterogènes) est très important pour la réponse du fardeau à la fragmentation, le rapport de forces entre sélection et dérive l'est également. Ce rapport dépend notamment de la taille des dèmes.

Lorsque les populations sont fragmentées en petits dèmes, l'importance de la dérive augmente relativement à celle de la sélection à l'échelle locale. Pour les mutations

inconditionnellement délétères, ceci a deux conséquences. (i) Les petites populations profitent davantage de l'effet de rescousse génétique d'une migration accrue vis-à-vis des mutations inconditionnellement délétères (Willi et Fischer 2005): lorsque les populations ne sont pas très petites, la diminution du fardeau dû à ces mutations quand la migration augmente est très faible. (ii) De même, les grandes populations sont beaucoup moins sensibles à la diminution de la migration efficace liée à l'augmentation de la proportion de locus soumis à une sélection spatialement hétérogène: le fardeau dû aux mutations inconditionnellement délétères augmente beaucoup moins avec la proportion de ces locus que dans le cas des petites populations.

La variation du rapport de forces local entre sélection et dérive a également des conséquences pour le fardeau associé aux locus impliqués dans les patrons d'adaptation locale. Dans les petites populations, le fardeau associé à ces locus peut, de façon contre-intuitive, diminuer quand la migration augmente et ce, d'autant plus que le nombre de locus soumis à une sélection hétérogène est élevé (chapitre 2). En effet, les résultats du deuxième chapitre montrent que de faibles valeurs du taux de migration peuvent, lorsque les populations locales sont de taille réduite, minimiser le fardeau moyen pour un locus d'adaptation locale. Au contraire, quand la taille des populations locales est plus élevée, le fardeau associé aux locus d'adaptation locale augmente toujours avec la migration. De façon surprenante et malgré un certain nombre d'hypothèses et de valeurs de paramètres communes, ce résultat n'a pas été retrouvé dans le chapitre 1. L'effet de la migration sur le fardeau génétique était alors qualitativement comparable dans les grandes et petites populations et conduisait toujours à son augmentation dans un paysage hétérogène. Ceci conduit à s'interroger sur l'importance des choix effectués en modélisation. Une autre étude théorique (Alleaume-Benharira et al 2006) a prédit que des taux de migration intermédiaire permettent de minimiser le fardeau génétique dans des petites population en dépit d'une sélection divergente entre sites (voir aussi Holt et al 2003). Cette étude est basée sur un modèle assez différent des nôtres (modèle de migration de proche en proche entre populations de tailles variables avec sélection sur un trait quantitatif unilocus dont l'optimum change graduellement dans l'espace). Dans les petites populations périphériques et maladaptées, un taux faible de migration permet de contrer les effets de la dérive sans compromettre trop fortement la divergence entre populations et leur adaptation locale.

Plus généralement, ce que ces divers résultats théoriques suggèrent est que le fardeau génétique est le produit des interactions conflictuelles et complexes entre dérive, sélection et migration. L'équilibre entre ces trois forces s'établit différemment en fonction de la nature des

locus concernés (effets spécifiques à l'environnement ou pas) et de la taille des populations. Parce qu'elle s'oppose à la fois à l'action de la dérive et de la sélection divergente, la migration a des effets antagonistes sur le fardeau génétique, observable à l'échelle d'un seul locus (Alleaume-Benharira et al 2006) ou bien entre locus (chapitre 2).

Qu'a-t-on appris en étudiant l'évolution transitoire du fardeau en réponse à la fragmentation ? La dynamique temporelle de cette réponse dépend fortement des effets et de la récessivité des mutations, de la taille des populations et de l'existence d'un fardeau de migration avant fragmentation. On retrouve une partie des conclusions sur le fardeau à l'équilibre, à savoir que (i) le fardeau total augmente en général avec la diminution de la migration dans les très petites populations, indépendamment de l'existence d'un fardeau de migration initial, (ii) il a une dynamique plus complexe dans les grandes populations et a tendance à diminuer après la fragmentation quand une partie des locus sont sous sélection spatialement hétérogène. Nous avons aussi vu que les patrons temporels de la réponse du fardeau à la fragmentation changent également en fonction du génome et du type des mutations: après la fragmentation, la purge du fardeau de migration se fait de façon rapide lorsque les mutations sont récessives et d'effets forts alors que cette purge est très lente lorsqu'elles sont codominantes et d'effets faibles. De plus, l'augmentation du fardeau de mutations inconditionnellement délétères dûe à l'expression des mutations récessives est un phénomène rapide alors que la purge et la fixation ont des effets visibles seulement sur le long terme. Enfin, les interactions entre locus soumis respectivement à une sélection hétérogène ou homogène ont des effets plus forts sur la dynamique transitoire du fardeau que sur sa valeur à l'équilibre.

