



Conséquences de l'autogamie sur le potentiel évolutif des populations approches empirique, expérimentale et théorique

Josselin Clo

► To cite this version:

Josselin Clo. Conséquences de l'autogamie sur le potentiel évolutif des populations approches empirique, expérimentale et théorique. Génétique des plantes. Montpellier SupAgro, 2020. Français. NNT : 2020NSAM0023 . tel-04069510

HAL Id: tel-04069510

<https://theses.hal.science/tel-04069510>

Submitted on 14 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Science de l'évolution et de la biodiversité

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau
Portée par

Unité de recherche AGAP

Conséquences de l'autogamie sur le potentiel évolutif des populations : approches empirique, expérimentale et théorique

Présentée par Josselin CLO

Le 07 décembre 2020

Sous la direction de Joëlle RONFORT
et Laurène GAY

Devant le jury composé de

Ruth SHAW, Professor, Université du Minnesota
Thomas BATAILLON, Associate Professor, Université d'Aarhus
Mathilde DUFAÏ, Professeur, Université de Montpellier
Yvonne WILLI, Associate Professor, Université de Bâle
Joëlle RONFORT, Directrice de recherche, INRAE
Laurène GAY, Chargée de recherche, INRAE
Diala ABU AWAD, Post-doctorante, Université technique de Munich

Rapporteur
Rapporteur
Présidente
Examinatrice
Directrice de thèse
Co-encadrante de thèse
Invitée



UNIVERSITÉ
DE MONTPELLIER



“Either inbreeding or selection, never both at the same time”

Sir R. A. Fisher

REMERCIEMENTS

On dit souvent que la thèse est faite de hauts et de bas, mais je dois dire que grâce aux différentes personnes que j'ai pu côtoyer durant ma thèse, cela a été trois merveilleuses années (avec quelques creux certes).

Je veux remercier Vincent Girard, Francesca Rossi et Sandrine Maurice, qui, durant mes différents stages de L3, m'ont fait découvrir le monde de la recherche. Je remercie tout particulièrement Sandrine, à qui je dois mon initiation à l'étude des systèmes de reproduction. Dire qu'à la base je devais travailler sur des moustiques, et que par hasard je me suis retrouvé à compter des grains de pollen au microscope. Comme quoi, un sujet de recherche peut se jour à un grain de pollen de silène gynodioïque.

Je veux remercier mon super comité de thèse, Céline Devaux, Ophélie Ronce, Patrice David, Sylvain Glémin, Denis Roze, Jean-Marc Lacape et Xavier Perrier. Nos différentes discussions m'ont fortement aidé dans la rédaction de mes différents chapitres de thèse. Je remercie également Ruth Shaw, Yvonne Willi, Mathilde Dufaÿ et Thomas Bataillon pour avoir accepté d'évaluer mon manuscrit et mon oral de thèse.

Je remercie l'équipe GE²POP, qui m'a apportée la bonne humeur pendant ces 3 années. Un grand merci à Sylvain, Jacques, Martin, Pierre, Nathalie, Hélène, Vincent, Murielle T, Jean-Marie, Morgane, Murielle L, Véro, Audrey, Christine, Denis, Philippe, Pascal et Karine. Merci également aux post-docs et thésards qui ont partagé cette aventure : Mathilde, Elsa, Margaux, Germain, Yacine, Pauline et Céline. Courage à ceux qui n'ont pas encore fini !

Merci à mes autres copains thésards qui n'étaient pas dans l'équipe, mais avec qui cela est toujours un plaisir de discuter : Claire, Maurine, Cécile, Maxime, Pierre, Antoine, Fiona, Agathe, courage à tous ! Un mot particulier pour Geoffroy, mon ami de toujours, c'est bientôt fini pour toi aussi !

Je voudrais remercier Sandrine Maurice, Céline Devaux, Pierrick Labbé, Vincent Ranwez et Jacques David pour m'avoir permis d'enseigner la biologie évolutive à différents niveaux universitaires. Ce fut une expérience plus que formatrice, et j'ai pu me rendre compte que l'enseignement est un exercice compliqué, mais finalement assez gratifiant.

J'ai eu la chance d'organiser et participer à de nombreux événements scientifiques, qui m'ont permis de me faire réfléchir à autre chose qu'à ma génétique quantitative habituelle. Merci aux copains de l'organisation du séminaire MEE (de loin la meilleure édition selon un panel non-objectif), Maria, Auriane et Maximilien. Merci au groupe de travail *Evolvabilty* de l'université d'Oslo, les deux séjours en Norvège ont été un vrai plaisir et des supers expériences scientifiques. Merci Christophe et Thomas pour les invitations, et Agnes, Øystein, David, Michael, Jacqueline, Frietson, Masahito, Arnaud, Geir, Jane, Filip et tous les autres pour ces discussions et ces maux de tête. Je n'imaginais pas que la variance additive pouvait faire autant souffrir ! Je remercie particulièrement Øystein, qui a accepté de se lancer dans ce dernier chapitre un peu fou de génétique quantitative avec dominance directionnelle. Tu n'as pas su voir le coupe-gorge que je te tendais, mais le résultat final me laisse penser que ce n'est pas notre dernière collaboration.

Un grand merci à mes encadrantes. Diala, même si tu n'es pas officiellement mon encadrante, tu as eu un rôle important dans cette thèse, un grand merci, même si ce fichu modèle nous a donné du fil à retordre ! Le pire dans cette histoire, c'est qu'on pense encore à faire des simulations ensemble... ;-) Et comment vous remercier, Joëlle et Laurène, tant vous m'avez apporté ? Vous m'avez aidé à devenir le jeune chercheur indépendant que je suis aujourd'hui. Vous avez laissé libre cours à mon imagination (parfois un peu trop débordante), et permis de réaliser pleins de choses en plus de ma thèse. Je pense qu'on a formé une bonne équipe, et cette thèse est la récompense de notre super collaboration. Grâce à vous j'ai cherché l'épistasie, et j'ai trouvé la dominance.

Et enfin, un merci inconditionnel à ma famille. Papa et maman, merci d'avoir tout fait pour que je puisse réussir mes études. Madeline, Maximilien mes frère et sœur adorés. Les petits et grands Angélique, Joseph et Aaron pour avoir agrandi la famille. Mamie, qui pense toujours que je vais à l'école. Et enfin les trois amours qui partage ma vie au quotidien et que j'aime profondément, Elmer, Maud et le petit bout de chou qui pousse sereinement dans le ventre de sa maman. Vous ensoleillez ma vie, et si cette thèse a été aussi facile, c'est aussi grâce à l'amour que vous me donnez tous les jours et qui me donne une force incommensurable.

Grâce à tous ces bons moments, je pars continuer mon apprentissage du monde de la recherche l'esprit tranquille, sûr de mes forces et de ce qu'il me reste à acquérir.

A vous tous, merci !

Table des matières

Introduction	4
I – Distribution et évolution de l'autofécondation en populations naturelles.....	4
1 – Des extrêmes bien représentés, et un continuum entre les deux	4
2 – L'évolution vers les extrêmes : hypothèses génétiques et écologiques.....	6
3 – L'autofécondation et l'hypothèse du cul de sac évolutif.....	9
II – Adaptation et systèmes de reproduction	10
1 – Systèmes de reproduction et adaptation à partir de nouvelles mutations	10
2 – Systèmes de reproduction et adaptation à partir de la diversité génétique préexistante.....	13
III – Variance génétique et autofécondation : une introduction à la génétique quantitative des espèces autogames.....	16
1 – L'effet de l'autofécondation sur le niveau de variance génétique, sous hypothèse de pure additivité.....	16
Les attendus théoriques	16
Confrontation avec les données empiriques	18
2 – L'effet de la dominance sur la décomposition de la variance génétique et sur la réponse à la sélection en régime autogame	20
Décomposer la variance génétique en présence de dominance et d'autofécondation	20
Intégrer ces nouvelles composantes de la variance génétique dans les modèles de réponse à la sélection.....	22
IV – Objectifs et présentation de la thèse	26
1 – Principaux objectifs.....	26
2 – Chapitre 1 : Méta-analyse de l'effet de l'autofécondation sur la variance génétique au sein des populations naturelles d'angiospermes.....	27
3 – Chapitre 2 : Le résidu d'allogamie et les covariances génétiques pourraient faciliter l'adaptation des espèces autogames aux changements environnementaux	27
4 – Chapitre 3 : Etude empirique sur l'importance de la dominance dans l'isolement reproductif entre lignées consanguines dans les populations naturelles de <i>Medicago truncatula</i>	28
5 – Chapitre 4 : Au-delà de l'additivité, de l'importance de considérer l'effet de la dominance sur le niveau et la structure de la diversité génétique des populations.....	28
Bibliographie	29
Chapitre 1 : Méta-analyse de l'effet de l'autofécondation sur la variance génétique au sein des population naturelle d'angiospermes.....	40
Présentation générale.....	41
How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species	41

Matériel supplémentaire	55
Chapitre 2 : Le résidu d'allogamie et les covariances génétiques pourraient faciliter l'adaptation des espèces autogames aux changements environnementaux	58
Présentation générale.....	59
Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations	59
Matériel supplémentaire	73
Chapitre 3 : Etude empirique sur l'importance de la dominance dans l'isolement reproductif entre lignées consanguines dans les populations naturelles de <i>Medicago truncatula</i>	88
Présentation générale.....	89
Fitness consequences of hybridization between natural inbred lines of the selfing species <i>Medicago truncatula</i> : investigating the role of dominance and genetic incompatibilities	89
Crosses	95
Greenhouse experiment.....	97
Matériel supplémentaire	120
Chapitre 4 : Au-delà de l'additivité, de l'importance de considérer l'effet de la dominance sur le niveau et la structure de la diversité génétique des populations.....	132
Présentation générale.....	133
Genetics of quantitative traits with directional dominance in partially selfing species: implications for evolution in changing reproductive environment	133
Matériel supplémentaire	170
Discussion générale et perspectives	178
Résumé des principaux résultats	179
Discussion : Repenser l'hypothèse du cul-de-sac évolutif	181
Les raisons des transitions unidirectionnelles du système de reproduction	181
Le rôle des mutations délétères dans la transition de l'autogamie vers l'allogamie.....	181
Les transitions de l'autogamie vers l'allogamie freinées par la dépression de croisement	183
Les coûts potentiels d'une transition de l'autogamie vers l'allogamie.....	184
Le taux de diversification négatif des populations autogames	186
Les rôles des mutations dans la probabilité d'adaptation et l'extinction des populations.....	186
L'adaptation des populations autogames à partir de la variation génétique pré-existante	188
L'autofécondation n'affecte pas ou peu le niveau de variance génique en populations naturelles	189
Les expériences de sélection artificielles, et l'effet de l'autofécondation sur la variance additive réalisée	189
L'accumulation de variance via les associations génétiques, et les contraintes à sa remobilisation	192

Hypothèses alternatives : extinctions globales vs locales, et l’hypothèse du suicide évolutif	194
Perspectives	196
Bibliographie	197
Annexe 1 - Digest: How mutational bias could explain the maintenance of sex	202
Annexe 2 - Very local gene dispersal in the pseudometallicolous herb <i>Noccaea caerulea</i> in the south of France	205
Annexe 3 - How mating systems influence evolvability	207

Introduction

I – Distribution et évolution de l'autofécondation en populations naturelles

Le mode de reproduction est un trait d'histoire de vie majeur d'une espèce, car il détermine la manière dont les gènes sont transmis d'une génération à la suivante. Le mode de reproduction renvoie généralement à la dichotomie entre espèces sexuées et asexuées, alors que le terme système de reproduction renvoie plutôt aux différents modes de reproduction sexuée. Une large gamme de systèmes de reproduction existe au sein du vivant, que ce soit dans le règne animal ou végétal, avec des espèces où la reproduction passe uniquement par l'union de deux gamètes provenant d'individus différents, comme chez les espèces auto-incompatibles, polyandres ou polygynes, et d'autres espèces où la reproduction allie à la fois l'union de gamètes issus d'individus différents et un degré plus ou moins élevé d'autofécondation (union de deux gamètes provenant d'un même individu hermaphrodite). Au cours de cette thèse, je me suis intéressé tout particulièrement à l'autofécondation, et à son impact sur l'organisation et le potentiel évolutif des populations.

1 – Des extrêmes bien représentés, et un continuum entre les deux

Les organismes hermaphrodites, qui possèdent à la fois les organes reproducteurs mâle et femelle sur un même individu (ou fleur dans le cas des plantes), sont très répandus chez les animaux et végétaux. On estime que les pourcentages d'espèces hermaphrodites sont de 72% chez les angiospermes (Richards, 1997), et 33% chez les animaux (hors insectes, Jarne & Auld, 2006). Cette caractéristique permet théoriquement l'autofécondation, bien qu'il existe de nombreux systèmes d'évitement de la consanguinité, comme des systèmes d'auto-incompatibilité moléculaires (Allen & Hiscock, 2008; Goubet *et al.*, 2012), une séparation spatiale des organes mâles et femelles (appelée herkogamie, Opedal, 2018), ou une maturation décalée dans le temps de ces mêmes organes (respectivement protandrie ou protogynie selon que ce sont les organes mâles ou femelles qui sont les plus précocement matures, Lloyd & Webb, 1986).

On distingue généralement trois grands systèmes de reproduction en fonction du taux d'autofécondation estimé au sein des populations. On parle généralement d'espèce allogame lorsque le taux d'autofécondation est en moyenne inférieur à 20%, d'espèces à régime de reproduction mixte lorsque le taux d'autofécondation est compris entre 20% et 80% et d'espèce autogame lorsque le taux d'autofécondation est supérieur à 80%. La distribution des taux d'autofécondation, estimés à l'échelle des populations, est bimodale, chez les plantes comme chez les animaux (figure 1), avec une haute fréquence d'espèces allogames, et dans une moindre mesure d'espèces autogames, et un continuum d'espèces à régime mixte en faible fréquence (figure 1, Schamske & Lande, 1985; Goodwillie *et al.*, 2005; Jarne & Auld, 2006). De plus, si les taux d'allofécondation peuvent atteindre 1 (notamment dans le cas des espèces auto-incompatibles), les taux d'autofécondation restent inférieurs à 1 dans un grand nombre de populations (seulement 1% des populations ont un taux d'autofécondation de 1 dans la méta-analyse de Winn *et al.* 2011).

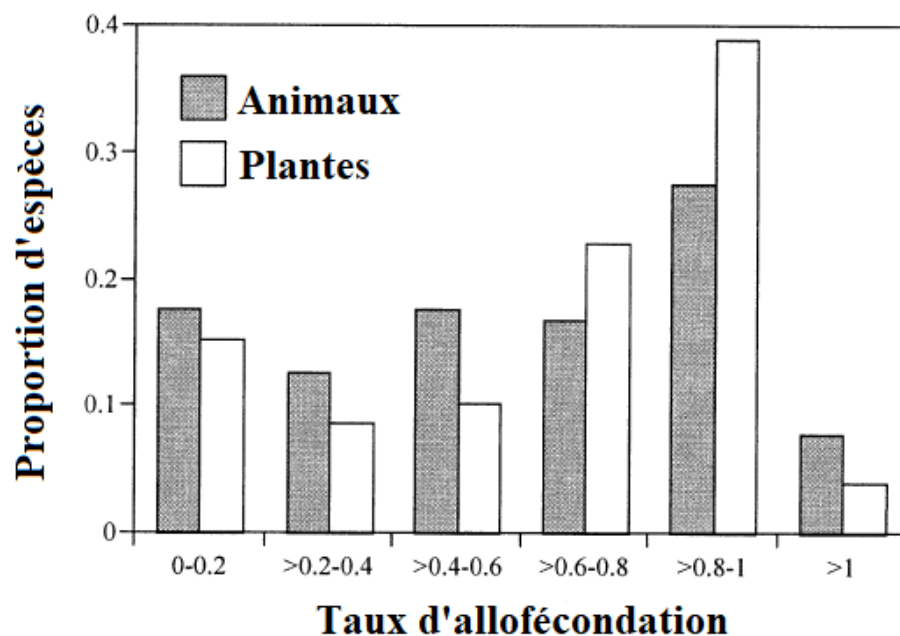


Figure 1. Distribution des taux d'allofécondation au sein des espèces animales (barres grises, 142 espèces) et végétales (barres blanches, 342 espèces). Figure adaptée de Jarne & Auld (2006).

Bien que l'on parle d'espèce autogame, allogame ou à régime mixte, le taux d'autofécondation peut très fortement varier au sein d'une même espèce, dans le temps (d'une génération à l'autre) ou l'espace (entre populations) (Schemske & Lande, 1985; Shaw *et al.*, 1998; Whitehead *et al.*, 2018). Par exemple, Barrett & Husband (1990) ont caractérisé la variation du taux d'autofécondation au sein de 32 populations sud-américaines de l'espèce *Eichhornia paniculata*, et ont montré que le taux variait de 4% à 99.8%. De nombreux facteurs peuvent expliquer cette forte variation au sein d'une espèce (voir Barrett & Eckert, 1990 pour une revue). On retrouve généralement la taille de la population (Barrett & Husband, 1990), la densité en plantes qui joue un rôle majeur sur le comportement des pollinisateurs (Karron *et al.*, 1995; Devaux *et al.*, 2014), ou encore la variation du niveau d'herkogamie entre individus ou populations (Opedal, 2018).

2 – L'évolution vers les extrêmes : hypothèses génétiques et écologiques

Théoriquement, cette bi-modalité dans la distribution des taux d'autofécondation est attendue, et plusieurs mécanismes ont été proposés afin de l'expliquer. Historiquement, les modèles d'évolution de l'autofécondation se sont d'abord tournés vers des hypothèses génétiques. Fisher (1941) a d'abord proposé que l'apparition d'un gène permettant à une population ancestrale strictement allogame de s'autoféconder devrait entraîner une rapide fixation, car les gènes d'autofécondation ont un avantage de transmission de 50% par rapport aux individus allogames : une plante transmet deux copies de ses gènes dans les graines qu'elle produit en autofécondation contre une seule en allofécondation. Cependant, une autre force majeure peut contrer cet avantage de transmission : la dépression de consanguinité, définie comme la baisse de valeur sélective des individus issus d'autofécondation comparé aux individus issus de croisements. Si le niveau de dépression de consanguinité est supérieur à 50%, alors l'évolution vers l'autofécondation est impossible (Lande & Schemske, 1985). Au contraire, si dans une population, la dépression de consanguinité est inférieure à 50%, l'autofécondation pourrait augmenter jusqu'à atteindre 100% (Lande & Schemske, 1985). Néanmoins, d'autres facteurs peuvent faire varier le seuil de dépression de consanguinité pour lequel une population devrait évoluer (en théorie) vers l'autofécondation complète.

L'un de ces facteurs est la purge des allèles délétères. L'excès d'homozygotie dans une descendance consanguine rend la purge plus efficace que dans une descendance issue de

croisements (Lande & Schemske, 1985). Ceci peut générer un déséquilibre de liaison négatif entre les allèles augmentant le taux d'autofécondation et les allèles délétères. Dans ce cas, on retrouve les allèles modificateurs du taux d'autofécondation dans les fonds génétiques « purgés », ce qui peut permettre l'évolution vers l'autogamie pour des niveaux de dépression de consanguinité initiaux supérieurs à 50% (Figure 2, Charlesworth *et al.*, 1990; Awad & Roze, 2020). Le phénomène est d'autant plus fort que la sélection purificatrice est forte et le taux de recombinaison est faible, permettant de maintenir le déséquilibre de liaison pendant plusieurs générations (Charlesworth *et al.*, 1990; Uyenoyama & Waller, 1991; Epinat & Lenormand, 2009).

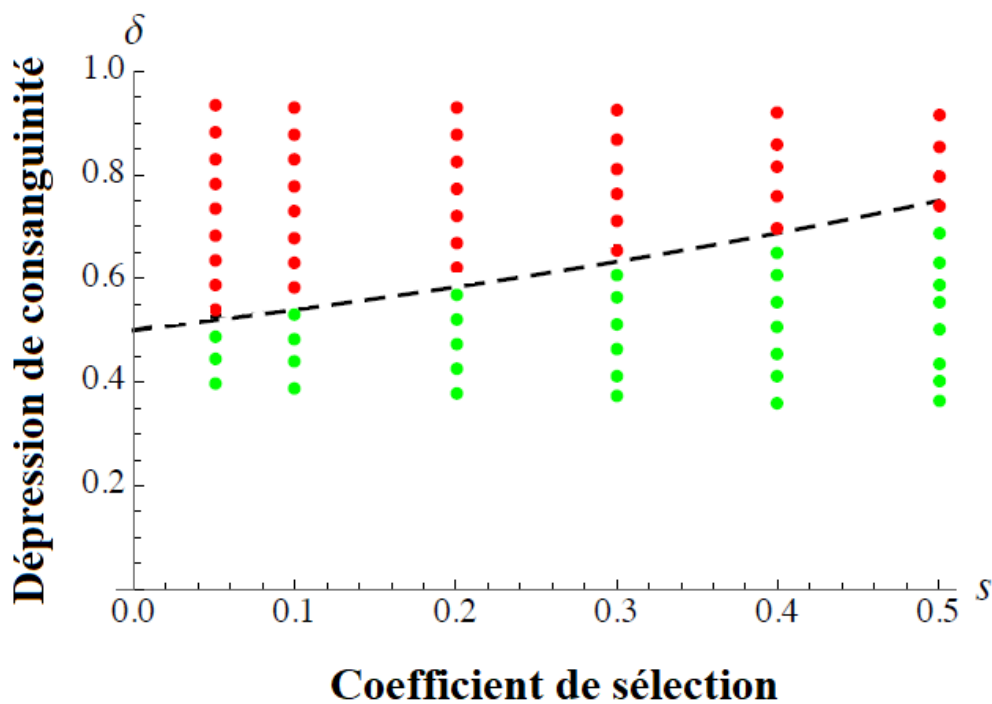


Figure 2. Evolution de l'autofécondation pour une population allogame ayant un niveau de dépression de consanguinité (δ), en fonction du coefficient de sélection contre les mutations délétères (s). La droite pointillée représente le niveau théorique de dépression de consanguinité maximal pour évoluer vers l'autofécondation quasi-strict, d'après l'équation [37] d'Abu Awad & Roze (2020). Les points représentent des résultats de simulations. Les points verts correspondent au cas où la population allogame ancestrale évolue vers l'autofécondation, les points rouges indiquent les cas où la population ancestrale reste allogame. Figure modifiée de Abu Awad & Roze (2020).

La limite de ces modèles est qu'ils permettent de prédire l'évolution vers l'autofécondation complète, ou le maintien de l'allofécondation, mais sont incapables d'expliquer (1) la forte proportion de régimes mixtes, et (2) le maintien d'un résidu d'allofécondation comme observé chez de nombreuses espèces très autogames, comme *Arabidopsis thaliana*, *Medicago truncatula* ou *Caenorhabditis elegans*. Plusieurs modèles théoriques expliquent ces deux patrons. Je vais présenter une liste non-exhaustive de modèles qui reprennent les grandes hypothèses avancées.

Des modèles expliquant le maintien de taux d'autofécondations intermédiaires ont été proposés en écologie. Par exemple, en simulant des populations plus ou moins denses en individus, Holsinger (1991) a montré qu'un taux d'autofécondation intermédiaire peut être maintenu dans des populations peu denses, dans lesquelles la probabilité de succès du pollen exporté est limitée. Dans ces populations, un taux intermédiaire d'autofécondation permet en effet une assurance de reproduction. Johnston (1998) a également montré qu'un taux intermédiaire peut être maintenu quand (1) une augmentation du taux d'autofécondation réduit la probabilité de fertiliser les ovules d'autres plantes (on parle de *seed discounting*), ou (2) quand l'autofécondation limite la production de pollen d'une plante (on parle alors de *pollen discounting*, phénomène fréquemment observé, Slotte *et al.*, 2012), limitant ainsi l'avantage reproductif décrit par Fisher (1941).

Le maintien d'un résidu d'allofécondation chez les espèces préférentiellement autogames est également soutenu par plusieurs modèles théoriques. Charlesworth *et al.* (1991) ont observé qu'un effet synergétique (épistatique) des mutations délétères entraîne un taux d'autofécondation évolutivement stable inférieur à 1. Bien qu'ils ne puissent l'expliquer concrètement, les auteurs pensent que cela vient du comportement de la valeur sélective moyenne des populations dans cette zone de très forts taux d'autofécondation (Charlesworth *et al.*, 1991). Dans un modèle plus classique, sans interactions épistatiques entre mutations délétères, Kamran-Disfani & Agrawal (2014) ont montré qu'un faible taux d'allogamie résiduelle permet de limiter la création de déséquilibre de liaison négatif entre mutations délétères et avantageuses, en augmentant le taux de recombinaison efficace, cet avantage empêchant les populations d'atteindre l'autofécondation complète. Enfin David *et al.* (1993) ont montré qu'un faible niveau d'allogamie résiduelle durant une phase de sélection directionnelle permet une meilleure réponse à la sélection dans le temps par rapport aux populations strictement autofécondantes. En effet, le résidu d'allofécondation permet le

brassage de la diversité génétique de la population tandis que l'autogamie permet le maintien et la sélection efficace des combinaisons génétiques favorables (David *et al.* 1993).

3 – L'autofécondation et l'hypothèse du cul de sac évolutif

Malgré les différents avantages liés à l'autofécondation, comme une purge plus efficace des mutations délétères (Lande & Porcher, 2015; Abu Awad & Roze, 2018), l'avantage de transmission (Fisher, 1941), ou l'assurance de reproduction (Baker, 1955), l'autofécondation n'est pas le système de reproduction majoritaire (Figure 1 : environ 15% et 18% respectivement pour les angiospermes et les animaux, Jarne & Auld, 2006). De plus, la majorité des transitions qui ont été documentées s'opèrent de l'allofécondation vers l'autofécondation (Igic & Busch, 2013). Enfin, Goldberg *et al.* (2010) ont mis en évidence, chez les Solanacées, un taux de diversification plus faible des espèces auto-compatibles (potentiellement capables de s'autoféconder, Figure 3) par rapport aux auto-incompatibles (allogames strictes, Figure 3). Ceci suggère que l'autofécondation pourrait être un cul-de-sac évolutif (*evolutionary dead-end*, Stebbins, 1957). Ainsi, il semblerait que les avantages à court terme associés à l'autofécondation soient compensés par de lourds désavantages sur des pas de temps plus longs, qui mèneraient les populations à l'extinction (Takebayashi & Morrell, 2001).

L'hypothèse selon laquelle l'autogamie préférentielle serait un cul-de-sac évolutif s'appuie sur deux grands mécanismes. Le premier stipule que les populations autogames, du fait de leurs plus faibles tailles efficaces (Glémin & Ronfort, 2013), et de leur taux de recombinaison efficace plus faible (Nordborg, 2000), sont plus à même de fixer des mutations délétères. Ces dernières peuvent entraîner l'extinction des populations ; on parle de fonte mutationnelle, ou *mutational meltdown*. Ce phénomène est bien détaillé et jugé plausible théoriquement (Abu Awad & Billiard, 2017), mais peu de données soutiennent ce processus (voir Escobar *et al.* 2010 & Willi, 2013 pour des cas d'études). Néanmoins, l'extinction d'une population par fonte mutationnelle est théoriquement extrêmement rapide, tellement rapide que ce n'est pas un mécanisme plausible pour expliquer le patron observé à l'échelle phylogénétique (Abu Awad & Billiard, 2017). Le second mécanisme avancé est que les populations autogames ont une capacité réduite à s'adapter à des changements environnementaux par rapport aux populations allogames (Stebbins, 1957), hypothèse que je vais développer dans la section suivante.

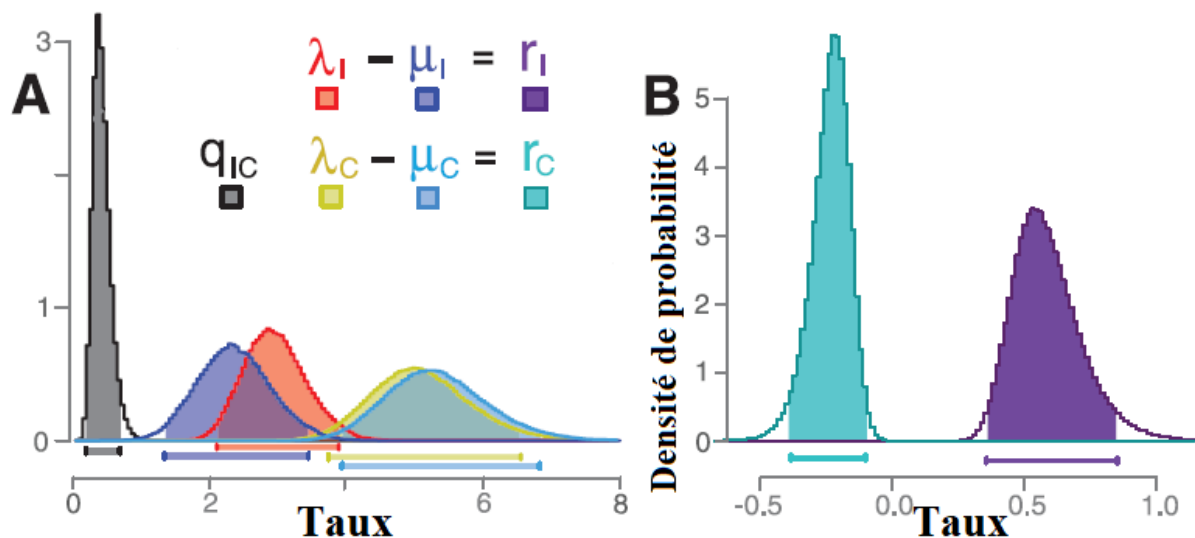


Figure 3. Distribution postérieure et intervalles de crédibilité (à 95%) des taux de spéciation (λ), extinctions (μ), diversification (r), et de transitions de l'auto-incompatibilité vers l'auto-compatibilité (q), estimés à partir de données phylogénétiques obtenues sur 356 espèces de la famille des solanacées. Le suffixe « I » désigne les estimations pour les espèces auto-incompatibles (41% des espèces), et le suffixe « C » dénote les espèces auto-compatibles (57% des espèces). La partie **A** décrit les paramètres du modèle, alors que la partie **B** représente les estimations des taux de diversifications (r). Figure adaptée de Goldberg *et al.* (2010).

II – Adaptation et systèmes de reproduction

L'adaptation à un nouvel environnement peut se faire, via l'apparition de nouvelles mutations (Hartfield & Glémin, 2014, 2016), ou à partir de la variation génétique préexistante dans la population (Glémin & Ronfort, 2013). Ces deux mécanismes ont été étudiés en détails, notamment via des modèles théoriques.

1 – Systèmes de reproduction et adaptation à partir de nouvelles mutations

Haldane (1927) fut le premier à décrire l'effet de l'autofécondation sur l'adaptation à partir de nouvelles mutations, avec un modèle mono-locus. Il a montré que les populations autogames sont plus à même de fixer des mutations avantageuses récessives, alors que les

populations allogames fixent plus facilement les mutations avantageuses dominantes (Figure 4). Ce résultat, nommé « le tamis d'Haldane », a été confirmé analytiquement (Pollak, 1987; Caballero & Hill, 1992; Charlesworth, 1992; Glémin & Ronfort, 2013), et si la dérive est négligeable par rapport à la force de la sélection naturelle, la probabilité de fixation d'une mutation avantageuse P_{new} est égale à

$$P_{new} = \frac{2(h_b + f - h_b f)s_b}{1 + f},$$

avec h_b étant le coefficient de dominance, s_b le coefficient de sélection de la mutation avantageuse, et f le coefficient de consanguinité de la population d'étude.

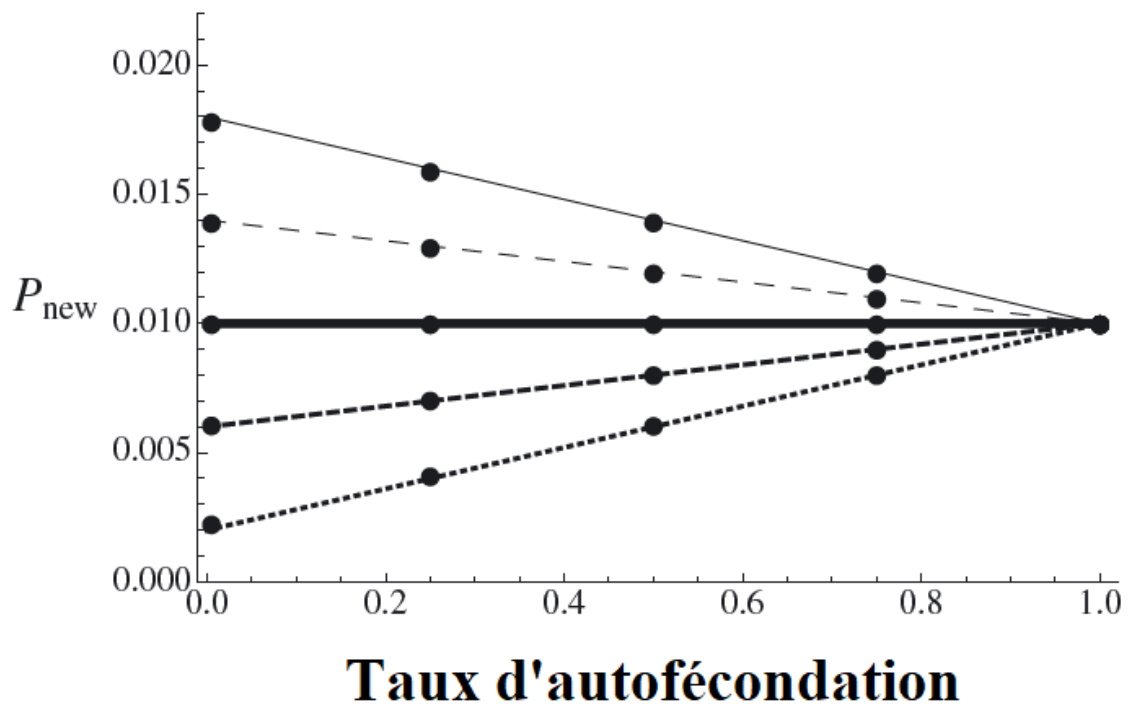


Figure 4. Probabilité de fixation d'une mutation avantageuse en fonction du taux d'autofécondation. De haut en bas, le coefficient de dominance h de la mutation avantageuse est égal à 0.9, 0.7, 0.5, 0.3 et 0.1. Le coefficient de sélection de l'allèle avantageux s est fixé à 0.01. Les points noirs représentent des résultats de simulations. Figure modifiée de Glémin & Ronfort (2013).

Il apparaît, à partir de cette équation, que plus l'avantage sélectif conféré par l'allèle avantageux est fort, plus sa probabilité de fixation est forte (il a moins de chance d'être perdu par dérive). De plus, plus le taux d'autofécondation augmente, moins le coefficient de dominance de la mutation avantageuse influe sur la probabilité d'adaptation (figure 4).

Un autre résultat intéressant est que la vitesse de fixation d'une nouvelle mutation avantageuse est plus rapide chez les espèces autogames (Glémin, 2012; Glémin & Ronfort, 2013). Dans le cas où la mutation est récessive, cela s'explique par le fait que l'autofécondation permet de mettre rapidement la mutation à l'état homozygote. Quand la mutation est dominante, ceci peut s'expliquer par le fait que l'autofécondation va augmenter la variance héritable (au sens de Cockerham & Weir, 1984) de la valeur sélective introduite par la nouvelle mutation sur tout ou partie de la phase de fixation de l'allèle (Figure 5), augmentant ainsi sa vitesse de fixation (Fisher, 1958).

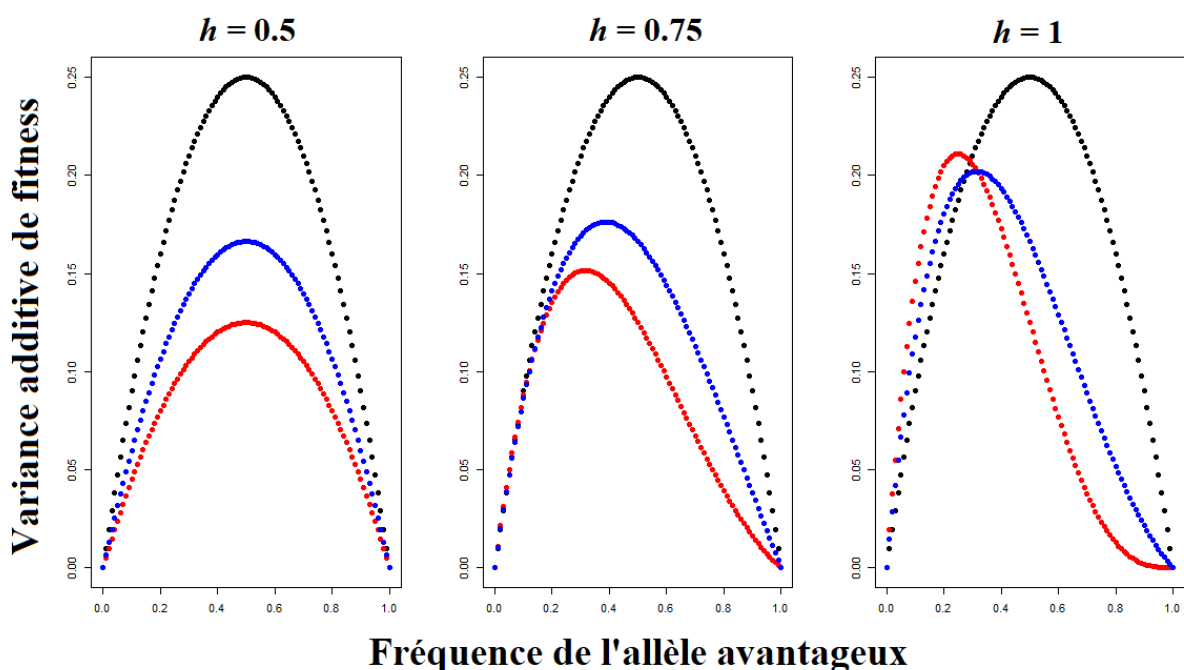


Figure 5. Variance additive de valeurs sélectives générée par une nouvelle mutation avantageuse durant sa phase de fixation, pour 3 coefficients de dominance ($h = 0.25$ & 0 sont similaires à $h = 0.75$ et 1, les pics sont juste inversés), et 3 niveaux d'autofécondation (0%, 50% et 100% d'autofécondation respectivement pour les courbes rouges, bleues et noires). Les valeurs sélectives des différents génotypes sont : $W(AA) = 0$, $W(Aa) = h$ et $W(aa) = 1$. Josselin Clo, résultats non-publiés.

Plus récemment Hartfield & Glémin (2014, 2016) ont étendu ces résultats mono-locus pour obtenir des prédictions bi-locus sur l'effet d'un balayage sélectif d'un allèle avantageux à un locus sur la probabilité de fixer une mutation délétère ou une seconde mutation avantageuse à un second locus. Du fait de la perte de la diversité durant le premier balayage sélectif, et de leur taux de recombinaison efficace plus faible, il y a plus de risques dans les populations autogames qu'une mutation délétère soit en déséquilibre de liaison avec la mutation avantageuse (Hartfield & Glémin, 2014). Ceci a pour effet de diminuer le gain de valeur sélective chez les populations autogames par rapport aux allogames, et de diminuer le seuil de dominance des mutations avantageuses pour lesquelles les populations autogames sont avantagées par rapport aux allogames (Hartfield & Glémin, 2014). De manière similaire, si l'on considère cette fois l'apparition d'une seconde mutation avantageuse pendant un premier balayage sélectif, les populations autogames sont défavorisées par rapport aux populations allogames (Hartfield & Glémin, 2016). En effet, si la seconde mutation est moins avantageuse que la première, elle a plus de chance d'être perdue en régime consanguin. Du fait de l'efficacité réduite de la recombinaison efficace, si la seconde mutation n'apparaît pas dans le fond génétique déjà sélectionné, il y a très peu de chance que les deux mutations soient associées dans un même haplotype par recombinaison, alors que cette probabilité est forte chez les populations allogames. On parle d'interférence de sélection (Hartfield & Glémin, 2016). Le deuxième effet majeur de ces interférences entre mutations avantageuses est que les populations allogames sont plus à même de fixer une ou plusieurs autres mutations faiblement avantageuses, indépendamment de leurs coefficients de dominance, annulant ainsi le tamis de Haldane pour les populations recombinantes (Hartfield & Glémin, 2016). On voit donc que les prédictions mono-locus qui indiquaient que l'adaptation à partir de nouvelles mutations pouvait être facilitée dans les populations autogames ne sont plus valables, voire sont inversées lorsque l'on considère les dynamiques simultanées de plusieurs mutations à différents locus.

2 – Systèmes de reproduction et adaptation à partir de la diversité génétique préexistante

Les modèles présentés précédemment font l'hypothèse que les populations n'ont pas de diversité génétique, et que l'adaptation ne peut se faire que *via* la fixation de nouvelles mutations. Or la plupart des populations naturelles ont de la diversité génétique pour toutes sortes de caractères quantitatifs (Mittell *et al.*, 2015; Wood *et al.*, 2016). Cette diversité, que

l'on appelle variation génétique préexistante, est considérée comme un indicateur important du potentiel adaptatif des populations (Orr & Betancourt, 2001; Hermisson & Pennings, 2005; Barrett & Schluter, 2008; Pritchard *et al.*, 2010; Matuszewski *et al.*, 2015). L'avantage de l'adaptation à partir de la diversité génétique déjà présente dans la population est que les allèles qui deviennent avantageux sont présents à plus haute fréquence dans la population (par rapport à une nouvelle mutation, en fréquence $1/2N$ lors de son apparition), augmentant ainsi la variance additive de la valeur sélective, et donc la vitesse d'adaptation (Figure 5). Cela augmente également la probabilité d'adaptation, puisque la probabilité de perdre la diversité nouvellement avantageuse est moindre par rapport à une nouvelle mutation qui apparaît à une fréquence $1/2N$ (Innan & Kim, 2004).

Une fois encore, on s'attend à ce que le système de reproduction de la population étudiée ait un effet sur la probabilité d'adaptation des populations. Le seul modèle ayant regardé l'effet de l'autofécondation sur l'adaptation à partir de la variation génétique pré-existante est celui de Glémin & Ronfort (2013). Dans ce modèle, les auteurs considèrent que l'allèle qui va contribuer à l'adaptation à un changement environnemental est déjà présent, ce dernier pouvant être délétère ou neutre avant le changement environnemental. Ils montrent que les populations autofécondantes ont moins de diversité préexistante, et que la probabilité d'adaptation décroît linéairement avec le système de reproduction (Glémin & Ronfort, 2013, Figure 6). En effet, la purge des mutations délétères plus efficace en autogamie (Figure 6B) réduit la diversité présente pour des allèles initialement délétères tandis que la baisse de taille efficace des populations autogames, et les effets d'autostop et de sélection d'arrière-plan (Smith & Haigh, 1974; Charlesworth *et al.*, 1993) associés à la consanguinité réduisent la diversité neutre maintenue dans les populations (Figure 6A) ; elles contiennent donc dans les deux cas moins de diversité pré-existante que les populations allogames.

Empiriquement, les données sur les marqueurs moléculaires supportent l'idée que les populations autogames ont une diversité génétique réduite par rapport aux espèces à régimes mixtes ou allogames (Barrett & Husband, 1990; Hamrick & Godt, 1996; Glémin *et al.*, 2006). Néanmoins, la diversité moléculaire calculée sur la base de marqueurs neutres n'est pas forcément représentative de la variance génétique des caractères quantitatifs (Cheverud *et al.*, 1994; Reed & Frankham, 2001).

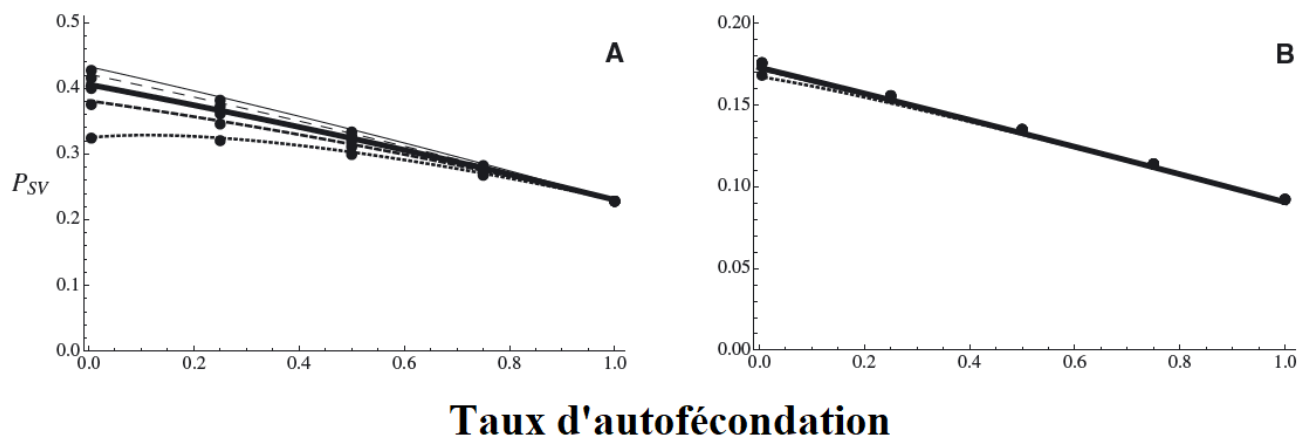


Figure 6. Probabilité d'adaptation à partir de la variation génétique préexistante, en fonction du taux d'autofécondation. De haut en bas, le coefficient de dominance h de la mutation avantageuse est égal à 0.9, 0.7, 0.5, 0.3 et 0.1, $N = 20.000$, taux de mutation = 10^{-6} . Dans le premier cas (A), la mutation est initialement neutre avant le changement. Dans le second cas (B), la mutation est initialement délétère ; et le coefficient de sélection de l'allèle s est égal à 0.01. Les points noirs représentent des résultats de simulations. Figure modifiée de Glémin & Ronfort (2013).

Ainsi, à travers les modèles d'adaptation à partir de nouvelles mutations ou à partir de la variation génétique pré-existante, il apparaît que l'autofécondation réduit la probabilité d'adaptation des populations (Glémin & Ronfort, 2013; Hartfield & Glémin, 2014, 2016). Si l'adaptation par nouvelles mutations est bien détaillée, avec des attendus analytiques mono- et bi-locus complémentaires, la théorie sur l'adaptation à partir de variation génétique pré-existante reste limitée, avec uniquement des attendus mono-locus. Or, nous savons que les caractères reliés à la valeur sélective des individus, sur lesquels la sélection va agir lors de processus d'adaptation, sont polygéniques, et les interactions entre locus, et l'effet de l'autofécondation sur ces dernières sont primordiales à comprendre pour quantifier le potentiel évolutif des populations.