III Modéliser, c'est choisir

Dans les petites populations, la migration peut avoir un effet bénéfique sur l'adaptation locale (Alleaume-Benharira et al 2006, chapitre 2), ce que ne montre pas le modèle du chapitre 1. L'exploration des paramètres étant toujours limitée avec une approche par simulations, il est difficile d'exclure la simple idée que nous n'ayons pas étudié les combinaisons de paramètres permettant de mettre ce patron en évidence. Ceci donne néanmoins l'occasion de réfléchir plus généralement aux implications de nos différents choix de modélisation. Dans le chapitre 1, nous nous sommes focalisés sur le fardeau engendré par la sélection stabilisante sur un trait phénotypique unique. Il est bien entendu que le fardeau génétique total dépend de la

sélection sur de multiples traits, certains sous sélection homogène à travers l'espace, d'autres sous sélection hétérogène. Dans le chapitre 2, nous avons adopté une vision plus globale de la valeur sélective en modélisant des génomes entiers et complexes. Mais à la différence du premier chapitre, nous avons modélisé une relation directe entre le génotype et la valeur sélective, sans passer par l'expression de traits phénotypiques, comme ce qui est classiquement fait en génétique des populations. Ces choix différents sont fondamentalement arbitraires et ont été motivés essentiellement par notre désir de pouvoir comparer nos résultats aux autres résultats théoriques sur les mêmes thématiques (adaptation locale d'une part et fardeau de mutation d'autre part, voir l'introduction). Il s'avère que ces choix ne sont pas pour autant neutres.

Avec la première approche, les effets des allèles sur la valeur sélective, leurs interactions de dominance et d'épistasie ne sont pas fixés a priori mais émergent de la relation non linéaire entre phénotype et valeur sélective qu'implique le modèle gaussien. Ce modèle est intéressant car il engendre donc également des prédictions sur la distribution et l'évolution de ces interactions entre allèles (voir par exemple les travaux de Martin et Lenormand 2006a et b, 2007) mais il est parfois difficile à interpréter et à comparer à des résultats de génétique des populations classique du fait de la nature très dynamique de ces interactions dans une population génétiquement diversifiée (par exemple le coefficient de dominance d'un allèle change en permanence en fonction du contexte génétique dans lequel il s'exprime). Inversement, les simplifications du modèle utilisé dans le chapitre 2 peuvent représenter autant de contraintes artificielles. Par exemple, le fait d'attribuer le même coefficient de sélection et le même degré de récessivité à toutes les mutations dans le génome est critiquable. Choisir une distribution appropriée pour les effets des mutations conditionnellement ou inconditionnellement délétères reste difficile étant donné que les estimations empiriques de ces paramètres sont encore rares. En excluant certaines formes d'épistasie, le modèle du chapitre 2 (qui est celui le plus classiquement utilisé pour étudier le fardeau de mutation) ne permet pas d'inclure l'effet de mutations compensatrices. En utilisant un formalisme alternatif du type sélection gaussienne sur un grand nombre de traits phénotypiques plus proche du modèle du chapitre 1, Poon et Otto (2000) ont montré que l'existence de telles mutations pouvait considérablement limiter l'évolution du fardeau génétique dans les petites populations. Leurs résultats donnent une piste d'interprétation pour expliquer l'absence d'un effet positif de la migration pour le fardeau d'adaptation locale dans les petites populations observé dans le chapitre 1. Le fardeau génétique

évoluant dans ces petites populations serait atténué par l'existence de mutations compensatrices et l'effet de rescousse de la migration serait donc moins visible. Une comparaison plus quantitative des fardeaux liés à la mutation, la dérive et la migration dans ces deux cadres théoriques reste à entreprendre.