III – Variance génétique et autofécondation : une introduction à la génétique quantitative des espèces autogames

La génétique quantitative (l'étude de la variance génétique des caractères quantitatifs) des espèces autofécondantes est complexe, car les effets de l'autogamie sur le niveau et la structure de la variance génétique, ainsi que sur la prédiction de la réponse à la sélection sont complexes. Cela a mené à la création de deux grands types de modèles de génétique quantitative d'espèces consanguines : les modèles caractérisant l'effet de l'autofécondation sur le niveau de variance génétique de caractères quantitatifs sous sélection, et les modèles de prédiction de réponse à la sélection en fonction du système de reproduction. Ici, nous décrirons ces deux grandes classes de modèles, et ce qu'ils nous apprennent de l'effet de l'autofécondation sur le potentiel évolutif des populations. Ces différents modèles font appel à des notions de génétique quantitative qui sont résumées dans l'encadré 1 (visible après la section bibliographie).

1 – L'effet de l'autofécondation sur le niveau de variance génétique, sous hypothèse de pure additivité

Les attendus théoriques

Les modèles étudiant l'effet du système de reproduction sur la variance des caractères quantitatifs ont grandement évolué au cours du temps. Néanmoins, un point commun de tous ces modèles est qu'ils considèrent une architecture purement additive des allèles présents aux différents locus contrôlant le caractère quantitatif sous étude. Ainsi, dans tous les modèles présentés dans cette section, la décomposition de la variance génétique d'une population ayant un coefficient de consanguinité f s'écrit comme

$$V_{G,f} = (1 + f)V_{genique} + COV_{DL} \quad (1),$$

avec $V_{genique}$ étant la variance des effets additifs des allèles (qui sont indépendants des autres allèles ou locus), et COV_{DL} représente l'effet des covariances génétiques sur la variance génétique (voir encadré 1 pour plus de détails).

Historiquement, le premier à avoir étudié l'effet de l'autofécondation sur le niveau de variance génétique des caractères quantitatifs est Wright (1922). En considérant un caractère quantitatif neutre, l'effet de l'autofécondation est d'augmenter l'homozygotie, ce qui augmente la variance additive du caractère d'un facteur $(1 + f)$ (Wright, 1922). Par la suite, d'autres modèles ont été proposés en considérant que les caractères quantitatifs sont sous sélection stabilisante et en ne négligeant pas les covariances génétiques. Ces modèles ont montré théoriquement que l'autofécondation doit mener à une diminution (plus ou moins graduelle) du niveau de variance génétique (Charlesworth & Charlesworth, 1995; Lande & Porcher, 2015 ; Abu Awad & Roze, 2018, figure 7). Néanmoins, l'origine de la diminution dépend d'un paramètre clé : le taux de mutation associé au caractère quantitatif (Abu Awad & Roze, 2018).

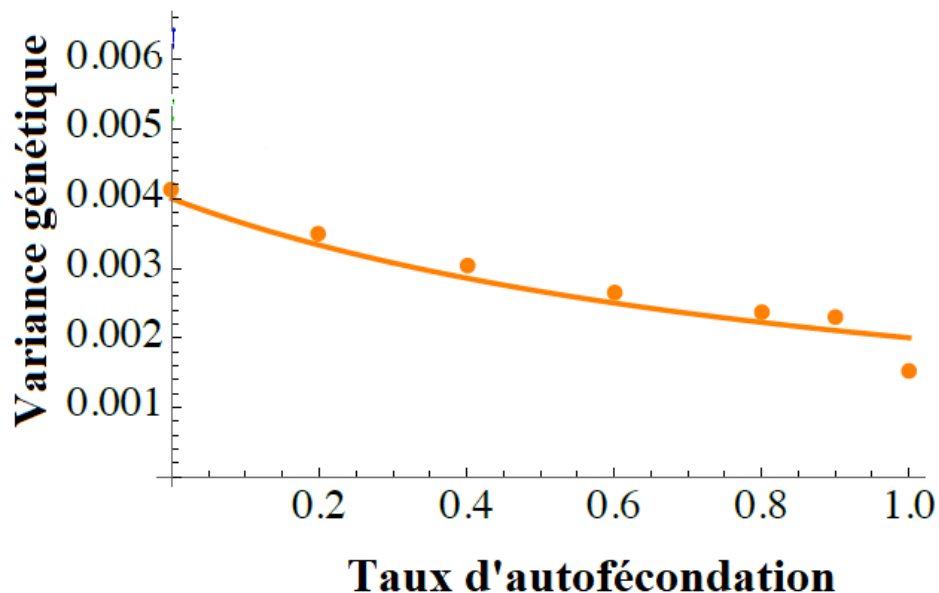


Figure 7. Effet du taux d'autofécondation sur le niveau de variance génétique d'un caractère quantitatif à l'équilibre mutation-sélection, sous hypothèse de pure additivité. Les paramètres sont $N = 50\,000$, $L = 10\,000$, et $U = 0.1$. Figure adaptée de Abu Awad & Roze (2018).

Lorsque le taux de mutation est faible, les mutations arrivent assez lentement pour que la sélection purificatrice puisse les éliminer progressivement, la purge étant plus efficace pour les populations autogames puisque les mutations, généralement récessives (Manna *et al.*, 2011) sont exposées dans les individus homozygotes (Lande & Porcher, 2015; Abu Awad & Roze,

2018). Ainsi, la variance génétique des populations autogames ($V_{G,f}$) est plus faible que celle des populations allogames et des populations à régime mixte du fait d'une importante diminution de la variance génique ($V_{genique}$).

Lorsque le taux de mutation devient important, la cadence d'apparition des mutations délétères est plus importante, et la purge est moins efficace, plus lente. Dans ce cas, des associations statistiques entre allèles ayant des effets opposés sur le phénotype peuvent se former. Lorsque les mutations associées se compensent, elles peuvent permettre de rester proche de l'optimum phénotypique. Du fait de leur taux de recombinaison efficace plus faible (Nordborg, 2000), les populations autogames sont plus à même de maintenir ce déséquilibre de liaison négatif que les populations allogames ou à régimes mixtes. Cela a pour effet de structurer les populations autogames en génotypes multi-locus répétés. Ces génotypes sont des lignées consanguines phénotypiquement similaires mais ayant fixé des combinaisons alléliques différentes (on parle de redondance phénotypique). Une structure de population en lignées répétées a effectivement pu être observée au sein de populations d'espèces très autogames sur la base de marqueurs moléculaires (Siol *et al.*, 2008; Jullien *et al.*, 2019). L'autre effet de ces associations génétiques est de diminuer la pression de sélection sur les mutations délétères en déséquilibre de liaison. Ainsi, lorsque les taux de mutation sont élevés, l'autofécondation a un faible effet sur le niveau de variance génique ($V_{genique}$), mais les associations génétiques entre locus (COV_{DL} , qui est par construction négatif) diminuent la variance génétique de la population ($V_{G,f}$).

Confrontation avec les données empiriques

Empiriquement, la relation négative entre le taux d'autofécondation et la variance génique est en revanche débattue. Tout d'abord, pour comparer des estimations de variance géniques entre populations et caractères, il est nécessaire de passer par des variables standardisées. Deux variables sont utilisables, l'héritabilité au sens strict (h^2), qui standardise la variance génique par la variance phénotypique (Falconer & Mackay, 1996), et le coefficient de variance génétique (CV_A), qui standardise la racine carrée de la variance génique par la moyenne du trait (Houle, 1992). Houle (1992) soutient que le CV_A est un meilleur proxy de la variance génique, car il existe une corrélation positive entre la variance génique d'un caractère et sa variance environnementale, ce qui pourrait limiter le pouvoir informatif de l'héritabilité.

Plusieurs méta-analyses ont tenté de concilier les attendus théoriques avec les données empiriques. Charlesworth & Charlesworth (1995, sur le CV_A , Figure 8A) et Geber & Griffen (2003, sur h^2) ont trouvé un effet du système de reproduction sur le niveau de variance génique, bien que le taux d'autofécondation n'explique qu'une faible fraction de la variabilité dans les deux études ($R^2 \leq 0.18$ dans les deux études). Au contraire, Opedal *et al.* (2017, sur le CV_A , Figure 8B) n'ont pas trouvé d'effet du taux d'autofécondation sur le niveau de variance génique de caractères floraux liés à l'herkogamie.

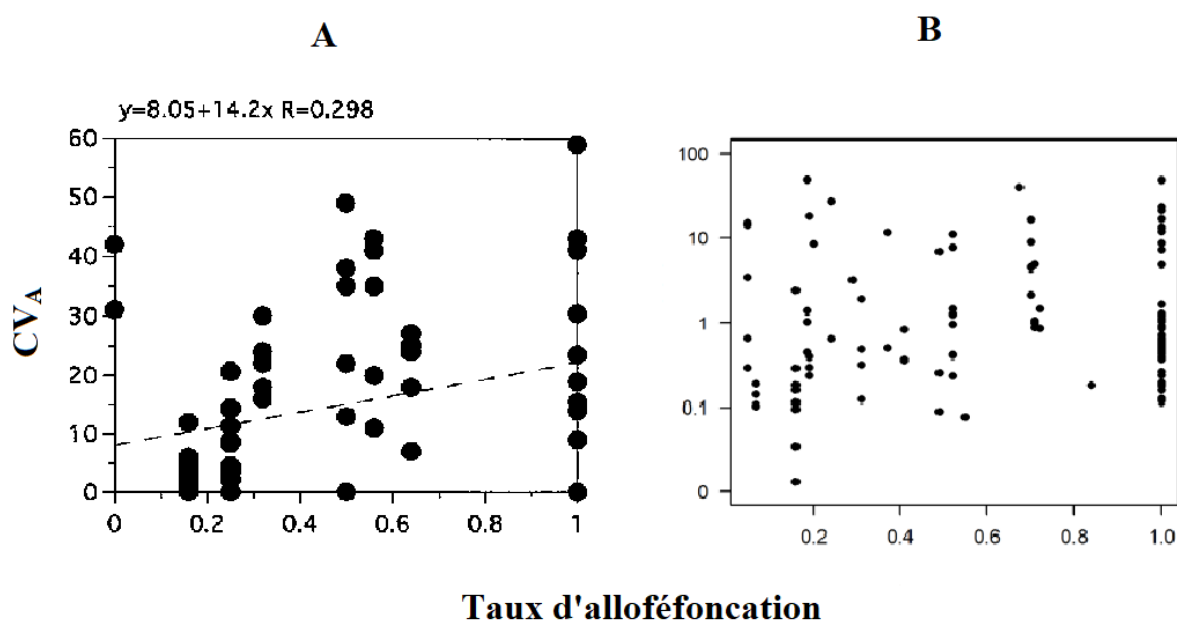


Figure 8. Niveau de variance génique de caractères quantitatifs (utilisant le CV_A comme unité standardisée) en fonction du taux d'allofécondation des espèces/populations associées. **A.** Figure adaptée de Charlesworth & Charlesworth (1995). **B.** Figure adaptée de Opedal *et al.* (2017).

Ces deux scénarios de taux de mutation prédisent donc tous les deux une diminution de la variance additive pour les populations autogames. Cependant, les conséquences pour la réponse à la sélection à court ou moyen terme peuvent être très différentes. Si la variance additive est essentiellement constituée de la variance génique, alors la réponse à court terme restera plus faible pour les populations autogames par rapport aux populations allogames ou à

régimes mixtes. Dans le cas contraire, si les covariances génétiques contribuent beaucoup à la variance additive, elles pourraient représenter une source de variance mobilisable par les événements de recombinaison entre lignées. Ceci pourrait donc augmenter le potentiel évolutif des populations sur les court et moyen termes (Lande & Porcher, 2015; Abu Awad & Roze, 2018). Néanmoins, ces prédictions théoriques sont basées sur l'hypothèse que les effets des allèles sont purement additifs à l'échelle phénotypique, ce qui est peu réaliste sur la base des études empiriques décomposant la variance génétique des caractères quantitatifs (voir Wolak & Keller, 2014 pour une revue). Nous allons voir dans la section suivante que considérer uniquement la composante additive de la variance génétique est probablement une vision biaisée du potentiel évolutif des populations.

2 – L'effet de la dominance sur la décomposition de la variance génétique et sur la réponse à la sélection en régime autogame

Décomposer la variance génétique en présence de dominance et d'autofécondation

En présence de dominance, la décomposition de la variance génétique est plus complexe. Dans ces conditions, la variance génétique d'une population ayant un coefficient de consanguinité f s'écrit :

$$V_{G,F} = (1 + f) V_{\text{génétique}} + (1 - f)V_{\text{DR}} + f V_{\text{DI}} + 4f \text{COV}_{\text{ADI}} + f(1 - f)H^* + \text{COV}_{\text{DL}} \quad (2).$$

La variance génétique et les covariances génétiques sont définies comme précédemment; V_{DR} et V_{DI} représentent la variance des déviations entre les valeurs génétiques des individus et les valeurs additives des allèles qu'ils portent à un locus donné, respectivement sous hypothèse de panmixie et d'autofécondation totale. Le terme COV_{ADI} représente la covariance entre l'effet additif d'un allèle et sa déviation de dominance à l'état homozygote. Le terme H^* représente la somme sur tous les locus des déviations de dominance homozygote (aussi égal à la dépression de consanguinité au locus donné), le tout mis au carré. Toutes ces composantes sont détaillées dans l'encadré 1. Ces nouveaux termes compliquent grandement la génétique quantitative des populations autogames, tant dans l'estimation des composantes de la variance génétique (Cockerham & Weir, 1984; Shaw *et al.*, 1998; Kelly & Arathi, 2003), que dans la prédiction de la réponse à la sélection (Wright & Cockerham, 1985; Kelly, 1999a; b).

L'estimation de toutes les composantes de la variance génétique demande des dispositifs expérimentaux compliqués, notamment en terme de pedigree. En plus des covariances habituelles entre pleins-frères et demi-frères qui permettent d'estimer les variances génique et de dominance dans une population panmictique (V_{DR}), il est nécessaire de calculer des covariances entre paires d'apparentés partageant des allèles à l'état homozygote (permettant de calculer V_{DI} et H^*), et des paires d'apparentés avec un individu ayant un allèle à l'état hétérozygote et l'autre individu ayant le même allèle à l'état homozygote (permettant de calculer COV_{ADI}). Dans des dispositifs expérimentaux utilisant des couples d'apparentés ayant des coefficients de consanguinité différents (comme dans Shaw *et al.*, 1998), le calcul des covariances entre apparentés est compliqué (Cockerham & Weir, 1984), bien que des algorithmes récursifs permettant de calculer les coefficients d'apparentements nécessaires aux estimations existent (Cockerham, 1971). L'utilisation de lignées pures (ayant donc un f égal à 1) simplifie grandement les estimations (Kelly & Arathi, 2003), puisque les covariances entre apparentés sont simplifiées (Cockerham & Weir, 1984), mais cela peut demander de nombreuses générations d'autofécondation préliminaires pour créer des lignées pures, ce qui est contraignant pour des espèces à régime mixte ayant un temps de génération (pluri-) annuel. Du fait de ces difficultés empiriques, on trouve peu d'estimations dans la littérature, généralement faites sur des populations manipulées, et non *in natura* (voir Wolak & Keller, 2014 pour une revue). Dans les rares cas où ces composantes ont été estimées, on observe généralement qu'au moins une partie de ces nouvelles composantes contribue significativement à la variance génétique (Hoeschele & Vollema, 1993; Shaw *et al.*, 1998; Edwards & Lamkey, 2002; Kelly, 2003; Kelly & Arathi, 2003; Wardyn *et al.*, 2007; Marriage & Kelly, 2009). Parmi les grands patrons mis en évidence par ces études, la variance des déviations de dominance homozygote (V_{DI}) dans les populations autogames est généralement forte, supérieure à la variance de dominance d'une population panmictique (V_{DR}), et souvent équivalente ou supérieure à la variance génique, ce qui est attendu lorsque les fréquences alléliques sont déséquilibrées (Walsh & Lynch, 2018). La covariance entre l'effet additif d'un allèle et sa déviation de dominance homozygote (COV_{ADI}) est assez variable entre traits et espèces, mais lorsque cette composante contribue, elle est généralement fortement négative (Shaw *et al.*, 1998; Edwards & Lamkey, 2002; Kelly, 2003; Kelly & Arathi, 2003). Le carré de la somme des déviations de dominance homozygote (H^*) est généralement faible et non-significatif, ce qui implique que la dépression de consanguinité pour les traits étudiés est due à l'action conjointe d'un grand nombre de locus à faibles effets (Shaw *et al.*, 1998).

Ainsi, ces nouvelles composantes ne semblent pas négligeables et contribuent à la variance génétique des populations. Lorsque les populations sont consanguines, ces (co-)variances participent à la covariance entre parents et enfants, dont dépend la réponse à la sélection. Si la réponse à la sélection est facilement prédictible pour les espèces strictement allo- ou autogames (il suffit d'incorporer respectivement la variance génique ou la variance génétique totale dans l'équation du sélectionneur ou dans le théorème fondamental de la sélection naturelle, Fisher, 1958; Falconer & Mackay, 1996), prédire la réponse pour des régimes intermédiaires n'est pas trivial. Deux grand modèles ont tenté de caractériser la réponse à la sélection pour les espèces à régime mixte, le *Covariance Genetic Model* (CGM, Wright & Cockerham, 1985), et le *Structured Linear Model* (SLM, Kelly, 1999a). Je me focaliserai uniquement sur le SLM, puisque le CGM n'est qu'une simplification de ce dernier (Kelly & Williamson, 2000).

Intégrer ces nouvelles composantes de la variance génétique dans les modèles de réponse à la sélection

L'un des problèmes majeurs des taux d'autofécondation intermédiaires est que le niveau de consanguinité peut être variable entre les individus. Si le trait étudié montre de la dépression de consanguinité, ou si les individus fortement consanguins ont une valeur sélective plus faible, alors on s'attend à ce que les individus issus d'évènements récents d'allofécondation soient choisis plus souvent par la sélection (naturelle ou artificielle). Dans ce cas, la régression parent-enfant n'est plus linéaire (Kelly, 1999a; b, Figure 9), or cette hypothèse est cruciale pour la validité des équations classiques de réponse à la sélection (comme l'équation du sélectionneur). Afin de palier à ce problème, Kelly (1999) a proposé de décomposer les populations en cohortes de différents niveaux de consanguinité, les individus d'une même cohorte partageant le même niveau de consanguinité (et donc le même F). Au sein de chaque cohorte, la régression parent-enfant est supposée linéaire (d'où le nom, *structured linear model*), et la réponse à la sélection est prédite à partir des variances et covariances au sein des cohortes, et les transitions d'une cohorte à une autre.

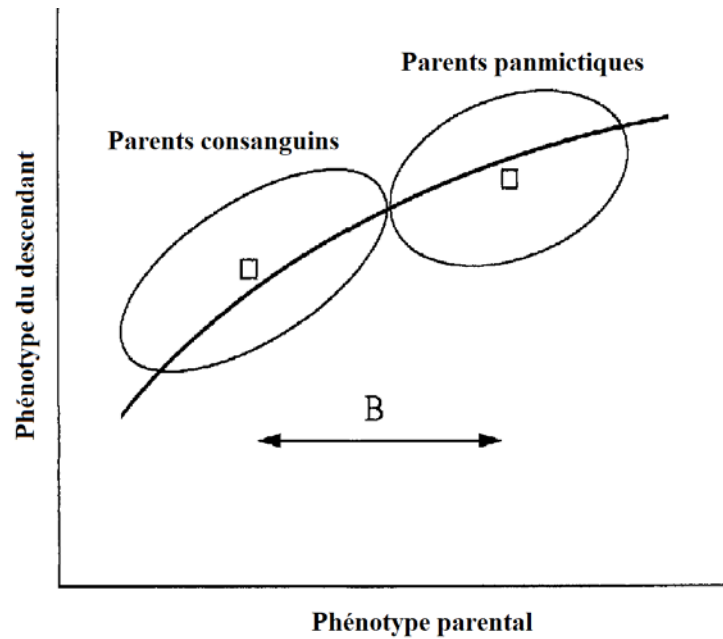


Figure 9. Relation entre les phénotypes parentaux et ceux de leurs descendants, dans une population constituée d'individus purement consanguins ($F = 1$) et d'individus panmictiques ($F = 0$). Au sein de chaque groupe, la régression parent-enfant est linéaire. Mais dans la population globale (considérant les individus indépendamment de leur histoire de consanguinité) la régression n'est plus linéaire. Ici B définit la dépression de consanguinité, la différence de moyenne entre les individus purement consanguins et les individus issus d'allopécundation (défini par les carrés blancs). Figure adaptée de Kelly (1999).

Le modèle se décompose ainsi : considérons un caractère quantitatif Z , la moyenne de ce caractère à la génération t est égale à

$$Z(t) = \mu_o(t) + F(t)(\mu_l(t) - \mu_o(t))$$

ou

$$Z(t) = \mu_o(t) + F(t)B(t),$$

où $\mu_o(t)$ et $\mu_l(t)$ sont les moyennes d'une population respectivement strictement allogame et autogame, ayant les mêmes fréquences alléliques que la population d'étude à la génération (t) , $B(t)$ est la différence entre la moyenne autogame et allogame à cette même

génération, et $F(t)$ est le coefficient de consanguinité moyen de la population au temps t . Dans le modèle de Kelly, chaque cohorte i (i étant le nombre de génération d'autofécondation dont sont issus les individus, Figure 10) a ses propres valeurs de F_i (avec $F_i = 1 - \frac{1}{2^i}$), $\mu_{0,i}$ et $B_{i,i}$. A l'échelle de la population $F(t)$, $B(t)$ et $\mu_{0(t)}$ sont les valeurs des différentes cohortes pondérées par leur fréquence (par exemple, si λ_i est la fréquence de la cohorte i , alors $\mu_{0(t)} = \sum \lambda_i \mu_{0,i}(t)$). Pour prédire le changement de valeur moyenne du trait, il faut pouvoir prédire les changements dans chaque cohorte, et les mouvements d'une cohorte à l'autre. A chaque évènement de reproduction, une proportion s de chaque cohorte i se reproduit par autofécondation, et alimente la cohorte $i+1$, tandis qu'une proportion $1-s$ alimente la cohorte 0 (Figure 10).

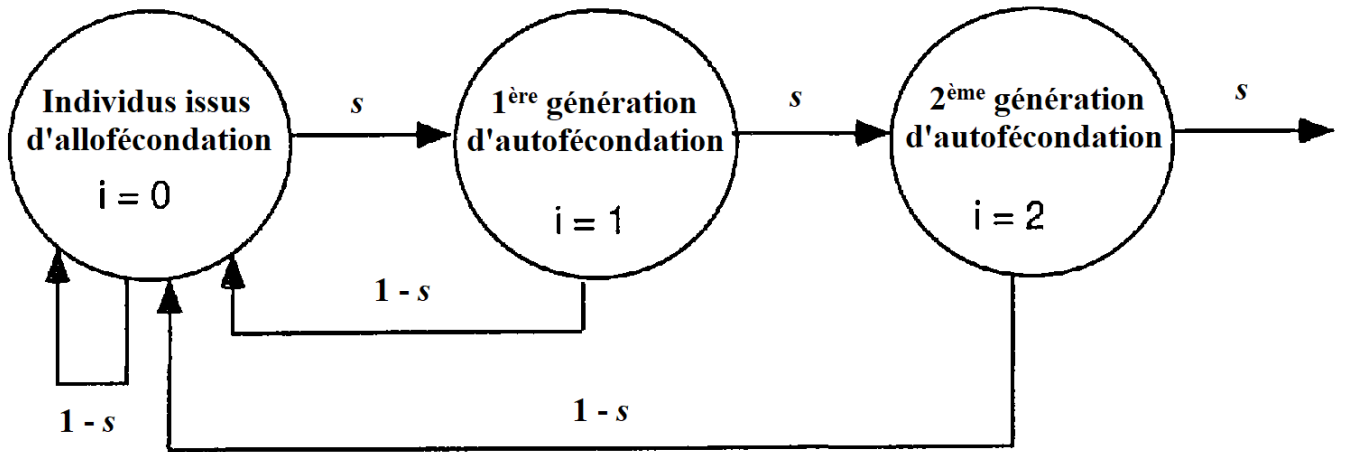


Figure 10. Visualisation de la structure en cohortes d'autofécondation du SLM. Les individus issus d'allofécondation (avec une probabilité $1-s$) rentrent automatiquement dans la cohorte 0. Les descendants d'individus de la cohorte i issus d'autofécondation se retrouvent dans la cohorte $i+1$ (avec une probabilité s). Figure adaptée de Kelly (1999).

Soit $W_i(t)$ la valeur sélective moyenne des individus de la cohorte d'autofécondation i et $W(t)$ celle de la population. La fréquence de la cohorte $i+1$ à la génération $t+1$ est alors égale à

$$\lambda_{i+1}(t+1) = s \left(\frac{W_i(t)}{W(t)} \right) \lambda_i(t),$$

et donc

$$F(t + 1) = \sum_i \lambda_i(t + 1)F_i.$$

Sous hypothèse de linéarité au sein de chaque cohorte,

$$\mu_{0,i+1}(t + 1) = \mu_{0,i}(t) + \frac{(1 + F_i)V_{g\u00e9n\u00e9rique} + 2F_iCOV_{ADI}}{V_{Z,i}(t)}S_i(t)$$

et

$$B_{i+1}(t + 1) = B_i(t) + \frac{(1 + F_i)COV_{ADI} + F_iV_{DI}}{V_{Z,i}(t)}S_i(t),$$

avec $V_{Z,i}(t)$ et $S_i(t)$ \u00e9tant respectivement la variance ph\u00e9notypique et le diff\u00e9rentiel de s\u00e9lection de la cohorte i \u00e0 la g\u00e9n\u00e9ration t . La cohorte 0 rassemble les individus \u00e9tant issus d'allof\u00e9condation, quelle que soit leur cohorte d'origine. Ainsi, \u00e0 la g\u00e9n\u00e9ration $t+1$, les valeurs de la cohorte 0 se calculent comme

$$\lambda_0(t + 1) = (1 - s),$$

$$\mu_{0,0}(t + 1) = \mu_0(t) + \sum \lambda_i(t) \frac{(1 + F_i)V_{g\u00e9n\u00e9rique} + 2F_iCOV_{ADI}}{V_{Z,i}(t)}S_i(t)$$

et

$$B_0(t + 1) = B(t) + \sum \lambda_i(t) \frac{(1 + F_i)COV_{ADI} + F_iV_{DI}}{V_{Z,i}(t)}S_i(t).$$

L'application de ce mod\u00e8le est compliqu\u00e9e. En plus de l'estimation complexe des diff\u00e9rentes composantes de la variance g\u00e9n\u00e9tique, l'application du SLM est co\u00fbteuse en param\u00e8tres, qui peuvent \u00eatre difficiles \u00e0 estimer. Tout d'abord, le nombre de cohortes et leurs proportions relatives au sein de la population demande un g\u00e9notypage massif au sein de la population d'\u00e9tude, et un grand nombre de marqueurs (microsatellites, SNPs...) permettant de caract\u00e9riser l'histoire d'autof\u00e9condation des individus. Deuxi\u00e8mement, en plus des mesures r\u00e9alis\u00e9es sur le trait \u00e9tudi\u00e9, il est n\u00e9cessaire de caract\u00e9riser la valeur s\u00e9lective des individus afin

de pouvoir prédire la contribution de chaque cohorte aux changements de valeurs à la génération suivante. Ces difficultés techniques expliquent sûrement pourquoi aucune application du modèle de Kelly n'existe dans la littérature, et pourquoi les études tentant de caractériser le potentiel adaptatif des populations se limitent à des estimations de la variance génique.

Il apparaît ainsi que les différents modèles de génétique quantitative regardant l'effet du système de reproduction sur le niveau de variance génétique, et ceux regardant la réponse à la sélection, ne soient pas facilement réconciliables. Les premiers modèles font généralement des hypothèses fortes sur l'architecture des traits quantitatifs pour parvenir à des attendus analytiques, mais l'intégration de ces résultats dans les modèles de prédiction de réponse à un événement de sélection directionnelle n'est pas aisée du fait des hypothèses faites. Les seconds, à l'inverse, supposent une architecture plus réaliste des caractères quantitatifs, et permettent de comprendre plus finement comment décomposer la variance et la structure des populations pour prédire précisément la réponse à la sélection. Cependant, les paramètres nécessaires pour appliquer ces méthodes empêchent de prédire mathématiquement et empiriquement l'effet du système de reproduction sur le niveau et la vitesse d'adaptation des populations.

IV – Objectifs et présentation de la thèse

1 – Principaux objectifs

L'objectif de cette thèse est de comprendre plus finement le potentiel évolutif des populations autogames, en étudiant théoriquement et empiriquement le niveau de diversité génétique pour des caractères polygéniques au sein de ces espèces, sa structure, et la manière dont elle peut être mobilisée pour répondre à la sélection. En m'appuyant sur une méta-analyse, une étude expérimentale et des modèles théoriques, j'ai cherché à confirmer et approfondir les connaissances sur l'adaptation des populations autogames. Pour développer cela, ma démarche s'est structurée en 4 axes majeurs :

2 – Chapitre 1 : Méta-analyse de l'effet de l'autofécondation sur la variance génétique au sein des populations naturelles d'angiospermes.

Dans ce premier chapitre, j'ai voulu mettre à jour les données empiriques disponibles chez les plantes sur la relation entre variance génique et taux d'autogamie, et déterminer si ces données sont en accord avec les prédictions théoriques. Pour cela, j'ai réalisé une méta-analyse en compilant un grand nombre de mesures d'héritabilité (sens strict et large) estimées dans des populations naturelles d'espèces angiospermes pour lesquelles nous disposions d'informations sur le système de reproduction (taux d'autofécondation de l'espèce ou des populations). À l'aide de ce jeu de données, j'ai caractérisé la relation entre système de reproduction et variance génique, en tentant de limiter le plus possible les biais dus à l'imprécision des taux d'autofécondation et aux autres composantes de la variance génétique qui deviennent non négligeables avec la consanguinité. Dans un second temps, j'ai testé l'hypothèse d'un niveau de variance génique plus variable entre populations d'une même espèce autogame, par rapport aux allogames. Enfin, j'ai examiné, de manière indirecte, si une mesure plus intégrative du potentiel évolutif des espèces autogames (qui comprend à la fois la variance génique et les composantes de dominance) différait de la variance génique des espèces strictement allogames, en comparant les héritabilités au sens large chez les espèces autogames avec les héritabilités au sens strict observées chez les espèces allogames.

3 – Chapitre 2 : Le résidu d'allogamie et les covariances génétiques pourraient faciliter l'adaptation des espèces autogames aux changements environnementaux

Dans ce second chapitre, je me suis intéressé à la possible contribution des covariances génétiques dans le potentiel évolutif des espèces autogames. Les modèles ayant étudié l'adaptation des populations en fonction de la variance génétique préexistante et du système de reproduction ont utilisé un modèle mono-locus. Les études théoriques de la relation variance génique – système de reproduction ont souligné l'importance des covariances génétiques dans la structuration et le niveau de diversité génétique, et ont supposé que ces covariances pourraient représenter un stock de variance exploitable lors d'une phase de sélection directionnelle. J'ai donc étendu un modèle théorique de trait quantitatif sous sélection stabilisante pour y incorporer une phase de réponse à un changement environnemental. J'ai dans un premier temps quantifié la contribution de ces covariances génétiques à la variance

génétique totale dans les populations à l'équilibre mutation-sélection-dérive, en fonction du taux d'autofécondation des populations. Dans un second temps, j'ai étudié les dynamiques de valeur sélective et de variance génétique pendant la phase de sélection directionnelle afin de comparer les dynamiques et mécanismes d'adaptation des populations en fonction de leur système de reproduction.

4 – Chapitre 3 : Etude empirique sur l'importance de la dominance dans l'isolement reproductif entre lignées consanguines dans les populations naturelles de *Medicago truncatula*.

Si l'allogamie résiduelle permet de remobiliser de la variance génétique (voir chapitre 2), elle pourrait également révéler des incompatibilités génétiques accumulées entre lignées autogames. Dans ce troisième chapitre, je me suis intéressé aux conséquences d'évènements d'hybridation entre génotypes multi-locus de populations naturelles de l'espèce préférentiellement autogame *Medicago truncatula*. En effet, le taux de recombinaison efficace étant extrêmement faible chez les espèces autogames, on peut s'attendre à ce que les lignées consanguines accumulent des incompatibilités génétiques (par exemple des inversions chromosomiques, ou des incompatibilités Dobzhansky-Müller) entre populations, voire même au sein d'une même population. Si ces incompatibilités étaient révélées lors de croisements (allofécondation résiduelle), elles limiteraient le potentiel évolutif de ces populations. Afin de tester cette hypothèse, j'ai étudié les conséquences d'évènements d'hybridation entre lignées pures de *Medicago truncatula* issues de populations différentes, ou de la même population. J'ai caractérisé l'architecture génétique de la dépression d'allofécondation, et j'ai proposé un mécanisme basé sur l'effet de la dominance directionnelle sur un caractère sous sélection, qui permet d'expliquer de manière parcimonieuse les patrons de dépression d'allofécondation observés dans notre étude.

5 – Chapitre 4 : Au-delà de l'additivité, de l'importance de considérer l'effet de la dominance sur le niveau et la structure de la diversité génétique des populations

Les résultats des chapitres précédents ayant souligné le rôle de la dominance chez les espèces autogames, j'ai, dans ce dernier chapitre, développé un modèle théorique de génétique

quantitative relâchant l'hypothèse classique de pure additivité à l'échelle phénotypique. Ce modèle simule un trait quantitatif sous sélection stabilisante codé par L loci bi-alléliques ayant des effets additifs, et des relations de dominance entre allèles à un locus, et permet de considérer l'effet du système de reproduction des populations. J'ai dans un premier temps regardé la contribution des composantes de dominance à la variance génétique, et estimé la part de variance expliquée par la composante additive par rapport à la composante de dominance. Dans un second temps, j'ai étendu le modèle structuré linéairement (SLM) de Kelly dans le théorème fondamental de la sélection naturelle, et j'ai regardé si une simplification de ce dernier suffisait à prédire correctement le changement de valeur sélective d'une génération à l'autre dans les populations partiellement autofécondantes. J'ai également quantifié le potentiel évolutif immédiat des populations en fonction de leur système de reproduction, afin de prédire si le taux d'autofécondation aurait un impact sur l'avenir des populations face à un changement environnemental lorsque l'on considère une architecture génétique plus réaliste. Enfin, j'ai examiné si l'évolution de l'allogamie vers l'autogamie restait avantageuse lors d'une phase de sélection directionnelle lorsque la dominance et les nouvelles composantes génétiques qui en découlent contribuent au potentiel évolutif des populations.

Ces chapitres ont fait l'objet de quatre manuscrits, acceptés ou soumis à des revues scientifiques, et je vais structurer mon manuscrit de thèse autour de ces différents écrits.

Bibliographie

- Abu Awad, D. & Billiard, S. 2017. The double edged sword: The demographic consequences of the evolution of self-fertilization. *Evolution* **71**: 1178–1190.
- Abu Awad, D. & Roze, D. 2018. Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution* **72**: 751–769.
- Allen, A.M. & Hiscock, S.J. 2008. Evolution and phylogeny of self-incompatibility systems in angiosperms. In: *Self-incompatibility in flowering plants*, pp. 73–101. Springer.
- Awad, D.A. & Roze, D. 2020. Epistasis, inbreeding depression and the evolution of self-fertilization. *Evolution*. Wiley Online Library.
- Baker, H.G. 1955. SELF-COMPATIBILITY AND ESTABLISHMENT AFTER “LONG-DISTANCE” DISPERSAL. *Evolution* **9**: 347–349.

- Barrett, R.D. & Schluter, D. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**: 38–44.
- Barrett, S.C. & Eckert, C.G. 1990. Current issues in plant reproductive ecology. *Isr. J. Plant Sci.* **39**: 5–12. Brill.
- Barrett, S.C. & Husband, B.C. 1990. Variation in outcrossing rates in *Eichhornia paniculata*: the role of demographic and reproductive factors. *Plant Species Biol.* **5**: 41–55. Wiley Online Library.
- Bulmer, M. 1971. The effect of selection on genetic variability. *Am. Nat.* **105**: 201–211. University of Chicago Press.
- Bulmer, M.G. 1974. Linkage disequilibrium and genetic variability. *Genet. Res.* **23**: 281–289. Cambridge University Press.
- Caballero, A. & Hill, W.G. 1992. Effective size of nonrandom mating populations. *Genetics* **130**: 909–916.
- Charlesworth, B. 1992. Evolutionary Rates in Partially Self-Fertilizing Species. *Am. Nat.* **140**: 126–148.
- Charlesworth, B., Morgan, M.T. & Charlesworth, D. 1991. Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. *Genet. Res.* **57**: 177–194. Cambridge University Press.
- Charlesworth, B., Morgan, M.T. & Charlesworth, D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289.
- Charlesworth, D. & Charlesworth, B. 1995. Quantitative genetics in plants: the effect of the breeding system on genetic variability. *Evolution* **49**: 911–920.
- Charlesworth, D., Morgan, M.T. & Charlesworth, B. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution* **44**: 1469–1489. Wiley Online Library.
- Chelo, I.M., Afonso, B., Carvalho, S., Theologidis, I., Goy, C., Pino-Querido, A., *et al.* 2017. Partial selfing eliminates inbreeding depression while maintaining genetic diversity. *bioRxiv* 210005.
- Cheverud, J., Routman, E., Jaquish, C., Tardif, S., Peterson, G., Belfiore, N., *et al.* 1994. Quantitative and molecular genetic variation in captive cotton-top tamarins (*Saguinus oedipus*). *Conserv. Biol.* **8**: 95–105. Wiley Online Library.
- Cockerham, C.C. 1971. Higher order probability functions of identity of alleles by descent. *Genetics* **69**: 235. Genetics Society of America.
- Cockerham, C.C. & Weir, B.S. 1984. Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics* 157–164. JSTOR.
- David, J.L., Savy, Y. & Brabant, P. 1993. Outcrossing and selfing evolution in populations under directional selection. *Heredity* **71**: 642–651. Nature Publishing Group.

- Devaux, C., Lepers, C. & Porcher, E. 2014. Constraints imposed by pollinator behaviour on the ecology and evolution of plant mating systems. *J. Evol. Biol.* **27**: 1413–1430. John Wiley & Sons, Ltd.
- Edwards, J.W. & Lamkey, K.R. 2002. Quantitative genetics of inbreeding in a synthetic maize population. *Crop Sci.* **42**: 1094–1104.
- Epinat, G. & Lenormand, T. 2009. The evolution of assortative mating and selfing with in-and outbreeding depression. *Evol. Int. J. Org. Evol.* **63**: 2047–2060. Wiley Online Library.
- Escobar, J.S., Cenci, A., Bolognini, J., Haudry, A., Laurent, S., David, J., *et al.* 2010. An integrative test of the dead-end hypothesis of selfing evolution in Triticeae (Poaceae). *Evol. Int. J. Org. Evol.* **64**: 2855–2872.
- Escobar, J.S., Nicot, A. & David, P. 2008. The Different Sources of Variation in Inbreeding Depression, Heterosis and Outbreeding Depression in a Metapopulation of *Physa acuta*. *Genetics* **180**: 1593.
- Falconer, D.S. & Mackay, T.F.C. 1996. Introduction to quantitative genetics. Essex. UK Longman Group.
- Fisher, R.A. 1941. AVERAGE EXCESS AND AVERAGE EFFECT OF A GENE SUBSTITUTION. *Ann. Eugen.* **11**: 53–63.
- Fisher, R.A. 1958. *The genetical theory of natural selection*, 2nd ed. Dover Publication, New York.
- Geber, M.A. & Griffen, L.R. 2003. Inheritance and natural selection on functional traits. *Int. J. Plant Sci.* **164**: S21–S42.
- Glémin, S. 2012. Extinction and fixation times with dominance and inbreeding. *Theor. Popul. Biol.* **81**: 310–316.
- Glémin, S., Bazin, E. & Charlesworth, D. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proc. R. Soc. B Biol. Sci.* **273**: 3011–3019. The Royal Society London.
- Glémin, S. & Ronfort, J. 2013. Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evol. Int. J. Org. Evol.* **67**: 225–240.
- Goldberg, E.E., Kohn, J.R., Lande, R., Robertson, K.A., Smith, S.A. & Igić, B. 2010. Species Selection Maintains Self-Incompatibility. *Science* **330**: 493.
- Goodwillie, C., Kalisz, S. & Eckert, C.G. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annu Rev Ecol Evol Syst* **36**: 47–79. Annual Reviews.
- Goubet, P.M., Bergès, H., Bellec, A., Prat, E., Helmstetter, N., Mangenot, S., *et al.* 2012. Contrasted patterns of molecular evolution in dominant and recessive self-incompatibility haplotypes in *Arabidopsis*. *PLoS Genet* **8**: e1002495. Public Library of Science.

- Haldane, J.B.S. 1927. A mathematical theory of natural and artificial selection, part V: selection and mutation. In: *Mathematical Proceedings of the Cambridge Philosophical Society*, pp. 838–844. Cambridge University Press.
- Hamrick, J.L. & Godt, M.W. 1996. Effects of life history traits on genetic diversity in plant species. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **351**: 1291–1298. The Royal Society London.
- Hartfield, M. & Glémin, S. 2014. Hitchhiking of Deleterious Alleles and the Cost of Adaptation in Partially Selfing Species. *Genetics* **196**: 281.
- Hartfield, M. & Glémin, S. 2016. Limits to Adaptation in Partially Selfing Species. *Genetics* **203**: 959.
- Hermisson, J. & Pennings, P.S. 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**: 2335–2352.
- Hoeschele, I. & Vollema, A.R. 1993. Estimation of variance components with dominance and inbreeding in dairy cattle. *J. Anim. Breed. Genet.* **110**: 93–104.
- Holsinger, K.E. 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. *Am. Nat.* **138**: 606–622. University of Chicago Press.
- Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* **130**: 195–204. Genetics Soc America.
- Igic, B. & Busch, J.W. 2013. Is self-fertilization an evolutionary dead end? *New Phytol.* **198**: 386–397.
- Innan, H. & Kim, Y. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. *Proc. Natl. Acad. Sci.* **101**: 10667–10672.
- Jarne, P. & Auld, J.R. 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. *Evolution* **60**: 1816–1824. Wiley Online Library.
- Johnston, M.O. 1998. Evolution of intermediate selfing rates in plants: pollination ecology versus deleterious mutations. In: *Mutation and Evolution*, pp. 267–278. Springer.
- Jullien, M., Navascués, M., Ronfort, J., Loridon, K. & Gay, L. 2019. Structure of multilocus genetic diversity in predominantly selfing populations. *Heredity* **123**: 176–191.
- Kamran-Disfani, A. & Agrawal, A.F. 2014. Selfing, adaptation and background selection in finite populations. *J. Evol. Biol.* **27**: 1360–1371.
- Karron, J.D., Thumser, N.N., Tucker, R. & Hesseuauer, A.J. 1995. The influence of population density on outcrossing rates in *Mimulus ringens*. *Heredity* **75**: 175–180. Nature Publishing Group.
- Kelly, J.K. 2003. Deleterious mutations and the genetic variance of male fitness components in *Mimulus guttatus*. *Genetics* **164**: 1071–1085. Genetics Soc America.

- Kelly, J.K. 1999a. Response to selection in partially self-fertilizing populations. I. Selection on a single trait. *Evolution* **53**: 336–349.
- Kelly, J.K. 1999b. Response to selection in partially self-fertilizing populations. II. Selection on multiple traits. *Evolution* **53**: 350–357.
- Kelly, J.K. & Arathi, H.S. 2003. Inbreeding and the genetic variance in floral traits of *Mimulus guttatus*. *Heredity* **90**: 77.
- Kelly, J.K. & Williamson, S. 2000. Predicting response to selection on a quantitative trait: a comparison between models for mixed-mating populations. *J. Theor. Biol.* **207**: 37–56.
- Kopp, M. & Matuszewski, S. 2014. Rapid evolution of quantitative traits: theoretical perspectives. *Evol. Appl.* **7**: 169–191. Wiley Online Library.
- Lande, R. 1977. The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* **86**: 485–498.
- Lande, R. & Porcher, E. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* **200**: 891–906.
- Lande, R. & Schemske, D.W. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**: 24–40.
- Lloyd, D.G. & Webb, C.J. 1986. The avoidance of interference between the presentation of pollen and stigmas in angiosperms I. Dichogamy. *N. Z. J. Bot.* **24**: 135–162. Taylor & Francis.
- Manna, F., Martin, G. & Lenormand, T. 2011. Fitness Landscapes: An Alternative Theory for the Dominance of Mutation. *Genetics* **189**: 923.
- Marriage, T.N. & Kelly, J.K. 2009. Inbreeding depression in an asexual population of *Mimulus guttatus*. *J. Evol. Biol.* **22**: 2320–2331. John Wiley & Sons, Ltd.
- Mittell, E.A., Nakagawa, S. & Hadfield, J.D. 2015. Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.* **18**: 772–778.
- Nordborg, M. 2000. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics* **154**: 923–929.
- Opedal, Ø.H. 2018. Herkogamy, a Principal Functional Trait of Plant Reproductive Biology. *Int. J. Plant Sci.* **179**: 677–687. The University of Chicago Press.
- Opedal, Ø.H., Bolstad, G.H., Hansen, T.F., Armbruster, W.S. & Pélabon, C. 2017. The evolvability of herkogamy: Quantifying the evolutionary potential of a composite trait. *Evolution* **71**: 1572–1586. Wiley Online Library.
- Orr, H.A. & Betancourt, A.J. 2001. Haldane’s sieve and adaptation from the standing genetic variation. *Genetics* **157**: 875–884.

- Pollak, E. 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. *Genetics* **117**: 353–360.
- Pritchard, J.K., Pickrell, J.K. & Coop, G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**: R208–R215.
- Reed, D.H. & Frankham, R. 2001. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**: 1095–1103. Wiley Online Library.
- Richards, A.J. 1997. *Plant breeding systems*. Garland Science.
- Schemske, D.W. & Lande, R. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution* **39**: 41–52. Wiley Online Library.
- Shaw, R.G., Byers, D.L. & Shaw, F.H. 1998. Genetic components of variation in *Nemophila menziesii* undergoing inbreeding: morphology and flowering time. *Genetics* **150**: 1649–1661.
- Siol, M., Prosperi, J.M., Bonnin, I. & Ronfort, J. 2008. How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in *Medicago truncatula*. *Heredity* **100**: 517.
- Slotte, T., Hazzouri, K.M., Stern, D., Andolfatto, P. & Wright, S.I. 2012. Genetic architecture and adaptive significance of the selfing syndrome in *Capsella*. *Evol. Int. J. Org. Evol.* **66**: 1360–1374. Wiley Online Library.
- Smith, J.M. & Haigh, J. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**: 23–35.
- Stebbins, G.L. 1957. Self fertilization and population variability in the higher plants. *Am. Nat.* **91**: 337–354.
- Takebayashi, N. & Morrell, P.L. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am. J. Bot.* **88**: 1143–1150.
- Uyenoyama, M.K. & Waller, D.M. 1991. Coevolution of self-fertilization and inbreeding depression I. Mutation-selection balance at one and two loci. *Theor. Popul. Biol.* **40**: 14–46. Elsevier.
- Walsh, B. & Lynch, M. 2018. *Evolution and selection of quantitative traits*. Oxford University Press.
- Wardyn, B.M., Edwards, J.W. & Lamkey, K.R. 2007. The genetic structure of a maize population: The role of dominance. *Crop Sci.* **47**: 467–474.
- Whitehead, M.R., Lanfear, R., Mitchell, R.J. & Karron, J.D. 2018. Plant mating systems often vary widely among populations. *Front. Ecol. Evol.* **6**: 38. Frontiers.