D'autres hypothèses des modèles peuvent également avoir une grande importance. Le choix des paramètres mutationnels, par exemple, est crucial. Dans le chapitre 1, on a vu que, pour un même niveau de variance introduite à chaque génération, la façon de modéliser les effets des mutations a un impact important sur la variance génétique d'équilibre: lorsque les mutations sont très nombreuses et d'effets faibles, la variance génétique est beaucoup moins élevée dans les populations que lorsque les mutations sont plus rares et d'effets plus forts. Ceci souligne l'importance d'une construction et d'une calibration raisonnables des modèles théoriques appuyées sur des données empiriques. Il est important à ce stade de continuer à confronter ces deux types de formalismes (modèles de type *h-s* versus sélection gaussienne sur phénotypes) aux données empiriques afin de déterminer lequel des deux est le plus adéquat pour décrire l'évolution du fardeau génétique.

IV Quelques implications pratiques

De nombreux programmes de conservation des espèces ont échoué à cause de problèmes d'adaptation locale (Storfer 1999). L'ensemble de mon travail de thèse a fourni quelques résultats qui pourraient permettre d'aider les gestionnaires à savoir s'ils doivent se soucier, ou ignorer, les questions d'adaptation locale dans leurs programmes. Bien entendu, de tels résultats demandent à être confirmés et restent à considérer avec précaution. Ainsi, lorsque les populations sont de très petite taille, l'apport de variabilité génétique a peu de chances de diminuer l'adaptation locale, puisque ces populations sont déjà maladaptées (chapitre 2). De manière générale, le choix d'augmenter la migration par le biais de grains de pollen ou de graines ne semble pas avoir d'importance pour les patrons d'adaptation locale, sauf si l'environnement est très hétérogène (chapitre 1). Enfin, les résultats obtenus soulignent l'importance à accorder aux objectifs précis des programmes de conservation: si la variabilité génétique peut être vue comme une source de maladaptation, elle est néanmoins fondamentale pour l'évolution des populations. Dans la perspective des changements globaux en cours, le maintien de la variabilité génétique des populations peut être important pour l'adaptation future

des espèce. Dans un environnement hétérogène, les objectifs ne seront pas du tout les mêmes selon que l'on veut optimiser l'adaptation locale (et donc le maintien à court terme des populations) ou le niveau de variabilité génétique (et donc la capacité d'évolution et de maintien à long terme des populations). Dans le premier cas, les résultats du chapitre 1 suggère qu'un taux de migration très faible serait à privilégier, tandis qu'une optique de maintien d'une forte variabilité génétique requiert un taux de migration fort (2 migrants par génération, ou davantage dans les paysages très hétérogènes).

V Perspectives

Tout au long de ce travail, je n'ai pas pris en compte d'éventuelles rétroactions de la génétique sur la capacité des populations à maintenir un effort reproducteur constant. Pourtant, les études s'intéressant aux conséquences de la fragmentation des habitats montrent très souvent un lien entre fragmentation et succès de reproduction dans les populations (Leimu et al 2006, Aguilar et al 2006) avec potentiellement des conséquences démographiques. Si l'évolution du fardeau affecte celle de la taille des populations, l'importance relative des fardeaux de dérive et de migration pourraient s'en trouver affectée. Plus généralement, nos résultats ne permettent pas de juger directement des risques d'extinction encourus par les populations subissant un épisode de fragmentation. Dès lors, la pleine prise en compte des rétroactions entre la démographie et la génétique permettrait de bâtir des scénarios plus réalistes d'un point de vue biologique, dont les résultats seraient de meilleurs guides pour les politiques de gestion et de conservation de la biodiversité. De plus, j'ai choisi de modéliser l'altération des habitats par la seule disparition des flux génétiques entre populations, afin de m'affranchir des effets génétiques de la perte d'habitat. Or, les habitats naturels subissent de nos jours non seulement une fragmentation, mais également une destruction croissante. De ce fait, une étude conjointe de la perte d'habitat et de la diminution des flux migratoires entre les populations restantes constituerait un pas en avant vers davantage de réalisme biologique.

RÉFÉRENCES BIBLIOGRAPHIQUES

Aguilar, R., L. Ashworth, L. Galetto, and M. A. Aizen. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. Ecology Letters 9:968-980.

Alleaume-Benharira, M., I. R. Pen. & O. Ronce. 2006. Geographical patterns of adaptation within a species range: interactions between drift and gene flow. Journal of Evolutionary Biology 19: 203-215.

Alvarez-Buykka, E. R. & R. Garcia-Barrios. 1991. Seed and forest dynamics: a theoretical framework and an example from the neotropics. American Naturalist 137:133-154.

Antonovics, J. & A. D. Bradshaw. 1970. Evolution in closely adjacent plant populations .8. Clinal patterns at a mine boundary. Heredity. 25: 349-362.

Bacles, C. F. E., A. J. Lowe & R. A. Ennos. 2006. Effective seed dispersal across a fragmented landscape. Science 311:628.

Baldi, G., J. P. Guerschman & J. M. Paruelo. 2000. Characterizing fragmentation in temperate South America grasslands. Agriculture Ecosystems & Environment 116:197-208.

Barton, N. H. 1999. Clines in polygenic traits. Genetical Research 74:223-236.

Bataillon T. 2003. Shaking the 'deleterious mutations' dogma ? Trends in Ecology & Evolution 18:315-317.

Bataillon, T. & M. Kirkpatrick. 2000. Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Genetical Research 75:75-81.

Bennett, K. D., P. C. Tzedakis & K. J. Willis. 1991. Quaternary refugia of north european trees. Journal of biogeography 18:103-115.

Bogaert, J., P. Van Hecke & R. ceulemans. 2002. The Euler number as an index of spatial integrity of landscapes: evaluation and proposed improvement. Environmental Management 29:673-682.

Bolnick, D. I. & P. Nosil. 2007. Natural selection in populations subject to a migration load. Evolution 61:2229-2243.

Bouchy, P., K. Theodorou & C. Couvet. 2005. Metapopulation viability: influence of migration. Conservation Genetics 6:75-85.

Bourguet D. 1999. The evolution of dominance. Heredity 83:1-4.

Brown, J. H. & A. Kodric-Brown. 1977. Turnover rates in insular biogeography: effect of immigration and extniction. Ecology 58:445-449.

Byers, D. L. & D. M. Waller. 1999. Do plant populations purge their genetic load? Effects of

population size and mating history on inbreeding depression. Annual Review Of Ecology And Systematics 30:479-513.

Cayuela, L., J. M. Rey Benayas & C. Echeverria. 2006. Clearance and fragmentationof tropical montane forests in the Highland of Chiapas, Mexico (1975-2000). Forest Ecology and Management 226:208-218.

Charlesworth B. 1998. Adaptive evolution: The struggle for dominance. Current biology 8:R502-R504.

Cheddadi, R., G. G. Vendramin, T. Litt, L. François, M. Kageyama, S. Lorentz, J.-M. Laurent, J.-L. de Beaulieu, L. Sadori, A. Jost & D. Lunt. 2006. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. Global Ecology and Biogeography 15:271-282.

Cherry, J. L. & J. Wakeley. 2003. A diffusion approximation for selection and drift in a subdivided population. Genetics 163:421-428.

Colas, B., I. Olivieri & M. Riba. 1997. *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink to extinction: A demographic and genetic study. Proceedings of the National Academy of Sciences of the united States of America 94:3471-3476.

Couvet D. 2002. Deleterious effects of restricted gene flow in fragmented populations. Conservation Biology 16:369-376.

Crow, J. F. & M. Kimura. 1970. An introduction to population genetics theory. New York: Harper & Row.

Cushman, S. A. 2006. Effects of habitat loss and fragmentation on amphibians: A review and prospectus. Biological Conservation 128:231-240.

Echeverria, C., D. Coomes, J. Salas, J. M. Rey-Benayas, A. Lara, A. & A. Newton. 2006. Rapid deforestation and fragmentation of Chilean Temperate Forests. Biological Conservation 130:481-494.