- Willi, Y. 2013. MUTATIONAL MELTDOWN IN SELFING ARABIDOPSIS LYRATA. *Evolution* **67**: 806–815.
- Winn, A. A., E. Elle, S. Kalisz, P.-O. Cheptou, C. G. Eckert, C. Goodwillie, M. O. Johnston, D. A. Moeller, R. H. Ree, R. D. Sargent, and M. Vallejo-Marín. 2011. Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution* **65**:3339-3359.
- Wolak, M.E. & Keller, L.F. 2014. Dominance genetic variance and inbreeding in natural populations. *Quant. Genet. Wild* 104–127. Oxford University Press Oxford.
- Wood, J.L., Yates, M.C. & Fraser, D.J. 2016. Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evol. Appl.* **9**: 640–657.
- Wright, A.J. & Cockerham, C.C. 1985. Selection with partial selfing. I. Mass selection. *Genetics* **109**: 585–597. Genetics Soc America.
- Wright, S. 1922. Coefficients of inbreeding and relationship. *Am. Nat.* **56**: 330–338. Science Press.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97. Genetics Society of America.

Encadré 1 : Introduction à la génétique des espèces consanguines : théorie et exemple

La variance génétique d'une population ayant un coefficient de consanguinité f est égale à (Cockerham & Weir, 1984)

$$V_{G,F} = (1 + f) V_{\text{génétique}} + (1 - f) V_{\text{DR}} + f V_{\text{DI}} + 4f \text{COV}_{\text{ADI}} + f(1 - f)H^* + \text{COV}_{\text{DL}}.$$

La variance génique ($V_{\text{génétique}}$, plus classiquement appelée variance additive) représente la variance des effets additifs des allèles (qui sont indépendants des autres allèles ou locus). La variance de dominance d'une population panmictique (V_{DR}) représente la variance des déviations entre les valeurs génétiques et les valeurs additives des individus à un locus donné, sous hypothèse de panmixie. La variance de dominance d'une population purement consanguine (V_{DI}) représente la même variance des déviations entre les valeurs génétiques et additives des individus à un locus donné, mais sous l'hypothèse d'autofécondation complète. Le terme COV_{ADI} représente la covariance entre l'effet additif d'un allèle et sa déviation de dominance à l'état homozygote. Ce paramètre décrit la hausse de probabilité qu'un allèle soit transmis à l'état homozygote dans la descendance à cause de la consanguinité, ce qui a pour effet de créer une covariance entre la contribution additive d'un allèle chez les individus homozygotes, et la déviation de dominance de ce même génotype (Wolak & Keller, 2014). Le terme H^* représente la somme sur tous les locus des déviations de dominance homozygotes (H^* est aussi la dépression de consanguinité au locus donné), le tout mis au carré. Si la dépression de consanguinité est due à l'effet conjoint de nombreux locus à faible effets, alors H^* est attendu très petit (Shaw *et al.*, 1998). Enfin, COV_{LD} représente l'effet des covariances génétiques (ou effet Bulmer) sur la variance génétique.

L'estimation des différents termes nécessite de calculer les valeurs additives des allèles aux locus, et toutes les formes de déviations de dominance entre les valeurs additives et génotypiques. En génétique quantitative, les différentes quantités sont calculées par rapport à la population ancestrale panmictique, et les valeurs génétiques, additives et de dominance sont des quantités relatives qui dépendent de la moyenne de la population d'étude. Ainsi, il est nécessaire de centrer les valeurs génotypiques qui vont nous servir de base pour le calcul des autres quantités (Tableau 1.)

Tableau 1. Considérons un locus à deux allèles, A1 et A2, ayant respectivement des fréquences p_1 et p_2 . Génotypes et leurs fréquences (sous hypothèse de panmixie), valeurs phénotypiques et génotypiques associées. La valeur moyenne μ est la moyenne phénotypique et est égale à la somme des produits entre les valeurs phénotypiques et leurs fréquences.

Génotype	Fréquence	Valeur phénotypique	Valeur génotypique
A1A1	p_1^2	P_{11}	$G_{11} = P_{11} - \mu$
A1A2	$2 p_1 p_2$	P_{12}	$G_{12} = P_{12} - \mu$
A2A2	p_2^2	P_{22}	$G_{22} = P_{22} - \mu$

Les valeurs additives des allèles A1 et A2 (α_1 et α_2) sont les effets moyens des allèles sur les valeurs génotypiques (G) et sont calculés comme

$$\alpha_1 = p_1 G_{11} + p_2 G_{12} ,$$

et

$$\alpha_2 = p_1 G_{12} + p_2 G_{22} .$$

Les déviations de dominance (δ) associées aux différents génotypes peuvent être calculées avec

$$\delta_{11} = G_{11} - 2\alpha_1 ,$$

$$\delta_{12} = G_{12} - \alpha_1 - \alpha_2,$$

et

$$\delta_{22} = G_{22} - 2\alpha_2.$$

A partir de ces différentes quantités et des fréquences alléliques, il est possible de calculer les différentes composantes de la variance génétique (Tableau 2).

Tableau 2. Estimations des différentes composantes de la variance génétique. Dans les différentes formules, n_k est le nombre d'allèles au $k^{\text{ème}}$ locus, L est le nombre de loci affectant le caractère phénotypique, $p_{i,k}$ est la fréquence du $i^{\text{ème}}$ allèle au locus k , $\alpha_{i,k}$ est l'effet additif du $i^{\text{ème}}$ allèle au locus k , et $\delta_{ij,k}$ est la déviation de dominance au locus k associé au génotype ij .

Composante de la variance	Estimation
$V_{\text{génétique}}$	$2 \sum_k^L \sum_i^{n_k} p_{i,k} \alpha_{i,k}^2$
V_{DR}	$\sum_k^L \sum_i^{n_k} \sum_j^{n_k} p_{i,k} p_{j,k} \delta_{ij,k}^2$
H^*	$\sum_k^L \left(\sum_i^{n_k} p_{i,k} \delta_{ii,k} \right)^2$
σ_{DI}^2	$\sum_k^L \left(\sum_i^{n_k} p_{i,k} \delta_{ii,k}^2 - \left(\sum_i^{n_k} p_{i,k} \delta_{ii,k} \right)^2 \right)$
σ_{ADI}	$\sum_k^L \sum_i^{n_k} p_{i,k} \alpha_{i,k} \delta_{ii,k}$

Ainsi, toutes les composantes de la variance génétique sont calculables pour toutes les populations, indépendamment de leur taux d'autofécondation, mais la contribution de

chaque composante va dépendre du coefficient de consanguinité moyen de la population, et des fréquences alléliques (Walsh & Lynch, 2018). Une application numérique est proposée ci-dessous.

Considérons un caractère phénotypique tel que l'espérance de vie d'un généticien quantitatif spécialisé dans les espèces consanguines. L'espérance de vie est déterminée par un locus à deux allèles, A1 et A2. Les homozygotes A1A1 vivent en moyenne 25 ans (le temps de finir leur thèse), alors que les hétérozygotes A1A2 et les homozygotes A2A2 vivent en moyenne 80 ans (A2 est dominant sur A1). Dans la région de Montpellier l'allèle A1 est trouvé en fréquence 0.3, et l'allèle A2 en fréquence 0.7. A partir de ces informations, il est donc possible de calculer les valeurs génotypiques, et les effets additifs et de dominance (Tableau 3.)

Tableau 3. Estimation des différentes quantités génétiques pour l'exemple numérique. Ici, $\alpha_1 = -11.55$ et $\alpha_2 = 4.95$.

Génotype	Fréquence	Valeur génotypique	Valeur additive	Déviation de dominance
A1A1	0.09	-50.05	-23.10	-26.95
A1A2	0.42	4.95	-6.60	11.55
A2A2	0.49	4.95	9.90	-4.95

On peut ensuite calculer les composantes de la variance génétique :

$$V_{genique} = 2(0.3)(-11.55^2) + 2(0.7)(4.95^2) = 114.34 ,$$

$$V_{DR} = 0.09(-26.95^2) + 0.42(11.55^2) + 0.49(-4.95^2) = 133.40 ,$$

$$H^* = (0.3(-26.95) + 0.7(-4.95))^2 = 133.40 ,$$

$$V_{DI} = 0.3(-26.95^2) + 0.7(-4.95)^2 - 133.40 = 101.64 ,$$

$$COV_{ADI} = 0.3(-11.55)(-26.95) + 0.7(4.95)(-4.95) = 76.23 .$$

Chapitre 1 : Méta-analyse de l'effet de l'autofécondation sur la variance génétique au sein des population naturelle d'angiospermes

This article is published in *Evolution*.

Clo, J., Gay, L., & Ronfort, J. (2019). How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species. *Evolution*, 73(8), 1578-1590.

Présentation générale

Self-fertilization is thought to be an evolutionary dead-end (Stebbins 1957). The main argument highlighted by Stebbins is that selfing populations harbour less genetic diversity than their more outcrossing counterparts do. If our understanding of the effect of inbreeding on the genetic variance has changed over the last decades (Lande 1977, Charlesworth & Charlesworth 1995, Lande & Porcher, 2015; Abu Awad & Roze, 2018), it is now theoretically accepted that the Stebbins' argument is true when considering the effect of self-fertilization on the amount of additive of quantitative traits under stabilizing selection (Lande & Porcher, 2015; Abu Awad & Roze, 2018). Due to a more efficient purifying selection against deleterious mutations, and the magnified Bulmer effect due to reduced effective recombination rates, selfing populations are expected to harbour less genetic diversity (Lande & Porcher, 2015; Abu Awad & Roze, 2018). Empirically, however, this signal is not well supported. The analyses are nevertheless made on limited datasets (Charlesworth & Charlesworth 1995, Geber & Griffen, 2003, Oepdal et al. 2017), potentially limiting the statistical power to detect an effect of the mating system.

In this paper, I collected a large number of narrow-sense heritability estimates (1853) of natural populations of a wide range of angiosperm species (84) in order to test the effect of selfing on the genic variance. In a second time, I tried to estimate more carefully the effect of the mating system on the true adaptive potential of populations, by comparing narrow-sense heritabilities for outcrossing populations and broad-sense heritabilities for predominantly selfing populations. Finally, I tried to see if the amount of genetic variance among populations is more variable in selfing populations than in mixed-mating or outcrossing ones.

I have shown that if on average self-fertilization leads to a decrease of the adaptive potential of populations, the mating system explains overall a weak to moderate proportion of the total genetic variance. Taking into account the total genetic variance for selfing populations leads to a less drastic diminution of the adaptive potential, even if selfers are still expected to respond less efficiently to selection than their mixed-mating and outcrossing counterparts. I did not find an effect of the mating system on the among-populations amount of genetic variance. All the results and their implications for the evolution of selfing are discussed in the following article.



How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species

Josselin Clo,^{1,2} Laurène Gay,¹ and Joëlle Ronfort¹

¹AGAP, Montpellier University, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

²E-mail: josselin.clo@gmail.com

Received September 4, 2018

Accepted May 13, 2019

Most theoretical works predict that selfing should reduce the level of additive genetic variance available for quantitative traits within natural populations. Despite a growing number of quantitative genetic studies undertaken during the last two decades, this prediction is still not well supported empirically. To resolve this issue and confirm or reject theoretical predictions, we reviewed quantitative trait heritability estimates from natural plant populations with different rates of self-fertilization and carried out a meta-analysis. In accordance with models of polygenic traits under stabilizing selection, we found that the fraction of additive genetic variance is negatively correlated with the selfing rate. Although the mating system explains a moderate fraction of the variance, the mean reduction of narrow-sense heritability values between strictly allogamous and predominantly selfing populations is strong, around 60%. Because some nonadditive components of genetic variance become selectable under inbreeding, we determine whether self-fertilization affects the relative contribution of these components to genetic variance by comparing narrow-sense heritability estimates from outcrossing populations with broad-sense heritability estimated in autogamous populations. Results suggest that these nonadditive components of variance may restore some genetic variance in predominantly selfing populations; it remains, however, uncertain how these nonadditive components will contribute to adaptation.

KEY WORDS: Genetic variance, heritability, mating system, meta-analysis, quantitative traits.

Mating systems in plants are very diverse, especially among angiosperms, where hermaphroditism is frequent (~70%; Richards 1997). An important fraction of these species (~40%) regularly self-fertilizes, some (~15%) self-fertilize at a very high frequency (>90%), but apparently never exclusively (Igic and Kohn 2006). Moreover, a continuum exists between strictly allogamous and predominantly selfing species (Darwin 1876, 1878; Igic and Kohn 2006; Barrett et al. 2014) and frequent transitions from obligate outcrossing to selfing have occurred independently during flowering plant evolution (Wright et al. 2013). Two advantages of selfing have been put forward to explain these frequent transitions: (1) reproductive assurance, whereby selfing assures reproduction by eliminating the dependency on pollinating agents (Baker 1955), and (2) transmission advantage,

as the probability of transmission of a maternal gene in a selfed offspring is twice as high as that in an outcrossed offspring (Fisher 1941; Nagylaki 1976; Lloyd 1979). Despite these advantages, the small proportion of predominantly self-fertilizing species (Barrett et al. 2014; Hartfield 2016) and their generally recent origin (Goldberg et al. 2010) suggest that self-fertilization could be an evolutionary “dead-end” (Stebbins 1957; Takebayashi and Morrell 2001; Igic and Busch 2013). Stebbins’ original argument for the dead-end hypothesis is that selfing results in the loss of genetic diversity, which consequently limits the capacity of selfing populations to adapt to new or changing environments, thus increasing their probability to become extinct (Stebbins 1957). Here, we propose to use a review of the literature to explore this prediction.

Empirical data on molecular markers strongly support the prediction that selfing reduces populations' genetic diversity. Indeed, intraspecific (see, for example, Barrett and Husband 1990) and interspecific (Hamrick and Godt 1996; Glémin et al. 2006) comparisons have confirmed that selfing populations have less molecular genetic diversity than allogamous ones. However, molecular marker-based estimates of diversity may not be representative of quantitative genetic variation (Cheverud et al. 1994; Reed and Frankham 2001). Several theoretical models have been proposed to predict the effect of selfing on quantitative genetic variation within populations. Lande (1977) considered a model of stabilizing selection acting on a single polygenic trait controlled by additive loci in a partially selfing population. Assuming a Gaussian distribution of allelic effects at each locus, he found that, compared to an outcrossing population, the increase in variance due to the homozygosity generated by selfing is exactly compensated by the "purging" effect of selfing (selection against and elimination of unfavorable alleles). He also showed that the decrease in genetic variance due to selfing and caused by linkage disequilibria between alleles at different loci with opposite effects on the trait (the Bulmer effect; Bulmer 1971) is exactly compensated by the reduction in the efficacy of selection due to linkage. He concluded that the equilibrium additive genetic variance available in a population is not affected by selfing. More recently, Lande and Porcher (2015) extended this model to multiple selected traits, assuming either a Gaussian distribution of allelic effects at each locus or an infinitesimal model. They showed that, below a threshold selfing rate, the level of genetic variance should not be affected by self-fertilization (for the same reasons as Lande 1977). However, above this threshold selfing rate, the genetic variance declines drastically due to purifying selection and strong compensatory associations between alleles at different loci. Using an alternative assumption where quantitative traits under stabilizing selection are controlled by additive biallelic loci, Charlesworth and Charlesworth (1995) have shown that the genetic variance at equilibrium between mutation and selection in completely selfing populations should be one quarter that maintained in random mating ones. More recently, Abu Awad and Roze (2018) have generalized this model to take into account any selfing rate, and found a different pattern than Lande and Porcher (2015). In their model, the predicted level of genetic variance generally decreases with increasing selfing rates, with a sudden acceleration of the decrease for very high selfing rates. In all these models, when the selfing rate is very high, the effective recombination rate is so weak that selection acts on genotypes and not at the level of single loci, hence very little genetic diversity is maintained at equilibrium (Neher and Shraiman 2009; Lande and Porcher 2015; Abu Awad and Roze 2018).

Compared to molecular data, empirical evidence for a decrease in quantitative-genetic variation under selfing is more limited. A few studies comparing outcrossing and selfing populations

of closely related taxa (within species or genus) have found that outcrossing populations exhibit higher quantitative genetic variance than ones with higher selfing rates (Carr and Fenster 1994; Fenster and Carr 1997; Bartkowska and Johnston 2009). At a larger scale, Charlesworth and Charlesworth (1995) reviewed the literature to test their theoretical predictions (described above) using heritability values as well as the genetic coefficient of variation as proxies of the amount of quantitative genetic variance. The data available were scarce (37 studies) and, although theoretical predictions were supported qualitatively, the effect of the selfing rate was significant only for the genetic coefficient of variation and when the dataset was restricted to studies controlling for maternal effects (i.e., estimating paternal plant components of variance; Fig. 4 in Charlesworth and Charlesworth 1995). Hence, so far there is no strong empirical consensus on how self-fertilization affects the genetic variance available for quantitative traits within natural populations. More data have been published for numerous plant species since this literature review (Lavergne and Molofsky 2007; Leinonen et al. 2013; Wood et al. 2016), making it timely to conduct a new and more precise study.

All the above-mentioned theoretical models assume that the phenotypic value of a quantitative trait is determined by purely additive genetic effects, and thus neglect dominance and epistasis. This assumption is relevant to predict the evolutionary fate of a random mating population under selection because the additive genetic variance (V_A) is a major predictor of a population's ability to respond to selection (following the Breeders' equation; Falconer and Mackay 1996). Under selfing however, some non-additive components of genetic variation (also called "quadratic components" of genetic variance, see Cockerham and Weir 1984; Walsh and Lynch 2018, chapter 11) become nonnegligible due to the effect of dominance in conjunction with inbreeding. Neglecting these additional components could bias estimations of levels of additive genetic variance, as they contribute to the resemblance between relatives. In addition, Kelly (1999) showed that two of these above-mentioned nonadditive components, the homozygous dominance variance and the covariance of additive and dominance effects in inbred individuals (respectively, V_{DI} and COV_{ADI}), are needed to predict the evolution of a quantitative trait under partial selfing. However, the way they modify the response to selection compared to strictly outcrossing populations is not trivial (Shaw et al. 1998; Kelly 1999). Selfing also reduces the efficiency of recombination and may thus allow selection to act on gene combinations. As a result, epistasis may also be an important parameter of the response to selection in mixed mating and predominantly selfing populations (Cheverud and Routman 1995; Monnahan and Kelly 2015). Taking into account these particularities of self-fertilizing populations, we propose to further explore Stebbin's dead-end hypothesis, interpreting the quantitative genetic parameters we collected in terms of adaptive potential.

In this article, we provide an updated review of the experimental data available in the literature on the level of quantitative genetic variance in natural populations of angiosperm species. We use heritability estimates and meta-analysis methods to determine whether the currently available data on quantitative traits variation in plants confirms theoretical predictions and previous empirical observations that self-fertilization reduces the amount of additive genetic variance available for polygenic traits within natural populations. We also use the collected data to make predictions about the contribution to heritability estimates of nonadditive components of genetic variance in selfing populations, and their possible impact on adaptive potential. Finally, and as suggested by previous studies on molecular genetic diversity that found that selfers exhibit more among-populations variation in isozyme diversity than outcrossers (Schoen and Brown 1991), we look for differences in among-population variation of heritability values between mating regimes.

Materials and Methods

DATABASES COLLECTION

Our goal was to compare levels of genetic variance between traits, populations, or species, which requires some standardization of the genetic variance (to obtain a unit-free variable). This can be achieved using heritability estimates (h^2 , broad- or narrow-sense) or coefficients of genetic variance (CV_G or CV_A , the ratio of additive or total genetic standard deviation to the mean of the trait). Houle (1992, 1998) has argued that the coefficient of genetic variance is a better measure of a trait's evolvability than heritability because it is independent of environmental effects on the trait. Yet, the coefficient of genetic variation or parameters allowing its calculation are rarely reported and we therefore used h^2 instead of CV_A or CV_G as a proxy for the level of genetic variance maintained within populations.

To build our database, we collected the data previously reported in other meta-analyses on heritability (Charlesworth and Charlesworth 1995; Geber and Griffen 2003; Laverne and Molofsky 2007; Leinonen et al. 2013; Wood et al. 2016). In addition, we did a bibliographic search using Google Scholar, looking for articles including the keywords "heritability," "quantitative genetic parameters," " F_{ST} - Q_{ST} comparison," "natural population," and "plant." We limited our search to data from natural populations, and we removed any heritability estimates that were obtained by averaging over results from multiple populations; each heritability estimate we used is from data from a single population.

From each selected study, we extracted the species name, its botanic family, the type of trait (see below), the estimated heritability of the trait, and, when available, the standard error of the estimated value. Traits were separated into three classes: "morphological traits" like plant height or leaf size, "life history traits"

like flowering time or growth rate, and "others" like synthesis of chemicals compounds. We also recorded the associated P -value, the type of estimated heritability (narrow- or broad-sense), and the estimation method (animal model, half- and full-sib families, parent-offspring regression).

Selfing rate estimates were obtained from three different sources. In 47% of the collected heritability values, selfing rate estimates of the studied populations were reported in the same paper as heritability estimates. For 20% of our dataset, we used a database of selfing rates (see Winn et al. 2011) for different plant species estimated using progeny array analyses (Ritland 2002). These selfing rates are precise, but generally estimated in populations different from those used to estimate heritabilities. For 33% of our dataset, the article did not provide an estimate of the selfing rate and the species was absent from the above-mentioned database. In these cases, we searched for estimates of the inbreeding coefficient (F_{IS}) provided in other studies and used the well-known formula: $s = 2F_{IS}/(1 + F_{IS})$ (Hedrick and Cockerham 1986) to estimate the species' selfing rate. Due to potential imprecision of the two last strategies, we considered the selfing rate first as a continuous variable and then as a factor, as described in the next section.

STATISTICAL MODELS

Following previous studies, we assumed a Gaussian distribution of the estimation error (Stirling et al. 2002; Ricklefs and Cadena 2008; Wood et al. 2016), and we verified that this was correct by checking the normality of the residuals of the mixed model fitted on heritability estimates. We performed a Bayesian meta-analysis using the MCMCglmm package (Hadfield 2010) in R (R Development Core Team 2013). We used the following model:

$$h_{ijklmn}^2 = \mu + s_i + \text{trait} - \text{type}_j + \text{method}_k + \text{species}_l + \text{article_in_species}_{lm} + \text{pop_in_study}_{lmn} + \varepsilon_{ijklmn} \quad (1)$$

where h_{ijklmn}^2 is the heritability level, μ its mean, trait-type $_j$ is the fixed effect of the trait type j , method $_k$ is the fixed effect of the method of estimation k (animal model, half- and full-sib families, or parent-offspring regression). We also included the following random effects: species $_l$ is the effect of the species l . As some species were represented several times through different studies, "article_in_species $_{lm}$ " is the effect of the article m within the species l . "Pop_in_study $_{lmn}$ " is the effect of the population n within study m , within the species l , and ε_{ijklmn} is the residual error. Finally, s_i is the fixed effect of the mating system, and, as mentioned above, has been treated sequentially as a continuous variable and then as a factor to account for the imprecision of populations' selfing rate and for the potential bias in heritability estimates for mixed mating populations (see below). We defined three classes of selfing rates: complete outcrossing (including

dioecious or self-incompatible species, or populations with a selfing rate estimated at 0); mixed mating ($0 < s < 0.8$) and predominant selfing ($s \geq 0.8$).

As recommended in Charlesworth and Charlesworth (1995), we also performed the analysis described above using a single trait. We extracted heritabilities for “height at maturity” because we had sufficient data for this trait. This analysis is probably more precise because heritability values may vary between traits depending on the trait’s sensitivity to the environment of measurement. Restricting the data to a single trait may also reduce variability due to different levels and kinds of selection affecting the trait under study. For this analysis, we used the following model, in a Bayesian framework:

$$h_{iklmn}^2 = \mu + s_i + \text{method}_k + \text{species}_l + \text{article_in_species}_{lm} + \text{pop_in_study}_{lmn} + \varepsilon_{iklmn} \quad (2)$$

with the same factors as in model 1.

We tested the robustness of our results by performing 10,000 bootstraps on the whole dataset and for the analysis on height at maturity for narrow-sense heritability estimates. Simulated datasets were created by sampling the same number of observations, with replacement, as in the original dataset. We computed the robustness of our results as the number of times the analysis showed a significant effect of the selfing rate on heritability estimates.

For all analyses, we used the default priors in MCMCglmm (weakly informative; see Hadfield 2010 for more details). We verified a posteriori that it does not influence the results by permuting selfing rate associated to h^2 estimates, and verified that the priors do not bias the results toward a significant relationship between selfing rates and heritability values. MCMC chains were run for 6,000,000 iterations with a burn period of 1,000,000 and thinning interval of 50. Model parameter estimations were thus based on sampling the posterior distribution 100,000 times. We reported the 95% credible interval (Bayesian equivalent of confidence interval) and the P_{MCMC} , calculated as two times the posterior probability that the parameter is greater or smaller than zero, whichever is the smallest (Hadfield 2010). Model convergence and mixing were verified by visual examination of posterior traces and autocorrelation values.

In meta-analyses, there is a lack of uniformity across studies in parameter estimations (Nakagawa and Santos 2012). Using the standard deviation of the estimates (SD) is a way to consider this variability. In our dataset, this information was available for 27.2% of the heritability data and a preliminary analysis showed that SD values associated with the different methods of h^2 estimation differed significantly: the animal model gives more precise estimations (mean SD = 0.145) than half- and full-sib families (mean SD = 0.205), both being more precise than parent–offspring re-

gression (mean SD = 0.424). We thus used the mean SD values per method to weight h^2 estimates in our analyses. Nevertheless, as this weighting method makes the strong hypothesis that error around heritability estimates do not vary within estimation method, and as using this approach does not change the results compared to unweighted analyses, we present only results from the unweighted models. Results from the weighted method are available in the Supporting Information Materials.

EFFECT OF SELF-FERTILIZATION ON ADDITIVE GENETIC VARIANCE

We first used a dataset composed only of narrow-sense heritability estimates, and performed a regression analysis following model 1. In a second stage, we used the same dataset and treated the selfing rate as a factor to compare obligate outcrossing and predominantly selfing populations and to avoid potential biases due to estimation error in mixed mating populations. Indeed, and as mentioned in the introduction, in the case of nonnegligible rates of selfing, the covariance between relatives is no longer only described by additive (V_A), dominance (V_D), and interaction (V_I) components of genetic variance. Additional nonadditive components (sum of squared locus-specific inbreeding depressions (ι^*), V_{DI} , and COV_{ADI} ; see Walsh and Lynch 2018, chapter 11) affect the covariance between relatives and their relative contribution to the total genetic variance varies with the selfing rate (Cockerham and Weir 1984; Walsh and Lynch 2018). These components are unfortunately rarely taken into account in quantitative genetic analyses dealing with partially selfing populations (never in our dataset), which may result in biased estimates of additive genetic variance and heritability values. Fully outbred populations are not affected ($\iota^* = 0$, $V_{DI} = 0$, and $COV_{ADI} = 0$), and predominantly selfing populations are expected to be little affected when narrow-sense heritability is estimated using crosses between fully inbred lines (Walsh and Lynch 2018), as in our database. For mixed mating populations however, it is unclear whether and how neglecting these components could bias the estimation of additive genetic parameters. We thus expect that in the analyses treating the selfing rate as a factor, this potential bias should only affect the mixed mating class, and that the estimations for the outcrossing and predominantly selfing classes should be accurate and comparable.

Few narrow-sense heritability estimates were available for predominantly selfing populations, as a large proportion of estimates available for such species were broad-sense heritabilities (see Table 1). To take advantage of the full dataset and as a secondary test, we propose to correct broad-sense heritability estimates for predominantly selfing populations to make them comparable to narrow-sense ones and increase the sample size. Our correction was based on the following procedure. To infer the heritability of a trait in a predominantly selfing population, researchers generally use inbred lines or genotypes, and estimate

Table 1. Sample size in each class of mating system (and the total) for each dataset (all traits or restricted to height at maturity) and the different analyses.

		Narrow-sense heritability	Narrow-sense and corrected broad-sense estimates	Composition of genetic variance ¹
All traits	Strictly allogamous	998	998	998
	Mixed mating	761	851	275
	Autogamous	94	763	672
	Total	1853	2612	1945
Height at maturity	Strictly allogamous	48	48	–
	Mixed mating	45	50	–
	Autogamous	6	52	–
	Total	99	150	–

¹This analysis includes narrow-sense heritability estimates for outcrossing populations, and broad-sense estimates for mixed mating and predominantly selfing populations.

the genetic variance V_G as the between-lines variance of phenotypic trait values. Under the assumption of purely additive effects of loci on phenotypic values, $V_G = (1 + F_{IS}) \times V_A$, where F_{IS} is the inbreeding coefficient (Mitchell-Olds 1986; Mitchell-Olds and Bergelson 1990). It appears that, within this simplified framework, approximating V_A using V_G overestimates the true V_A value by a factor $(1 + F_{IS})$ in inbreeding populations compared to outcrossing ones. In highly selfing populations, F_{IS} can range from 0.5 (after a single generation of selfing) to 1 (after recurrent selfing for many generations). We thus decided to correct our broad-sense estimates of heritability in predominantly selfing populations by using the lowest bound of the expected F_{IS} (0.5) and divided broad-sense heritability values by 1.5. We then reiterated the Bayesian analysis with model 1 on the dataset composed of narrow-sense h^2 estimates for outcrossing and mixed mating populations, and corrected broad-sense H^2 estimates for predominantly selfing populations ($s \geq 80\%$). This correction increases the number of studies for predominantly selfing populations (Table 1). However, it may have introduced an autocorrelation problem, as F_{IS} in a population is not independent of its selfing rate. We will thus not interpret the statistical significance of those results, but only consider how they quantitatively and qualitatively differ or not from the analysis restricted to narrow-sense heritabilities.

CONSIDERATION OF NONADDITIVE COMPONENTS OF GENETIC VARIANCE IN SELFING POPULATIONS

As detailed in the introduction, the genetic parameters involved in a population's response to selection are not necessarily the same under predominant outcrossing and predominant selfing. If additive genetic variance is the major predictor of the adaptive potential of outcrossing populations, in inbreeding ones, V_{DI} and COV_{ADI} may also play a significant role (Kelly 1999). As a result, narrow-sense heritability should not totally reflect the adaptive potential and the evolvability of an inbreeding population.

Broad-sense heritability estimates are expected to include the dominance terms mentioned above. We thus propose to compare narrow-sense heritability of obligate outcrossing populations to broad-sense heritability estimates from mixed mating and predominantly selfing populations using the Bayesian statistical model detailed in model (1) with the selfing rate treated as a factor. This analysis will determine whether nonadditive components of genetic variance could compensate the loss of additive genetic variance predicted in selfing populations. In addition, we tested whether the mean narrow- and broad-sense heritability estimates differ significantly from one another within mixed mating and predominantly selfing classes (using Bayesian method).

EFFECT OF THE MATING SYSTEM ON THE AMONG POPULATIONS VARIABILITY IN LEVELS OF GENETIC DIVERSITY

To determine whether the level of genetic diversity maintained in predominantly selfing populations is more variable among populations than for outcrossing species, we used either the whole database or the narrow-sense estimates only, and we selected all the papers in which heritability has been simultaneously estimated for at least two populations. We computed the coefficient of variation of heritability values between populations (CV) for each trait. Then, we calculated the mean CV over all traits for each study. We used the same three selfing classes as defined above, and ran a nonparametric Kruskal–Wallis rank test.

Results

DATABASE COMPOSITION

We identified 106 articles meeting our research and selection criteria (associated dataset available at <https://doi.org/10.15454/OMDNEH> on the INRA data portal). They were published between 1965 and 2017, most of them after 2000. The full database

included 2836 estimates of h^2 , representing 334 populations of 84 angiosperm species. Among those, 1772 were estimated for morphological traits, 860 for life history traits, and 204 for “other” type of trait. For the trait “height at maturity,” 150 heritability estimates were available from 31 species and 108 populations. The distribution of heritability values as a function of the selfing rate was U-shaped (see Figure S1), and similar to the distribution of selfing rates over angiosperms (Igic and Kohn 2006). Table 1 summarizes the sample size of the different datasets we defined to analyze the relationship between heritability and the mating system. Using bootstrap re-sampling on the narrow-sense heritability dataset, we computed a robustness of 77.6% and 91.6% when combining all the traits and for height at maturity only, respectively.

In all the analyses we performed, the fixed effect “method_i” was not significant (whatever the dataset, $P_{\text{MCMC}} > 0.1$). Results for the fixed effect “trait-type_j” showed that heritability levels for “morphological” and “life history” traits are not significantly different from one another ($P_{\text{MCMC}} > 0.1$). Finally, the random effect “species_i” explained around 2% of the observed variance, “study_in_species_{in}” explained between 30% and 40% of the observed variance, whereas the random effect “pop_in_study_{inm}” explained around 5% of the variance.

EFFECT OF SELF-FERTILIZATION ON THE LEVEL OF ADDITIVE GENETIC VARIANCE

We first restricted our analysis to narrow-sense heritability estimates ($n = 1853$ for all traits, $n = 99$ for height at maturity, Table 1). Using the selfing rate as a continuous variable, we found a significant negative relationship between the selfing rate and heritability values, both when combining all the traits ($h^2 = 0.379 - 0.172 \times s$, 95% credible interval on slope: $[-0.326; -0.031]$; Fig. 1A) and for the trait “height at maturity” ($h^2 = 0.450 - 0.411 \times s$, 95% credible interval on slope: $[-0.599; -0.180]$; Fig. 1B). Results were qualitatively similar when selfing was treated as a factor: outcrossing populations exhibited significantly higher levels of heritability compared to predominantly selfing populations, for both “all traits” (Fig. 1C; Table 2) and the trait “height at maturity” (Fig. 1D; Table 2). Among those analyses, we can assume that the most accurate estimations of the effect of selfing were the ones obtained when selfing was treated as a factor. Indeed, and as explained in the Materials and Method section, this allowed us to isolate the mixed mating category, which is the mating strategy the most affected by potential biases due to the omission of nonadditive components of genetic variance and imprecisions in the estimation of the selfing rates. Using the selfing rate as a factor, the mean reduction in heritability level between obligate outcrossing and predominantly selfing populations reached 65% (Table 2). Yet, it is important to note that this reduction is highly variable (Table 2). In mixed mating populations, the heritability

level when considering all traits was not significantly different from the one observed in obligate outcrossing populations, but significantly higher than the one observed in predominantly selfing populations (Fig. 1C). For the trait “height at maturity,” mixed mating populations displayed intermediate levels of heritability, not significantly different from the two other mating classes (Fig. 1D).

Correcting broad-sense heritability values estimated in predominantly selfing populations and adding them to the narrow-sense datasets slightly reduced the average effect of the selfing rate as well as the corresponding credible interval, but we observed qualitatively similar results. The negative relationship between selfing rates and heritability levels remained significant, whether selfing was treated as a continuous variable (Fig. 2A, B) or as a factor (Fig. 2C, D). For mixed mating populations, the pattern of variation was the same as the one observed using only narrow-sense heritability values (Fig. 2C, D) with only small quantitative differences.

Whether we added corrected broad-sense heritabilities to narrow-sense heritabilities, for analyses combining all the traits, the mating system explained a significant but small amount of variation in heritability levels ($R^2 < 6\%$; Table 2). As expected, the proportion of variance explained was higher when we considered only the trait “height at maturity” ($9.9\% \leq R^2 \leq 22.2\%$; Table 2).

CONSIDERATION OF NONADDITIVE COMPONENTS OF GENETIC VARIANCE

As the total genetic variance is expected to contain nonadditive variance components that may contribute to the evolution of quantitative traits in partially selfing populations, we compared narrow-sense heritabilities in outcrossing populations to broad-sense heritabilities in both mixed mating and predominantly selfing populations. Contrary to our previous results, we found no significant differences between the three mating strategies: $h^2 = 0.361$ (95% credible interval: 0.305–0.423) for outcrossing, $H^2 = 0.302$ (0.180–0.423) for mixed mating, and $H^2 = 0.332$ (0.231–0.428) for selfing populations, $P_{\text{MCMC}} > 0.3$ in all pairwise comparisons. Furthermore, narrow- and broad-sense estimates do not differ significantly in mixed mating populations ($P_{\text{MCMC}} > 0.05$), but do in predominantly selfing ones ($P_{\text{MCMC}} = 0.023$).

AMONG POPULATION GENETIC VARIABILITY IN HERITABILITY VALUES

Forty and thirty-six articles reported heritability estimates for at least two populations (respectively, for analyses on narrow-sense estimates only and on narrow-sense estimates for outcrossing populations and broad-sense ones for inbred populations) and were suitable for comparing among populations variability in heritability levels between mating systems. The between-populations

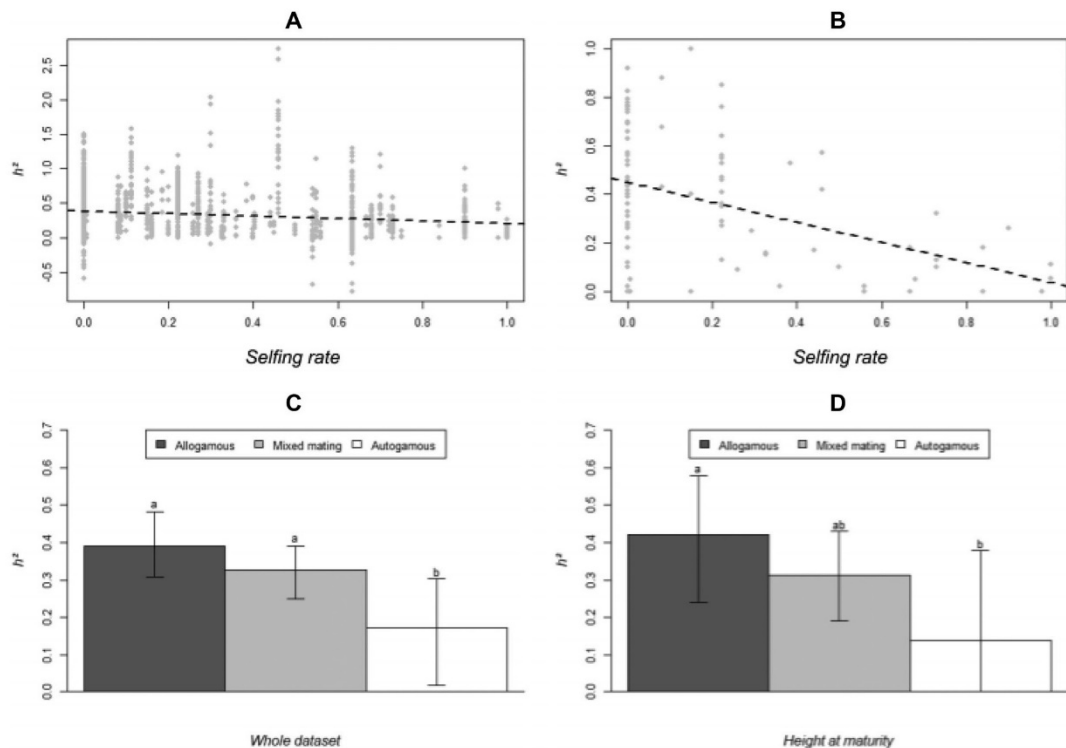


Figure 1. Relationship between narrow-sense heritability (h^2) and selfing rate. (A) Analysis on the whole dataset using selfing rate as a continuous variable. The dashed line represents the regression on narrow-sense heritabilities ($n = 1853$, $h^2 = 0.379-0.172*s$). (B) Analysis on the height at maturity using selfing rate as a continuous variable. The dashed line represents the regression on narrow-sense heritabilities ($n = 99$, $h^2 = 0.450-0.411*s$). (C) Analysis on the whole dataset using selfing rate as a factor (gray scale). (D) Analysis on the height at maturity using selfing rate as a factor (gray scale). For (C) and (D), errors bars stand for credible intervals associated with estimations, and different letters indicate a significant difference of heritability level between mating classes.

coefficient of variation of heritability estimates (CV) was not significantly different between the three mating classes (Fig. 3; Kruskal–Wallis rank test, $df = 2$, P -value > 0.5 in both cases).

Discussion

SELFING REDUCES THE LEVEL OF ADDITIVE GENETIC VARIANCE

Although the idea that self-fertilization reduces the level of additive genetic variation is widespread, strong empirical evidence was still lacking (Takebayashi and Morrell 2001; Igic and Busch 2013; Lande and Porcher 2015). In this article, we update the literature review on quantitative genetic data from natural plant populations, and confirm theoretical predictions that predominantly selfing populations should exhibit less genetic diversity than outcrossing ones for quantitative traits controlled by purely additive loci under stabilizing selection (Charlesworth and Charlesworth 1995; Lande and Porcher 2015; Abu Awad and Roze 2018). The

results also suggest a lower decrease in genetic variance for mixed mating populations, as predicted in Lande and Porcher (2015) and Abu Awad and Roze (2018). However, this result should be interpreted with caution due to the potential bias caused by the omission of V_{DI} and COV_{ADI} when estimating narrow-sense heritability in mixed mating populations. Despite this potential bias, our results lend support to the mechanisms suggested by theoretical works (Lande and Porcher 2015; Abu Awad and Roze 2018) to explain the predicted reduction in genetic variance, notably like the expected purging effect of selfing against deleterious alleles, and the predicted emergence of strong and stable linkage disequilibria between loci under strong inbreeding and stabilizing selection.

Our analysis benefits from several improvements compared to previous clues collected by Charlesworth and Charlesworth (1995). First, our results are based on a larger dataset (1853 estimations of narrow-sense heritability versus 251) that includes more species (84 vs. 37), and covering a more continuous spectrum of selfing rates. We also confirm their pioneering results in

Table 2. Summary of the effect of the mating system on heritability values, depending on the way selfing was treated in the model (factor or covariable) and on the dataset used.

			P_{MCMC}	Selfing effect ¹ (95% credible interval)	Percent diminution ² (95% credible interval)	R^2
Selfing rate as a factor	All traits	Narrow-sense h^2	0.004	−0.220 (−0.372 to −0.093)	56.8% (27.3–100%)	0.028
		Narrow-sense + corrected broad-sense h^2	–	−0.164 (−0.281 to −0.062)	44.4% (18.6–69.2%)	0.052
	Height at matu- rity	Narrow-sense h^2	0.043	−0.280 (−0.524 to −0.049)	66.7% (16.7–100%)	0.099
		Narrow-sense + corrected broad-sense h^2	–	−0.244 (−0.400 to −0.030)	58.7% (23.6–100%)	0.161
Selfing rate as a continuous variable	All traits	Narrow-sense h^2	0.011	−0.172 (−0.326 to −0.031)	45.5% (9.7–81.2%)	0.023
		Narrow-sense + corrected broad-sense h^2	–	−0.160 (−0.260 to −0.062)	42.4% (20.0–65.8%)	0.047
	Height at matu- rity	Narrow-sense h^2	0.004	−0.411 (−0.599 to −0.180)	91.5% (53.7–100%)	0.177
		Narrow-sense + corrected broad-sense h^2	–	−0.295 (−0.441 to −0.111)	70.7% (38.0–100%)	0.222

¹Equal to the slope of regression or the fixed effect of predominantly selfing mating system for analysis where selfing was treated as a covariable or a factor.

²Relative diminution of heritability level observed between completely outcrossing populations and completely selfing populations (when mating system is treated as a variable), or between the two extreme selfing classes (when treated as a factor). Bayesian 95% credible intervals were computed using posterior distributions of the above-mentioned diminutions.

less restrictive conditions: Charlesworth and Charlesworth (1995) only found a significant effect of the mating system when reducing their dataset to genetic coefficients of variation inferred from paternal half-sib analyses (12 studies), whereas our results were significant using our complete dataset, and using heritability as a proxy of additive genetic variance. Finally, as we pointed out in the introduction several factors may cause inaccuracies for mixed mating populations. We reiterated our analyses after splitting the selfing rates into classes to account as much as possible for potential biases related to the dataset, and our conclusions remained unchanged, giving strong support to our results. All these improvements allow us (1) to confirm previous expectations and to detect for the first time a clear and significant effect of self-fertilization on the amount of additive genetic variance, and (2) to quantify the contribution of nonadditive components of genetic variance in selfing populations.