Etter, A., C. McAlpine, S. Phinn, D. Pullar & H. Possingham. 2006. Characterizing a tropical deforestation wave: a dynamic spatial analysisof a deforestation hotspot in the Colombian Amazon. Global Change Biology 12:1409-1420.

Etterson, J. R. & R. G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. Science. 294: 151-154.

Etterson, J. R. 2004a. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. 1. Clinal pattems of selection along an environmental gradient in the great plains. Evolution. 58: 1446-1458.

Etterson, J. R. 2004b. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate

change. II. Genetic architecture of three populations reciprocally planted along an environmental gradient in the great plains. Evolution. 58: 1459-1471.

Fischer, M., M. Hock & M. Paschke. 2003. Low genetic variation reduces cross-compatibility and offspring fitness in populations of a narrow endemic plant with a self-incompatibility system. Conservation Genetics 4:325-336.

Fisher R. A. 1928. The possible modification of the response of the wilt type to recurrent mutations. The American Naturalist 62:115-126.

Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. Evolution 57:2678-2687.

Glémin, S., J. Ronfort & T. Bataillon. 2003. Patterns of inbreeding depression and architecture of the genetic load in subdivided populations. Genetics 165:2193-2212.

Goldringer, I., C. Prouin, M. Rousset, N. Galic & I. Bonnin. 2006. Rapid differentiation of experimental populations of wheat for heading time in response to local climatic conditions. Annals of Botany 98: 805-817.

Hardy, O. J., S. C. Gonzalez-Martinez, H. Fréville, G. Boquien, A. Mignot, B. Colas & I. Olivieri. 2004. Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae) I. Pattern of pollen dispersal. Journal of evolutionary Biology 17:795-806.

Higgins, K. & M. Lynch. 2001. Metapopulation extinction caused by mutation accumulation. Proceedings of the National Academy of Sciences of the United States of America 98:2928-2933.

Holt, R. D., R. Gomulkiewicz & M. Barfield. 2003. The phenomenology of niche evolution via quantitative traits in a 'black-hole' sink. Proceedings of the Royal Society of London Biological Sciences 270:215-224.

Honnay, O., D. Adriaens, E. Coart, H. Jacquemyn & I. Roldan-Ruiz. 2007. Genetic diversity within and between remnant populations of the endangered calcareous grassland plant *Globularia bisnagarica* L. Conservation Genetics 8:293-303.

Hooftman, D. A. P., M. Van Kleunen & M. Diemer. 2003. Effects of habitat fragmentation on the fitness of two common wetland species, *Carex davalliana* and *Succisa pratensis*. Oecologia 134:350-359.

Hu, X. S. & B. L. Li. 2001. Assessment of the ratio of pollen to seed flow in a cline for genetic variation in a quantitative trait. Heredity 87:400-409.

Hu, X. S. & B. L. Li. 2003. On migration load of seeds and pollen grains in a local population. Heredity 90:162-168.

Hu, X. S. 2006. Migration load in males and females. Theoretical Population Biology 70:183-

200.

Hufford, K. M. & S. Mazer. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. Trends in Ecology & Evolution 18:147-155.

Hutchinson, G. E. 1957. Population studies – Animal Ecology and Demography – Concluding remarks. Cold Spring Harbour Symposium.

Ingvarsson, P. K. 2001. Restoration of the genetic variation lost – the genetic rescue hypothesis. trends in Ecology & Evolution 16:62-63.

Jimenez-Ambriz, G., C. Petit, I. Bourrie, S. Dubois, I. Olivieri, I. & O. Ronce. 2007. Life history variation in the heavy metal tolerant plant *Thlaspi caerulescens* growing in a network of contaminated and noncontaminated sites in southern France: role of gene flow, selection and phenotypic plasticity. New Phytologist 173: 199-215.

Joshi J., B. Schmid, M. C. Caldeira, P.G. Dimitrakopoulos, J. Good, R. Harris, A. Hector, K. Huss-Danell, A. Jumpponen, A. Minns, C. P. H. Mulder, J. S. Pereira, A. Prinz, M. Scherer-Lorenzen, A. -S. D. Siamantziouras, A. C. Terry, A. Y. Troumbis & J. H. Lawton. 2001. Local adaptation enhances performance of common plant species. Ecology Letters 4:536-544.