We observed a mean reduction of narrow-sense heritability of around 60% between strictly outcrossing populations and

predominantly selfing ones (for the most accurate analyses, with selfing treated as a factor). This value is close to the reduction reported for quantitative traits in Geber and Griffen (2003, around 48%), to Charlesworth and Charlesworth's (1995) empirical observation (around 60% of reduction of the genetic coefficient of variation), and theoretical prediction (75% between obligate outcrossing and selfing populations; equations 1B and 3A in Charlesworth and Charlesworth 1995). However, the 95% credible interval for the estimated reduction in genetic variance is large, showing that, if some selfing populations should have close to zero genetic diversity, some may have a nonnegligible amount of genetic variance. Moreover, and despite this strong effect, the mating system only explains between 2.3% and 22.2% of the variation in heritability estimates available in our datasets. This range is also similar to previous comparisons between selfing and outcrossing plant species (for quantitative variation: around 9% in Charlesworth and Charlesworth 1995 and 18% for multiple factors including mating system in Geber and Griffen 2003; between

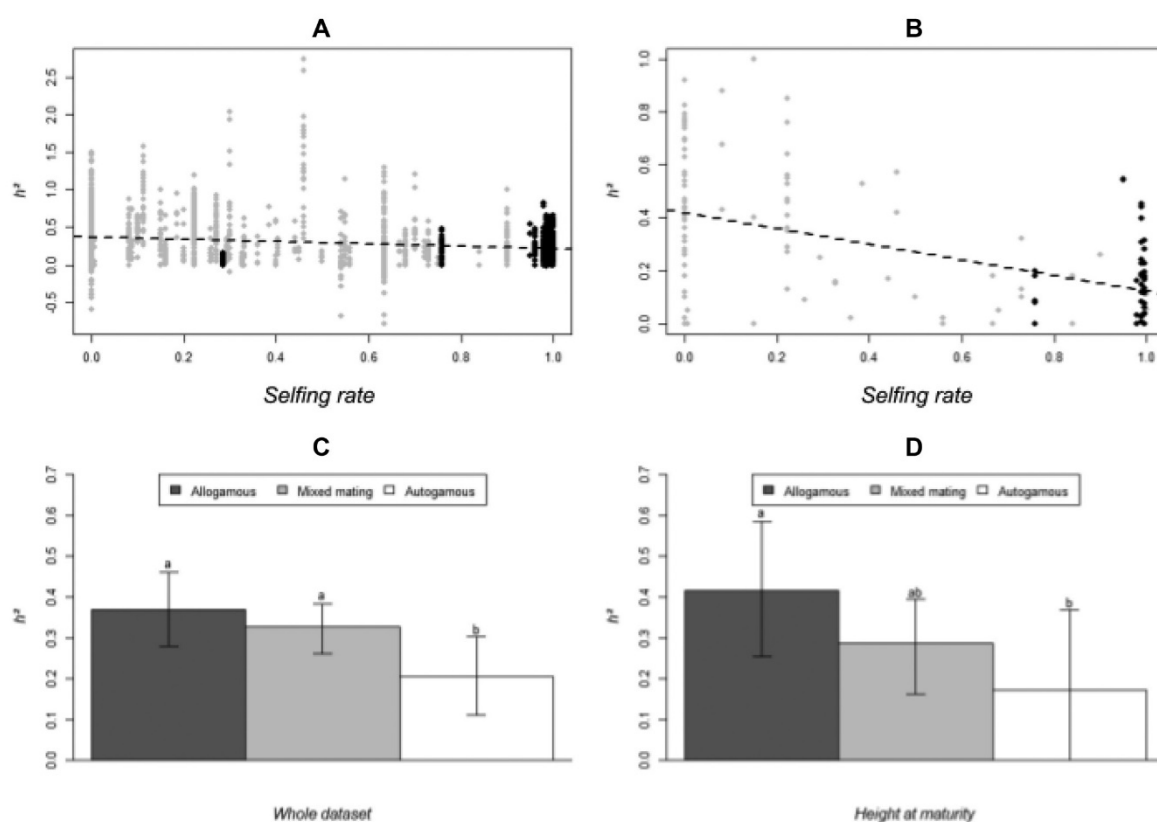


Figure 2. Relationship between narrow-sense (gray points) and corrected broad-sense heritability (black points) and selfing rate. (A) Analysis on the whole dataset using selfing rate as a continuous variable. The dashed line represents the regression on heritabilities ($n = 2612$, $h^2_{cor} = 0.377-0.160*s$). (B) Analysis on the height at maturity using selfing rate as a continuous variable. The dashed line represents the regression on heritabilities ($n = 150$, $h^2_{cor} = 0.417-0.295*s$). (C) Analysis on the whole dataset using selfing rate as a factor (gray scale). (D) Analysis on the height at maturity using selfing rate as a factor (gray scale). For (C) and (D), error bars stand for credible intervals associated with estimations, and different letters indicate a significant difference of heritability level between mating classes.

8.3 and 15.6% for nucleotide diversity variation in Glémin et al. 2006). This strong variation in the magnitude of the effect of selfing, and the limited explained variance suggest that other factors, not considered here, modulate the levels of genetic variation available for quantitative traits within natural populations. Theoretical models show that the strength and kind of selection affect the level of genetic variance maintained within populations (Charlesworth and Charlesworth 1995; Abu Awad and Roze 2018). We are probably mixing traits and populations under different kinds and strength of selection in our dataset, which could contribute to the large variance in the effect size and to the limited variance explained by the selfing rate. Other properties of populations and/or species should also modulate the level of genetic variance, either in interaction or not with the mating system, such as the effective population size (Glémin 2007), variation in life span, or dispersal abilities. Even though our dataset was not suitable to account

for these effects, the different analyses we conducted and the large dataset we gathered well support the conclusion that selfing reduces the level of additive genetic variance maintained in populations.

A major difference between theoretical models and empirical data is that the former generally assume purely additive effects of loci on phenotypic traits. However, it has been shown that V_A includes more than just additive genotypic values. Functional epistasis can also greatly contribute to this genetic component of variance (Cheverud and Routman 1995; Carter et al. 2005; Monahan and Kelly 2015); if there is epistasis, the effect of an allele will depend on the genetic background at other loci. Because selfing decreases the efficiency of recombination and thus favors the maintenance of allelic combinations between loci, functional epistasis may be more common in inbreeding populations. This could compensate the loss of additive genetic variance expected under

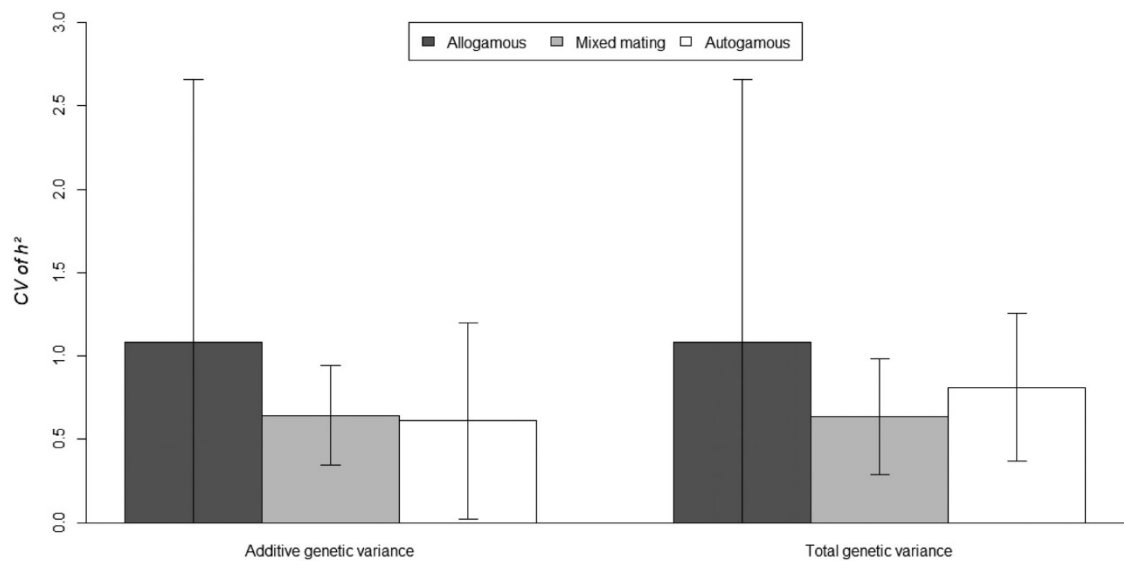


Figure 3. Influence of the mating system (gray scale) on the among populations coefficient of variation of heritability averaged over traits and estimated using articles containing at least two populations, for additive (on the left) and total genetic variance (on the right). Error bars stand for standard deviation associated.

inbreeding due to the purging process. On the hand, the strong diminution of additive variance observed between outcrossing and predominantly selfing populations in our dataset suggests that, if there is a contribution of functional epistasis to additive genetic variance, it is not sufficient to compensate the loss of diversity due to purging.

We have detailed how omitting the nonadditive genetic variance components V_{DI} and COV_{ADI} could lead to biased estimations of narrow-sense heritability for mixed mating populations. The few studies that have tried to take these nonadditive components into account found either no contribution to genetic variance (Shaw and Woolliams 1999; Abney et al. 2000), or a significant contribution of at least one of these components (generally V_{DI} ; Shaw et al. 1998; Edwards and Lamkey 2002; Kelly and Arathi 2003; Tambarussi et al. 2018). In latter cases, quantitative genetic analyses ignoring these components would have led to an overestimation of narrow-sense heritability values (Tambarussi et al. 2018). In our data, narrow- and broad-sense heritability estimates are not significantly different in the mixed mating class, suggesting that these nonadditive terms either inflate narrow-sense heritability or that they are negligible. Yet, no generalization concerning the direction of this estimation bias can be made, because if t^* and V_{DI} are positive by construction, COV_{ADI} can be either positive or negative, depending on allele frequencies (Walsh and Lynch 2018). To account for this potential bias, we excluded mixed mating populations from some analyses and compared levels of heritability only between predominantly selfing and predomi-

nantly outcrossing populations. The conclusions remained unchanged, which proves that our results are robust to this potential bias.

CONSIDERATIONS ABOUT NONADDITIVE COMPONENTS OF GENETIC VARIANCE IN SELFING POPULATIONS

Under moderate to high levels of selfing, V_A could be a poor predictor of a population's ability to respond to selection (Heywood 2005), because nonadditive components of variance become important for the prediction of the response to selection (Kelly 1999). These components include V_{DI} and COV_{ADI} , and potentially V_I (even if its significance is debated; Cheverud and Routman 1995; Hansen 2013; Monnahan and Kelly 2015), and are expected to be at least partially encompassed into broad-sense heritability estimates in mixed mating and predominantly selfing populations. Indeed, when comparing narrow-sense estimates for outcrossing populations and broad-sense estimates of heritability for mixed mating and selfing populations, the effect of the selfing rate on heritability values is no longer significant. In predominantly selfing populations, the difference between narrow and broad-sense heritability levels is significant, suggesting a nonnegligible contribution of the above-mentioned nonadditive components to the genetic variance of quantitative traits under strong inbreeding. Nevertheless, broad-sense heritability estimates also include additional nonadditive terms not involved in the adaptive response (notably V_D), and can be biased upwards, especially if there

are nongenetic maternal plant effects. With the empirical data collected, it is difficult to draw conclusions about the contribution of V_{DI} and COV_{ADI} to genetic variance, as these terms are practically never estimated in experiments, especially for predominantly selfing populations. Overall, our results clearly show that the composition of genetic variation differs as a function of the mating system: outcrossing species have more heritable additive effects, whereas selfing species have more variation in nonadditive components, and some of these components could be selectable.

IMPLICATIONS FOR THE ADAPTATION OF SELFING POPULATIONS TO ENVIRONMENTAL CHANGES

Standing genetic variation is known to play a key role in adaptation processes, at least in the short-term response to selection (Orr and Betancourt 2001; Hermisson and Pennings 2005; Barrett and Schluter 2008; Pritchard et al. 2010; Glémin and Ronfort 2013). Our results provide evidence for a decreased level in the amount of additive genetic variance available for quantitative traits in predominantly selfing populations, although this effect is highly variable between populations, as reflected by the large credible intervals and low to moderate R^2 of our models. However, our results also show that genetic variability (V_G) in selfing populations is restored by nonadditive components of genetic variance, notably the potentially selectable V_{DI} and COV_{ADI} . Yet, the way these components could modify the response to selection seems to be quite complex. Both Shaw et al. (1998) and Kelly (1999) have shown theoretically that a significant contribution of V_{DI} and COV_{ADI} to genetic variance generally greatly slows down the response to selection in a selfing population compared to an obligate outcrossing one. More empirical works, involving different populations with different selfing rates, are needed to understand at which level of inbreeding these nonadditive components become nonnegligible, and whether they could contribute positively to adaptation. Therefore, despite evidence for reduced additive genetic variance in selfing populations, our results remain insufficient to conclude on Stebbins' hypothesis, that is, that selfing populations have less genetic diversity to respond to environmental changes compared to random mating populations. Nevertheless, a directional selection experiment with freshwater snails conducted by Noël et al. (2017) has shown that, even with similar level of genetic variance, selfing lineages lost genetic variance in a few generations, leading to a slower long-term response to selection compared to outcrossing lineages.

This uncertainty about the adaptive potential of selfing populations could potentially be resolved by considering other properties of selfing populations. For instance, genetic associations among loci may hide large amounts of additive genetic variance in predominantly selfing populations, mainly due to reduced effective recombination rates allowing the persistence of

beneficial interactions between loci that improve fitness (Lande and Porcher 2015; Abu Awad and Roze 2018). This variance could be released by residual allogamy in predominantly selfing populations, thereby improving short-term evolutionary responses following an environmental change. Genetic correlations between traits (which are expected to be stronger in inbred populations due to lower effective recombination rate) could also modulate the way populations evolve in response to an environmental change, depending on whether the major axis of the G matrix is aligned or not with that of selection (Allard 1975; Bulmer 1985; Etterson and Shaw 2001; Hansen et al. 2019).

CONSIDERATIONS ABOUT HERITABILITY ESTIMATES

Whether heritability is the best descriptor of the level of genetic variance is questionable. Most quantitative genetic experiments are conducted in controlled environments: the environmental variance is artificially reduced, and this can lead to an overestimation of heritability (Falconer and Mackay 1996; Geber and Griffen 2003; but see Weigensberg and Roff 1996). Furthermore, under recurrent selfing or inbreeding, the excess of homozygosity could lead to a less efficient control of the organism's development (breakdown of developmental canalization) and could thereby inflate the phenotypic variance (Lerner 1954; Noël et al. 2017). Using the genetic coefficient of variation of quantitative traits would be preferable, as genetic variance is not standardized by a component depending on environmental variance (Houle 1992, 1998). Yet, genetic variances and mean phenotypic values are required to infer the genetic coefficient of variation and our literature review showed that they are all too rarely reported.

Our meta-analysis highlighted that the method of estimation affects the precision of h^2 estimates and suggested that the animal model gives more precise heritability estimates than other methods (see Lynch and Walsh 1998 for details about the different methods of estimation). As mentioned above, it is important to take into account the variability in the dataset in meta-analyses due to the fact that observations come from different experiments and studies (Nakagawa and Santos 2012). However, the SE values associated with h^2 estimates are rarely reported (27% of our data, 47% in Wood et al. 2016). We believe that, in a "bigdata" era, where global analyses become tractable, using information from multiple experiments, reporting the error associated with estimated parameters is essential to improve the precision of future meta-analyses.

In conclusion, we have shown that predominantly selfing populations exhibit lower levels of additive genetic variance for quantitative traits, confirming previous theoretical works and strengthening previously reported empirical observations. Our analysis, however, also showed that this reduction in additive genetic variation is compensated by the contribution of nonadditive components of genetic variance. Yet, based on

our current limited knowledge, it remains uncertain whether these nonadditive components will contribute positively to the adaptive potential of selfing populations. In addition, the ways in which selfing affects the structure of genetic variance within and between populations, the response to selection, and the interaction with the environment are quite complex. Thus, more empirical evidences are needed to understand when and why self-fertilization is a “dead-end” mating system.

AUTHOR CONTRIBUTIONS

J.C. carried out the statistical analyses. J.C., L.G., and J.R. collected the data and wrote the manuscript.

ACKNOWLEDGMENTS

We would like to thank D. Abu Awad, D. Charlesworth, P. David, E. Noël, and M. Siol for their comments on previous versions of this manuscript, and B. Fontez for discussions about the statistics. We also thank Yaniv Brandvain, Jeremiah Busch, and the three anonymous reviewers for their helpful comments that improved the manuscript. This work has been conducted with the help of the data and calculation center, South Green of the CIRAD-UMR AGAP. This work was supported by project SEAD (ANR-13-ADAP-0011).

DATA ARCHIVING

The complete dataset has been submitted to INRA dataportal: <https://doi.org/10.15454/OMDNEH>

LITERATURE CITED

- Abney, M., M. S. McPeck, and C. Ober. 2000. Estimation of variance components of quantitative traits in inbred populations. *Am J Hum Genet* 66:629–650.
- Abu Awad, D., and D. Roze. 2018. Effects of partial selfing on the equilibrium genetic variance, mutation load and inbreeding depression under stabilizing selection. *Evolution* 72:751–769.
- Allard, R. W. 1975. The mating system and microevolution. *Genetics* 79:115–126.
- Baker, H. G. 1955. Self-compatibility and establishment after ‘long-distance’ dispersal. *Evolution* 9:347–349.
- Barrett, S. C., and B. C. Husband. 1990. Variation in outcrossing rates in *Eichhornia paniculata*: The role of demographic and reproductive factors. *Plant Species Biol.* 5:41–55.
- Barrett, R. D., and D. Schluter. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38–44.
- Barrett, S. C., R. Arunkumar, and S. I. Wright. 2014. The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philos. Trans. R. Soc. B* 369:20130344.
- Bartkowska, M. P., and M. O. Johnston. 2009. Quantitative genetic variation in populations of *Amsinckia spectabilis* that differ in rate of self-fertilization. *Evolution* 63:1103–1117.
- Bulmer, M. G. 1971. The effect of selection on genetic variability. *Am. Nat.* 105:201–211.
- . 1985. The mathematical theory of quantitative genetics. Clarendon Press, Gloucestershire.
- Carr, D. E., and C. B. Fenster. 1994. Levels of genetic variation and covariation for *Mimulus* (Scrophulariaceae) floral traits. *Heredity* 72:606.
- Carter, A. J., J. Hermisson, and T. F. Hansen. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor. Popul. Biol.* 68:179–196.
- Charlesworth, D., and B. Charlesworth. 1995. Quantitative genetics in plants: The effect of the breeding system on genetic variability. *Evolution* 49:911–920.
- Cheverud, J., E. Routman, C. Jaquish, S. Tardif, G. Peterson, N. Belfiore, and L. Forman. 1994. Quantitative and molecular genetic variation in captive cotton-top tamarins (*Saguinus oedipus*). *Conserv. Biol.* 8:95–105.
- Cheverud, J. M., and E. J. Routman. 1995. Epistasis and its contribution to genetic variance components. *Genetics* 139:1455–1461.
- Cockerham, C. C., and B. S. Weir. 1984. Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics* 40:157–164.
- Darwin, C. 1876. The different forms of flowers on plants of the same species. John Murray, London.
- . 1878. The effects of cross and self-fertilization in the vegetal kingdom. John Murray, London.
- Edwards, J. W., and K. R. Lamkey. 2002. Quantitative genetics of inbreeding in a synthetic maize population. *Crop Sci.* 42:1094–1104.
- Etterson, J. R., and R. G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. *Science* 294:151–154.
- Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Pearson Education Limited, Edimbourg, U.K.
- Fenster, C. B., and D. E. Carr. 1997. Genetics of sex allocation in *Mimulus* (Scrophulariaceae). *J. Evol. Biol.* 10:641–661.
- Fisher, R. A. 1941. Average excess and average effect of a gene substitution. *Ann. Eugen.* 11:53–63.
- Geber, M. A., and L. R. Griffen. 2003. Inheritance and natural selection on functional traits. *Int. J. Plant Sci.* 164:S21–S42.
- Glémin, S. 2007. Mating systems and the efficacy of selection at the molecular level. *Genetics* 177:905–916.
- Glémin, S., and J. Ronfort. 2013. Adaptation and maladaptation in selfing and outcrossing species: New mutations versus standing variation. *Evolution* 67:225–240.
- Glémin, S., E. Bazin, and D. Charlesworth. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proc. R. Soc. Lond. B* 273:3011–3019.
- Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith, and B. Igić. 2010. Species selection maintains self-incompatibility. *Science* 330:493–495.
- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *J. Statist. Softw.* 33:1–22.
- Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. *Philos. Trans. R. Soc. B* 351:1291–1298.
- Hansen, T. F. 2013. Why epistasis is important for selection and adaptation. *Evolution* 67:3501–3511.
- Hansen, T. F., T. M. Solvin and M. Pavlicev. 2019. Predicting evolutionary potential: A numerical test of evolvability measures. *Evolution* 73:689–703.
- Hartfield, M. 2016. Evolutionary genetic consequences of facultative sex and outcrossing. *J. Evol. Biol.* 29:5–22.
- Hedrick, P. W., and C. C. Cockerham. 1986. Partial inbreeding: Equilibrium heterozygosity and the heterozygosity paradox. *Evolution* 40:856–861.
- Hermisson, J., and P. S. Pennings. 2005. Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics* 169:2335–2352.
- Heywood, J. S. 2005. An exact form of the breeder's equation for the evolution of a quantitative trait under natural selection. *Evolution* 59:2287–2298.
- Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130:195–204.

- . 1998. How should we explain variation in the genetic variance of traits? *Genetica* 102:241.
- Igic, B., and J. W. Busch. 2013. Is self-fertilization an evolutionary dead end? *New Phytol.* 198:386–397.
- Igic, B., and J. R. Kohn. 2006. The distribution of plant mating systems: Study bias against obligately outcrossing species. *Evolution* 60:1098–1103.
- Kelly, J. K. 1999. Response to selection in partially self-fertilizing populations. I. Selection on a single trait. *Evolution* 53:336–349.
- Kelly, J. K., and H. S. Arathi. 2003. Inbreeding and the genetic variance in floral traits of *Mimulus guttatus*. *Heredity* 90:77.
- Lande, R. 1977. The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* 86:485–498.
- Lande, R., and E. Porcher. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* 200:891–906.
- Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc. Natl. Acad. Sci. USA* 104:3883–3888.
- Leinonen, T., R. S. McCaums, R. B. O'hara, and J. Merilä. 2013. QST–FST comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nat. Rev. Genet.* 14:179–190.
- Lerner, I. M. 1954. *Genetic homeostasis*. Oliver & Boyd, Edinburgh.
- Lloyd, D. G. 1979. Some reproductive factors affecting the selection of self-fertilization in plants. *Am. Nat.* 113:67–79.
- Lynch, M., and B. Walsh. 1998. *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, MA.
- Mitchell-Olds, T. 1986. Quantitative genetics of survival and growth in *Impatiens capensis*. *Evolution* 40:107–116.
- Mitchell-Olds, T., and J. Bergelson. 1990. Statistical genetics of an annual plant, *Impatiens capensis*. I. Genetic basis of quantitative variation. *Genetics* 124:407–415.
- Monnahan, P. J., and J. K. Kelly. 2015. Epistasis is a major determinant of the additive genetic variance in *Mimulus guttatus*. *PLoS Genet.* 11:e1005201.
- Nagylaki, T. 1976. A model for the evolution of self-fertilization and vegetative reproduction. *J. Theor. Biol.* 58:55–58.
- Nakagawa, S., and E. S. Santos. 2012. Methodological issues and advances in biological meta-analysis. *Evol. Ecol.* 26:1253–1274.
- Neher, R. A., and B. I. Shraiman. 2009. Competition between recombination and epistasis can cause a transition from allele to genotype selection. *Proc. Natl. Acad. Sci. USA* 106:6866–6871.
- Noël, E., P. Jarne, S. Glémin, A. MacKenzie, A. Segard, V. Sarda, and P. David. 2017. Experimental evidence for the negative effects of self-fertilization on the adaptive potential of populations. *Curr. Biol.* 27:237–242.
- Orr, H. A., and A. J. Betancourt. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics* 157:875–884.
- Pritchard, J. K., J. K. Pickrell, and G. Coop. 2010. The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* 20:R208–R215.
- R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Reed, D. H., and R. Frankham. 2001. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55:1095–1103.
- Richards, A. J. 1997. *Plant breeding systems*. Garland Science, New York, NY.
- Ricklefs, R. E., and C. D. Cadena. 2008. Heritability of longevity in captive populations of nondomesticated mammals and birds. *J. Gerontol. Ser. A* 63:435–446.
- Ritland, K. 2002. Extensions of models for the estimation of mating systems using independent loci. *Heredity* 88:221.
- Schoen, D. J., and A. H. Brown. 1991. Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proc. Natl. Acad. Sci. USA* 88:4494–4497.
- Shaw, F. H., and J. A. Woolliams. 1999. Variance component analysis of skin and weight data for sheep subjected to rapid inbreeding. *Genet. Sel. Evol.* 31:43.
- Shaw, R. G., D. L. Byers, and F. H. Shaw. 1998. Genetic components of variation in *Nemophila menziesii* undergoing inbreeding: Morphology and flowering time. *Genetics* 150:1649–1661.
- Stebbins, G. L. 1957. Self-fertilization and population variability in the higher plants. *Am. Nat.* 91:337–354.
- Stirling, D. G., D. Réale, and D. A. Roff. 2002. Selection, structure and the heritability of behaviour. *J. Evol. Biol.* 15:77–289.
- Takebayashi, N., and P. L. Morrell. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am. J. Bot.* 88:1143–1150.
- Tambarussi, E. V., F. B. Pereira, P. H. M. da Silva, D. Lee, and D. Bush. 2018. Are tree breeders properly predicting genetic gain? A case study involving *Corymbia* species. *Euphytica* 214:150.
- Walsh, B., and M. Lynch. 2018. *Evolution and selection of quantitative traits*. Sinauer, Sunderland, MA.
- Weigensberg, I., and D. A. Roff. 1996. Natural heritabilities: Can they be reliably estimated in the laboratory? *Evolution* 50:2149–2157.
- Winn, A. A., E. Elle, S. Kalisz, P. O. Cheptou, C. G. Eckert, C. Goodwillie, M. O. Johnston, D. A. Moeller, R. H. Ree, R. D. Sargent, et al. 2011. Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution* 65:3339–3359.
- Wood, J. L., M. C. Yates, and D. J. Fraser. 2016. Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evol. Appl.* 9:640–657.
- Wright, S. I., S. Kalisz, and T. Slotte. 2013. Evolutionary consequences of self-fertilization in plants. *Proc. R. Soc. B* 280:20130133.

Associate Editor: Y. Brandvain
Handling Editor: M. R. Servedio

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Distribution of selfing rates associated with our heritability estimates for the complete dataset ($n = 2836$).

Table S1. Summary of the effect of mating system on heritability values for weighted analyses, depending on the dataset used, using selfing rate as a continuous variable.

Table S2. Summary of the effect of mating system on heritability values for weighted analyses, depending on the dataset used, using selfing rate as a factor.

Matériel supplémentaire

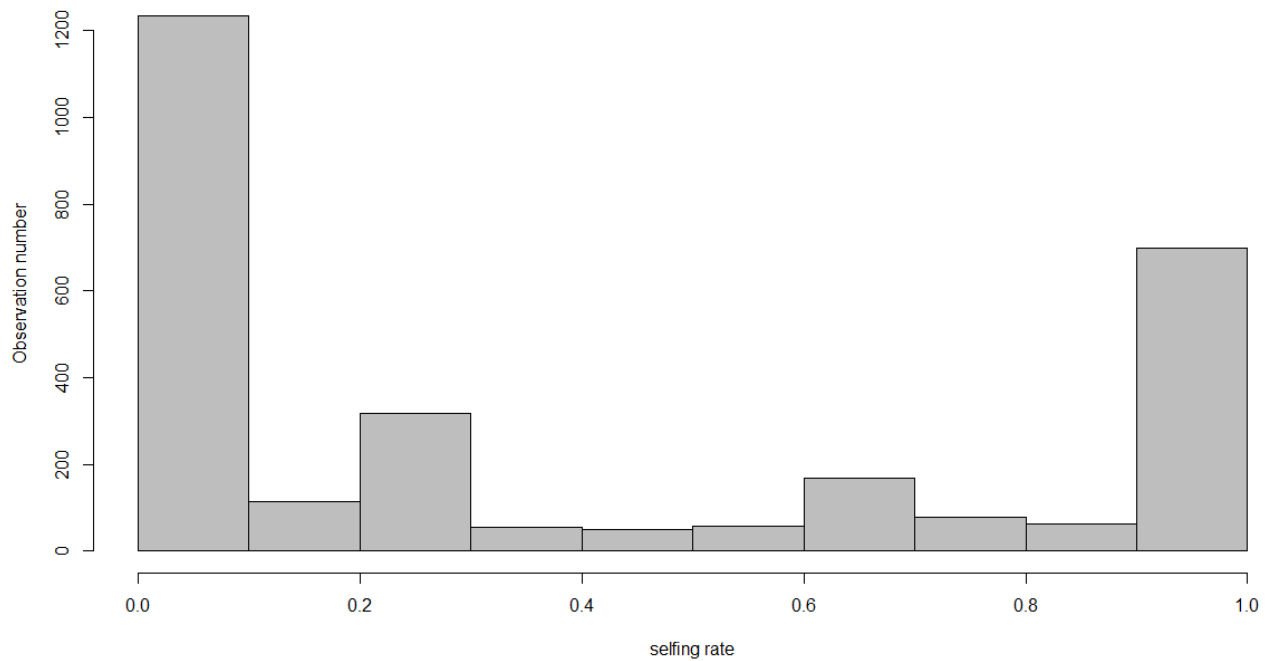


Figure S1. Distribution of selfing rates associated with our heritability estimates for the complete dataset (n=2612).

Table S1. Summary of the effect of mating system on heritability values for weighted analyses, depending on the dataset used, using selfing rate as a continuous variable.

		Equation (CI on slope)	P-value
All traits	Narrow-sense h^2	$h^2_A = 0.384 - 0.191*s$ (-0.332 – -0.029)	0.014
	broad-sense estimates corrected	$h^2 = 0.379 - 0.172*s$ (-0.271 – -0.072)	0.001
Height at maturity	Narrow-sense h^2	$h^2_A = 0.442 - 0.378*s$ (-0.588 – -0.141)	0.004
	broad-sense estimates corrected	$h^2 = 0.402 - 0.250*s$ (-0.398 – -0.101)	0.002

Table S2. Summary of the effect of mating system on heritability values for weighted analyses, depending on the dataset used, using selfing rate as a factor.

		Mean heritability values (CI)	P-value
All traits	Narrow-sense h^2	Outcrossing: 0.39 (0.32-0.46) Mixed mating: 0.32 (0.23-0.40) Selfing: 0.15 (0.00-0.31)	0.004
	broad-sense estimates	Outcrossing: 0.39 (0.32-0.45) Mixed mating: 0.31 (0.24-0.39) Selfing: 0.20 (0.09-0.30)	-
	corrected		
Height at maturity	Narrow-sense h^2	Outcrossing: 0.41 (0.30-0.52) Mixed mating: 0.32 (0.17-0.47) Selfing: 0.16 (0.00-0.40)	0.042
	broad-sense estimates	Outcrossing: 0.38 (0.27-0.48) Mixed mating: 0.31 (0.18-0.46) Selfing: 0.18 (0.01-0.35)	-
	corrected		

Chapitre 2 : Le résidu d'allogamie et les covariances génétiques pourraient faciliter l'adaptation des espèces autogames aux changements environnementaux

This article is published in *Journal of Evolutionary Biology*.

Clo, J., Ronfort, J., & Abu Awad, D. (2020). Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. *Journal of Evolutionary Biology*, 33: 1203-1215.

Présentation générale

In the previous chapter, I have shown that the effect of the mating system on additive genetic variance is not as strong as expected by theoretical models. Nevertheless, it seems that when the effect is observable, the mean decrease in additive variance is substantial (around 60%). Nevertheless, and as pointed out by the same theoretical works, the genetic associations generated by the reduced effective recombination rate could represent a bunch of variance that could boost the adaptive potential of selfing populations during a shift from stabilizing to directional selection (Lande & Porcher 2015, Abu Awad & Roze 2018).

In this second chapter, I explored if this prediction is true, by coding a change of the phenotypic optimum in previous models of stabilizing selection assuming pure additivity at the phenotypic scale. In a first time, I decomposed the genetic variance component of outcrossing ($s = 0$), mixed mating ($s = 0.5$) and selfing populations ($s = 0.95$) at mutation-selection-drift equilibrium, in within-loci component (the genic variance), and among-loci component (genetic associations), for different mutation rates, as it is the critical parameters influencing the storage of cryptic diversity. In a second time, I followed the dynamics of fitness, phenotype and genetic components of variance in order to better understand how populations adapt or not.

We found that, as expected from single-locus models, when the mutation rate is low, selfing populations do not store diversity through genetic associations. In this case, the lack of genic variance, due to a more effective purifying selection, leads to a weaker response to selection in selfing populations compare to outcrossing and mixed mating ones. When the mutation rate is higher, and when the among-loci component contributes substantially to the genetic variance, the mating system does not limit the adaptation of populations anymore. In this case, cryptic diversity is released in selfing populations by residual allogamy between fully inbred lines, allowing populations to track the new phenotypic optimum as efficiently than their more outcrossing counterparts. The genetic variance at the end of the adaptation process is similar that the one at mutation-selection-drift equilibrium, suggesting that selfing populations can face several environmental changes. All the results are described in the following manuscript.

Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations

Josselin CLO¹  | Joëlle Ronfort¹ | Diala Abu Awad^{1,2} 

¹AGAP, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France

²Department of Population Genetics, Technische Universität München, Freising, Germany

Correspondence

Josselin CLO, AGAP, CIRAD, INRAE, Institut Agro, Univ Montpellier, Arthur Young Street, 34000 Montpellier, France.
Email: josselin.clo@gmail.com

Funding information

Alexander von Humboldt-Stiftung; Agence Nationale de la Recherche; SEAD, Grant/Award Number: ANR-13-ADAP-0011; TUM University Foundation Fellowship

Abstract

Standing genetic variation is considered a major contributor to the adaptive potential of species. The low heritable genetic variation observed in self-fertilizing populations has led to the hypothesis that species with this mating system would be less likely to adapt. However, a non-negligible amount of cryptic genetic variation for polygenic traits, accumulated through negative linkage disequilibrium, could prove to be an important source of standing variation in self-fertilizing species. To test this hypothesis, we simulated populations under stabilizing selection subjected to an environmental change. We demonstrate that, when the mutation rate is high (but realistic), selfing populations are better able to store genetic variance than outcrossing populations through genetic associations, notably due to the reduced effective recombination rate associated with predominant selfing. Following an environmental shift, this diversity can be partially remobilized, which increases the additive variance and adaptive potential of predominantly (but not completely) selfing populations. In such conditions, despite initially lower observed genetic variance, selfing populations adapt as readily as outcrossing ones within a few generations. For low mutation rates, purifying selection impedes the storage of diversity through genetic associations, in which case, as previously predicted, the lower genetic variance of selfing populations results in lower adaptability compared to their outcrossing counterparts. The population size and the mutation rate are the main parameters to consider, as they are the best predictors of the amount of stored diversity in selfing populations. Our results and their impact on our knowledge of adaptation under high selfing rates are discussed.

KEYWORDS

adaptation, linkage disequilibrium, mutation rate, self-fertilization, standing genetic variance

1 | INTRODUCTION

Natural populations typically harbour genetic variation, especially at loci governing polygenic traits (Clo, Gay, & Ronfort, 2019; Mittell, Nakagawa, & Hadfield, 2015; Wood, Yates, & Fraser, 2016). This variation, known as standing genetic variation,

has been considered to be an important predictor for the adaptive potential of populations (Barrett & Schluter, 2008; Glémin & Ronfort, 2013; Hermisson & Pennings, 2005; Matuszewski, Hermisson, & Kopp, 2015; Orr & Betancourt, 2001; Pritchard, Pickrell, & Coop, 2010). Indeed, standing variation represents an easily accessible source of genetic variation, that is readily

The peer review history for this article is available at <https://publons.com/publon/10.1111/jeb.13660>

© 2020 European Society For Evolutionary Biology. Journal of Evolutionary Biology © 2020 European Society For Evolutionary Biology

available for adaptation to changing or heterogeneous conditions (Barrett & Schluter, 2008; Hermisson & Pennings, 2005). Compared to adaptation from de novo mutations, the probability of adapting from standing variation is higher simply because mutations already segregating in a population are expected to be present at higher frequencies (Barrett & Schluter, 2008; Innan & Kim, 2004). It has also been suggested that populations adapting from standing genetic variation can cope with more severe and more rapid environmental change, as they are able to cross larger distances in the phenotype space (Matuszewski et al., 2015). The amount of standing variation available in a population is thus expected to play a key role in adaptation, and any forces affecting it may greatly influence whether or not populations are able to survive new environments.

The mating system is a population characteristic that is known to greatly affect the amount and structure of genetic variation. For instance, both theoretical models (Abu Awad & Roze, 2018; Charlesworth & Charlesworth, 1995; Lande & Porcher, 2015) and empirical data (Charlesworth & Charlesworth, 1995; Clo et al., 2019; Geber & Griffen, 2003) have shown that, compared to outcrossing populations, self-fertilization reduces, on average, the amount of additive genetic variance for polygenic or quantitative traits under stabilizing selection. This diminution is due to more efficient purifying selection under selfing and to negative linkage disequilibria maintained between alleles at different loci: the so-called Bulmer effect (Abu Awad & Roze, 2018; Lande & Porcher, 2015). Due to this reduction in genetic variability, predominant selfing has been described as an evolutionary dead-end (Ilgic & Busch, 2013; Stebbins, 1957; Takebayashi & Morrell, 2001). However, recent theoretical work and empirical data have highlighted that cryptic genetic variability could contribute to the adaptive potential of natural populations. Cryptic genetic variation has been defined as the part of a population's standing genetic variation that has no effect on the phenotypic variation in a stable environment but can increase heritable variation in environmental conditions rarely experienced (Gibson & Dworkin, 2004; Paaby & Rockman, 2014). Such variability has been detected in both outcrossing (in sticklebacks, McGuigan, Nishimura, Currey, Hurwit, & Cresko, 2011, cavefish, Rohner et al., 2013, dung flies, Berger, Bauerfeind, Blanckenhorn, & Schäfer, 2011, gulls, Kim, Noguera, Tato, & Velando, 2013 or spadefoot toads, Ledon-Rettig, Pfennig, & Crespi, 2010) and selfing species (*Caenorhabditis elegans*, Milloz, Duveau, Nuez, & Félix, 2008; *Arabidopsis thaliana*, Queitsch, Sangster, & Lindquist, 2002). Two main mechanisms could explain the accumulation and the release of such variance: interactions between loci (Badano & Katsanis, 2002; Shao et al., 2008), and phenotypic plasticity (Anderson, Lee, Rushworth, Colautti, & Mitchell-Olds, 2013). In this paper, we focus on interactions between loci maintained at stabilizing selection.

In maintaining the population close to the phenotypic optimum, stabilizing selection disfavours genetic and phenotypic diversity (Abu Awad & Roze, 2018; Lande & Porcher, 2015). However, the

structure of the additive variance also strongly depends on the trait mutation rate and the prevalence of pleiotropy (Abu Awad & Roze, 2018; Charlesworth, 1990; Lande & Porcher, 2015). When the per-trait mutation rate is weak, associations between loci are negligible (mutations of strong effect arise slowly and are easily purged), but when the rate increases, the creation and maintenance of co-adapted gene complexes structure the additive variance into positive within-loci components and negative among-loci components of variance, reducing the observed level of additive variance (Abu Awad & Roze, 2018). The remobilization through recombination of this among-loci component of variance could boost the adaptability of populations undergoing an environmental shift (Le Rouzic & Carlborg, 2008). Indeed, if associations between loci are broken, segregating alleles could express some or all of their additive effects in new genetic backgrounds. Such remobilization is only possible if residual allogamy occurs (i.e. if the selfing rate is not equal to 1), which is expected to be common in natural plant populations (Clo et al., 2019; Kamran-Disfani & Agrawal, 2014). Classical models analysing the effect of selfing on adaptation from standing genetic variation have considered a single locus (Glémin & Ronfort, 2013), thus neglecting interactions among loci that could contribute to an important fraction of standing genetic variation. As self-fertilization reduces the effective recombination rate (Nordborg, 2000), allowing the maintenance of co-adapted gene complexes, the storage of genetic diversity through genetic associations should be more prevalent in selfing populations (as suggested in Abu Awad & Roze, 2018; Charlesworth, 1990; Lande & Porcher, 2015).

In this paper, we explore this hypothesis, using a quantitative genetics framework. We describe and quantify how, to what degree, and under which conditions populations accumulate genetic variation at polygenic traits. Because the development of tractable analytical predictions of the phenotype and genetic variance dynamics is not trivial and often requires strong hypotheses to be made, notably random mating and considering the system to be at linkage equilibrium (Barton & Turelli, 1987; Bürger, 1993; Keightley & Hill, 1989), we chose to use a simulation approach. Although it is known that directional dominance plays an important role in the fate of mutations, especially in the presence of self-fertilization, the interaction of both directional dominance and epistasis with mating systems is not well known. In order to avoid further complicating the interpretation of our results, we have considered a fully additive model, which despite its simplicity, has previously led to surprisingly accurate predictions (Manna, Martin, & Lenormand, 2011; Martin, Elena, & Lenormand, 2007). We show that, in models allowing genetic diversity to be stored through genetic associations and when adaptation is only possible from pre-existing standing genetic variation, predominantly selfing populations can adapt as well as their mixed-mating and outcrossing counterparts, when mutation rate is high (but realistic), despite initially low levels of observable genetic diversity. If the mutation rate is low, purifying selection impedes the storage of diversity through genetic associations, and our simulations confirm

the more classical results that selfing populations adapt less well than their outcrossing counterparts do.

2 | MATERIALS AND METHODS

2.1 | General assumptions

We consider the evolution of a quantitative trait Z in a population of size N , made of diploid individuals reproducing through partial self-fertilization, with a constant selfing rate, (and hence constant inbreeding level) σ . The phenotypic value z of an individual is determined by the additive action of L loci each with an infinite possible number of alleles and is given by

$$Z = g + e, \quad (1)$$

where g is the genetic component of the individual's phenotype and is given by $g = \sum_j^L (g_j^M + g_j^P)$, with g_j^M (respectively, g_j^P) the additive allelic effect at locus j inherited from the maternal (respectively, paternal) gamete. There are no dominance or epistasis at the phenotypic scale in this model, but both of which arise naturally at fitness scale, when considering stabilizing selection and a mean phenotype close to the optimum, the mean dominance $h \approx 0.25$ and epistasis effects are on average null for fitness. The random environmental effect, e , is drawn from a Gaussian distribution of mean 0 and variance V_E , and is considered to be independent from the genetic components of fitness. The trait initially undergoes stabilizing selection around an optimal phenotypic value, denoted Z_{opt} . The fitness value W_Z of an individual with phenotype z is thus described by the Gaussian function:

$$W_Z = e^{-d^2 / 2\omega^2}, \quad (2)$$

where d is the distance between the individual's phenotype z and the optimum trait value and ω^2 is the width of the fitness function and represents the strength of selection.

2.2 | Simulation model

We implement the model described above into an individual-based simulation model written in C++, a modified version of the 'continuum of alleles' program provided in Abu Awad and Roze (2018) available in Supporting information and online (<https://github.com/dialaAbAw/SelfingAdaptation>).

We consider a population of N diploid individuals, each represented by two linear chromosomes with L multi-allelic loci, coding for a single quantitative trait under selection. At the beginning of each simulation, all individuals are genetically identical and are at the phenotypic optimum (all loci carry alleles with effect 0 and $Z_{\text{OPT}} = 0$). The life cycle can be summarized by five successive events. First, there is a phenotype-dependent choice of the first parent (selection), followed by mating-type choice (selfing vs. outcrossing at rates

σ and $(1-\sigma)$, respectively) and then, in the case of outcrossing, phenotype-dependent choice of the second parent. Selection takes place as follows: if the ratio of the selected parent's fitness over the highest recorded fitness value in the current generation is higher than a number sampled in a uniform law comprised between 0 and 1, the individual is allowed to reproduce. Once the two parents are chosen, they each contribute a gamete, produced through uniformly recombining the parental chromosomes. The number of cross-overs is sampled from a Poisson distribution with parameter R , the map length. From Haldane's mapping function, the recombination rate

between two adjacent loci is $r = \frac{1}{2} \left[1 - \exp\left(-\frac{2R}{L-1}\right) \right]$. We choose parameters that ensure that $r \approx .5$, such that loci remain, at least physically, unlinked. This phase is then followed by the introduction of new mutations, the number of which is sampled from a Poisson distribution with parameter U (with $U = \mu L$, μ being the per locus mutation rate). The additive value of a new mutant allele is drawn from a normal distribution of mean 0 and variance a^2 .

Each simulation consists of two phases, the first being burn-in time to allow the population to attain mutation-selection-drift (M-S-D) equilibrium. The population is considered to be at M-S-D equilibrium when the average fitness value calculated over the last thousand generations does not differ by more than one per cent from the mean fitness calculated over the previous thousand generations. The second phase consists of following the population after a shift in the phenotypic optimum. After the shift, the haploid genomic mutation rate U is set to 0 so that the only source of genetic variability to reach the new optimum is the standing variation accumulated before the environmental change.

Each simulation consists of two phases, the first being burn-in time to allow the population to attain mutation-selection-drift (M-S-D) equilibrium. The population is considered to be at M-S-D equilibrium when the average fitness value calculated over the last thousand generations does not differ by more than one per cent from the mean fitness calculated over the previous thousand generations. The second phase consists of following the population after a shift in the phenotypic optimum. After the shift, the haploid genomic mutation rate U is set to 0 so that the only source of genetic variability to reach the new optimum is the standing variation accumulated before the environmental change.

2.3 | Simulation parameter values

Simulations were run for several parameter sets in order to evaluate the effects of each parameter on both the equilibrium conditions and on the ability of populations to adapt after the environmental change. The values chosen for the mutation rate U range from 0.005 to 0.1, reflecting the *per-trait* haploid genomic mutation rate found in the literature (Haag-Liautard et al., 2007; Keightley & Bataillon, 2000; Shaw, Geyer, & Shaw, 2002). We use parameter set values similar to those explored in Bürger, Wagner, and Stettinger (1989) and Ronce, Shaw, Rousset, and Shaw (2009), with the number of freely recombining loci under selection $L = 50$, and $a^2 = 0.05$, $V_E = 1$, $\omega^2 = 1, 9$ or 99 . In this model, the mean deleterious effect of mutations $\bar{s}$ is therefore equal to 0.0125, 0.0025 or 0.00025 ($\bar{s} = a^2 / 2(\omega^2 + V_E)$, Martin & Lenormand, 2006).

Before the shift in the optimum, Z_{OPT} is set to 0, and after M-S-D equilibrium, it is set to 2.5, so that the shift is of order $L a^2$, and ensures that adaptation is due to genetic changes and not environmental effects (Ronce et al., 2009). We follow populations only over 20 generations after the shift in the optimum, as we chose to follow the possibility of adaptation without any

contribution from de novo mutations. After this time limit, genetic variation starts to erode, and the observed results are mainly due to the effects of loss of diversity due to drift and not to the selection process. Although simulations were run over a large range of selfing rate values, throughout the manuscript we show results run principally for three rates of self-fertilization, $\sigma = 0, 0.5$ and 0.95 , representing outcrossing, mixed-mating and predominantly selfing, respectively. These three values were chosen because they were representative of the main patterns of adaptation we observed over the whole range of selfing rates (σ from 0 to 1). We also considered three population sizes $N = 250, 1,000$ or $10,000$, which allow to test the effect of drift.