Kawecki, T. J. & D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters. 7: 1225-1241.

Kawecki, T. J. 2000.Adaptation to marginal habitats: contrasting influence of the dispersal rate on the fate of alleles with small and large effects. Proceedings of the Royal Society of London Biological Sciences 267:1315-1320.

Keightley P. D. & M. Lynch. 2003. Toward a realistic model of mutations affecting fitness. Evolution 57:683-685.

Koenig, W.D. & M. V. Ashley. 2003. Is pollen limited? The answer is blowin' in the wind. Trends in Ecology & Evolution 18:157-159.

Leimu, R., P. Mutikainen, J. Koricheva & M. Fischer. 2006. How general are positive relationships between plant population size, fitness and genetic variation? Journal of Ecology 94:942-952.

Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends in Ecology & Evolution 17: 183-189.

Linhart, Y. B. & M. C. Grant. 1996. Evolutionary significance of local genetic differentiation in plants. Annual Review of Ecology Evolution and Systematics 27:237-277.

Liu, Y. F., Y. Wang & H. W. Huang. 2006. High interpopulation genetic differenciation and unidirectional linear migration patterns in *Myricaria laxiflora* (Tamaricaceae), an

endemic riparian plant in the Three Gorges valley of the Yangtze river. American Journal of Botany 93:206-215.

Lynch M., J. Blanchard, D. Houle, T. Kibota, S. Schultz, L. Vassilieva & J. Willis. 1999. Perspective: spontaneous deleterious mutations. Evolution 53:645-663.

Lynch, M., J. Conery & R. Burger. 1995. Mutation accumulation and the extinction of small populations. The American Naturalist 146:489-518.

Lythgoe, K. A. 1997. Consequences of gene flow in spatially structured populations. Genetical Research 69: 49-60.

MacNair M.R. 1987. Heavy-metal tolerance in plants: a model evolutionary system. Trends in Ecology & Evolution 2: 354-359.

Madriñán, L. F., A. Etter, G. D. Boxall & A. Ortega-Rubio. 2007. Tropical alluvial forest fragmentation in the eastern lowlands of Columbia (1939-1997). Land Degradation & Development 18:199-208.

Martin, G. & T. Lenormand. 2006a. A general multivariate extension of Fisher's geometrical model and the distributin of mutation fitness effects across species. Evolution 60:893-907.

Martin, G. & T. Lenormand. 2006b. The fitness effect of mutations across environments: A survey in light of fitness landscape models. Evolution 60:2413-2427.

Martin, G. & T. Lenormand. 2007. Distribution of epistasis in microbes fit predictions from a fi tness landscape model. Nature Genetics 39:555-560.

McNeilly, T. 1968. Evolution in closely adjacent plant populations .3. *Agrostis tenuis* on a small copper mine. Heredity. 23: 99-108.

Muller, H. J. 1950. Our load of mutations. The American Journal of Human Genetics 2:111-176.

Noël, F., E. Porcher, J. Moret, J., & N. Machon. 2006. Connectivity, habitat heterogeneity and population persistence in *Ranunculus nodiflorus*, an endangered species in France. New Phytologist 169:71-83.

Ohara, M., H. Tomimatsu, T. Takada & S. Kawano. 2006. Importance of life history studies for conservation of fragmented populations: A case study of the understory herb, *Trillium camschatcense*. Plant Species Biology 21:1-12.

Ouborg, N. J., Y. Piquot & J. M. Van Groenendael. 1999. Population genetics, molecular markers and the study of dispersal in plants. Journal Of Ecology 87:551-568 .

Paland S. & B. Schmid. 2003. Population size and the nature of genetic load in *Gentianella germanica*. Evolution 57:2242-2251.

Paschke, M., G. Bernasconi & B. Schmid. 2005. Effects of inbreeding and pollen donor provenance and diversity on offspring performance under environmental stress in the rare plant *Cochlearia bavarica*. Basic and Applied Ecology 6:325-338.