2.4 | Analysis of the effect of selfing on the equilibrium genetic variance

Following Turelli and Barton (1990), we decompose the genetic variance of a polygenic trait using the following equation:

$$V_A = 2 \sum_j^L (C_{jj} = C_{jj}^M) = 2 \sum_{j \neq k}^L (C_{jk} = C_{jk}^M), \quad (3)$$

with

$$C_{jj} = \frac{1}{2} E[(g_j^M - \bar{g})^2 + (g_j^P - \bar{g})^2], \quad (4)$$

and

$$C_{jk} = E[(g_j^M - \bar{g})(g_k^P - \bar{g})], \quad (5)$$

where $\bar{g}$ is the mean allelic effect on the phenotype at locus j and g_j^M (respectively, g_j^P) is the allelic effect at locus j inherited from the maternal (respectively, paternal) gamete. The sum of C_{jj} over loci represents the genic variance (noted V_{genic} , which is the genetic variance of a trait in a population harbouring the same allelic frequencies as the population under study, but without any genetic associations between loci) and is computed from simulation outputs following Equation 4. The sum of all values of C_{jk} represents the covariance in allelic effects at locus j on the maternally and paternally inherited chromosomes and represents the fraction of the genetic variance due to excess of homozygosity (noted V_{inbred}); we compute it following Equation 5. This quantity represents $F V_{\text{genic}}$, where F is the inbreeding coefficient of the population. The first term of Equation 3 ($2 \sum_j^L (C_{jj} + C_{jj}^M)$) represents the genetic variance due to within locus variation. The second term ($2 \sum_{j \neq k}^L (C_{jk} + C_{jk}^M)$) represents the component of the variance due to associations between loci (noted COV_{LD}) and is obtained by computing the sum of covariances between genotypic values among all pair of loci. This component is proportional to linkage disequilibrium (LD) and tends to be negative under stabilizing selection due to associations between alleles at different loci with compensatory effects (*i.e.* the allele on one locus is positive, the other negative, their effects on the phenotype thus cancel out when both are present). At any time in the simulated populations, $V_A = V_{\text{genic}} + V_{\text{inbred}} + \text{COV}_{\text{LD}}$.

As a check, estimates of additive genetic variance observed at M-S-D equilibrium in our simulations were compared to analytical predictions obtained by Bürger et al. (1989) (stochastic house-of-cards approximate) and by Abu Awad and Roze (2018) under similar assumptions. Comparisons and details are shown in Appendix S1.

2.5 | Analyses of the response to an environmental change

To analyse the population's response to an environmental change, we focus on the 20 first generations after the environmental change. We follow the temporal dynamics of the additive genetic variance of the trait and of its components. We also analyse changes with time of the mean trait value and population fitness, as well as the dynamics of the number of haplotypes and the proportion of new haplotypes generated after the environmental change.

In addition, and in order to determine whether the remobilization of COV_{LD} plays a role in the adaptive process of selfing populations, we compute for each mutation rate, the slope of the trait mean variation shortly after the environmental change (during the first five generations) as a function of the amount of additive variance available at M-S-D equilibrium. If remobilization of COV_{LD} is involved in the adaptive process of selfing populations, the initial slope, for an initially similar amount of additive variance, should be higher in selfing populations than in mixed-mating and outcrossing ones.

3 | RESULTS

3.1 | Level and structure of additive genetic variance at mutation-selection-drift equilibrium

In agreement with expectations from previous works, the additive genetic variance present at M-S-D equilibrium in our simulations is overall negatively correlated with the selfing rate (Figure 1), in part due to the higher efficiency of selfing in purging deleterious mutations. In general, reducing the strength of stabilizing selection increases the overall amount of additive variance, as mutations of small effect are not as easily purged from the population (Figure 1). The population size, however, has little qualitative effect on the relationship between selfing and V_A , at least as long as the mutation rate is not too high (Figure S1). As detailed in Appendix S1, the values of V_A obtained from our simulations were compared to expectations from analytical approximations from Abu Awad and Roze (2018) and the stochastic house-of-cards approximation (SHC, Bürger et al., 1989). We generally find a good agreement between our results and the SHC approximation for the smaller population sizes and outcrossing populations ($\sigma = 0$ or 0.5). Approximations from Abu Awad and Roze (2018) only hold for predominantly selfing populations, considering the whole genome as a single locus ($\sigma = 0.95$, equation 44 in Abu Awad & Roze, 2018), or for all mating systems when population size is large ($N = 10,000$).

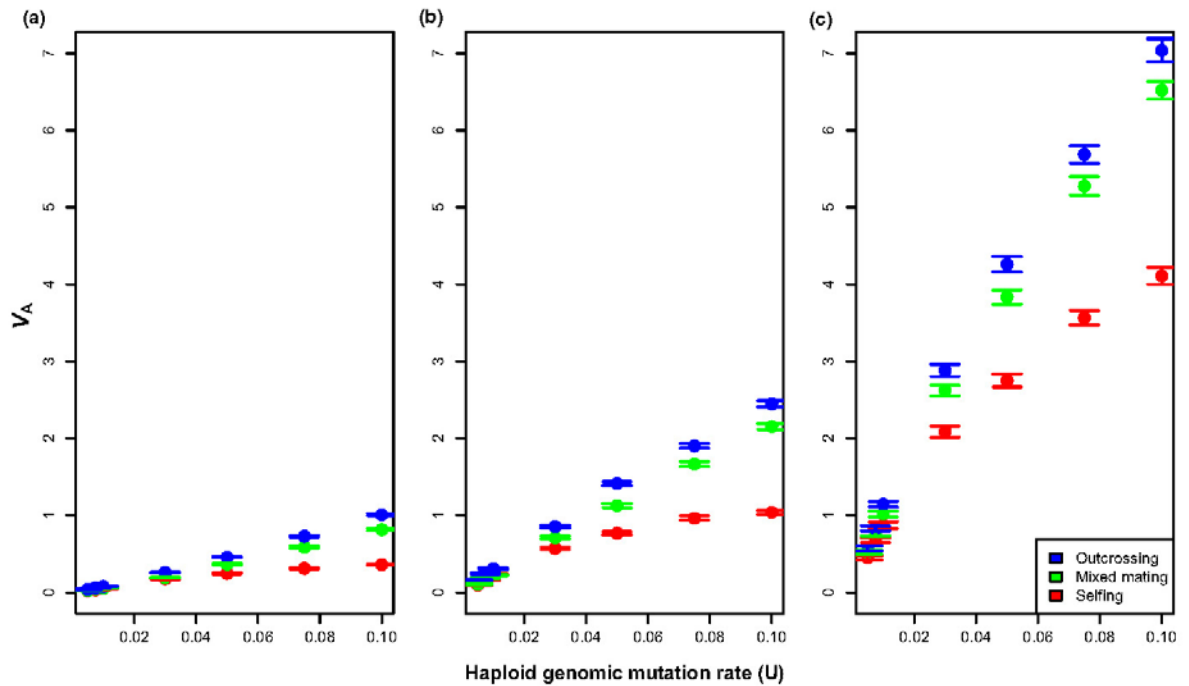


FIGURE 1 Additive genetic variance observed at mutation–selection–drift equilibrium in the simulated populations, for different mating systems, $N = 1,000$, and different strengths of selection (A: $\omega^2 = 1$ (strong selection) B: $\omega^2 = 9$ and C: $\omega^2 = 99$ (weak selection), from the left to the right). Error bars stand for the 95% confidence interval ($n = 100$ simulations)

and the strength of selection weak ($\omega^2 = 99$, equation D23). The lack of fit between our results and approximations from Abu Awad and Roze (2018) is probably due to the small number of loci considered in our simulations ($n = 50$ in our parameter set) and to the fact that we consider only a single trait, both parameters violate the assumptions for these analytical approximations to hold.

By decomposing the observed additive variance V_A into the components described in the Methods above (V_{genetic} , V_{inbred} , COV_{LD}), we find that for the range of parameters we chose to explore, the strong decrease in V_A observed under predominant selfing is mostly driven by a large but negative COV_{LD} (Figure 2). This confirms the expectation that recurrent selfing favours the maintenance of associations among loci. Under such associations, the efficiency of purifying selection is reduced, as the selection pressure on a single mutation is weakened by its positive association with other mutations at other loci, which in turn contributes to higher V_{genetic} and V_{inbred} in selfers compared to outcrossers, as observed in Figure 2. Our analysis also showed that the higher the mutation rate, the more genetic associations are maintained and the weaker the efficiency of purifying selection. The increase of both V_{inbred} and V_{genetic} can, in some cases, overcome the negative effect of COV_{LD} , as is the case for large populations ($N = 10,000$), leading to selfing populations that have close to, and sometimes even greater amounts of genetic variance than outcrossing or mixed-mating populations (Figure S2). As can be expected from classical population genetics models, smaller population size leads to a decrease in V_A , whereas higher mutation rates tend to increase it (Figure S1).

3.2 | Patterns of adaptation through standing genetic variation

After populations have reached mutation–selection–drift balance, an environmental change is induced and no new mutations are introduced. Population dynamics are followed over 20 generations, and if during this time the population reaches a similar level of fitness as that observed before the environmental change, then we consider that the population is able to adapt to the new optimum. When the strength of purifying selection is moderate to weak ($\omega^2 = 9$ or 99), all populations seem to harbour enough genetic variation to quickly and efficiently respond to an environmental change, independently of their selfing rate. As there were no observable dynamic changes in V_A and its components for these parameter values over the course of the first 20 generations following the environmental change, the results for these simulations are only presented in Figures S3–S8. In the following sections, we will only be concentrating on the case of $\omega^2 = 1$, representing the strongest selection pressure we explored for.

3.3 | Dynamics of phenotype and fitness during adaptation under strong selection

As illustrated in Figure 3, when the mutation rate is small ($U = 0.005$), and hence the amount of diversity stored through genetic

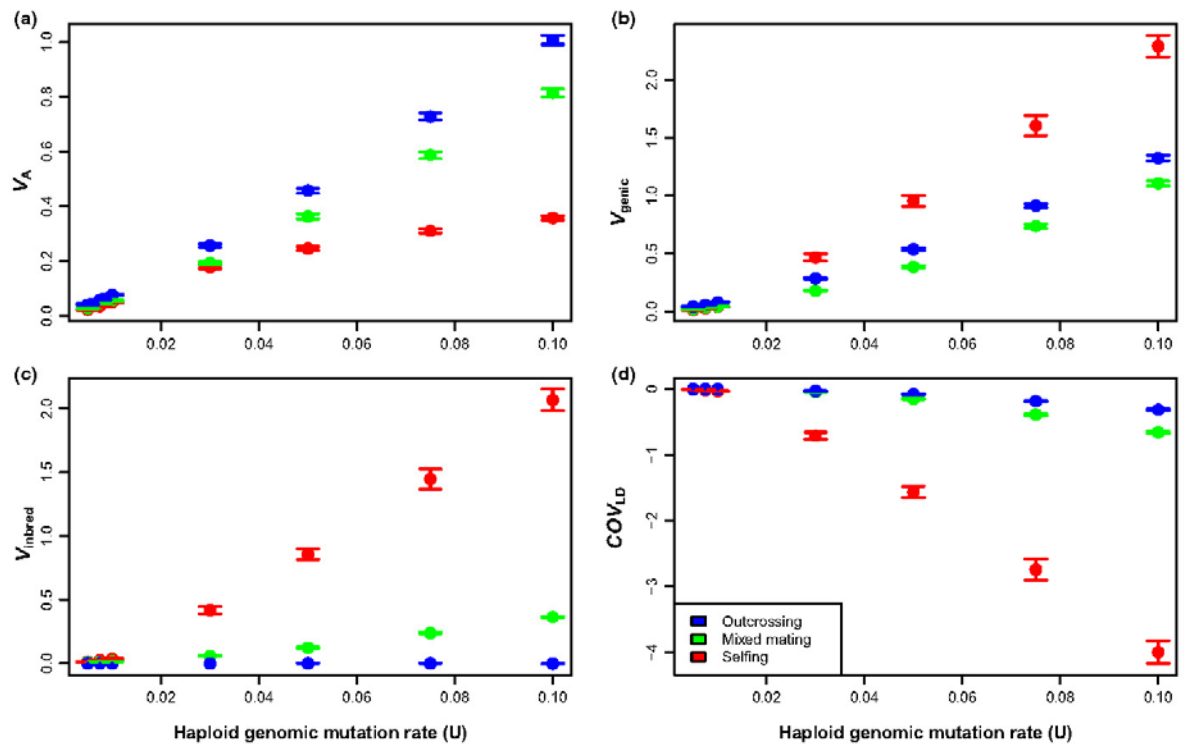


FIGURE 2 Total additive genetic variance and its three components estimated at MSD equilibrium, for different genomic mutation rates and mating systems, for $N = 1,000$ and $\omega^2 = 1$. (a) Total additive variance of the phenotypic trait, identical to Figure 1 first panel. (b) Genic variance for the phenotypic trait (V_{genic}). (c) Genetic variance due to inbreeding (V_{inbred}). (d) Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$)

associations for selfing population is also small, none of the populations are able to fully adapt during the 20 first generations, and population fitness remains low. In this case, the amount of within-loci genetic variation at equilibrium is the only source of diversity, and as outcrossing and mixed-mating populations have higher V_A , they reach phenotypic values closer to the new optimum, and have a significantly higher fitness than their selfing counterparts (Figure 3 for $U = 0.005$). For higher mutation rates ($U = 0.1$), dynamics of the phenotypic trait and fitness are extremely similar for all mating systems, despite lower genetic diversity in selfing populations. As shown in the following sections, when the mutation rate increases, the variance stored in selfing populations through genetic associations is partially released, increasing the adaptive potential of these populations. In such cases, selfing populations are able to reach (a) the new phenotypic optimum, and (b) level of fitness similar to those observed at M-S-D equilibrium, thus performing as well as outcrossing populations (Figure 3).

The above-mentioned patterns are similar irrespective of the population size. As mentioned in the first section of the results, the greater the population size, the greater the genetic diversity (V_A and COV_{LD} for selfing species) for a given mutation rate, increasing the range of mutation rates for which selfing populations are at least as well, if not better, adapted as outcrossing and mixed-mating

populations (Figures S9 and S10). This means that population size has a quantitative but not a qualitative effect.

3.4 | Dynamics of the additive variance and of its components

As the population undergoes directional selection, the dynamics of the observable additive variance and each of its components are telling of the processes driving adaptation to the new optimum. In small outcrossing populations ($N = 250$), the additive variance exhibits very small changes during the adaptation process, as do its components, whatever the strength of selection (Figures S9–S11). For larger population sizes ($N = 1,000$ or $10,000$), the additive variance increases during the first generations when the mutation rate is small to medium ($U \leq 0.05$, Figure 4a for $N = 1,000$, see Supporting information for $N = 10,000$). This increase of V_A stems from an increase in V_{genic} , indicative of the change in the phenotype being driven by an increased frequency of alleles that may have been neutral or deleterious before the environmental change (Bürger, 1993). For high mutation rates ($U = 0.1$), as for small populations, the additive genetic variance and its components remain constant through time (Figure 4). V_{inbred} does not vary (Figure 4c). Overall patterns are similar for mixed-mating populations.

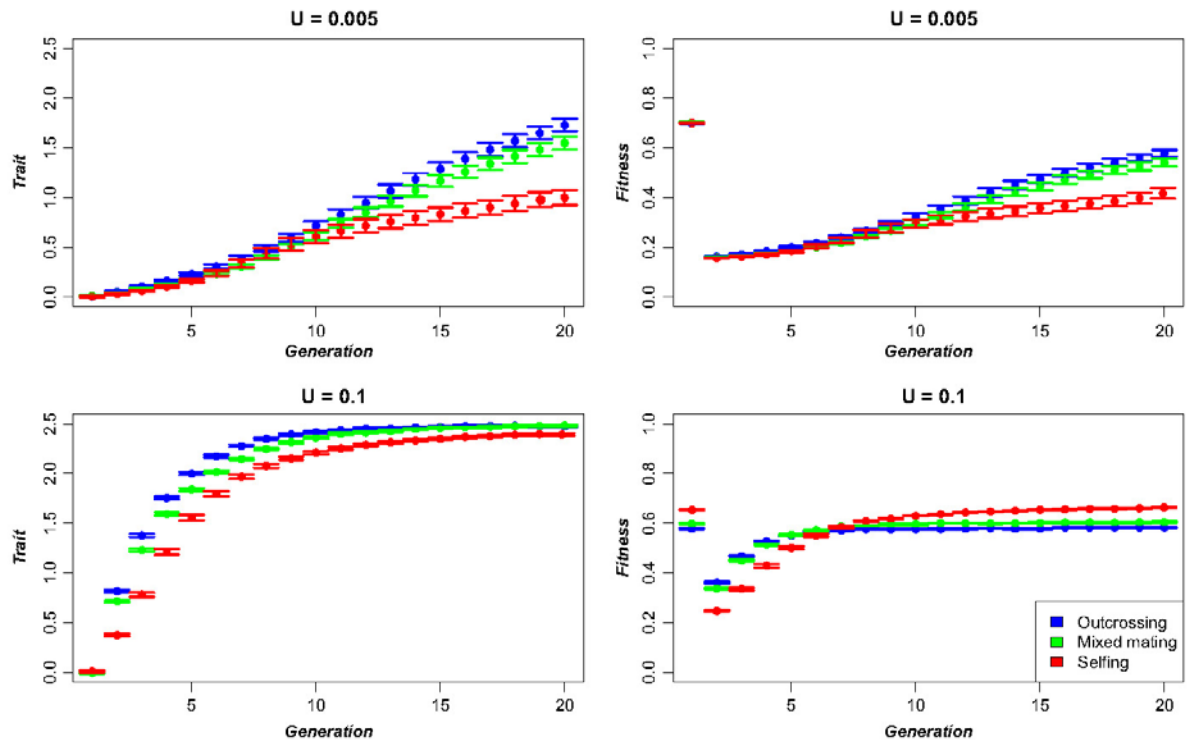


FIGURE 3 Dynamics of the mean trait value and of the mean fitness value of the population during the 20 generations following the environmental change, for different haploid mutation rates (top-down: $U = 0.005$; $U = 0.1$) and as a function of the mating systems. Other parameter values are $N = 1,000$ and $\omega^2 = 1$. Error bars stand for 95% confidence interval ($n = 100$)

For predominantly selfing populations, the initial dynamics strongly depend on 'how much' diversity the population has stored at M-S-D equilibrium (represented by COV_{LD}). In all cases, additive genetic variance increases during the adaptation process (Figure 5a). When genetic associations are rare (Figure 5, $U = 0.005$), adaptation is driven solely through the increase in frequency of beneficial alleles, leading to an increase of the genic variance and inbred variance (Figure 5b,c). In this case, COV_{LD} slightly decreases during the first generations (Figure 5d), probably due to the formation of linkage disequilibrium between alleles rising up in frequency. When among-loci associations are more substantial ($U > 0.005$), the dynamics observed are qualitatively different, and reflect another process of adaptation. The increase in additive variance is due to the 'release' of genetic diversity (COV_{LD} increases, Figure 5d) over the first five generations, whereby the genic and inbred variances remain constant or decrease (Figure 5b,c). The variation of V_{inbred} remains equal to FV_{genic} , indicating that, as observed for outcrossing populations, the inbreeding coefficient does not vary during adaptation.

The initial response to selection (increase in trait mean over the first five generations) tightly correlates with the initial amount of additive variance (Figure 6). Nevertheless, for a theoretical equal amount of additive variance at equilibrium (over all parameter sets combined), selfing populations have a stronger

initial response at the phenotypic level compared to more allogamous ones (Figure 6). This is due the remobilization of the diversity stored by genetic associations during the first few generations, increasing the additive variance and thus the initial response to selection in selfing populations. The larger the population size, the more the pattern observed for selfing populations differs from that observed in outcrossing and mixed-mating ones (Figure 6). When selection is moderate to weak, the release of the stored genetic variation is not necessary for selfing populations to adapt; thus, this increase in the strength of adaptation for selfers is no longer observed.

4 | DISCUSSION

In accordance with Stebbins' definition of the dead-end hypothesis (Stebbins, 1957), single-locus models predict that adaptation is less likely in selfing populations compared to outcrossing ones, notably due to the reduced standing genetic variation resulting from purging (Glémin & Ronfort, 2013). Considering a polygenic trait, and among-loci associations, we find that this is not always the case. Previous works have highlighted that stabilizing selection favours the build-up of associations between loci and, through this, favours the storage of genetic diversity especially in self-fertilizing

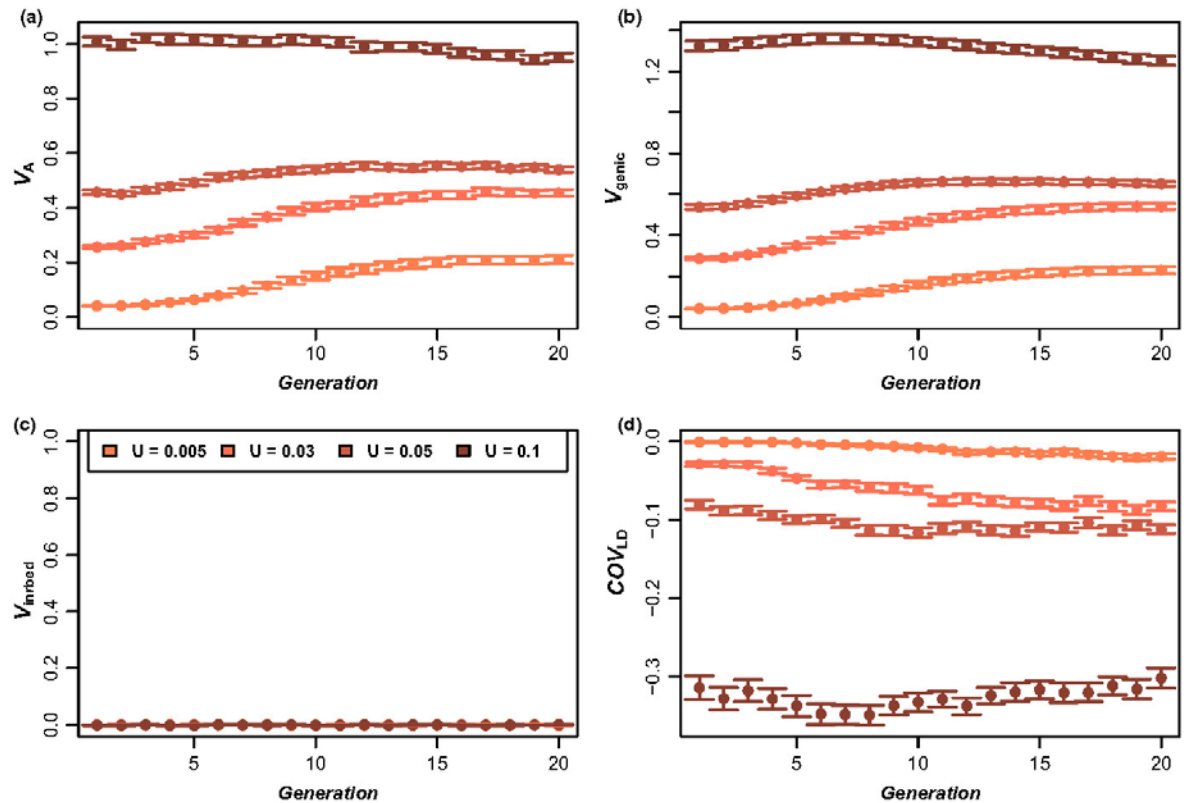


FIGURE 4 Dynamics of the additive genetic variance and its components in outcrossing populations ($\sigma = 0$), mutation rates are set at zero after generation 0. (a) Observed additive variance for the phenotypic trait. (b) Genic variance for the phenotypic trait (V_{genic}). (c) Genetic variance due to inbreeding (V_{inbred}). (d) Genetic covariance due to linkage disequilibrium (COV_{LD}), for different mutation rates ($U = 0.005$; $U = 0.03$; $U = 0.05$; $U = 0.1$). Error bars stand for 95% confidence interval ($n = 100$). Other parameter values are $N = 1,000$ and $\omega^2 = 1$

populations (Abu Awad & Roze, 2018; Charlesworth, 1990; Lande & Porcher, 2015). If the selective regime were to change from stabilizing to directional selection, due to a strong environmental effect, for example, our simulations show that some of this variance can be released through residual allogamy. Residual allogamy is primordial, as selfing populations are organized in multi-locus genotypes. The release of the diversity stored through genetic associations is only possible through rare outcrossing events between two lineages, resulting in fully heterozygote hybrids, which can potentially generate D^3 new genotypes (D being the number of differentiated loci among two selfing lineages) after one generation of selfing. Although the strength of the response to selection is positively correlated with the amount of observable standing genetic variation, it is possible to observe similar levels of adaptation in selfing compared to outcrossing or mixed-mating populations, for high but realistic mutation rates (Shaw et al., 2002). When the strength of selection is moderate to weak ($\omega^2 = 9$ or 99), populations seem to harbour enough genetic variation to quickly and efficiently respond to an environmental change, independently of the self-fertilization rate, and thus without remobilizing any variance accumulated through genetic associations.

4.1 | Genetic variance dynamics under directional selection: the role of genetic associations

Few theoretical studies have tried to predict the dynamic of the additive variance for a trait under directional selection (Barton & Turelli, 1987; Bürger, 1993; Keightley & Hill, 1989). From these works, under the assumption of random mating and no association between loci, it is predicted that V_A remains unchanged in the short term if population size is small ($N < 500$, Bürger, 1993) and/or if the inverse of the strength of selection, given by the term ω^2 , is large compared to the amount of available genetic variance (Barton & Turelli, 1987). This prediction is supported by our results for obligate outcrossing populations (Figures S11–S15), even when some negative linkage disequilibrium contributes significantly to additive variance (Figures 2 and 4). For selfing populations, the hypothesis of linkage equilibrium is no longer valid. As shown in our analysis and known from previous results (Abu Awad & Roze, 2018; Lande & Porcher, 2015), genetic associations between loci are significant in these populations, notably when the mutation rate is high. In this case, the breakup of genetic associations is necessary for adaptation when most of the genetic variance is hidden in among-loci associations, leading to an increase

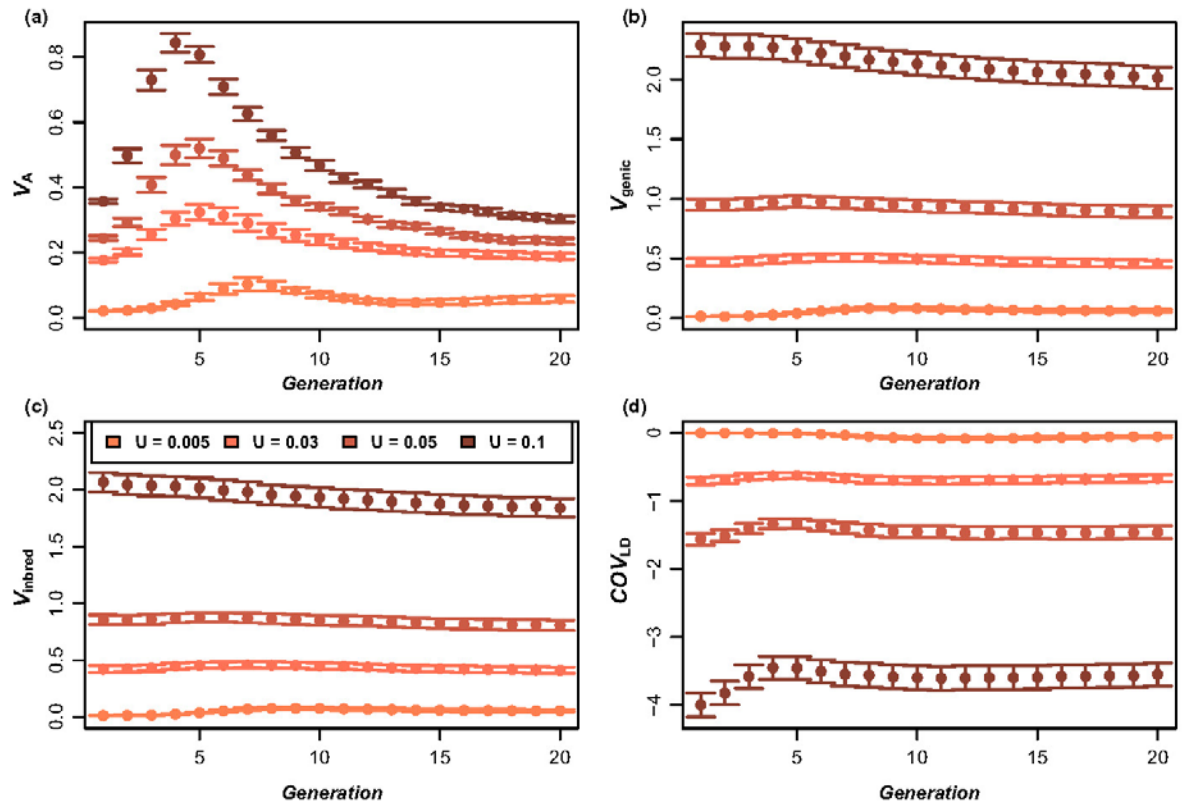


FIGURE 5 Dynamics of the additive genetic variance and its components in selfing populations ($\sigma = 0.95$), mutation rates are set at zero after generation 0. (a) Observed additive variance for the phenotypic trait. (b) Genic variance for the phenotypic trait (V_{genic}). (c) Genetic variance due to inbreeding (V_{inbred}). (d) Genetic covariance due to linkage disequilibrium (COV_{LD}), for different mutation rates ($U = 0.005$; $U = 0.03$; $U = 0.05$; $U = 0.1$). Error bars stand for 95% confidence interval ($n = 100$). Other parameter values are $N = 1,000$ and $\omega^2 = 1$

of observed additive variance. The mutation rates for which such associations contribute significantly to the genetic additive variance are in the range of mutation rates observed for phenotypic traits in the literature (Haag-Liautard et al., 2007; Keightley & Bataillon, 2000; Shaw et al., 2002), which suggests that these associations are probably common in natural populations. The more frequent observations of transgressive segregation (progeny of a cross being outside the parental phenotypic range) in inbred compared to outbred species also give support to this mechanism (Johansen-Morris & Latta, 2006; Rieseberg, Archer, & Wayne, 1999).

4.2 | De novo mutations versus standing genetic variation: rethinking adaptation in selfing species?

It has been a long accepted paradigm that the advantage procured by selfing is the more rapid fixation of de novo beneficial mutations, independently of their dominance coefficient, compared to outcrossing populations, where recessive beneficial mutations can be lost through drift before selection drives them to fixation, a process known as 'Haldane's sieve' (Haldane, 1927). From single-locus

theory, it is expected that adaptation through new mutations should be more likely in selfing species and should be more likely than adaptation from standing genetic variation (Glémin & Ronfort, 2013). However, recent works have suggested that the reduced effective recombination rate in selfing populations adds a disadvantage even when it comes to the fixation of new mutations. Unlike what is expected in outcrossing populations, the fixation of beneficial mutations in selfing populations can be hindered if they appear during a selective sweep triggered by a beneficial allele at another locus (Hartfield & Glémin, 2016). This observation, as well as the results presented here, shows that taking interactions between loci into account can strongly modify the expectations from single-locus models. The mutation rate seems to be the major parameter to consider when studying the dynamics of adaptation from a polygenic point of view, independently of whether adaptation occurs from de novo mutations (Hartfield & Glémin, 2016) or standing genetic variation.

In our simulations, we have considered a quantitative trait with a simple architecture, assuming pure additivity of genetic effects on the phenotype. Epistasis, and notably its directionality, is known to play a key role in adaptation (Hansen, 2013). Positive epistasis (several loci reinforcing each other's effects in the direction of selection)

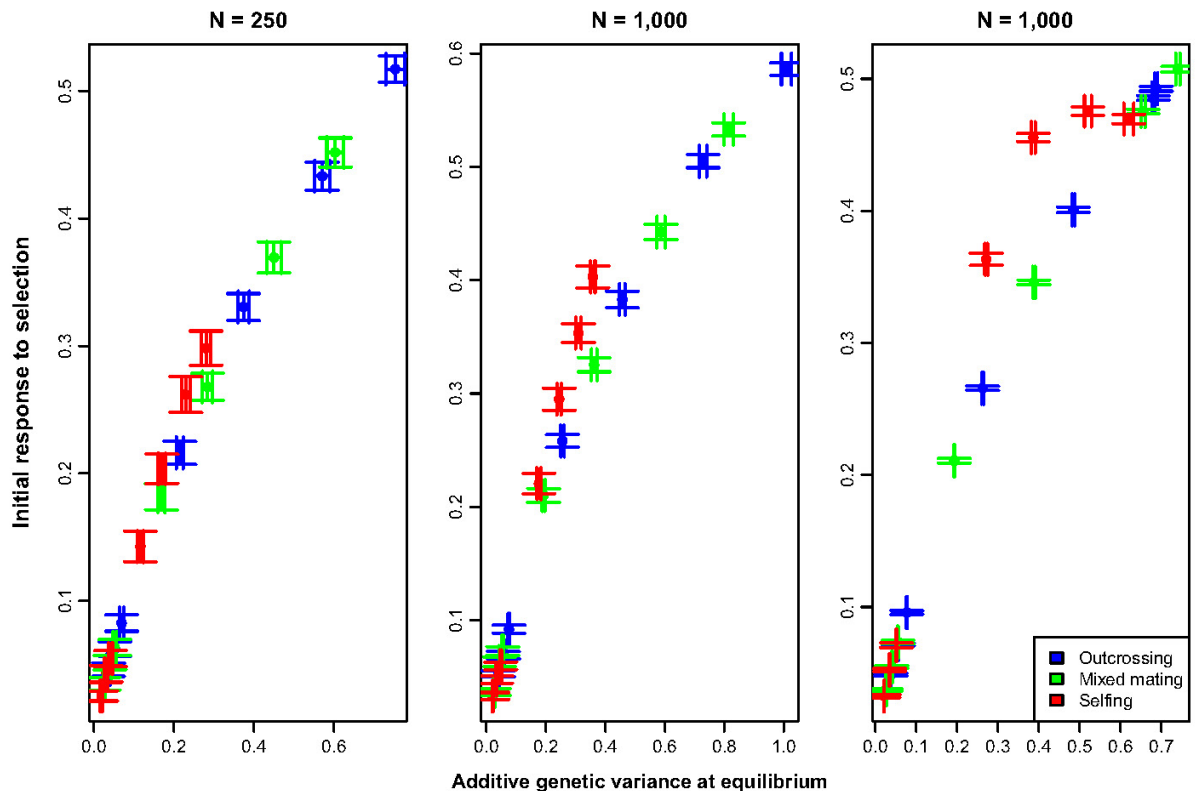


FIGURE 6 Initial response to selection in function of the amount of additive variance at equilibrium, as a function of the mating system, and for the three population sizes. Each point correspond to a haplotypic mutation rate (U ranging from 0.005 to 0.1, from the left to the right). Selection is strong ($\omega^2 = 1$). Error bars stand for 95% confidence interval ($n = 100$)

inflates the additive variance and thus the ability of a population to adapt to an environmental change (Carter, Hermisson, & Hansen, 2005; Monnahan & Kelly, 2015). On the contrary, negative epistasis (loci suppressing effects at other loci) reduces the additive variance of the character, thus limiting adaptive potential (Carter et al., 2005). Few empirical estimations of the directionality of epistasis are available in the literature (Le Rouzic, 2014; Monnahan & Kelly, 2015; Oakley, Ågren, & Schemske, 2015, all detecting positive epistatic interactions), despite numerous methods and the diversity of data used to infer it (Le Rouzic, 2014). How the directionality of epistatic interactions varies in relation to the mating system remains unknown in natural populations. Knowledge on this relationship may bring us closer to the understanding of the differences in patterns of adaptation observed between selfing and outcrossing populations.

4.3 | New insights into the role of standing genetic variation in the adaptation dynamics of selfing populations

The overwhelming success of selfing species in the domestication process and as invasive species has been attributed to mechanisms

other than their adaptive ability from standing genetic variation. For instance, the invasive success of selfing populations is attributed to their reproductive assurance, since a single individual is able to colonize a new environment (van Kleunen, Manning, Pasqualetto, & Johnson, 2008; Rambuda & Johnson, 2004), and to reduced gene flow, which is expected to limit maladapted gene exchanges between populations (Levin, 2010). Regarding domestication as an adaptation process, it has been suggested that domestication in selfing populations most probably relied on new mutations, due to the initially low genetic variance that would have been further reduced due to the bottleneck effect of domestication (Glémin & Bataillon, 2009). This idea is reinforced by the fact that selfing species are expected to quickly fix a rare beneficial mutation, independently of its dominance level (Ross-Ibarra, 2005). In their review on mating system variation in domesticated plant species, Glémin and Bataillon (2009) have suggested that the high frequency of self-fertilizing crop species could be related to an increase in the amount of additive variance during the domestication process of selection. This idea has, however, never been tested theoretically or empirically. Here we show that this increase in additive variance could indeed be an advantage when selfing species are faced with strong directional selection and selection occurs through standing genetic variation.

However, our results hold true only if the bottleneck related to domestication (or other invasion processes) is not too strong and if mutation rates are high enough to maintain enough diversity that can promote adaptation to new conditions.

4.4 | CONCLUSION AND PERSPECTIVES

In this work, we show that under stabilizing selection and if mutation rates are high enough (but realistic), selfing populations are able to accumulate genetic variation through negative linkage disequilibrium. This structure of genetic variation contributes to acting as a store of genetic variability, thanks to which adaptation under high self-fertilization rates is no longer constrained by the reduced observable additive genetic variance due to purging. For low mutation rates, the storage of diversity is impeded by stabilizing selection in selfing populations, leading to the more classical conclusion that adaptation in selfing species is limited by the available amount of genetic diversity. Our analysis thus shows that measuring the amount of additive variance available for quantitative traits is not always sufficient to make predictions about the adaptive potential of a population. Complementary analyses should be carried out when quantifying the short-term adaptability of a population. Such analyses could include looking for transgressive segregations or carrying out experimental evolutionary experiments in which directional selection is induced, and follow the dynamic of additive genetic variance. Following the components of additive variance is complicated, as it necessitates thorough knowledge of the genetic architecture of the quantitative trait under study (number of loci underlying phenotypic variation, their frequency, the linkage disequilibrium...), which is complicated, even for model species. More empirical evidence is required to determine how frequently diversity is stored through genetic associations in natural populations of selfing species and whether this property is sufficient to allow selfing species to adapt to a changing environment.

ACKNOWLEDGMENT

We would like to thank D. Roze for helpful discussions about the analyses and interpretations of the results, and E. Noël, P. David, O. Ronce, S. Glémin and C. Devaux for insightful discussions on the results. We also thank editors and reviewers for improving the manuscript. This work has been conducted with the help of the data and calculation centre South Green of the CIRAD-UMR AGAP. This work was supported by project SEAD (ANR-13-ADAP-0011), the TUM University Foundation Fellowship and the Alexander von-Humboldt Foundation.

ORCID

Josselin CLO  <https://orcid.org/0000-0002-3295-9481>

Diala Abu Awad  <https://orcid.org/0000-0002-2680-1223>

REFERENCES

- Abu Awad, D., & Roze, D. (2018). Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution*, 72, 751–769.

- Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I., & Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, 22, 699–708. <https://doi.org/10.1111/j.1365-294X.2012.05522.x>
- Badano, J. L., & Katsanis, N. (2002). Human genetics and disease: Beyond Mendel: An evolving view of human genetic disease transmission. *Nature Reviews Genetics*, 3, 779–789. <https://doi.org/10.1038/nrg910>
- Barrett, R. D., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, 23, 38–44. <https://doi.org/10.1016/j.tree.2007.09.008>
- Barton, N. H., & Turelli, M. (1987). Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genetical Research*, 49, 157–173.
- Berger, D., Bauerfeind, S. S., Blanckenhorn, W. U., & Schäfer, M. A. (2011). High temperatures reveal cryptic genetic variation in a polymorphic female sperm storage organ. *Evolution*, 65, 2830–2842. <https://doi.org/10.1111/j.1558-5646.2011.01392.x>
- Bürger, R. (1993). Predictions of the dynamics of a polygenic character under directional selection. *Journal of Theoretical Biology*, 162, 487–513. <https://doi.org/10.1006/jtbi.1993.1101>
- 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. <https://doi.org/10.1111/j.1558-5646.1989.tb02624.x>
- Carter, A. J., Hermisson, J., & Hansen, T. F. (2005). The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theoretical Population Biology*, 68, 179–196. <https://doi.org/10.1016/j.tpb.2005.05.002>
- Charlesworth, B. (1990). Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genetical Research*, 55, 199–221. <https://doi.org/10.1017/S0016672300025532>
- Charlesworth, D., & Charlesworth, B. (1995). Quantitative genetics in plants: The effect of the breeding system on genetic variability. *Evolution*, 49, 911–920. <https://doi.org/10.1111/j.1558-5646.1995.tb02326.x>
- Clo, J., Gay, L., & Ronfort, J. (2019). How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species. *Evolution*, 73(8), 1578–1590. <https://doi.org/10.1111/evo.13789>
- Geber, M. A., & Griffen, L. R. (2003). Inheritance and natural selection on functional traits. *International Journal of Plant Sciences*, 164, S21–S42. <https://doi.org/10.1086/368233>
- Gibson, G., & Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics*, 5, 681–690. <https://doi.org/10.1038/nrg1426>
- Glémin, S., & Bataillon, T. (2009). A comparative view of the evolution of grasses under domestication. *New Phytologist*, 183, 273–290. <https://doi.org/10.1111/j.1469-8137.2009.02884.x>
- Glémin, S., & Ronfort, J. (2013). Adaptation and maladaptation in selfing and outcrossing species: New mutations versus standing variation. *Evolution*, 67, 225–240. <https://doi.org/10.1111/j.1558-5646.2012.01778.x>
- Haag-Liautard, C., Dorris, M., Maside, X., Macaskill, S., Halligan, D. L., Charlesworth, B., & Keightley, P. D. (2007). Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature*, 445, 82–85. <https://doi.org/10.1038/nature05388>
- Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23, 607–615. <https://doi.org/10.1017/S0305004100011750>
- Hansen, T. F. (2013). Why epistasis is important for selection and adaptation. *Evolution*, 67, 3501–3511. <https://doi.org/10.1111/evo.12214>
- Hartfield, M., & Glémin, S. (2016). Limits to adaptation in partially selfing species. *Genetics*, 203, 959–974. <https://doi.org/10.1534/genetics.116.188821>

- Hermisson, J., & Pennings, P. S. (2005). Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, 169, 2335–2352. <https://doi.org/10.1534/genetics.104.036947>
- Igic, B., & Busch, J. W. (2013). Is self-fertilization an evolutionary dead end? *New Phytologist*, 198, 386–397. <https://doi.org/10.1111/nph.12182>
- Innan, H., & Kim, Y. (2004). Pattern of polymorphism after strong artificial selection in a domestication event. *Proceedings of the National Academy of Sciences*, 101, 10667–10672. <https://doi.org/10.1073/pnas.0401720101>
- Johansen-Morris, A. D., & Latta, R. G. (2006). Fitness consequences of hybridization between ecotypes of *Avena barbata*: Hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution*, 60, 1585–1595.
- Kamran-Disfani, A., & Agrawal, A. F. (2014). Selfing, adaptation and background selection in finite populations. *Journal of Evolutionary Biology*, 27, 1360–1371.
- Keightley, P. D., & Bataillon, T. M. (2000). Multigeneration maximum-likelihood analysis applied to mutation-accumulation experiments in *Caenorhabditis elegans*. *Genetics*, 154, 1193.
- Keightley, P. D., & Hill, W. G. (1989). Quantitative genetic variability maintained by mutation-stabilizing selection balance: Sampling variation and response to subsequent directional selection. *Genetical Research*, 54, 45–58. <https://doi.org/10.1017/S0016672300028366>
- Kim, S.-Y., Noguera, J. C., Tato, A., & Velando, A. (2013). Vitamins, stress and growth: The availability of antioxidants in early life influences the expression of cryptic genetic variation. *Journal of Evolutionary Biology*, 26, 1341–1352.
- Lande, R., & Porcher, E. (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics*, 200, 891–906.
- Le Rouzic, A. (2014). Estimating directional epistasis. *Frontiers in Genetics*, 5, 198. <https://doi.org/10.3389/fgene.2014.00198>
- Le Rouzic, A., & Carlborg, Ö. (2008). Evolutionary potential of hidden genetic variation. *Trends in Ecology & Evolution*, 23, 33–37. <https://doi.org/10.1016/j.tree.2007.09.014>
- Ledon-Rettig, C. C., Pfennig, D. W., & Crespi, E. J. (2010). Diet and hormonal manipulation reveal cryptic genetic variation: Implications for the evolution of novel feeding strategies. *Proceedings of the Royal Society B-Biological Sciences*, 277, 3569–3578. <https://doi.org/10.1098/rspb.2010.0877>
- Levin, D. A. (2010). Environment-enhanced self-fertilization: Implications for niche shifts in adjacent populations. *Journal of Ecology*, 98, 1276–1283. <https://doi.org/10.1111/j.1365-2745.2010.01715.x>
- Manna, F., Martin, G., & Lenormand, T. (2011). Fitness landscapes: An alternative theory for the dominance of mutation. *Genetics*, 189, 923–937. <https://doi.org/10.1534/genetics.111.132944>
- Martin, G., Elena, S. F., & Lenormand, T. (2007). Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nature Genetics*, 39, 555–560. <https://doi.org/10.1038/ng1998>
- Martin, G., & Lenormand, T. (2006). A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution*, 60, 893–907. <https://doi.org/10.1111/j.0014-3820.2006.tb01169.x>
- Matuszewski, S., Hermisson, J., & Kopp, M. (2015). Catch me if you can: Adaptation from standing genetic variation to a moving phenotypic optimum. *Genetics*, 200, 1255–1274. <https://doi.org/10.1534/genetics.115.178574>
- McGuigan, K., Nishimura, N., Currey, M., Hurwit, D., & Cresko, W. A. (2011). Cryptic genetic variation and body size evolution in threespine stickleback. *Evolution*, 65, 1203–1211. <https://doi.org/10.1111/j.1558-5646.2010.01195.x>
- Milloz, J., Duveau, F., Nuez, I., & Félix, M.-A. (2008). Intraspecific evolution of the intercellular signaling network underlying a robust developmental system. *Genes & Development*, 22, 3064–3075. <https://doi.org/10.1101/gad.495308>
- Mittell, E. A., Nakagawa, S., & Hadfield, J. D. (2015). Are molecular markers useful predictors of adaptive potential? *Ecology Letters*, 18, 772–778. <https://doi.org/10.1111/ele.12454>
- Monnahan, P. J., & Kelly, J. K. (2015). Epistasis is a major determinant of the additive genetic variance in *Mimulus guttatus*. *PLoS Genetics*, 11, e1005201. <https://doi.org/10.1371/journal.pgen.1005201>
- Nordborg, M. (2000). Linkage disequilibrium, gene trees and selfing: An ancestral recombination graph with partial self-fertilization. *Genetics*, 154, 923–929.
- Oakley, C. G., Ågren, J., & Schemske, D. W. (2015). Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. *Heredity*, 115, 73–82. <https://doi.org/10.1038/hdy.2015.18>
- Orr, H. A., & Betancourt, A. J. (2001). Haldane's sieve and adaptation from the standing genetic variation. *Genetics*, 157, 875–884.
- Paaby, A. B., & Rockman, M. V. (2014). Cryptic genetic variation: Evolution's hidden substrate. *Nature Reviews Genetics*, 15, 247. <https://doi.org/10.1038/nrg3688>
- Pritchard, J. K., Pickrell, J. K., & Coop, G. (2010). The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, 20, R208–R215.
- Queitsch, C., Sangster, T. A., & Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature*, 417, 618–624. <https://doi.org/10.1038/nature749>
- Rambuda, T. D., & Johnson, S. D. (2004). Breeding systems of invasive alien plants in South Africa: Does Baker's rule apply? *Diversity and Distributions*, 10, 409–416. <https://doi.org/10.1111/j.1366-9516.2004.00100.x>
- Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity*, 83, 363–372.
- Rohner, N., Jarosz, D. F., Kowalko, J. E., Yoshizawa, M., Jeffery, W. R., Borowsky, R. L., ... Tabin, C. J. (2013). Cryptic variation in morphological evolution: HSP90 as a capacitor for loss of eyes in cavefish. *Science*, 342, 1372–1375. <https://doi.org/10.1126/science.1240276>
- Ronce, O., Shaw, F. H., Rousset, F., & Shaw, R. G. (2009). Is inbreeding depression lower in maladapted populations? A quantitative genetics model. *Evolution*, 63, 1807–1819. <https://doi.org/10.1111/j.1558-5646.2009.00678.x>
- Ross-Ibarra, J. (2005). Quantitative trait loci and the study of plant domestication. In R. Mauricio (Ed.), *Genetics of adaptation* (pp. 197–204). Dordrecht, the Netherlands: Springer.
- Shao, H., Burrage, L. C., Sinasac, D. S., Hill, A. E., Ernest, S. R., O'Brien, W., ... Nadeau, J. H. (2008). Genetic architecture of complex traits: Large phenotypic effects and pervasive epistasis. *Proceedings of the National Academy of Sciences*, 105, 19910–19914. <https://doi.org/10.1073/pnas.0810388105>
- Shaw, F. H., Geyer, C. J., & Shaw, R. G. (2002). A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution*, 56, 453–463. <https://doi.org/10.1111/j.0014-3820.2002.tb01358.x>
- Stebbins, G. L. (1957). Self fertilization and population variability in the higher plants. *American Naturalist*, 91, 337–354. <https://doi.org/10.1086/281999>
- Takebayashi, N., & Morrell, P. L. (2001). Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany*, 88, 1143–1150. <https://doi.org/10.2307/3558325>
- Turelli, M., & Barton, N. H. (1990). Dynamics of polygenic characters under selection. *Theoretical Population Biology*, 38, 1–57. [https://doi.org/10.1016/0040-5809\(90\)90002-D](https://doi.org/10.1016/0040-5809(90)90002-D)
- van Kleunen, M., Manning, J. C., Pasqualetto, V., & Johnson, S. D. (2008). Phylogenetically independent associations between autonomous

self-fertilization and plant invasiveness. *American Naturalist*, 171, 195–201. <https://doi.org/10.1086/525057>

Wood, J. L., Yates, M. C., & Fraser, D. J. (2016). Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evolutionary Applications*, 9, 640–657. <https://doi.org/10.1111/eva.12375>

How to cite this article: CLO J, Ronfort J, Abu Awad D.

Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. *J Evol Biol*. 2020;00:1–13. <https://doi.org/10.1111/jeb.13660>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Matériel supplémentaire

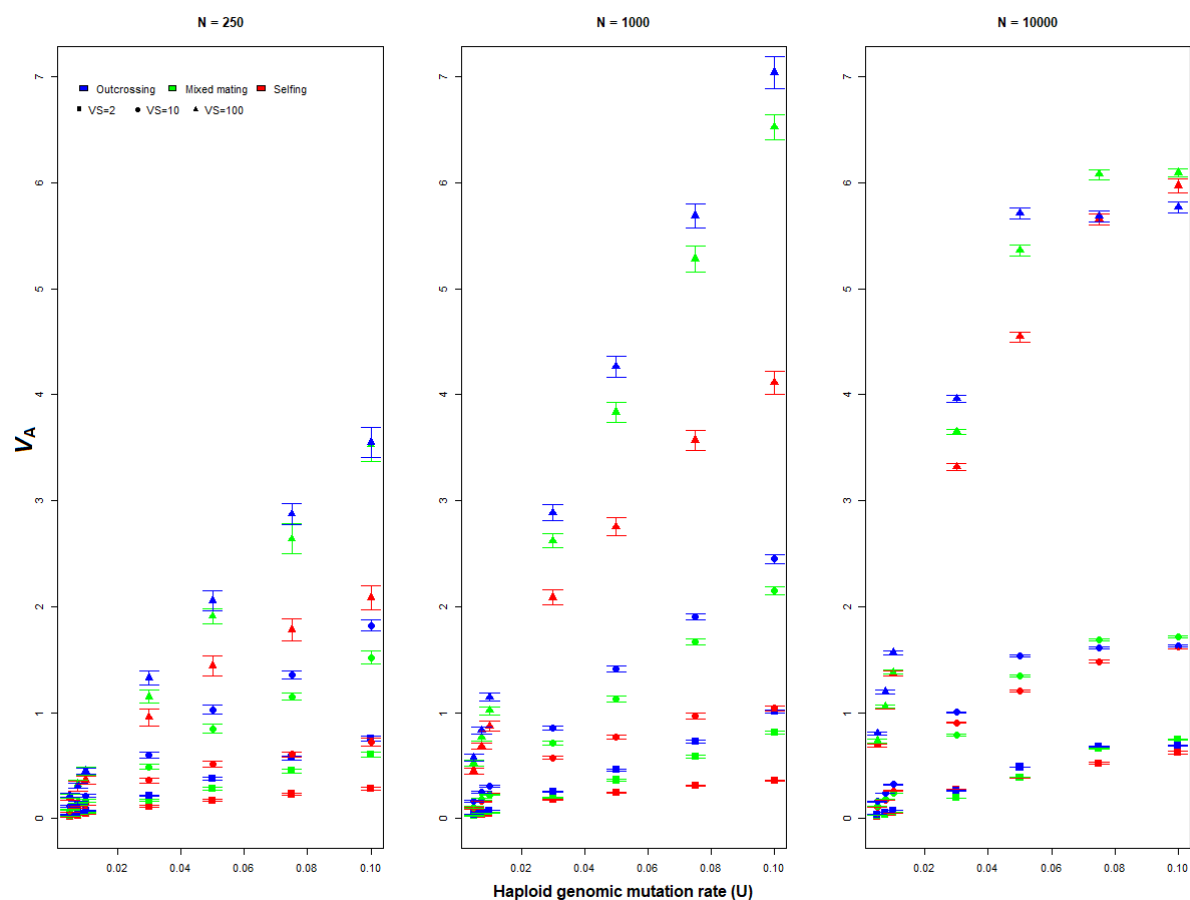


Figure S1. Additive genetic variance maintained at mutation-selection-drift equilibrium, in function of the mating system, the population size, and the strength of stabilizing selection. Error bars stand for 95% confidence interval ($n = 100$).

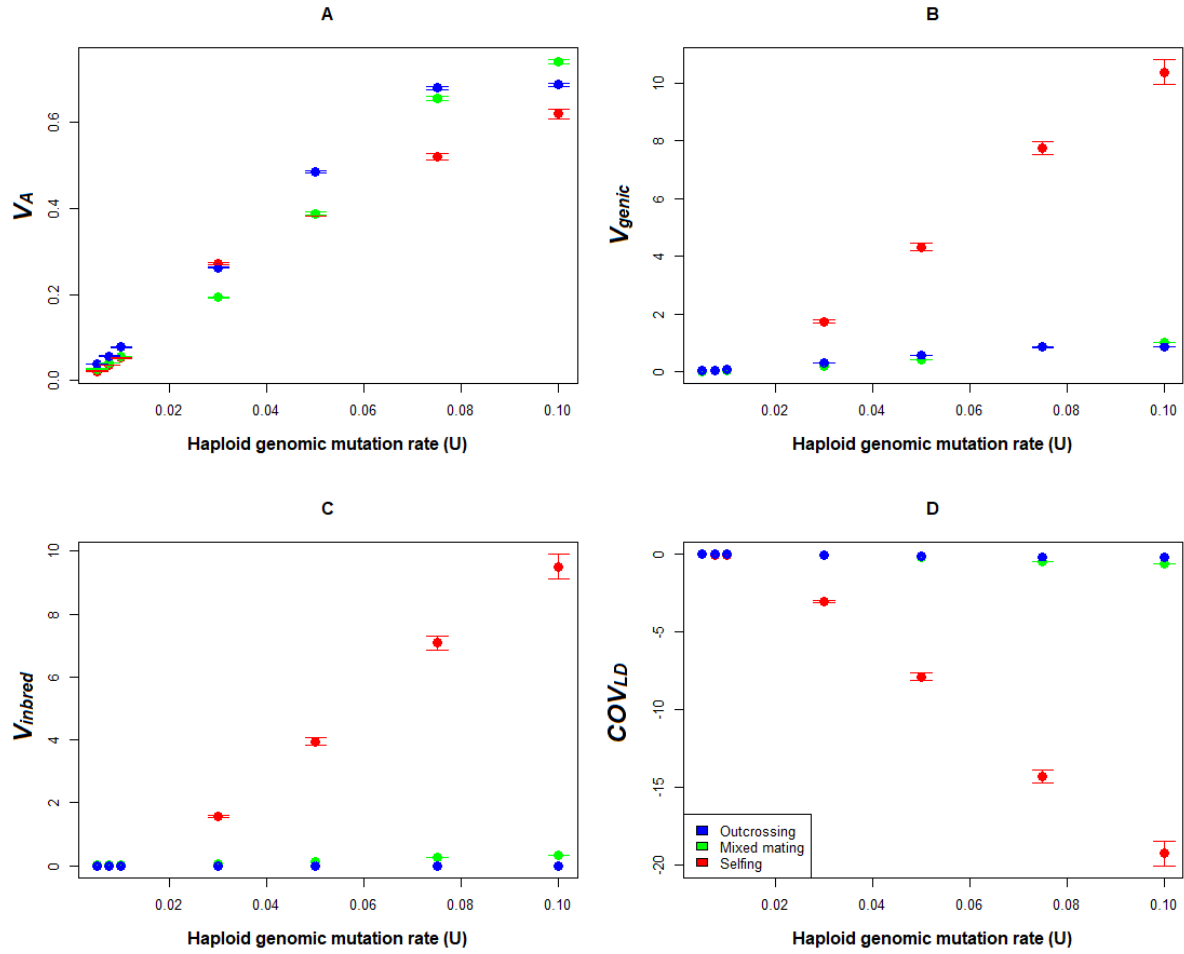


Figure S2. Additive genetic variance and its components as a function of the genomic mutation rate and the mating system, for $N=10,000$ and $\omega^2=1$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval (n = 100).

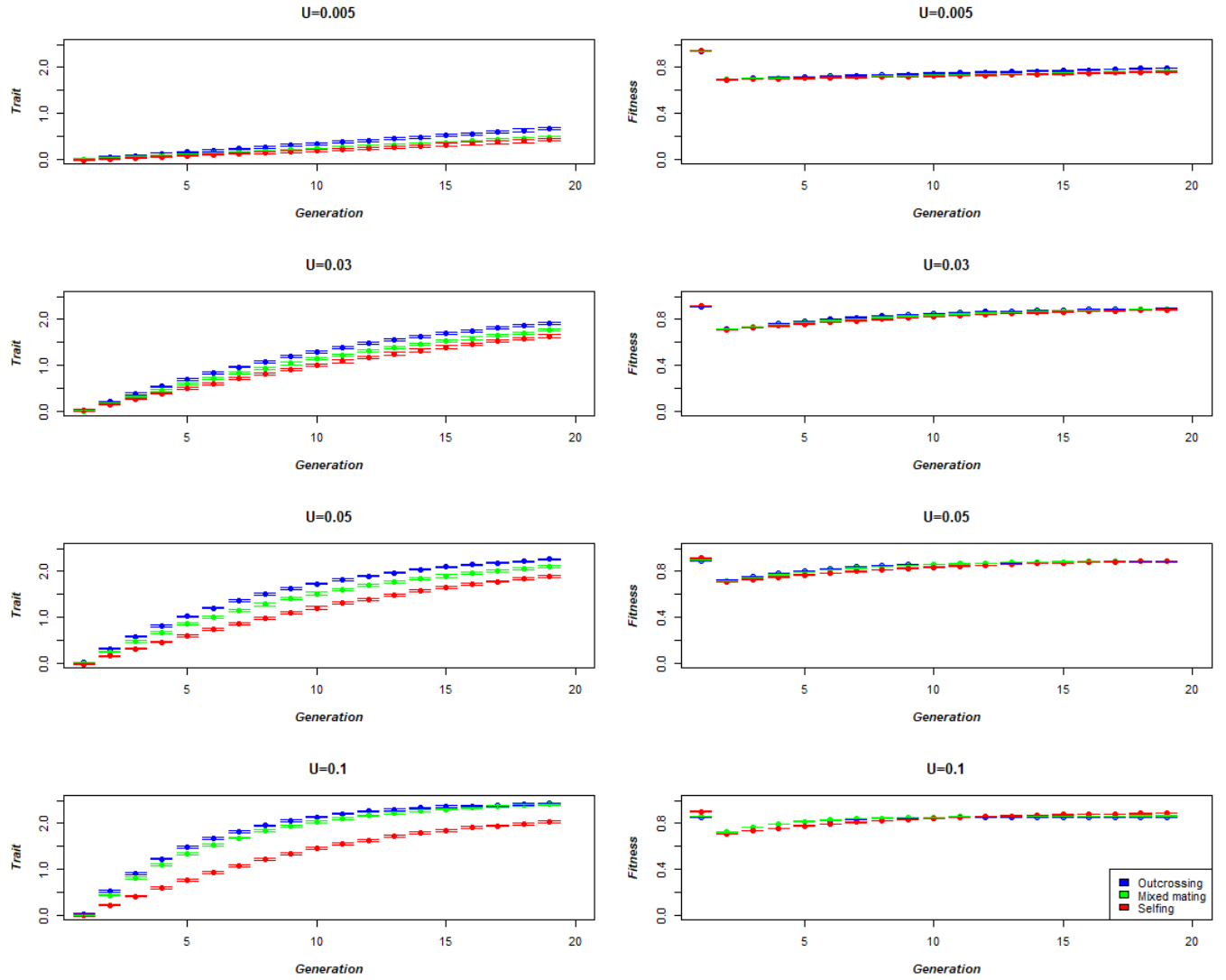


Figure S3. Dynamics of the trait and fitness, as a function of the trait haploid mutation rate and the mating system, for $N=1000$ and $\omega^2=9$. Error bars stand for 95% confidence interval ($n=100$).

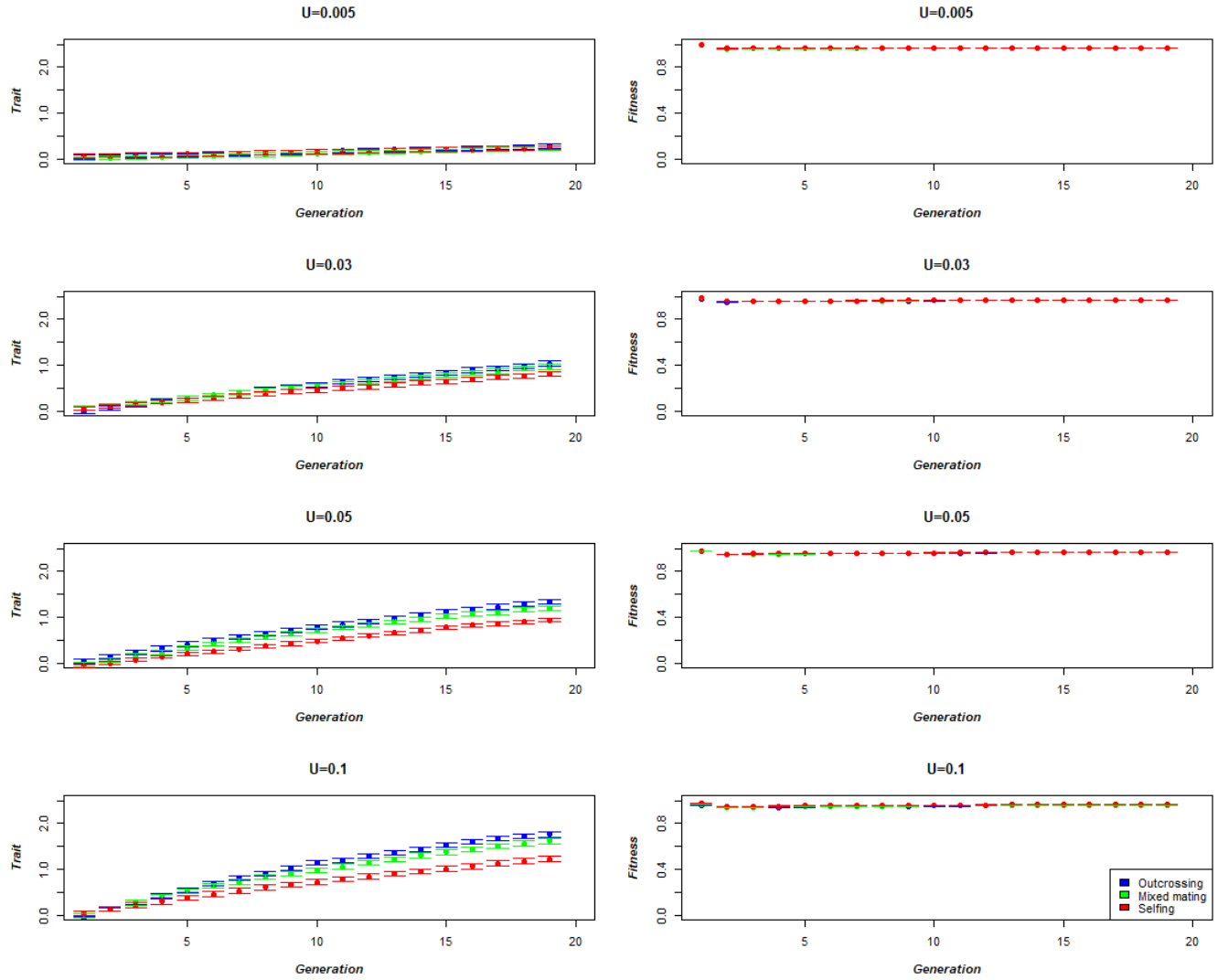


Figure S4. Dynamics of the trait and fitness, as a function of the trait haploid mutation rate and the mating system, for $N=1000$ and $\omega^2=99$. Error bars stand for 95% confidence interval ($n=100$).

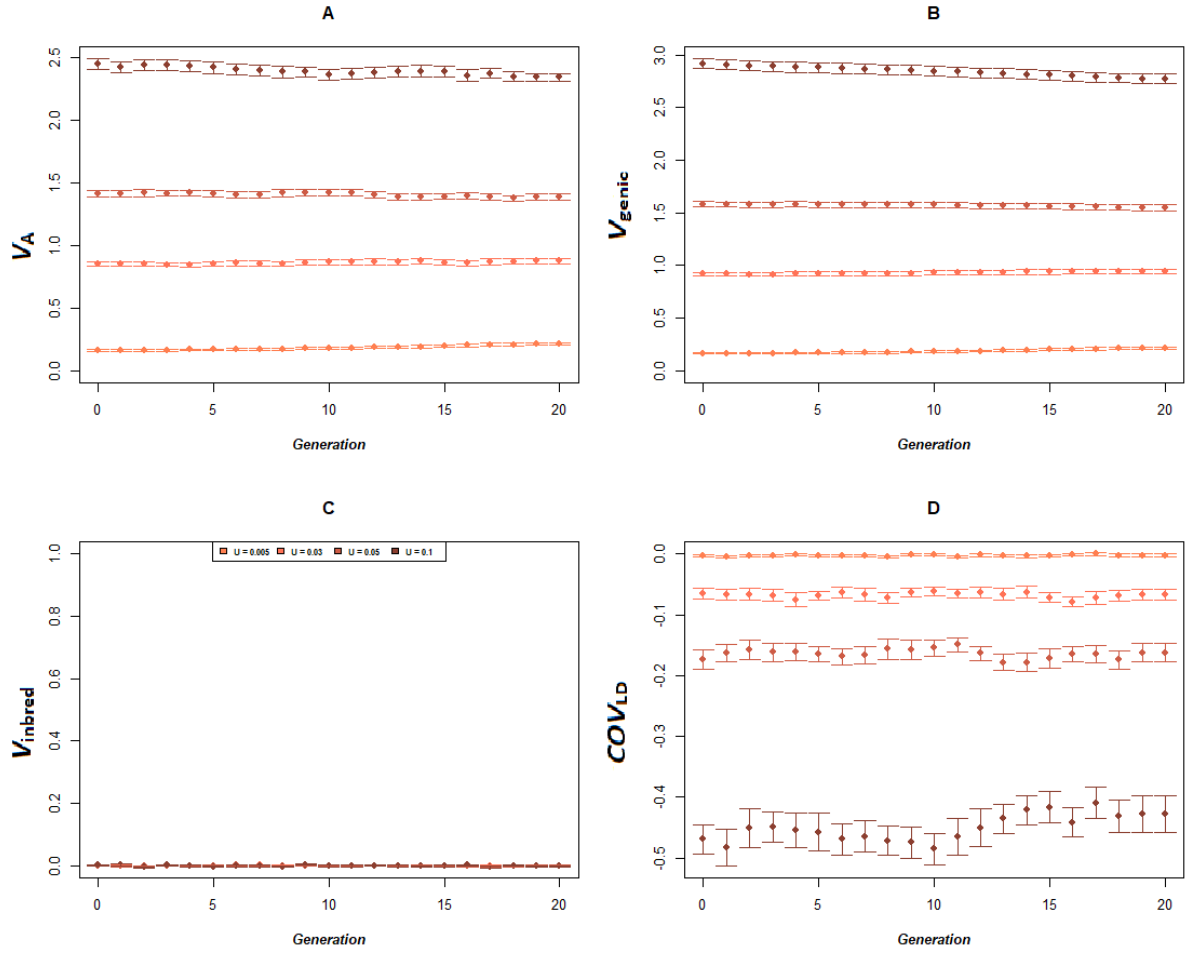


Figure S5. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=1000$ and $\omega^2=9$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

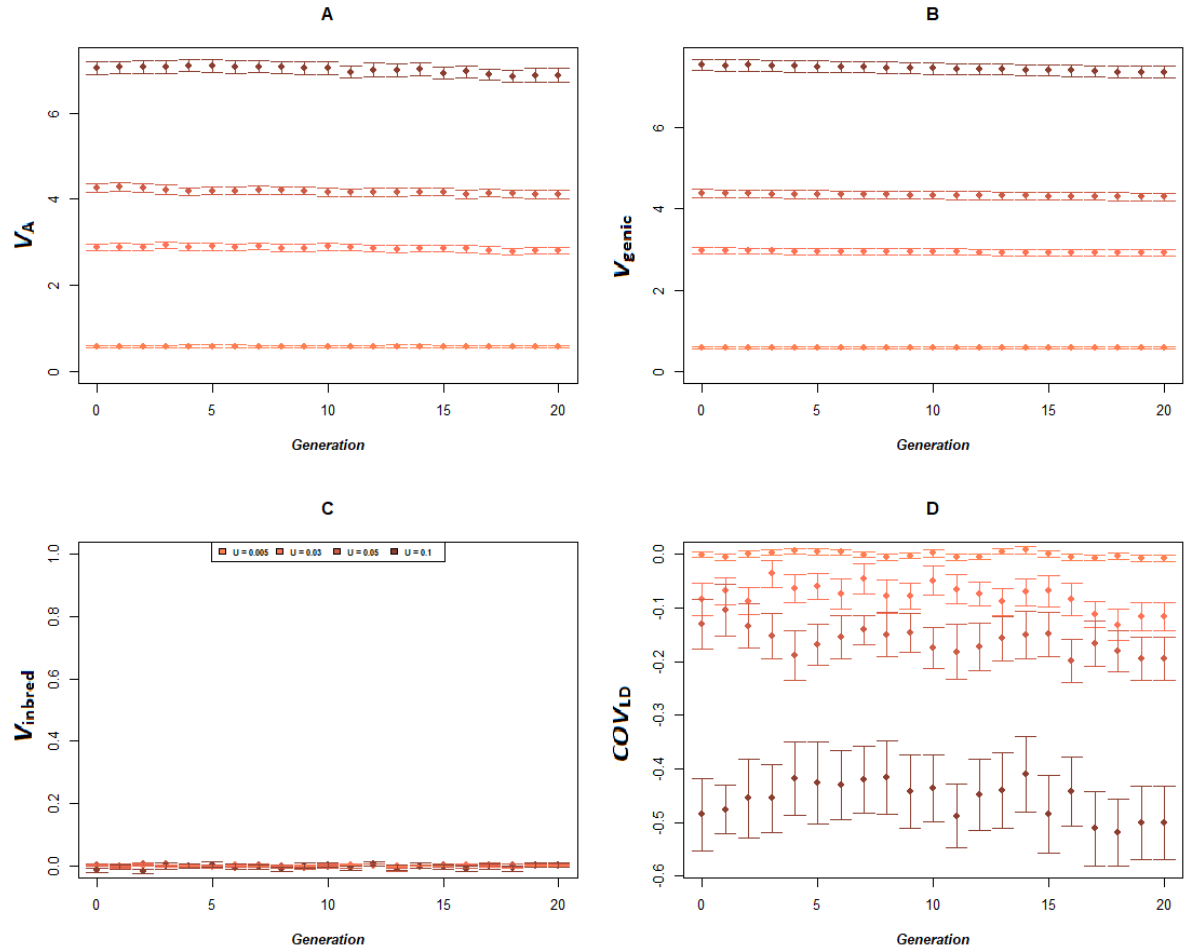


Figure S6. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=1000$ and $\omega^2=99$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

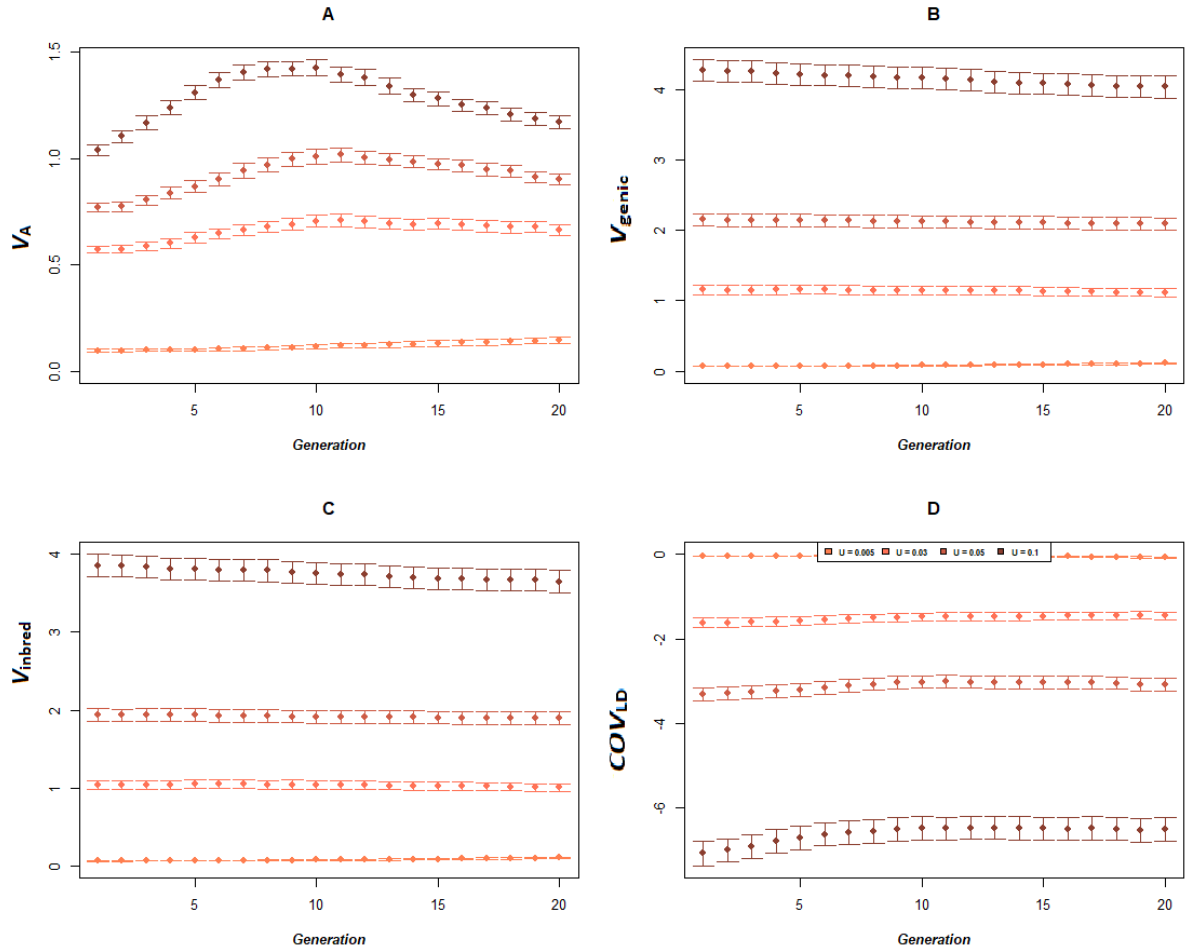


Figure S7. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for selfing populations of $N=1000$ and $\omega^2=9$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

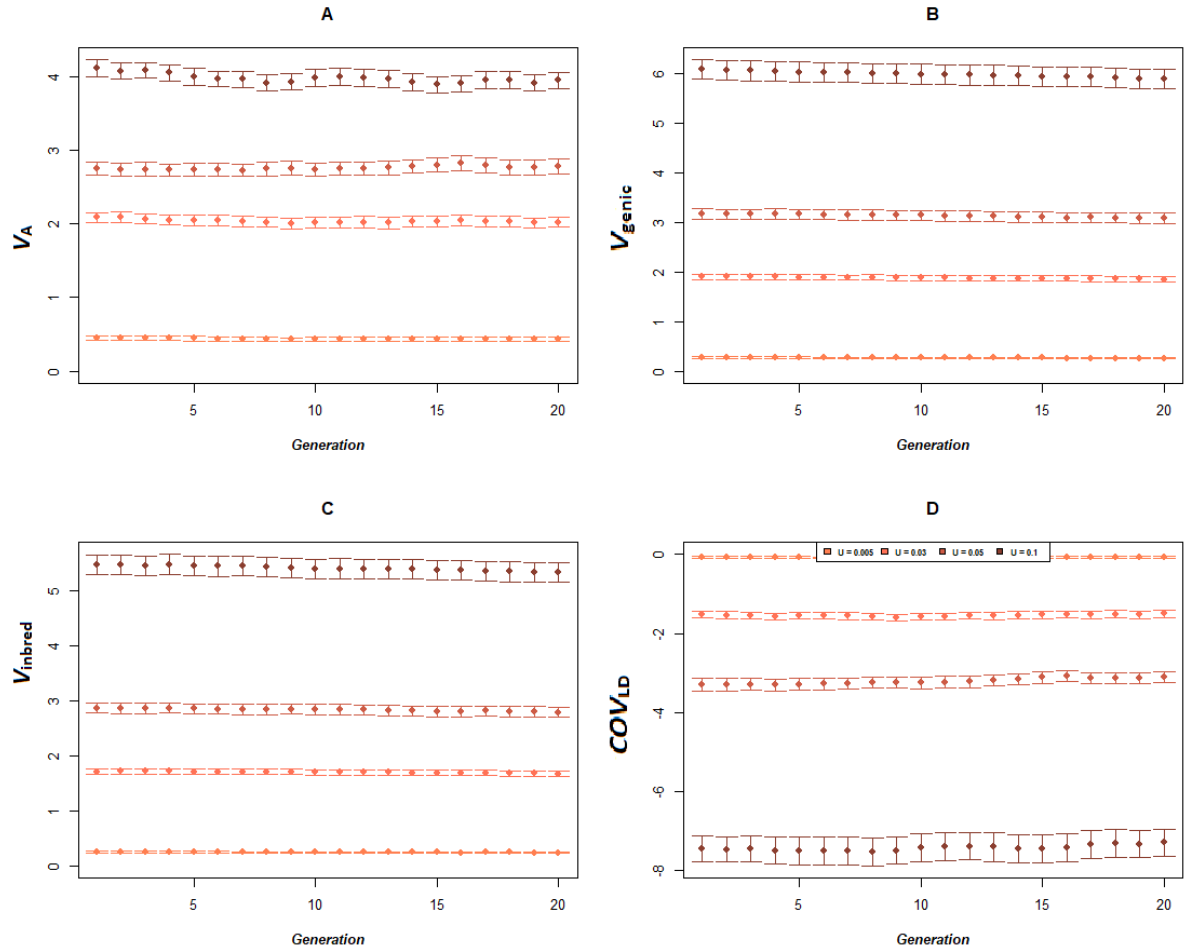


Figure S8. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for selfing populations of $N=1000$ and $\omega^2=99$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

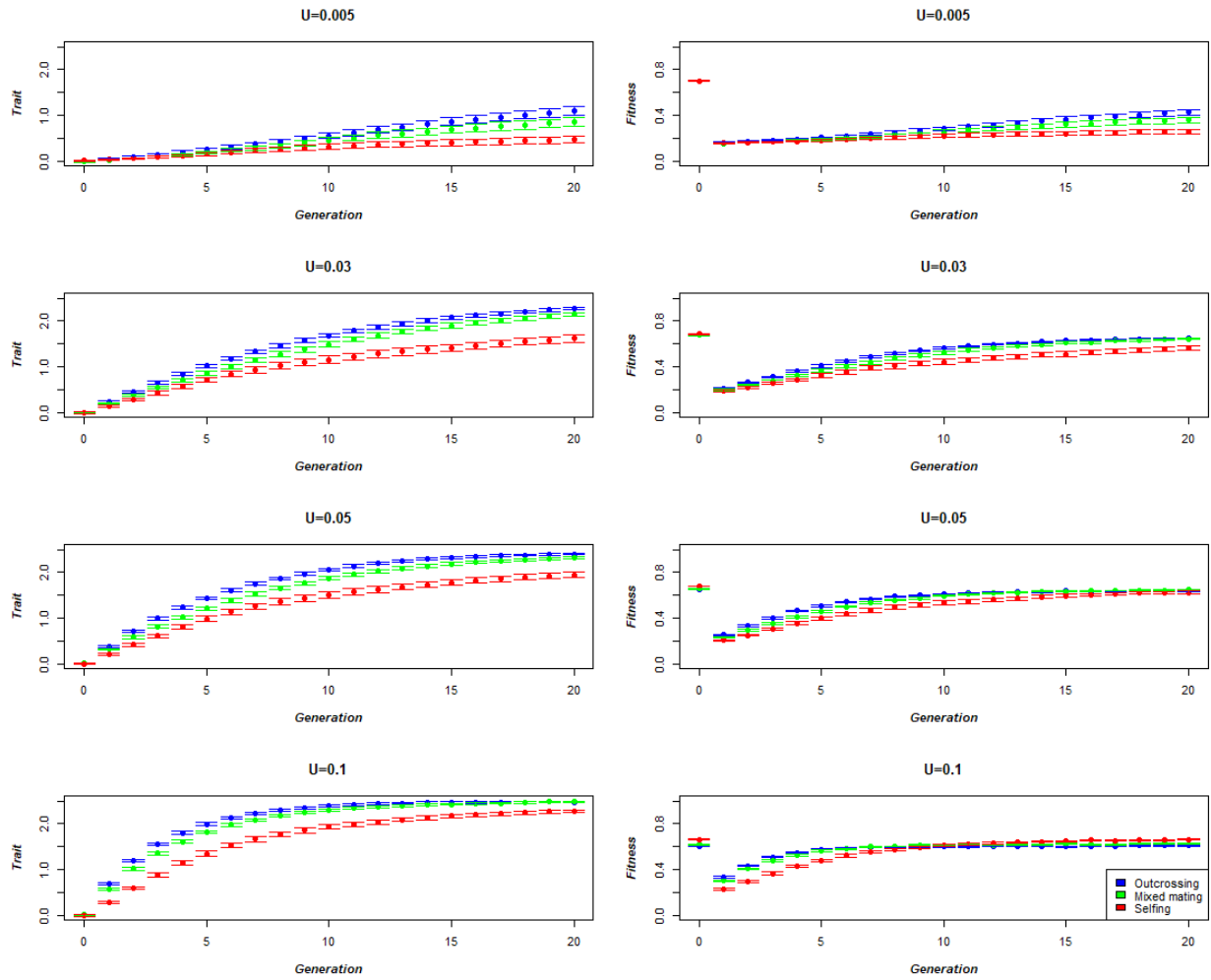


Figure S9. Dynamics of the trait and fitness, as a function of the trait haploid mutation rate and the mating system, for $N=250$ and $\omega^2=1$. Error bars stand for 95% confidence interval ($n=100$).

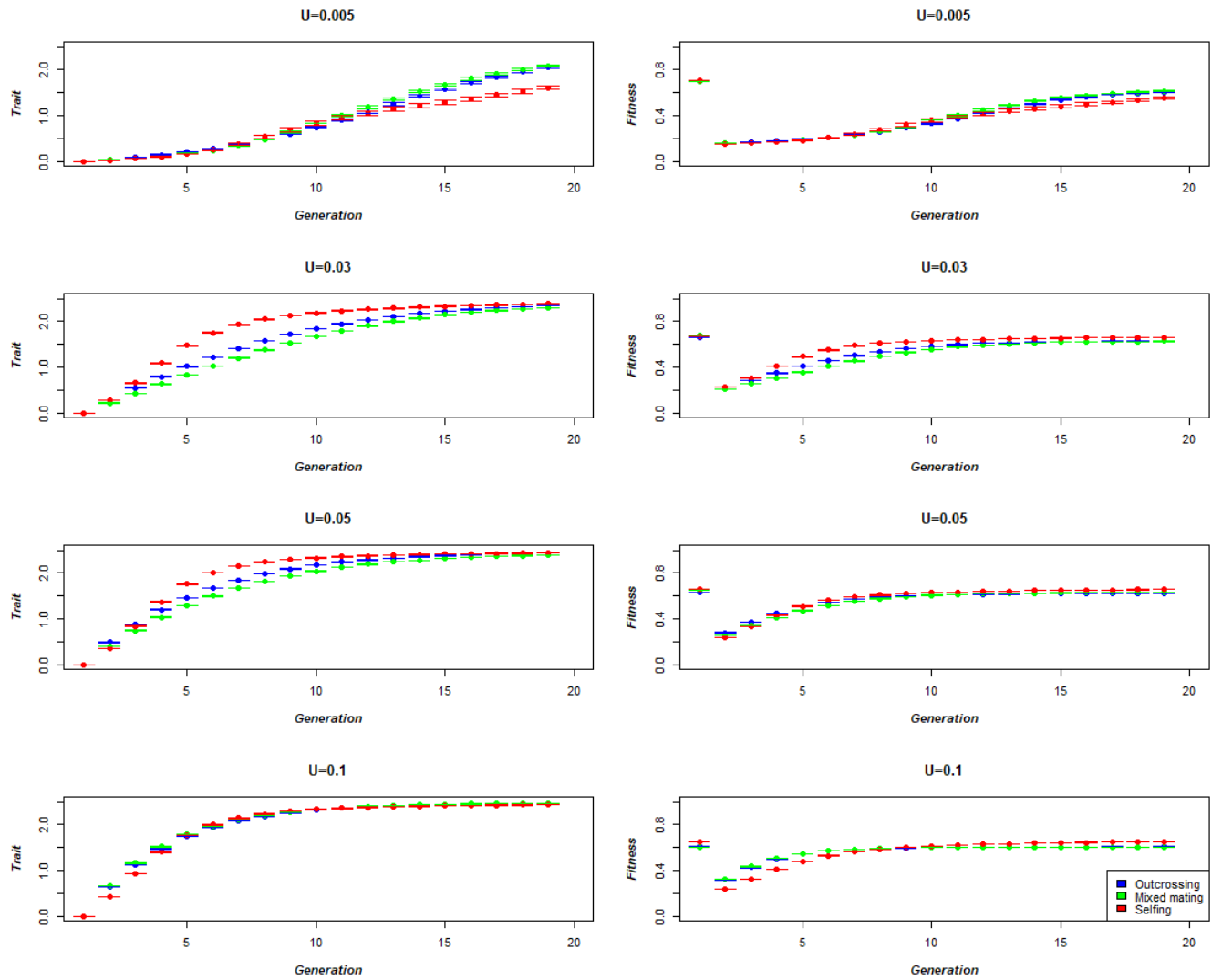


Figure S10. Dynamics of the trait and fitness, as a function of the trait haploid mutation rate and the mating system, for $N=10.000$ and $\omega^2=1$. Error bars stand for 95% confidence interval ($n=100$).