Phadnis N. & J. D. Fry. Widespread correlations between dominance and homozygous effects of mutations: implications for theories of dominance. Genetics 171:385-392.

Phillips, P. C. 1996. Maintenance of polygenic variation via a migration-selection balance under uniform selection. Evolution. 50: 1334-1339.

Poon, A. & S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54:1467-1479.

Rehfeldt, G. E., N. M. Tchebakova, Y. I. Parfenova, W. R. Wykoff, N. A. Kuzmina & L. I. Milyutin. 2002. Intraspecific responses to climate in *Pinus sylvestris*. Global Change Biology 8: 912-929.

Reinartz, J. A. & D. H. Les. 1994. Bottleneck-induced dissolution of self-incompatibility and breeding consequences in *Aster furcatus* (Asteraceae). American Journal of Botany 81:446-455.

Remold, S. K. & R. E. Lenski. 2004. Pervasive joint influence of epistasis and plasticity on mutational effects in *Escherichia coli*. Nature Genetics 36:423-426.

Richards, C. M. 2000. Inbreeding depression and genetic rescue in a plant metapopulation. The American Naturalist 155:383-394.

Rousset, F. 2004. Genetic structure and selection in subdivided populations. Éditeurs S. A. Levin & H. S. Horn. Princeton University Press, Princeton & Oxford.

Roze, D. & F. Rousset. 2003. Selection and drift in subdivided populations: A straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. Genetics 165:2153-2166.

Roze, D. & F. Rousset. 2004. Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. Genetics 167:1001-1015.

Savolainen, O., F. Bokma, R. Garcia-Gil, P. Komulainen & T. Repo. 2004. Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of *Pinus sylvestris* to climatic changes. Forest Ecology and Management. 197: 79-89.

Schmitt J. & S. E. Gamble. 1990. The effect of distance from the parental site on offspring performance and inbreeding depression in *Impatiens capensis*: a test of the local adaptation hypothesis. Evolution 44:2022-2030.

Schultz S. T., M. Lynch & J. Willis. 1999. Spontaneous deleterious mutations in *Arabidopsis*

thaliana. Proceedings of the National Academy of Sciences of the United States of America 96:11393-11398.

Shaw F. H., C. J. Geyer & R. G. Shaw. 2002. A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. Evolution 56:453-463.

Shaw R. G., F. H. Shaw & C. Geyer. 2003. What fraction of mutations reduces ditness ? A reply to Keightley and Lynch. Evolution 57:686-689.

Storfer, A. 1999. Gene flow and endangered species translocations: a topic revisited. Biological Conservation 87:173-180.

Tallmon, D. A., G. Luikart & R. S. Waples. 2004. The alluring simplicity and complex reality of genetic rescue. Trends in Ecology & Evolution 19:489-496.

Theodorou, K. & D. Couvet. 2006a. Genetic load in subdivided populations: interactions between the migration rate, the size and the number of subpopulations. Heredity 96:69-78.

Theodorou, K. & D. Couvet. 2006b. On the expected relationship between inbreeding, fitness, and extinction. Genetics Selection Evolution 38:371-387.

Tufto, J. 2000. Quantitative genetic models for the balance between migration and stabilizing selection. Genetical Research 76:285-293.

Whitlock, M. C. 2002. Selection, load and inbreeding depression in a large metapopulation. Genetics 160:1191-1202.

Whitlock, M. C. 2003. Fixation probability and time in subdivided populations. Genetics 164:767-779.

Willi, Y. & M. Fischer. 2005. Genetic rescue in interconnected populations of small and large size of the self-incompatible *Ranunculus reptans*. Heredity 95:437-443.

Willi, Y., J. Van Buskirk & M. Fischer. 2005. A threefold genetic Allee effect: population size affects cross-compatibility, inbreeding depression and drift load in the self-incompatible *Ranunculus reptans*. Genetics 169:2255-2265.

Wright S. 1969. Evolution and the genetics of populations-Volume 2 The theory of gene frequencies. Chicago :University of Chicago Press.

Yeaman, S. & A. Jarvis. 2006. Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. Proceedings of the Royal Society of London Biological Sciences 273:1587-1593.