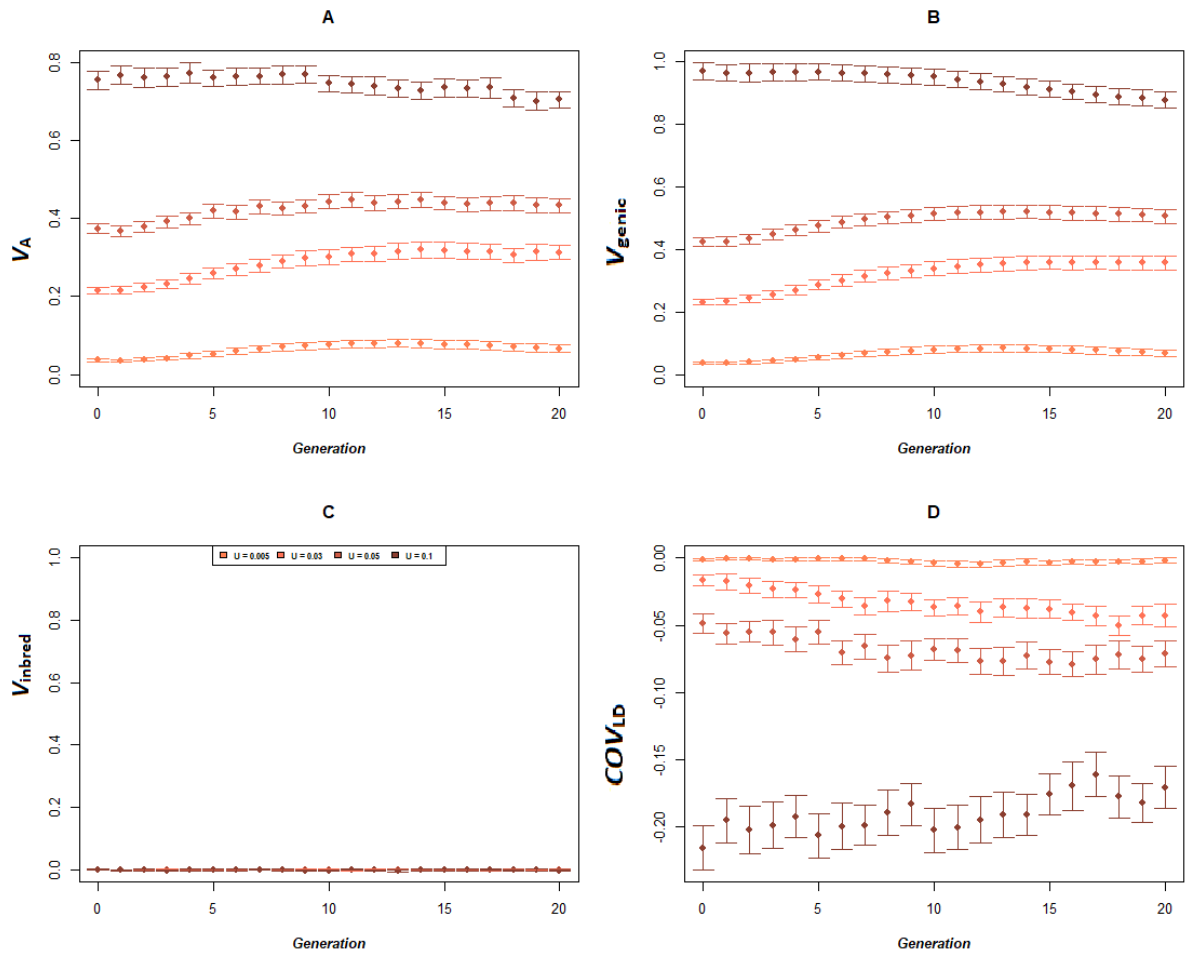


Figure S11. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=250$ and $\omega^2=1$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

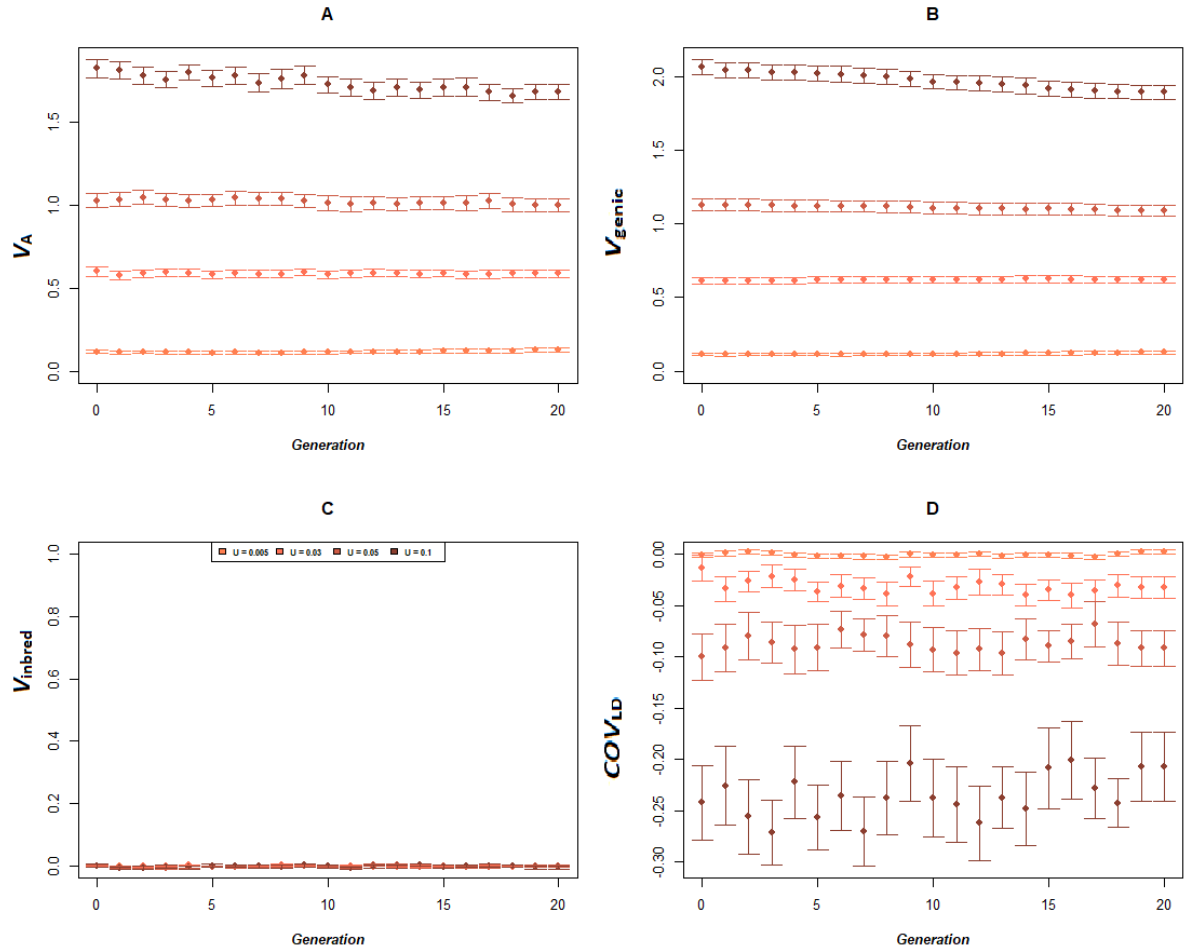


Figure S12. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=250$ and $\omega^2=9$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

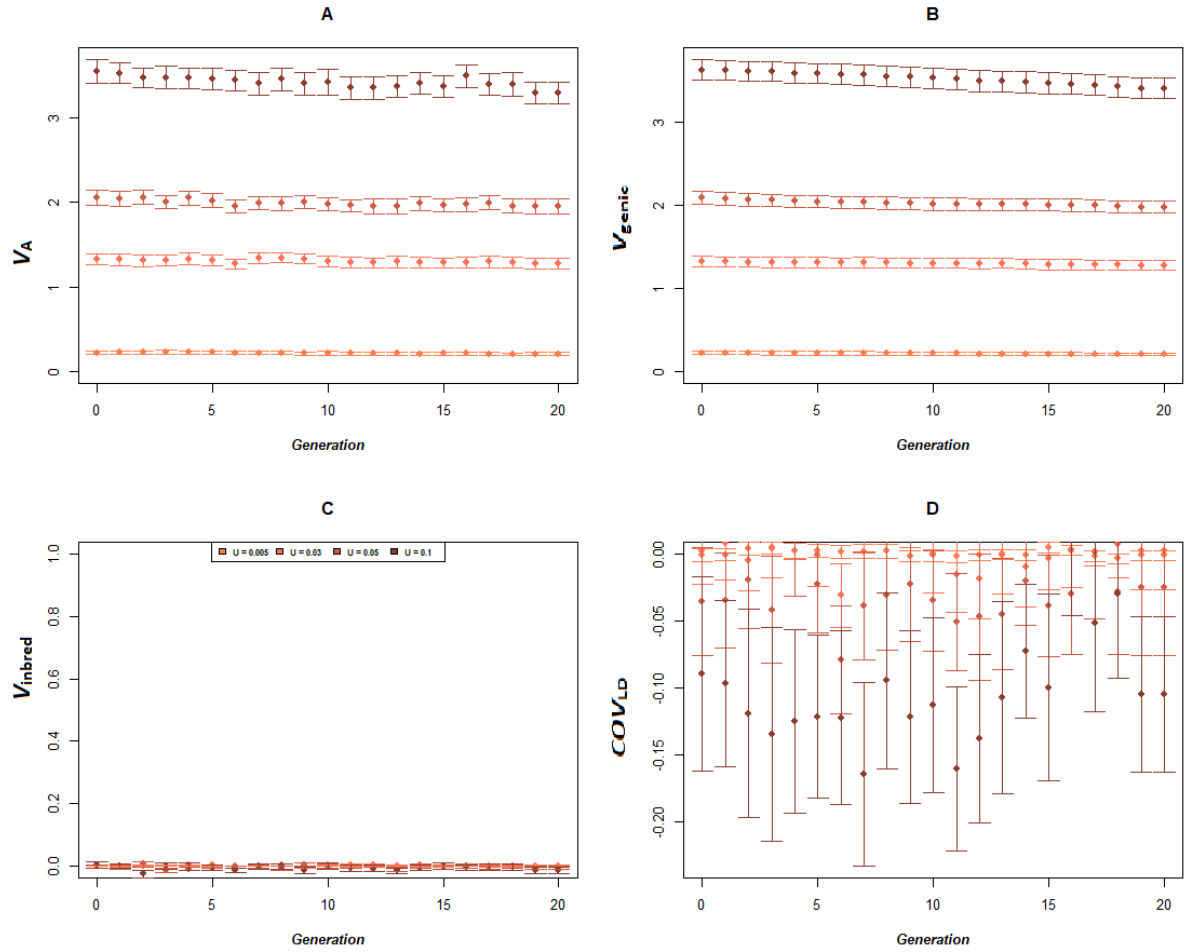


Figure S13. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=250$ and $\omega^2=99$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

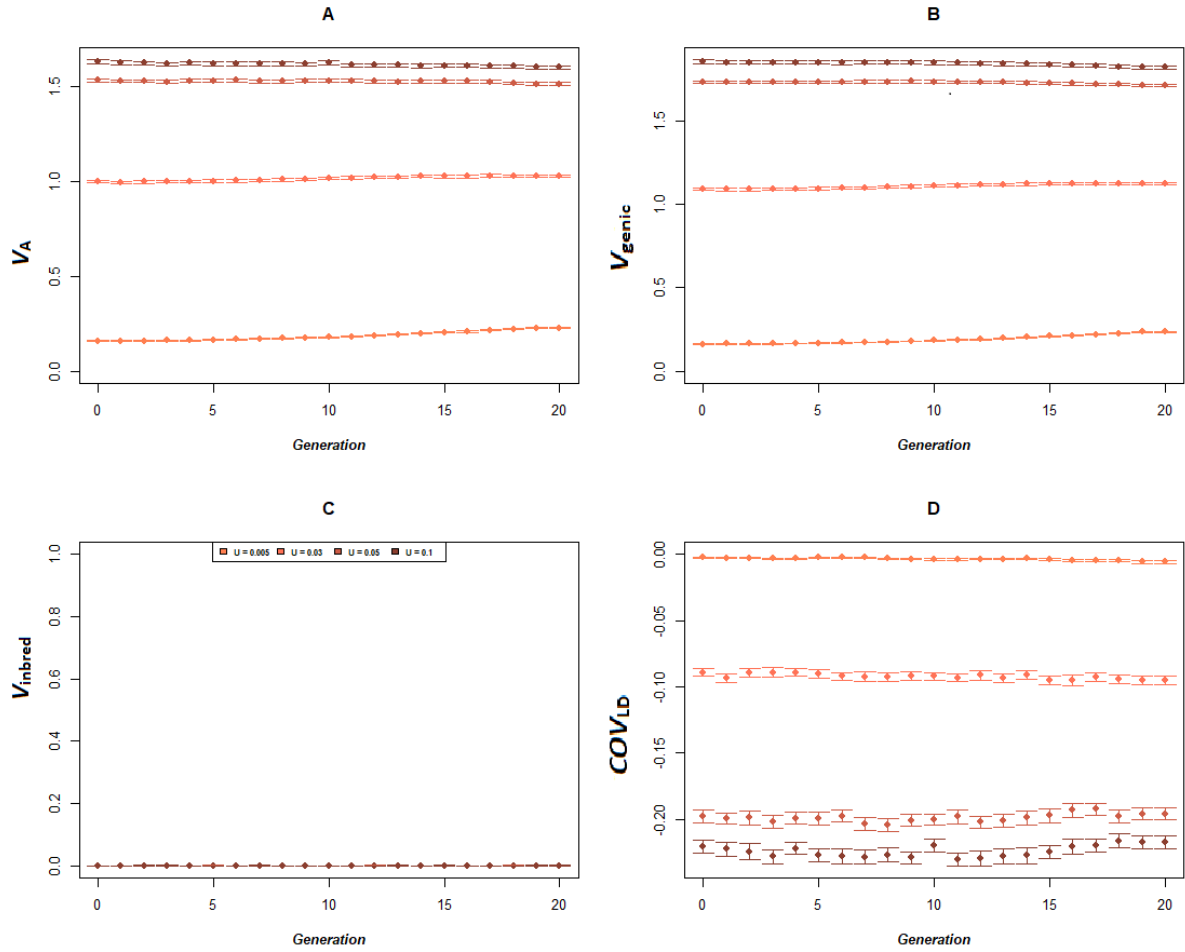


Figure S14. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=10.000$ and $\omega^2=9$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

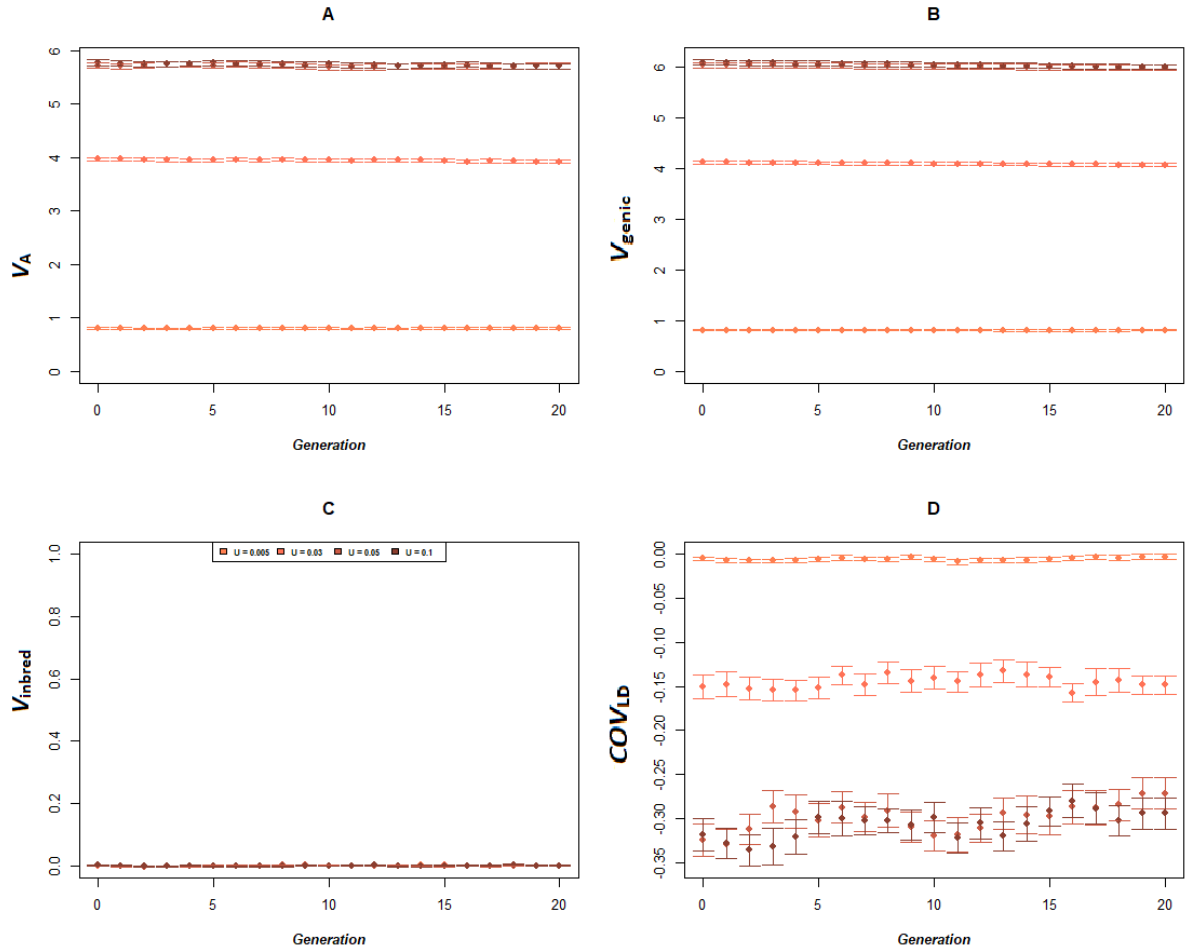


Figure S15. Dynamics of additive genetic variance and its components in function of the haplotypic trait mutation rate, for outcrossing populations of $N=10,000$ and $\omega^2=99$. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval ($n = 100$).

Chapitre 3 : Etude empirique sur
l'importance de la dominance dans
l'isolement reproductif entre lignées
consanguines dans les populations
naturelles de *Medicago truncatula*

This article is under review in *Heredity*.

Présentation générale

I have shown that residual allogamy can be of primordial importance when considering adaptation from standing genetic variation in predominantly selfing species. When two inbred lines hybrid, a large number of new genotypes can be created through F_1 hybrids, fuelling the adaptive process. Nevertheless, for such a mechanism to be possible, hybridization events between inbred lines need to have no (or positive) effects on hybrid fitness. By increasing homozygosity and reducing effective recombination rate (Nordborg, 2000), self-fertilization increases the probability of fixation of under-dominant loci, chromosomal rearrangements (Charlesworth, 1992), and epistatic incompatibilities (Coyne & Orr, 2004), at both within and between population scales (Wright *et al.*, 2013). If such isolation mechanisms occurs between lines, the remobilization of the genetic diversity could be impeded, limiting the adaptive potential of selfing populations.

In this third chapter, I studied the fitness consequences of hybridization between natural inbred lines of the predominantly selfing species, *Medicago truncatula*, at both within and among-population scales. I first searched for an effect of the spatial scale on the level of hybrids' performances. I then tried to decompose the genetic architecture of hybrids' performances by using family-based analyses. Finally, as the observed pattern of hybrid dysfunction was not compatible with the classical mechanisms of reproductive isolation, I explored an alternative hypothesis based on the effect of directional dominance on quantitative traits under stabilizing selection.

I found that hybridization events between lines is overall negative, even if the level of outbreeding depression is generally moderate (20% compared to mid-parent value). There was no significant effect of the spatial scale on the level of hybrids dysfunction. Family-based analyses revealed that outbreeding depression was most of the time due to a negative effect of dominance, even if I numerically proved that such analyses cannot totally rule out the contribution of fully dominant epistatic incompatibilities. I finally suggested that the negative effect of dominance is more likely due to an effect of directional dominance (that is fully expressed in F_1 hybrids), rather than under-dominant loci and/or chromosomal rearrangements.

The different results and their implications on the evolution of self-fertilization are described in the following manuscript.

**Fitness consequences of hybridization between inbred lines of the selfing
species *Medicago truncatula*: investigating the role of dominance and
genetic incompatibilities**

Josselin CLO¹, Joëlle RONFORT¹, Laurène GAY¹

¹ AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France.

For correspondence: Josselin Clo, josselin.clo@gmail.com, Unite Mixte de Recherche
AGAP, Montpellier SupAgro, 2 place Pierre Viala, 34060 Montpellier Cedex 02, France

Word count for the main text: 5269

Abstract: Studying the consequences of hybridization on plant performance can offer insight into the genetic histories and adaptive potential of populations, notably at local scales. Due to reduced effective recombination, predominantly selfing species are organized in highly homozygous multi-locus-genotypes (or lines) that accumulate genetic differentiation both among- and within-populations. This high level of homozygosity facilitates the dissection of the genetic basis of hybrid performance in highly selfing species, which gives insights into the mechanisms of reproductive isolation between lines. Here, we explored the fitness consequences of hybridization events between natural inbred lines of the predominantly selfing species *Medicago truncatula*, at both within- and among-populations scales. We found that hybridization was overall deleterious, independently of the spatial scale, even if the decrease in fitness remained moderate. Family-based analyses allowed us to determine that outbreeding depression was likely due to negative dominance effects, even if we showed that this method cannot totally rule out the contribution of fully dominant epistatic incompatibilities. As the patterns of outbreeding depression were not consistent with classical under-dominant mechanisms, we propose an alternative mode of action of dominance acting on fitness related traits, and show that it is supported by our data. We discuss how our results bring new insights into the role of dominance in hybridization events and into the evolution of self-fertilization.

INTRODUCTION

The evolutionary forces underlying divergence between populations, in particular the relative strength of genetic drift compared to natural selection, can be revealed by studying the consequences of hybridization (Lynch, 1991; Demuth and Wade, 2005; Fitzpatrick, 2008a). Strong genetic drift (e.g. due to demographic events or mating systems, (Barrett *et al.*, 2014) increases the probability of fixation of deleterious mutations (Kimura *et al.*, 1963). Hybridization between populations that have accumulated different recessive deleterious mutations is then expected to lead to heterosis, *i.e.* an increase in fitness in F_1 hybrids compared to the average fitness of their parents (Crow, 1948; Lynch, 1991; Glémin, 2003). However, if selection is much stronger than drift, divergent populations are expected to be adapted to their local conditions. In this case, hybridization could lead to a loss of fitness in F_1 and/or subsequent generations (in general the F_2 generation). Several non-exclusive hypotheses explaining the genetic architecture of outbreeding depression have been proposed, the most prominent of which is chromosomal rearrangement (Lande, 1985; Charlesworth, 1992; Kirkpatrick and Barton, 2006), under-dominance (Schierup and Christiansen, 1996), and negative epistatic interactions (Dobzhansky, 1937; Lynch, 1991; Demuth and Wade, 2005).

Predominantly selfing species are of particular interest when studying the relative strength of selection compared to drift, and the genetic architecture of heterosis and outbreeding depression. By decreasing the effective population size, selfing increases the strength of genetic drift (Glémin and Ronfort, 2013), and thereby the fixation probability of deleterious mutations (Kimura *et al.*, 1963) . On the other hand, by increasing homozygosity and reducing the effective recombination rate, self-fertilisation also increases the probability of fixation of under-dominant loci, chromosomal rearrangements (Charlesworth, 1992), and epistatic

incompatibilities (Coyne and Orr, 2004), at both within and between population scales (Wright *et al.*, 2013). In addition, by organizing populations in a mosaic of repeated fully homozygous multi-locus genotypes (thereafter MLG; (Siol *et al.*, 2008; Volis *et al.*, 2010; Jullien *et al.*, 2019), self-fertilization greatly simplifies experimental designs for clarifying the genetic architecture of hybrid fitness. Hybridization between MLGs produces F_1 hybrids that are fully heterozygous and the heterozygosity is divided by two at each consequent generation of selfing, providing the opportunity to test the relative role of dominance and different forms of epistasis in populations (Demuth and Wade, 2005; Fitzpatrick, 2008a).

Outbreeding depression and heterosis have been widely documented in predominantly selfing species (Rhode and Cruzan, 2005; Johansen-Morris and Latta, 2006; Dolgin *et al.*, 2007; Volis *et al.*, 2010; Gimond *et al.*, 2013; Oakley *et al.*, 2015). Explanations for heterosis always involve the masking of deleterious mutations while epistatic incompatibilities are generally thought to be the main cause of outbreeding depression. However, although outbreeding depression is frequently observed in predominantly selfing populations, the decrease in hybrid fitness is generally small to moderate (average between 20% and 30%, (Dolgin *et al.*, 2007; Gimond *et al.*, 2013; Oakley *et al.*, 2015), with hybrids often remaining within the parental range values. This observation does not support the hypothesis that hybrid breakdown is due to chromosomal rearrangements or epistatic incompatibilities, which are expected to have drastic effects on fitness. In most cases, these studies lack a proper testing of the mechanisms involved, despite the fact that selfing species facilitate distinction between the effects of dominance and different forms of epistasis. To our knowledge, a single study analysed the underlying mechanisms and led to the surprising conclusion that outbreeding depression observed after among-population crosses of *Arabidopsis thaliana* was caused by a negative effect of dominance (Oakley *et al.*, 2015). This suggests that incorporating dominance into classical

models of quantitative traits under selection could be illuminating (Schneemann *et al.*, 2020; Yamaguchi and Otto, 2020).

Little is known about the scales at which heterosis and outbreeding depression occur. As mentioned above, selfing populations are organised in inbred lines that may be as likely to fix genetic incompatibilities among them as geographically isolated lines. Nevertheless, few studies have tried to investigate how the local genetic structure affects hybrid performance compared to the among-population genetic structure, and the few studies that have studied hybridization at the within-population scale have yielded contradictory findings (Volis *et al.*, 2010; Gimond *et al.*, 2013). Furthering our knowledge on what happens at the population scale is of primary importance as it could notably explain why transitions from selfing to outcrossing are rare (Harkness *et al.*, 2019) and give insight into the short-term evolutionary potential of selfing populations (Clo *et al.*, 2020).

Overall, it appears that little is known about the mechanisms leading to outbreeding depression and heterosis in predominantly selfing species, how variable they are, and how variable the magnitude of hybrid performance between within- and among-population crosses of inbred lines is. In this paper, we investigate the effect of hybridization between naturally occurring inbred lines from the predominantly selfing species *Medicago truncatula*. We first quantify the occurrence and magnitude of heterosis and outbreeding depression at both within- and among-populations scales, and we compare hybrid performances between the two spatial scales. In a second time, we try to dissect the genetic architecture of heterosis and outbreeding depression using family-based analyses. Assuming that dominance could be a driver of outbreeding depression, at both within and between population scales, we further examine its underlying genetic architecture. The effect of dominance in hybrid failure is generally assumed to come from major effect loci, resulting in large negative effects of dominance. Here, we

examine an alternative architecture, leading to moderate outbreeding depression compared to mid-parent value.

MATERIAL AND METHODS

Model species and plant material

Medicago truncatula (Fabaceae) is an annual, diploid plant species commonly found in the Mediterranean basin. It reproduces predominantly through self-fertilization. Here, we sampled the most frequent MLGs (defined using 20 microsatellite markers) from five natural populations studied by Jullien *et al.* (2019): three populations collected in Corsica (CO1, CO2 and CO3), one in the south of France (FR1) and one in the south of Spain (SPA1). For each population, we chose two MLGs, except for the CO3 population, for which four MLGs were sampled, adding up to 12 parental lines (MLGs). The genetic differentiation among these MLGs is measured using the 20 microsatellite markers, and is defined as the number of alleles that differ between two MLGs, divided by the total number of alleles without missing data. It varies between 12% and 86% within-population and between 57% and 100% among-populations (see table S1 for details).

Crossing design and growth conditions

Crosses

Several plants representing each of these 12 MLGs were grown in the greenhouse in order to multiply the seed stock and to standardize any potential environmental maternal effects.

Flower buds were sampled before the flowers open. The flowers used as pollen donors were dissected by removing sepals and petals. For flowers used as pollen receivers, the immature anthers were carefully removed under a binocular scope and pollen from a donor was manually deposited on the stigma. The stems were then wrapped with cotton and placed in a tube with some water in order to maintain a moist environment and favour fertilization. Each hand-fertilized flower was then labelled. Once a fruit appeared, it indicated that the fertilization was successful and the tube was removed. Pods were left on the mother plants until fully ripe.

We performed controlled crosses using the 12 MLGs originating from the five populations selected. We performed all the possible within-population crosses (one for the four populations in which two MLGs are sampled, and six for the Corsican population in which four MLGs have been sampled), for a total of 10 within-population crosses. For among-populations crosses, we selected the most frequent MLG within each population and the two most frequent MLGs for the Corsican population, in which four MLGs had been sampled. With these six MLGs, all the possible among-population crosses were performed, leading to a total of 14 among-population crosses. The experimental design is summarized in the supplementary materials (Table S2). All the crosses were carried out using each MLG as a mother or pollen donor (but analyses showed that there was no effect of the direction of the cross). The 24 crosses were performed in spring between 2013 and 2015. In 2016, a subset of seeds from the F_1 families were germinated, seedlings were grown in an insect-proof greenhouse and self-fertilized to produce F_2 plants. Leaves from the F_1 mother plants were sampled for DNA extraction in order to confirm that the cross was successful given the parental MLGs, using four microsatellite loci (B12F1, MTIC59, TA34, FMT 11, (Jullien *et al.*, 2019)). The protocol for extractions and microsatellite genotyping was the same as detailed in (Jullien *et al.*, 2019). Individuals with mismatches (generally accidentally selfed) were removed from the analysis. One family produced from

within-population crosses (population CO1) had to be removed due to a technical error in the crossing design. The following analyses were thus performed on the remaining 23 families.

Greenhouse experiment

In autumn 2017, seeds from the 12 parental MLGs and from the 24 F_1 and F_2 families were weighed individually. They were scarified to ease germination and transferred to Petri dishes with water at room temperature for six hours, before being placed in a vernalisation chamber at 5 °C for five days.

After vernalisation, an average of 8 replicates from each parental MLG (hereafter noted “ P ”), 16 replicates from each F_1 family and 64 replicates from each F_2 family were transferred to the greenhouse, according to a randomized block design (eight blocks with one replicate per bloc for P , two for F_1 and eight replicates per block for F_2). Dormant seeds or seeds with germination problems were replaced. The germinated seeds were transplanted into one-litre pots, filled with a substrate made of 75% sand and 25% potting soil. Throughout the entire life cycle, plants were maintained under light water stress conditions, watered by soaking five minutes once a week. No light or temperature controls were imposed: day length varied from 9h.day⁻¹ to 15h.day⁻¹ in the region of Montpellier from November 2017 to June 2018, while recorded temperatures in the greenhouse varying on average between 4°C to 13°C at night, and between 20°C to 30°C during the day, over the same period.

All the plants that survived transplantation from Petri dishes to the greenhouse reached the flowering stage. The experiment was stopped when the fruits from all plants that had reached fructification were mature (i.e. their colour turned from green to brown). No correlations between flowering time and the different fitness proxies we measured were

detected, indicating that the experiment lasted long enough to allow late flowering individuals to end their reproductive cycle (Figure S1).

Traits and fitness measurements

For all the 1788 surviving plants, we measured a fitness component, the vegetative dry mass (g), and a fitness proxy, the total number of seeds. At the end of the flowering season, we collected the senescent plants and separated the vegetative biomass from the fruits. The biomass was dried in an oven and weighed to estimate the dry mass. The fruits produced were counted and threshed manually. During this step, some seeds were broken, we therefore counted the number and the total weight of unbroken seeds and measured the total weight of broken seeds. We then estimated the number of broken seeds by dividing the total weight of broken seeds by the mean weight of an unbroken seed.

Statistical analyses

Performance measure and its interaction with the spatial scale

For the two traits studied and for each family studied, we compute the relative performance (RP) of hybrids compared to the mid-parent values as

$$RP = \frac{\overline{W}_H - \overline{W}_P}{\overline{W}_P} ,$$

where $\overline{W}_H$ is the mean value of the trait of either the F_1 or F_2 progeny of a family, and $\overline{W}_P$ the mid-parent value. A negative RP indicates outbreeding depression, and a positive RP indicates heterosis. We test whether the means of F_1 and F_2 differ significantly from the mid-parent

values by performing likelihood ratio tests between linear mixed models, where the RP of F_1 or F_2 is either estimated or constrained to zero. In order to take the variance due to environmental effects into account, zones within the greenhouse and blocks within zones where plants were grown, are added as random effects within the models

We test if the average RP differs significantly between within- ($n = 9$ estimates) and among-population ($n = 14$ estimates) scales, for the two generations of hybrids, and for the two focal traits, using F -tests. For these analyses, we use all RP values, whether they differ significantly from zero or not (Gimond *et al.*, 2013).

Genetic architecture of hybrid performances

As mentioned in the Introduction, crosses in predominantly selfing species enable the dissection of the genetic architecture of hybrid performance. Indeed, following how the F_2 mean value differs from the mid-parent value, compared to F_1 mean value, provides insights into the contribution of dominance and/or epistasis to outbreeding depression and heterosis. Here, we use the theoretical framework introduced by Fitzpatrick (2008) to incorporate Dobzhansky-Müller incompatibilities (thereafter called DMI) into line-cross quantitative genetics analyses (Demuth and Wade, 2005). Nevertheless, it is important to note that the accuracy of Fitzpatrick's model strongly depends on the dominance of DMI. We demonstrate that this model can distinguish the effects of dominance from epistatic incompatibility when DMI are recessive, but is unable to distinguish between the two effects when DMI are dominant (Annexe 1). Indeed, when epistatic incompatibilities are dominant, epistatic effects decrease by the same order of magnitude as dominance effects between F_1 and F_2 generations (Annexe 1, Table 1). This may limit our ability to distinguish between models that point to dominance or dominant epistatic effects as causing outbreeding depression or heterosis.

We test if deviations of F_1 and F_2 offspring from mid-parent values are due to dominance, epistatic incompatibilities, or a combination of both, by decomposing the phenotypic variance of the studied traits into additive, dominance and epistatic components as follows: let $Z_{i,j}$ be the observed phenotypes of the j th individual of the i th family and

$$Z_{i,j,k,l} = \mu_i + I_{A,j} \cdot A + I_{D,j} \cdot D + (I_{R-DMI,j} \cdot R-DMI \text{ or } I_{D-DMI,j} \cdot D-DMI) + \text{Zone}_k + \text{block}_{lk} + \varepsilon_{i,j,k,l}$$

where μ_i is the mean of the i th family under pure additivity (also equal to the mid-parent value), A is the deviation due to additivity, D is the deviation due to dominance, $R-DMI$ and $D-DMI$ are respectively deviations due to recessive and dominant DMI. The contribution of each effect (I indices) depends on the generation of cross of the j th individual (parent, F_1 or F_2), and are detailed in Table 1. Zone_k and block_{lk} are random effects corresponding to the zone and the block within the zone in which the j th individual were grown in the greenhouse, and $\varepsilon_{i,j,k,l}$ is the residual error. We compare the null model, limited to additive effects and random effects, with models including one parameter (D , $R-DMI$ or $D-DMI$), or two parameters ($D + R-DMI$ or $D + D-DMI$), and test the significance by performing likelihood ratio tests.

We limited the analysis to families in which at least one generation of hybrids differed significantly from the mid-parent value (nine families for dry mass and five families for number of seeds).

Table 1. Contribution of different genetic effects (additivity, dominance or epistasis) to plant performance for each generation of cross.

Generation	I_A	I_D	I_{D-DMI}	I_{R-DMI}
Parental line 1	1	0	0	0
Parental line 2	-1	0	0	0
F_1	0	1	1	0
F_2	0	0.5	0.5625	1

Verbal and Theoretical investigation of the role of dominance on hybrid performance

To better understand the potential role of dominance in outbreeding depression, we explore the effect of directional dominance on hybrid phenotypes and fitness values. We consider a trait governed by L loci with additive and dominance effects. Let $\bar{P}_1, \bar{P}_2$ denote the mean phenotypic values of two MLGs sampled in a population and $\bar{P}$ the mid-parent phenotypic value. The mean phenotypic value of the F_1 hybrid is equal to

$$\bar{F}_1 = \bar{P} + \sum_i^L d_i$$

with d_i the biological dominance deviation of the heterozygote from the mid-value of both homozygotes, at the i th locus. If epistatic effects are negligible, $\bar{F}_1$ is expected to differ from $\bar{P}$ only if the biological dominance is directional on average ($\sum_i^L d_i \neq 0$). We will consider no under- and over- dominance at any given locus, such that the hybrids remain in the parental phenotypic range at each locus. If parental lines are phenotypically similar, but have fixed different sets of alleles (phenotypic redundancy, Figure 1A), then directional dominance should lead to hybrid phenotype and fitness lying outside the range of parental values (Figure 1A). In

this case, the relative performance of hybrids can greatly vary, but is expected to be far from the mid-parent value (Figure 1A). However, if we consider two parental lines that are phenotypically divergent (Figure 1B), the breeding values (α) of the parental lines are also expected to be very different, with the breeding values of the largest parental lines being on average higher at each of the L loci (due to dominance). In this case, the F_1 hybrids are expected to lie within the range of the parental values for both phenotype and fitness. Hybrids will be closest to the largest or smallest parental line, depending on the directionality of dominance (Figure 1B), and the performance of hybrid is expected to close to mid-parent value (Figure 1B).

With directional dominance, the probability that hybrid phenotypes will fall outside the parental range is therefore expected to increase with the phenotypic distance between MLGs. Note that this pattern is not expected with under-dominant loci, chromosomal rearrangement or epistatic incompatibilities, as hybrids should all be less fit than the two parental lines. Evidence for a positive correlation between parental similarity and the relative performance of hybrids therefore strongly suggests that directional dominance is involved in the genetic architecture of hybrid performances. We verified this using simulations that account for directional dominance effects (Figure 1C). Then, we examined whether directional dominance could play a role in the consequences of hybridization in our experiment by testing the correlation between the phenotypic similarity between parental MLGs (computed as $\min(P_1, P_2)/\max(P_1, P_2)$) and the phenotypic distance between hybrids and parents (measured by the ratio $\frac{\bar{F}_1}{\max(\bar{P}_1, \bar{P}_2)}$ if the trait shows heterosis, or $\frac{\bar{F}_1}{\min(\bar{P}_1, \bar{P}_2)}$ if it displays outbreeding depression).

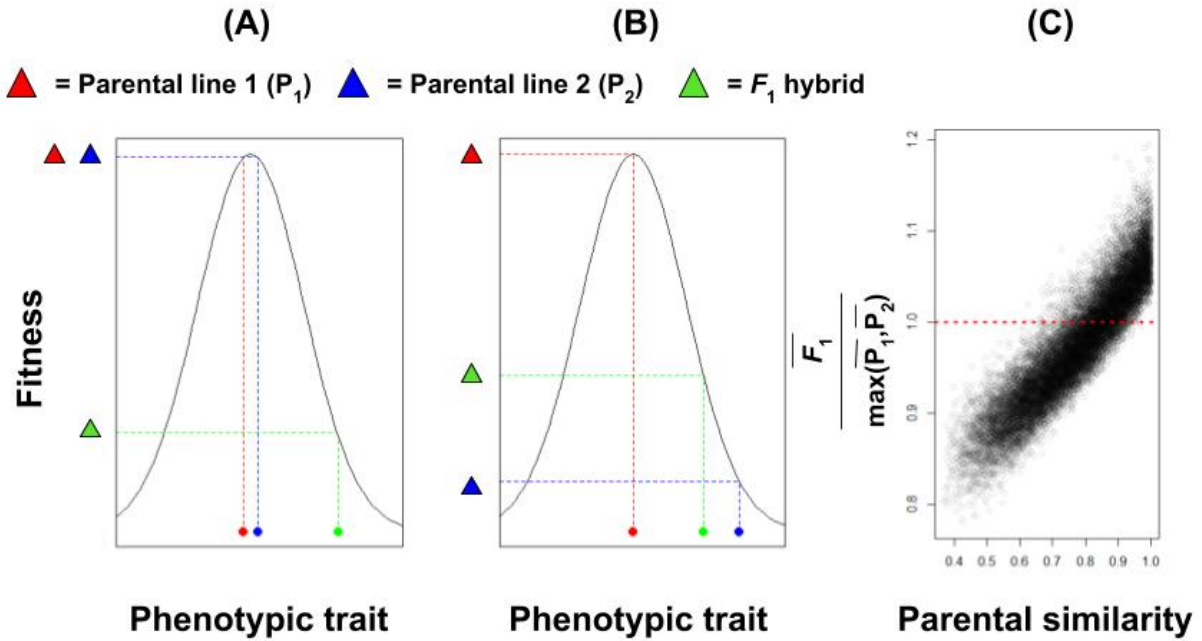


Figure 1. Schematic representation of the hypothesis of directional dominance and predictions of the consequences of hybridization events in predominantly selfing populations. Figures **A.** & **B.** represent how directional dominance affects the phenotypic values of hybrids for different levels of phenotypic differentiation between parental lines. In the hypothetical example A, both parental lines have similar phenotypic values, but have fixed different sets of alleles. The example B illustrates hybrid values when parental lines are phenotypically divergent. Figure **C.** illustrates how the probability that the F_1 mean value will fall outside the parental range changes in function of the phenotypic similarity of parental lines (computed as $\min(P_1, P_2)/\max(P_1, P_2)$). We simulated inbred lines that have different means $\bar{x}$ for a quantitative trait (ranging from 10 to 20, with an increment of 0.05). We then made all the possible pairwise crosses between lines to produce F_1 hybrids ($n = 20,100$). The trait is coded by L diploid loci ($L = 10$), and the contribution of the allele at the i th loci is drawn from a normal distribution of mean $\bar{x}/2L$, and of variance $0.3*\bar{x}/2L$, and dominance deviation was drawn from a uniform distribution ranging from the mid-parent value to the largest parent.

RESULTS

Heterosis, outbreeding depression and the effect of spatial scale

We find a strong negative correlation between vegetative dry mass and the number of seeds (Figure S2). Heterosis for vegetative dry mass should thus be associated with outbreeding depression for seed production. The untransformed means of parental values and hybrid lines for each family and trait are given in supplementary material (Figure S3 and S4). In the majority of families, hybrids remain within the parental range of values for both dry mass and number of seeds per individual (Figure S3 and S4).

For vegetative dry mass, we find a general increase in both the F_1 and the F_2 generations, the increase being smaller in the F_2 (Figure 2). Nevertheless, hybrids are significantly different from the mid-parent only in 9 families (7 where both generations of hybrids differ significantly, 2 with only the F_1 or the F_2 , Figure 2). The mean RP is equal to 0.195 (s.d. = 0.134) in F_1 and to 0.122 (s.d. = 0.110) in F_2 . There is a marginally significant effect of the generation on the mean RP ($F_{\text{value}} = 3.996$, $df = 1$, $p = 0.053$), but there is no effect of the spatial scale, whatever the generation ($F_{\text{value}} = 0.867$, $df = 2$, $p = 0.427$). Means RP values in F_1 are equal to 0.169 (s.d. = 0.144) at the within-population scale and to 0.211 (s.d. = 0.131) at the between -population scale. In the F_2 generations, mean RP are 0.089 (s.d. = 0.118) at the within population scale and 0.144 (s.d. = 0.104) at the between population scale in F_2 hybrids.

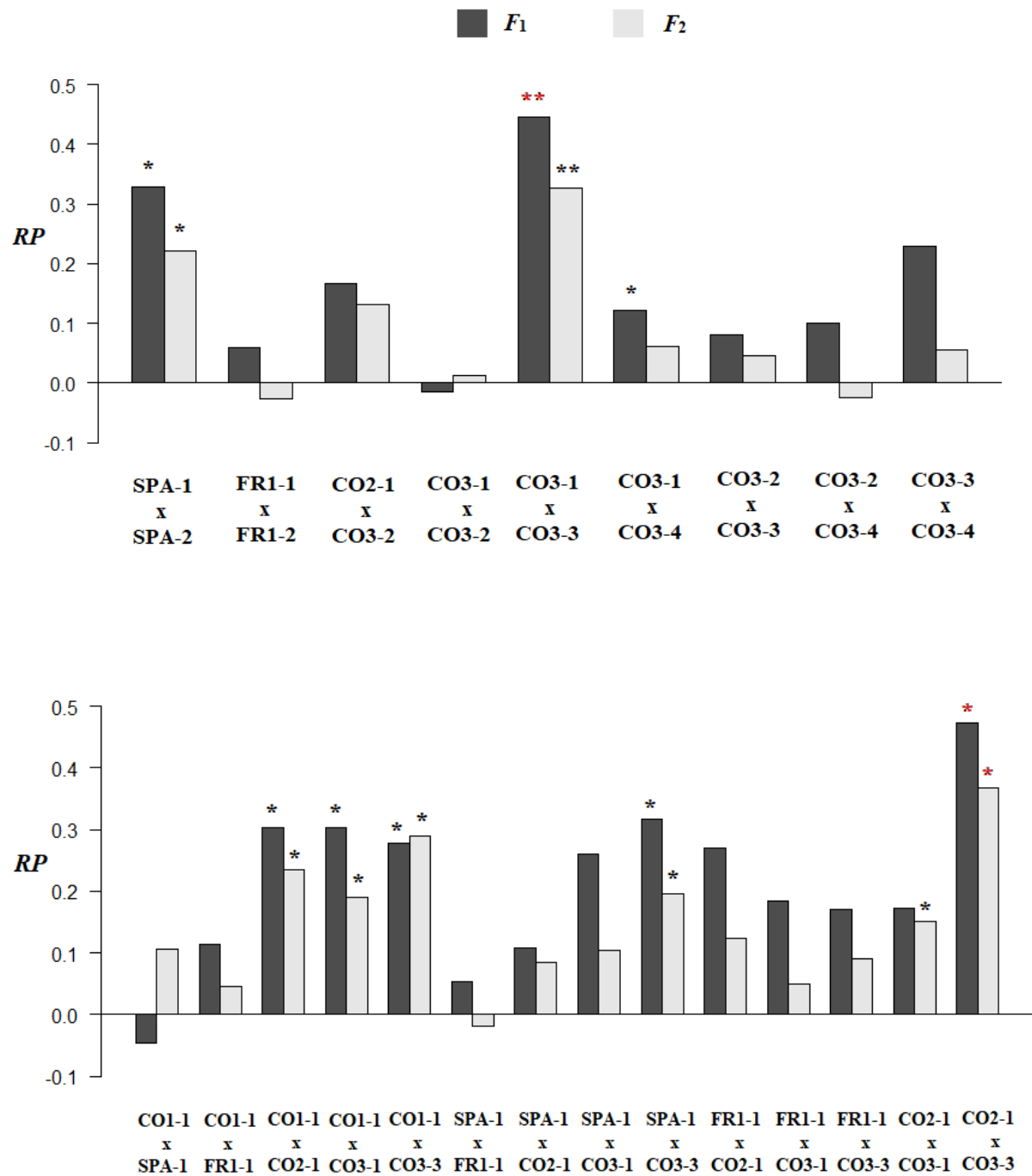


Figure 2. Relative performance (RP) of F_1 (dark grey) and F_2 (light grey) hybrids for the vegetative dry mass (g), for each family, at the within-population (top panel) and among-population scale (bottom panel). Significant difference between hybrids and the mid-parent (black asterisks) and best parent (red asterisks): * $p < 0.05$; ** $p < 0.01$.

For the number of seeds per individual, and in agreement with the negative relationship between this variable and the vegetative dry mass, we observe a general pattern of outbreeding depression in hybrids (Figure 3). This pattern is however not observed in crosses involving the CO3 MLGs, which generally lead to heterosis (Figure 3). Both heterosis and outbreeding depression are generally stronger in the F_1 hybrids compared to the F_2 generation (Figure 3). RP values are significant for 5 families (Figure 3). The mean RP is -0.181 (s.d. = 0.227) in F_1 , and raises to -0.099 (s.d. = 0.204) in F_2 . There is no significant effect of the generation on the mean RP ($F_{\text{value}} = 1.664$, $df = 1$, $p = 0.161$), and no effect of the spatial scale for either generation ($F_{\text{value}} = 0.919$, $df = 2$, $p = 0.407$). The mean RP in F_1 are equal to -0.127 (s.d. = 0.293) at the within-population scale and -0.217 (s.d. = 0.177) at the among-population scale, and are equal to -0.046 (s.d. = 0.271) at the within population scale and -0.134 (s.d. = 0.148) at the among population scale in F_2 hybrids.

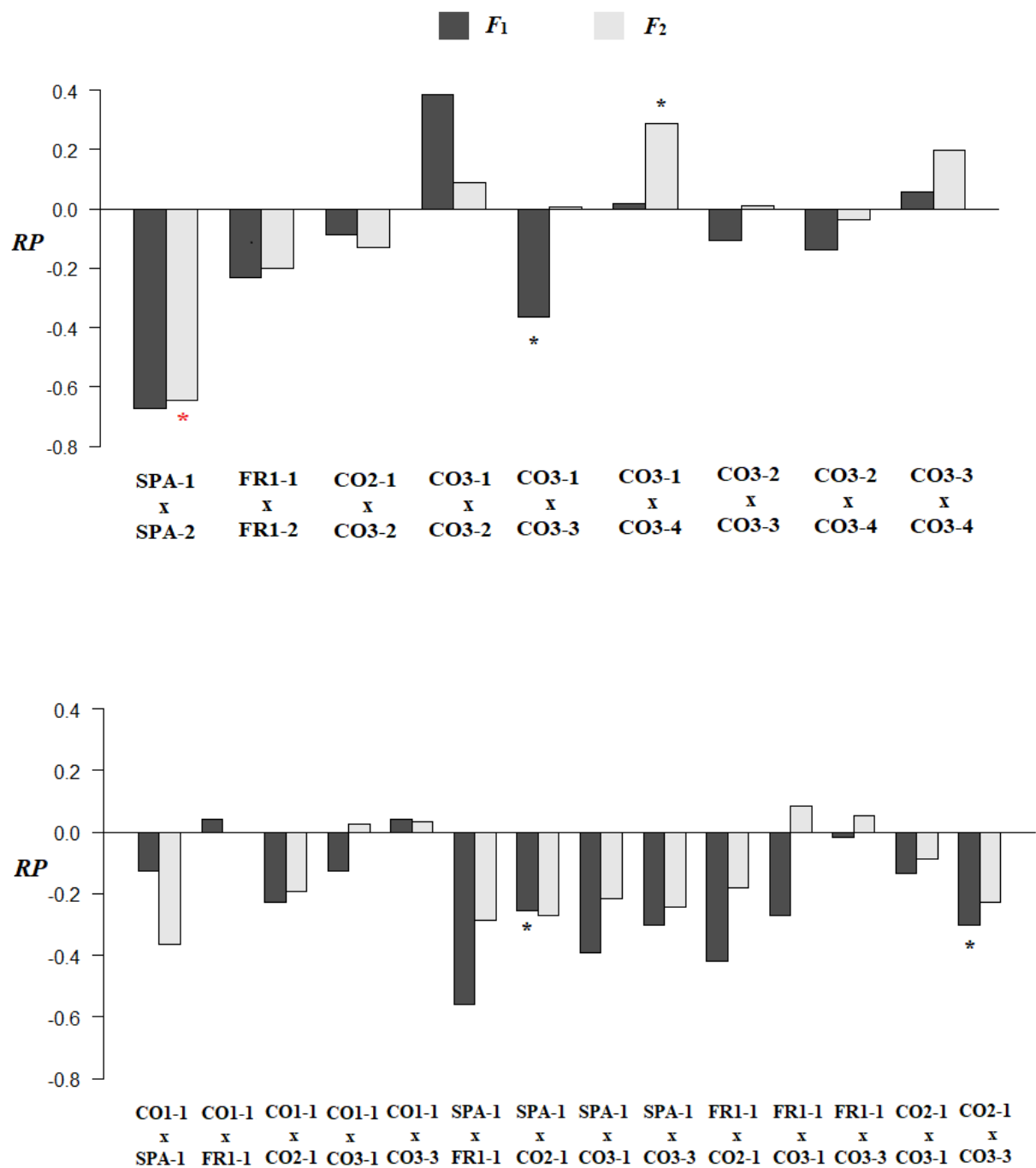


Figure 3. Relative performance (*RP*) of F_1 (dark grey) and F_2 (light grey) hybrids for the number of seeds, for each family, at the within-population (top panel) and among-population scale (bottom panel). Significant difference between hybrids and the mid-parent (black asterisks) and worst parent (red asterisks): * $p < 0.05$.

Genetic architecture of heterosis and outbreeding depression

Family-based analyses on both traits give very similar and concordant results (Table 2). For vegetative dry mass, we reject the fully additive model for 6 families out of the 9 exhibiting significant RP values. For seed production, we reject the fully additive model for all the families (Table 2). As explained above in the Material and Methods section, it is not possible to distinguish between hybrid dysfunction due to dominance or dominant DMI. For vegetative dry mass, we find that the increase observed in hybrids is always due to dominance or DMI incompatibilities (as the reduction between F_1 and F_2 of the deviation from mid-parent values is close to 0.5), except for the CO1-1-CO3-3 crosses, for which the two more complex models perform equally well. For the number of seeds produced per individual, the only case of significant heterosis is attributable to recessive DMI (Table 2), whereas outbreeding depression is due to either negative dominance or dominant DMI effect (Table 2), or a negative effect of dominance associated with a positive effect of either recessive or dominant DMI (Table 2, CO3-1 x CO3-3 family). Among the three families for which we observed a significant *RP* for both dry mass and seed number, two display an increase of biomass and a reduction of the number of seeds due to dominance or dominant DMI (SPA-1 x SPA-2 and CO2-1 x CO3-3 families, Table 1). The remaining family shows an increase in biomass due to either dominance or dominant DMI, and a reduction of the number of seeds due to a negative effect of dominance associated with a positive effect of either recessive or dominant DMI (CO3-1 x CO3-3 family, Table 2).

For most of the other families for which the difference between mid-parent and hybrids is not significant, we nevertheless detect a reduction between F_1 and F_2 of the deviation from mid-parent values by 0.5, which suggests that dominance or dominant DMI are also the most plausible mechanisms for explaining hybrid performance in these families (Figure S5).

Table 2. Summary of the family-based analyses to dissect the genetic architecture of hybrid performance for the studied traits and families where at least one generation of hybrids is significantly different from the mid-parent value. *D* indicates that the deviation between hybrid and mid-parent phenotypes is due to dominance and *D-DMI* and *R-DMI* to dominant or recessive DMI. “A” indicates that the best model is the purely additive one. “*” indicates significant relative performance.

Fitness proxy	Cross	Best model	RP – F1	RP - F2
Dry mass (g)	SPA-1 x SPA-2	<i>D</i> or <i>D-DMI</i>	0.33*	0.22*
	CO3-1 x CO3-3	<i>D</i> or <i>D-DMI</i>	0.44*	0.33*
	CO3-1 x CO3-4	A	0.12*	0.06
	CO1-1 x CO2-1	<i>D</i> or <i>D-DMI</i>	0.30*	0.23*
	CO1-1 x CO3-1	<i>D</i> or <i>D-DMI</i>	0.30*	0.19*
	CO1-1 x CO3-3	<i>D+D-DMI</i> or <i>D+R-</i>	0.28*	0.29*
	<i>DMI</i>			
	SPA-1 x CO3-3	A	0.32*	0.20*
	CO2-1 x CO3-1	A	0.17	0.15*
	CO2-1 x CO3-3	<i>D</i> or <i>D-DMI</i>	0.47*	0.37*
Number of seeds	SPA-1 x SPA-2	<i>D</i> or <i>D-DMI</i>	-0.67	-0.65*
	CO3-1 x CO3-3	<i>D+D-DMI</i> or <i>D+R-</i>	-0.37*	0.00
	<i>DMI</i>			
	CO3-1 x CO3-4	<i>R-DMI</i>	0.02	0.29*
	SPA-1 x CO2-1	<i>D</i> or <i>D-DMI</i>	-0.25*	-0.28
	CO2-1 x CO3-3	<i>D</i> or <i>D-DMI</i>	-0.32*	-0.23

Dominance and outbreeding depression

We find a significant correlation between the phenotypic similarity among parental lines and both the ratio $\frac{\bar{F}_1}{\max(\bar{P}_1, \bar{P}_2)}$ for dry mass and $\frac{\bar{F}_1}{\min(\bar{P}_1, \bar{P}_2)}$ for seed number (Figure 4). The relationship is positive for dry mass (Figure 4A, $F_{1,21} = 13.9$, $p = 0.001$, $R^2 = 0.37$), indicating that the more similar the parents, the higher the probability that hybrids are heavier than the heaviest parental line. We find a negative relationship for seed number (Figure 4B, $F_{1,21} = 22.7$, $p = 0.0001$, $R^2 = 0.50$), indicating that the more similar the parents, the higher the probability that hybrids are less fit than the worst parental line.

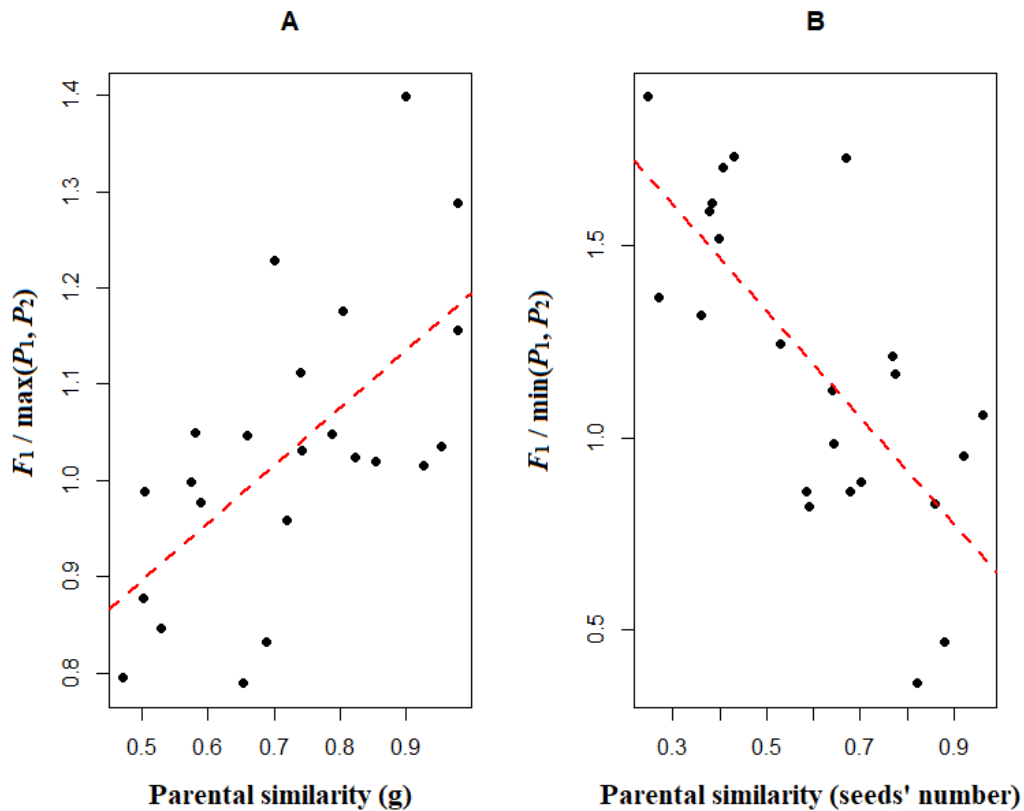


Figure 4. Relationship between the ratio of $\bar{F} / \max(\bar{P}_1, \bar{P}_2)$ for the dry mass (A) or $\bar{F}_1 / \min(\bar{P}_1, \bar{P}_2)$ for the number of seeds per individual (B), and the phenotypic similarity (computed as $\min(\bar{P}_1, \bar{P}_2) / \max(\bar{P}_1, \bar{P}_2)$) between the corresponding parental inbred lines.

DISCUSSION

In this study, we have demonstrated that hybridization events between multi-locus genotypes of *Medicago truncatula* are generally deleterious in terms of seed number and that this effect occurs for crosses done at both the within- and between -populations scales. We have dissected the genetic basis of hybrid performances, and we found that dominance is probably the main driver, even if it is not possible to distinguish between dominance and dominant Dobzhansky-Müller incompatibilities. Outbreeding depression caused by dominance is generally attributed to under-dominance and chromosomal rearrangement, in which case it is expected to have drastic effects on hybrids fitness, just as epistatic incompatibilities. In this experiment, we found that dominance could be the driver of outbreeding depression at the within-population scale, and generally lead to a moderate level of outbreeding depression (deviation of 20% on average compared to mid-parent values in F_1). This challenges the hypothesis that under-dominance and chromosomal rearrangement are the main drivers of hybrid dysfunction in *Medicago truncatula*. We propose an alternative explanation involving directional dominance that reconcile both the moderate level of outbreeding depression and the observation of mid- and worst-parent negative heterosis.

Outbreeding depression and dominance

Although comparing fitness across F_1 and F_2 generations of hybrids is not sufficient to distinguish the effects of dominance and dominant DMI, a series of arguments speak in favour of a leading role of dominance. First, it is commonly assumed that genetic incompatibilities should lead to sterility or sub-sterility, which is not the case in our crosses, and second, the loci underlying incompatibility are generally recessive, with minor effects in the F_1 hybrids while

leading to strong deleterious effects in F_2 progenies (Fishman and Willis, 2001; Presgraves, 2003; Tao and Hartl, 2003; Moyle and Graham, 2005; Moyle and Nakazato, 2008; Fitzpatrick, 2008b). Since we did not observe this pattern in our data, we can conclude that all the outbreeding depression patterns are not solely due to DMI incompatibilities and that dominance plays a role in hybrid dysfunction in *Medicago truncatula*.

Our study is not the first to attribute the origin of outbreeding depression in crosses between selfing lineages to dominance. Oakley *et al.* (2015) used a similar line cross design in natural accessions of *Arabidopsis thaliana*, and found that dominance is the main factor explaining outbreeding depression. Other studies found a decrease of performance in F_1 hybrids in predominantly selfing species (Dolgin *et al.*, 2007; Smith *et al.*, 2011; Gimond *et al.*, 2013); however without several generations of hybrids or the identification of candidate genes, it is impossible to confidently assign the observed outbreeding depression to dominance or negative epistatic incompatibilities. As shown in this study, the mode of action of dominance is also difficult to distinguish. The role of dominance has always been examined from the point of view of under-dominance or chromosomal rearrangement. Directional dominance can explain both the joint observation of mid- and worst-parent outbreeding depression, and the strong variability of outbreeding depression observed in *M. truncatula* crosses, as potentially in other predominantly selfing species. As a next step, quantifying the number of genes (major-effect loci or generalized effect across many loci), and their mode of action (directional dominance vs under-dominance vs epistatic incompatibilities) would allow to distinguish between all the above-mentioned models of outbreeding depression. Performing the same kind of experiments in natural populations rather than in a greenhouse would also be informative, because the directional dominance hypothesis assumes that traits are under stabilizing selection for an environmental optimum, and measuring the plants in a different environment (greenhouse) may alter the shape of the selection gradient.

Outbreeding depression in predominantly selfing species: within vs among population scales

As mentioned above, a loss of fitness following crosses between lines in predominantly selfing populations has been reported frequently at the among-population scale (Rhode and Cruzan, 2005; Heiser and Shaw, 2006; Johansen-Morris and Latta, 2006; Dolgin *et al.*, 2007; Gimond *et al.*, 2013; Oakley *et al.*, 2015; Vasseur *et al.*, 2019). Yet, selfing lineages are expected to accumulate genetic differentiation/incompatibilities even within populations, since genetic exchanges between plants are rare or absent. In this study, we found no significant difference between hybrids generated through crosses within- or among-populations, suggesting that genetic differentiation and incompatibilities can accumulate among selfing lineages even at local spatial scales. Few data comparing outbreeding depression at different spatial scales are available in the literature. Gimond *et al.* (2013) found that hybrids of *Caenorhabditis elegans* generated by crossing lines from the same population experience smaller outbreeding depression compared to among-populations hybrids. In the same species, Dolgin *et al.* (2007) found that within-population hybrids perform as badly as among-population hybrids, but performed no statistical tests. Rhode and Cruzan (2005) found a similar pattern in *Priquetia caroliniana*. Such results raise the question of defining what a selfing population is. In the extreme case, we could consider each MLG as a single population that interacts with other MLGs in a meta-population network. This result emphasises that endangered selfing populations require cautious management, as genetic rescue could be highly deleterious, even for individuals living in similar environments (but see Frankham, 2015, 2016). Finally, it is worth pointing out that the genetic architecture of hybrid performances seems uniform across spatial scales in *M. truncatula*, with dominance being the main driver of hybrid dysfunction at both scales.

Consequences for adaptation in predominantly selfing species

Self-fertilization is hypothesized to be an evolutionary dead-end (Takebayashi and Morrell, 2001; Iqbal and Busch, 2013). One argument to support this hypothesis is that phylogenetic transitions from outcrossing to predominant selfing are virtually irreversible (Barrett *et al.*, 1996; Escobar *et al.*, 2010; Harkness *et al.*, 2019). This is due to the fact that the evolution of selfing is associated with a diminution of inbreeding depression through the purging of deleterious mutations, which then allows evolution towards even higher selfing rates (Lande and Schemske, 1985). Predominant selfing leads to an organisation of populations in multi-locus-genotypes, which can be adapted to local conditions by combining different sets of alleles (Lande and Porcher, 2015; Abu Awad and Roze, 2018). Here we show that hybridization between those MLGs, both within and among populations, is mainly deleterious. This should lead to strong selection against a reversion from selfing to outcrossing. The weak additive variance for residual outcrossing found in a population of *M. truncatula* supports this idea (Jullien *et al.* unpublished data).

Nevertheless, residual allogamy could be of major importance for the adaptation of selfing populations to environmental changes. Using simulations, Clo *et al.* (Clo *et al.*, 2020) show that adaptation in predominantly selfing populations can rely on hidden genetic diversity stored into negative associations between loci (linkage disequilibrium). Rare outcrossing events occurring between differentiated MLGs are required to remobilize this genetic diversity. Yet, this theoretical model assumes a purely additive genetic determinism for the phenotype while the results of our experiment rather suggest that (directional) dominance and/or genetic incompatibilities are important within predominantly selfing populations. If hybrids between MLGs are unfit, they will be counter-selected, which will prevent the release of the hidden genetic diversity and could potentially lead to the extinction of the population. Conversely, if between-lines differentiation is due to directional dominance, and if the directionality of

dominance goes in the same direction as the environmental change, hybridization may be favourable and residual allogamy should boost adaptation (Harkness *et al.*, 2019; Clo *et al.*, 2020).

CONCLUSION

In this study, we have shown that dominance can be a major driver of phenotypic and fitness consequences of hybridization events between naturally occurring inbred lines of a predominantly selfing species. We have also shown that this pattern seems to be shared by most of the populations of *M. truncatula* used in this study. In addition, we demonstrated that hybridization consequences are as important in crosses involving MLGs from the same population and crosses involving MLGs from different populations. Our study stresses the limits of line-cross analyses to dissect the genetic architecture of hybrid performance, and the need to perform genomic and association genetics experiments for more accurate results. Altogether, our results bring to light the relevance of considering dominance as a driver of the evolution of self-fertilization and of the adaptive potential of inbred populations, but also the necessity to consider dominance in theoretical and experimental studies at both phenotypic and fitness scales.

Acknowledgments: We would like to thank D. Abu Awad, L. Marie-Orleach and T. Bataillon, for their comments on previous versions of this manuscript, M. Mousset, S. Glémin, P. David, O. Ronce and C. Devaux for helpful discussions. We also would like to thank D. Tauzin, P.

Noël, P. Joly and K. Loridon for the technical support. This work was supported by project SEAD (ANR-13-ADAP-0011).

Competing interests: We have no competing interests to declare.

Data archiving: The complete dataset and codes will be deposit on DRYAD upon acceptance.

BIBLIOGRAPHY

- Abu Awad D, Roze D (2018). Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution* **72**: 751–769.
- Barrett SCH, Arunkumar R, Wright SI (2014). The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**: 20130344.
- Barrett SCH, Harder LD, Worley AC (1996). The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **351**: 1271–1280.
- Charlesworth B (1992). Evolutionary Rates in Partially Self-Fertilizing Species. *The American Naturalist* **140**: 126–148.
- Clo J, Ronfort J, Abu Awad D (2020). Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. *Journal of Evolutionary Biology* **n/a**.
- Coyne JA, Orr HA (2004). *Speciation*. Sinauer Associates Sunderland, MA.
- Crow JF (1948). Alternative hypotheses of hybrid vigor. *Genetics* **33**: 477.
- Demuth JP, Wade MJ (2005). On the Theoretical and Empirical Framework for Studying Genetic Interactions within and among Species. *The American Naturalist* **165**: 524–536.
- Dobzhansky T (1937). *Genetics and the origin of species*. NY: Columbia Univ. Press. 364 p.

- Dolgin ES, Charlesworth B, Baird SE, Cutter AD (2007). INBREEDING AND OUTBREEDING DEPRESSION IN CAENORHABDITIS NEMATODES. *Evolution* **61**: 1339–1352.
- Escobar JS, Cenci A, Bolognini J, Haudry A, Laurent S, David J, *et al.* (2010). An integrative test of the dead-end hypothesis of selfing evolution in Triticeae (Poaceae). *Evolution: International Journal of Organic Evolution* **64**: 2855–2872.
- Fishman L, Willis JH (2001). Evidence for Dobzhansky-Muller incompatibilities contributing to the sterility of hybrids between *Mimulus guttatus* and *M. nasutus*. *Evolution* **55**: 1932–1942.
- Fitzpatrick BM (2008a). Hybrid dysfunction: population genetic and quantitative genetic perspectives. *The American Naturalist* **171**: 491–498.
- Fitzpatrick BM (2008b). Dobzhansky–Muller model of hybrid dysfunction supported by poor burst-speed performance in hybrid tiger salamanders. *Journal of Evolutionary Biology* **21**: 342–351.
- Frankham R (2015). Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. *Molecular ecology* **24**: 2610–2618.
- Frankham R (2016). Genetic rescue benefits persist to at least the F3 generation, based on a meta-analysis. *Biological Conservation* **195**: 33–36.
- Gimond C, Jovelín R, Han S, Ferrari C, Cutter AD, Braendle C (2013). OUTBREEDING DEPRESSION WITH LOW GENETIC VARIATION IN SELFING CAENORHABDITIS NEMATODES. *Evolution* **67**: 3087–3101.
- Glémin S (2003). How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57**: 2678–2687.
- Glémin S, Ronfort J (2013). Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evolution: International Journal of Organic Evolution* **67**: 225–240.
- Harkness A, Brandvain Y, Goldberg EE (2019). The evolutionary response of mating system to heterosis. *Journal of Evolutionary Biology* **32**: 476–490.
- Igic B, Busch JW (2013). Is self-fertilization an evolutionary dead end? *New Phytologist* **198**: 386–397.
- Johansen-Morris AD, Latta RG (2006). Fitness consequences of hybridization between ecotypes of *Avena barbata*: hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution* **60**: 1585–1595.
- Jullien M, Navascués M, Ronfort J, Loricón K, Gay L (2019). Structure of multilocus genetic diversity in predominantly selfing populations. *Heredity* **123**: 176–191.
- Kimura M, Maruyama T, Crow JF (1963). The mutation load in small populations. *Genetics* **48**: 1303.

- Kirkpatrick M, Barton N (2006). Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419–434.
- Lande R (1985). The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity* **54**: 323–332.
- Lande R, Porcher E (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* **200**: 891–906.
- Lande R, Schemske DW (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**: 24–40.
- Lynch M (1991). The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45**: 622–629.
- Moyle LC, Graham EB (2005). Genetics of hybrid incompatibility between *Lycopersicon esculentum* and *L. hirsutum*. *Genetics* **169**: 355–373.
- Moyle LC, Nakazato T (2008). Comparative genetics of hybrid incompatibility: sterility in two *Solanum* species crosses. *Genetics* **179**: 1437–1453.
- Oakley CG, Ågren J, Schemske DW (2015). Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. *Heredity* **115**: 73.
- Presgraves DC (2003). A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics* **163**: 955–972.
- Rhode JM, Cruzan MB (2005). Contributions of Heterosis and Epistasis to Hybrid Fitness. *The American Naturalist* **166**: E124–E139.
- Schierup MH, Christiansen FB (1996). Inbreeding depression and outbreeding depression in plants. *Heredity* **77**: 461–468.
- Schneemann H, De Sanctis B, Roze D, Bierre N, Welch JJ (2020). The geometry and genetics of hybridization. *bioRxiv*: 862235.
- Siol M, Prosperi JM, Bonnin I, Ronfort J (2008). How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in *Medicago truncatula*. *Heredity* **100**: 517.
- Smith LM, Bomblies K, Weigel D (2011). Complex evolutionary events at a tandem cluster of *Arabidopsis thaliana* genes resulting in a single-locus genetic incompatibility. *PLoS genetics* **7**.
- Takebayashi N, Morrell PL (2001). Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany* **88**: 1143–1150.
- Tao Y, Hartl DL (2003). GENETIC DISSECTION OF HYBRID INCOMPATIBILITIES BETWEEN *DROSOPHILA SIMULANS* AND *D. MAURITIANA*.: III. HETEROGENEOUS ACCUMULATION OF HYBRID INCOMPATIBILITIES,

DEGREE OF DOMINANCE, AND IMPLICATIONS FOR HALDANE—RULE.
Evolution **57**: 2580–2598.

Vasseur F, Fouqueau L, de Vienne D, Nidelet T, Violle C, Weigel D (2019). Nonlinear phenotypic variation uncovers the emergence of heterosis in *Arabidopsis thaliana*. *PLoS biology* **17**: e3000214.

Volis S, Shulgina I, Zaretsky M, Koren O (2010). Epistasis in natural populations of a predominantly selfing plant. *Heredity* **106**: 300.

Wright SI, Kalisz S, Slotte T (2013). Evolutionary consequences of self-fertilization in plants. *Proc R Soc B* **280**: 20130133.

Yamaguchi R, Otto SP (2020). Insights from Fisher's geometric model on the likelihood of speciation under different histories of environmental change. *Evolution* **n/a**.

Matériel supplémentaire

Annexe 1: Distinguish epistatic incompatibilities and dominance in line cross analyses.

In order to distinguish dominance effects from Dobzhansky-Müller incompatibilities (thereafter called DMI), we use the theoretical framework introduced by Fitzpatrick (2008) to incorporate DMI into line-cross analysis in quantitative genetics (Demuth and Wade 2005). We detail here how Fitzpatrick's model can efficiently distinguish the effects of dominance versus epistatic incompatibilities on hybrid performance when DMI are recessive, but is unable to distinguish between the two effects when DMI are dominant.

Let $\bar{P}$ denote the mid-parent value for a phenotypic trait, and D , e_{dd} , e_{da} , e_{da} , and e_{aa} be the deviations in F_1 and F_2 from mid-parent phenotypes due to respectively, dominance, dominance-by-dominance, dominance-by-additive, additive-by-dominance and additive-by-additive epistatic interactions. Assuming that incompatibilities are not present in parental lines, the mean of F_1 and F_2 are respectively

$$\bar{F}_1 = \bar{P} + D + e_{da}$$

and

$$\bar{F}_2 = \bar{P} + \frac{D}{2} + \frac{e_{dd}}{4} + \frac{e_{aa}}{16} + \frac{e_{ad} + e_{da}}{8}.$$

Hence, if epistatic interactions are negligible compared to dominance effects, the distance between the F_1 and the mid-parent values is divided by two in the F_2 . On the contrary, if dominance is negligible compared to epistatic interactions, the expectation changes depending whether DMI are recessive or dominant. Under partially recessive DMI (when one gene action

is dominant over the other, i.e. $e_{dd} \ll e_{da}, e_{da} < e_{aa}$, figure 1B in Fitzpatrick 2008), then $\bar{F}_1$ is expected to be close to $\bar{P}$, but $\bar{F}_2$ becomes increasingly different from $\bar{P}$ when additive-by-additive, dominance-by-additive and additive-by-dominance epistatic effects increase. In this case, the ratio $(\bar{F}_1 - \bar{P}) / (\bar{F}_2 - \bar{P})$ allows distinguishing outbreeding depression due to dominance from outbreeding depression due to partially recessive incompatibilities (Figure A1-A).

Nevertheless, the distinction becomes harder for dominant incompatibilities ($e_{dd} = e_{da}, e_{da} = e_{aa}$, figure 1A in Fitzpatrick 2008), as in this case the means become

$$\bar{F}_1 = \bar{P} + D + e_{dd}$$

and

$$\bar{F}_2 = \bar{P} + \frac{D}{2} + \frac{9}{16} e_{da}.$$

Hence, the distance between mid-parent and $\bar{F}_1$ phenotypes is still divided by two in F_2 if dominance is strong compared to epistasis, but is divided by 9/16 (=0.5625), when incompatibilities are strong compared to dominance (Figure A1-B). As the ratio $(\bar{F}_1 - \bar{P}) / (\bar{F}_2 - \bar{P})$ is very close under both scenarios, it prevents distinguishing between these two genetic architectures.

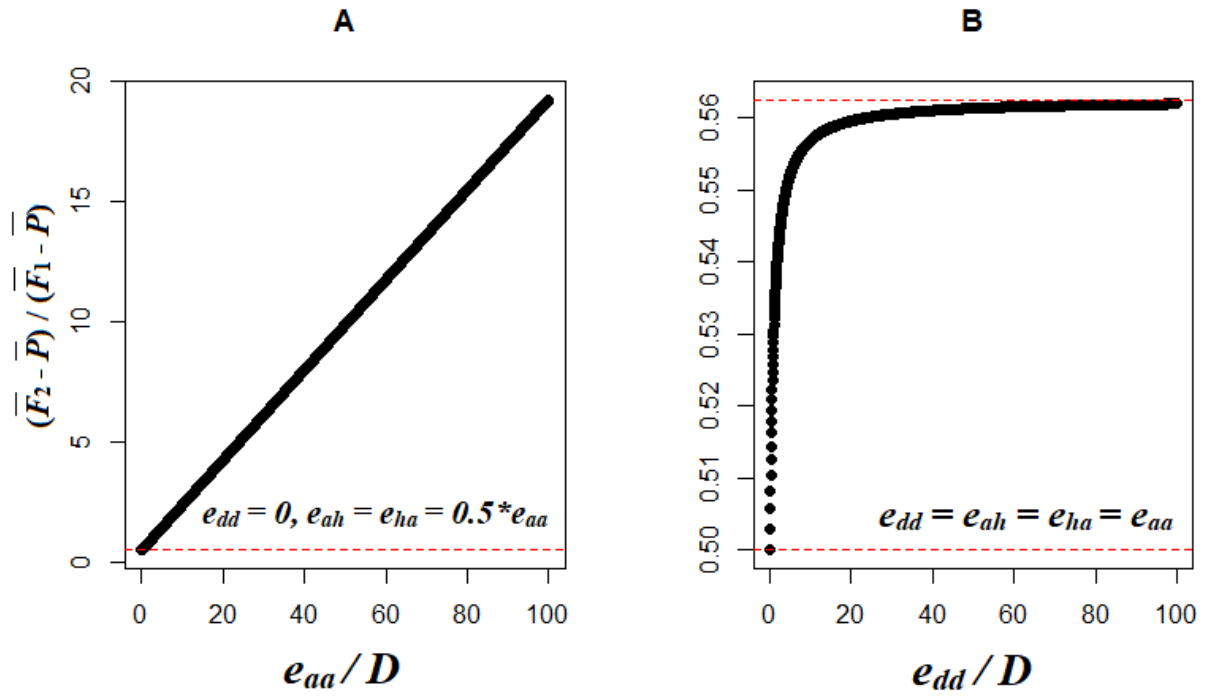


Figure A1. Distance between phenotypic values in F_2 and F_1 compared to mid-parent, as a function of the relative strength of epistasis compared to dominance, under two scenarios of Dobzhansky-Müller incompatibilities. D is the deviation from the mid-parent value due to dominance, and e_{dd} , e_{da} , e_{da} , and e_{aa} are deviations due to all epistatic forms. **A.** For partially recessive incompatibilities, the red line indicates a ratio of 0.5, corresponding to the case where deviations due to epistatic interactions are negligible compared to dominance. **B.** For dominant incompatibilities, the red lines are centred on 0.5 and 0.5625 respectively, the two limits of the ratio under this scenario, corresponding respectively to the case where deviations due to epistatic interactions are negligible compared to dominance, and *vice versa*.

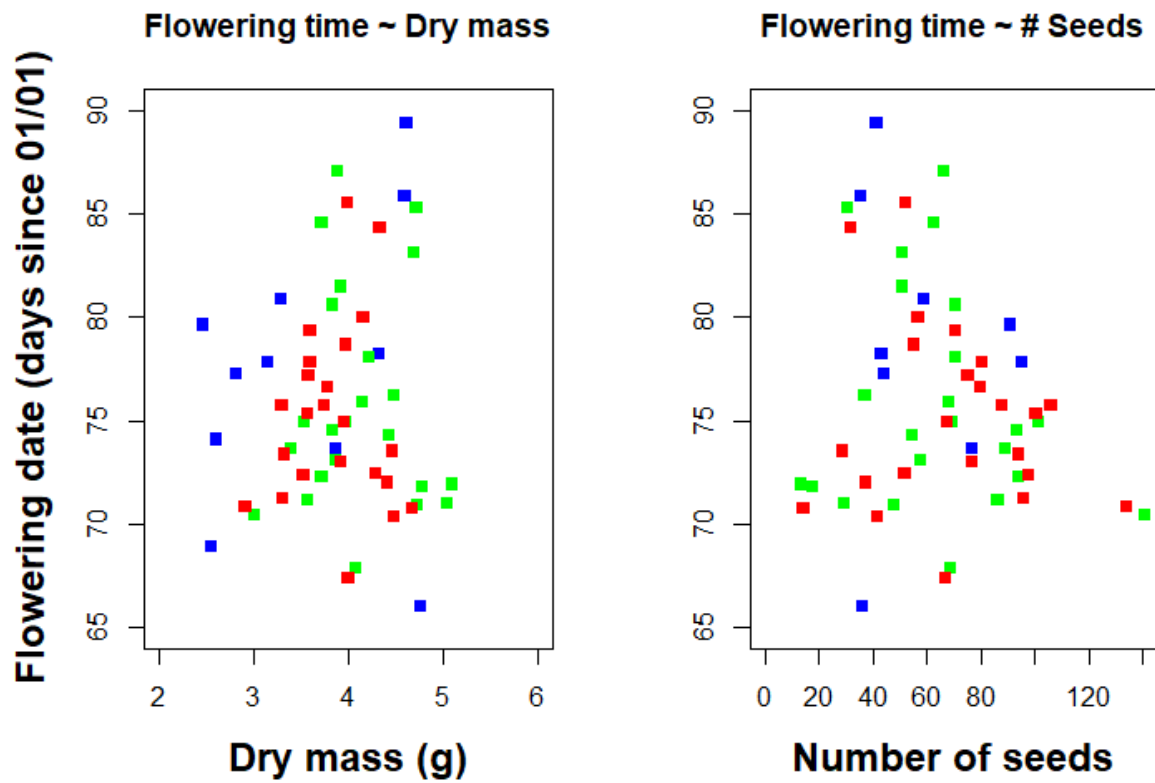


Figure S1. Relationship between either flowering date and the vegetative dry mass or the number of seeds. Blue points represents parental lines, green points are F_1 families, and red points are F_2 families. There are no significant relationships between flowering date and the two fitness proxies, for any generations ($p > 0.2$ in all cases).

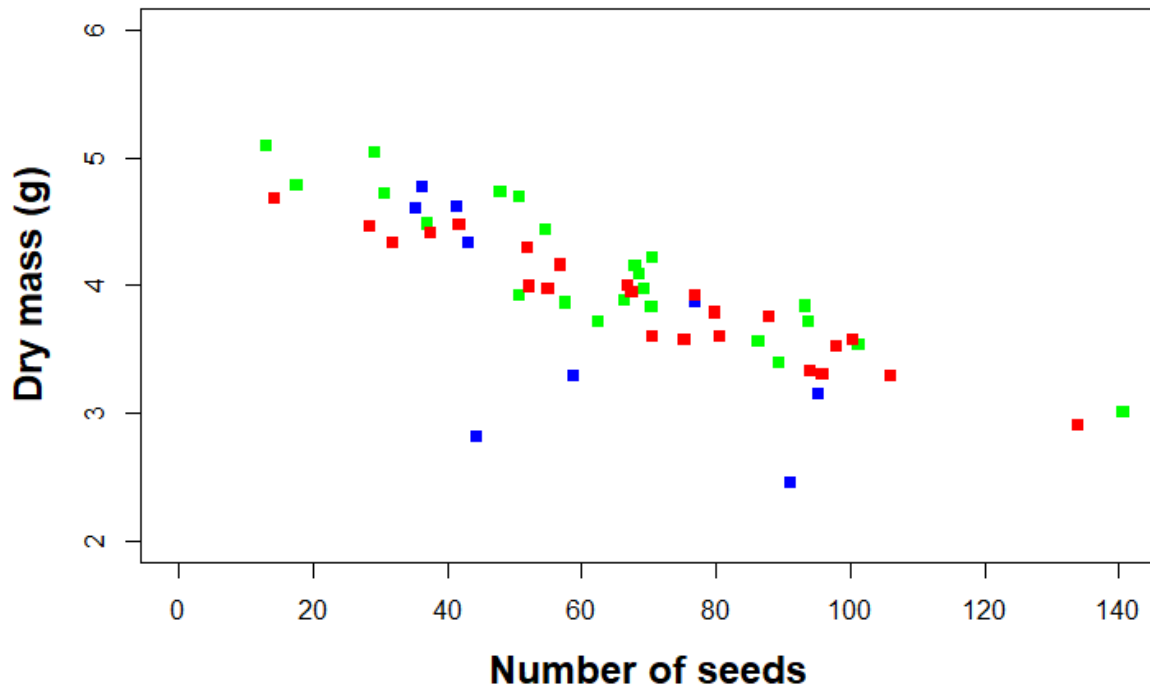


Figure S2. Relationship between the vegetative dry mass and the number of seeds. Blue points represents parental lines, green points are F_1 families, and red points are F_2 families. There are significant relationships between both variables in all generations ($p < 0.01$ in all cases), but the slope of regression does not differ between generations ($p > 0.3$).

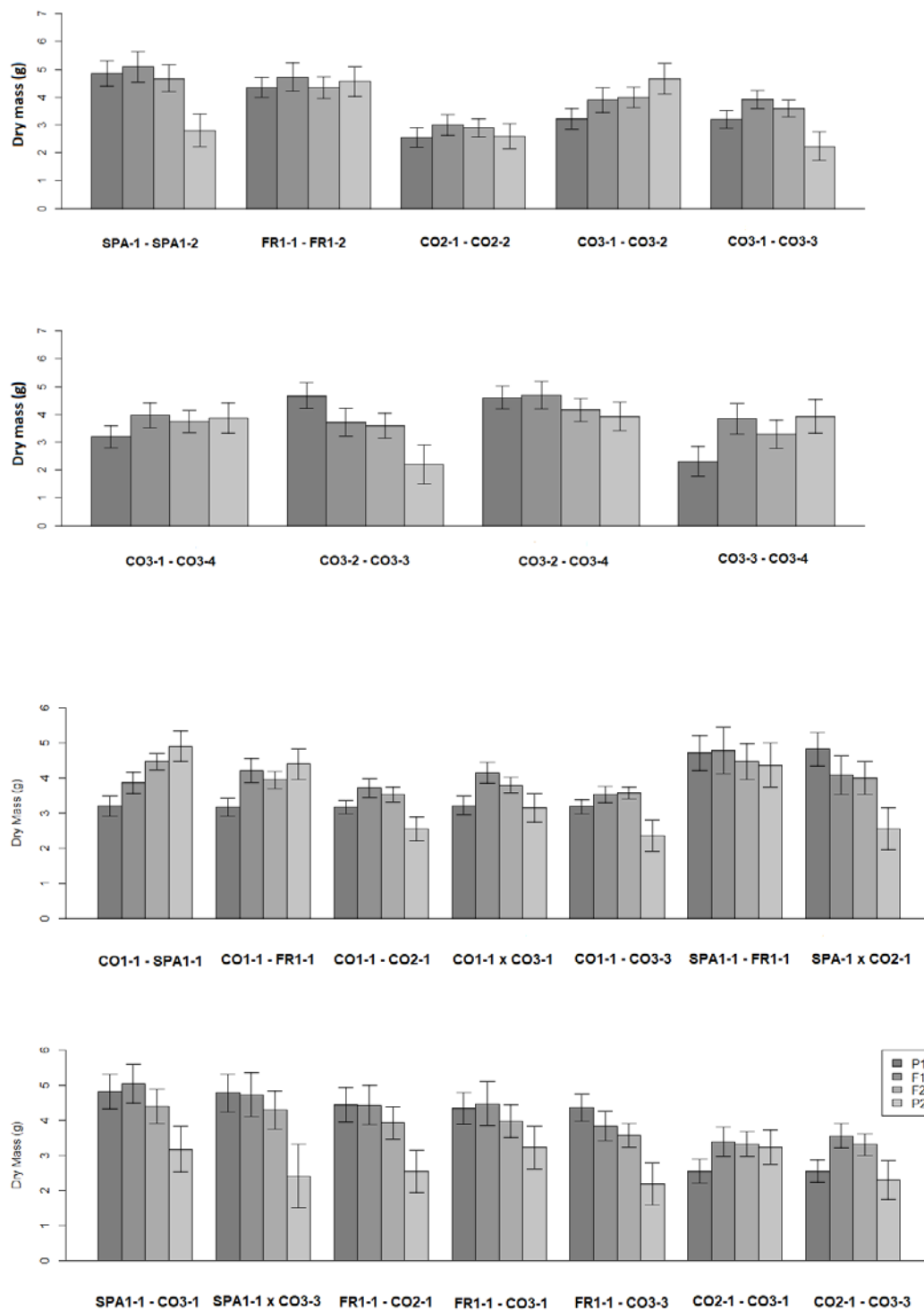


Figure S3. Dry mass (g) for each generation, for all families, for both within and among population crosses. Error bars stand for one standard deviation.

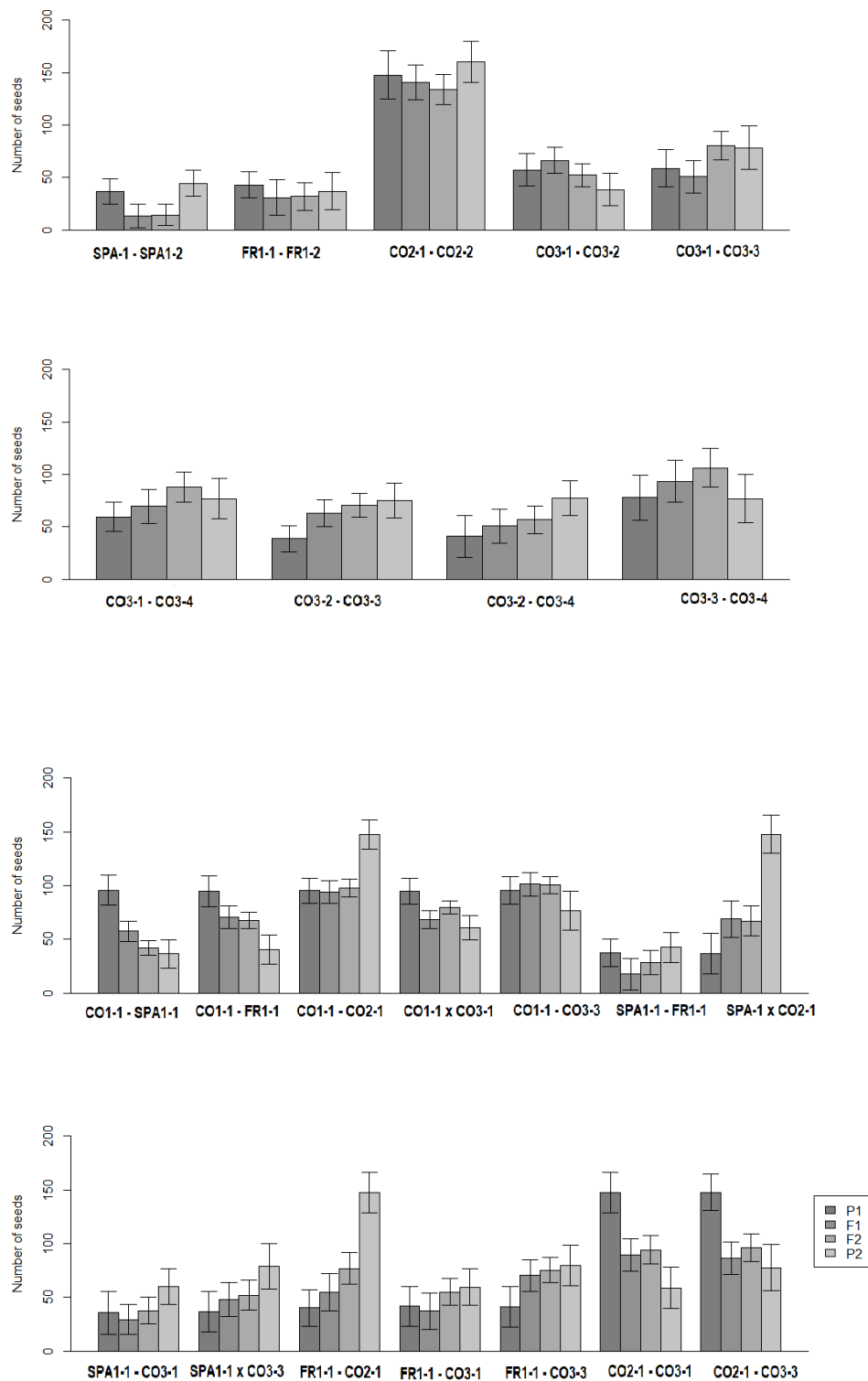


Figure S4. Number of seeds for each generation, for all families, for both within and among population crosses. Error bars stand for one standard deviation.

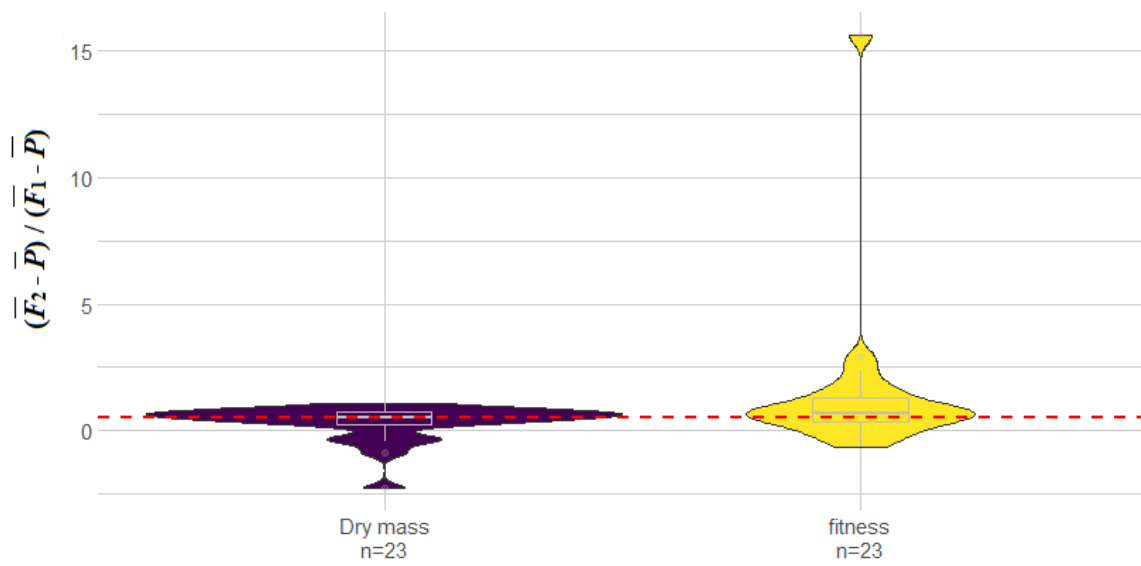


Figure S5. Violin plots and associated boxplots (in grey) of the D values for the dry mass and fitness. The red horizontal dashed line corresponds to $y = 0.5$, the theoretical value for which directional dominance is the main factor explaining the departure of $\bar{F}_1$ and $\bar{F}_2$ from the mid-parent value.

Table S1. Matrix of genetic differentiation between MLGs. The genetic differentiation among these MLGs was measured using the 20 microsatellite markers, as the number of alleles that differ between two MLGs divided by the total number of alleles without missing data

	CO3-1	CO3-2	CO3-3	CO3-4	FRA-1	FRA-2	CO2-1	CO2-2	SPA1-1	SPA1-2	CO1-1	CO1-2
CO3-1	0	0.14	0.5	0.71	0.71	0.78	0.78	0.78	0.93	0.93	0.75	0.92
CO3-2		0	0.57	0.71	0.57	0.71	0.78	0.78	0.93	0.93	0.67	0.83
CO3-3			0	0.86	0.71	0.78	0.78	0.78	0.86	0.86	0.91	1
CO3-4				0	0.78	0.78	0.71	0.78	0.86	0.86	0.91	1
FRA-1					0	0.44	0.81	0.81	0.82	0.82	0.77	0.77
FRA-2						0	0.94	0.88	0.82	0.82	0.77	0.92
CO2-1							0	0.12	0.89	0.89	0.69	0.84
CO2-2								0	0.89	0.89	0.69	0.84
SPA1-1									0	0.42	0.93	0.93
SPA1-2										0	0.87	0.87
CO1-1											0	0.54
CO1-2												0

Table S2. Summary of all the crosses performed between MLGs. Light grey represents self-fertilization events (Parental lines) on the diagonal. Mid-shaded grey represents within-population hybridization events between MLGs (10 families). Dark-shaded grey represents between-populations hybridization events between MLGs (14 families).

		pollen	CO1	CO1	SPA1	SPA1	FR1	FR1	CO2	CO2	CO3	CO3	CO3	CO3
			1	2	1	2	1	2	1	2	1	2	3	4
ovule														
CO1	1													
CO1	2													
SPA1	1													
SPA1	2													
FR1	1													
FR1	2													
CO2	1													
CO2	2													
CO3	1													
CO3	2													
CO3	3													
CO3	4													

Table S3. Summary of AICs for each model and each family for which at least one generation of hybrids (F_1 and F_2) is significantly different from mid-parent value, for the dry mass. Best model(s) have bold AIC(s).

Cross ID	A	A+D	A+D- DMI	A+R- DMI	A+D + D- DMI	A+D + R- DMI
SPA-1 x SPA-2	198	192	191	200	193	193
CO3-1 x CO3-3	241	231	230	243	231	231
CO3-1 x CO3-4	280	281	281	282	283	283
CO1-1 x CO2-1	306	297	296	305	297	297
CO1-1 x CO3-1	320	312	311	321	313	313
CO1-1 x CO3-3	277	273	272	274	270	270
SPA-1 x CO3-3	322	321	321	324	323	323
CO2-1 x CO3-1	260	260	260	262	261	261
CO2-1 x CO3-3	260	254	253	262	253	253

Table S4. Summary of AICs for each model and each family for which at least one generation of hybrids (F_1 and F_2) is significantly different from mid-parent value, for the number of seeds. Best model(s) have bold AIC(s).

Cross ID	A	A+D	A+D- DMI	A+R- DMI	A+D + D- DMI	A+D + R- DMI
SPA-1 x SPA-2	559	552	551	558	551	551
CO3-1 x CO3-3	925	920	921	921	918	918
CO3-1 x CO3-4	932	934	924	928	930	930
SPA-1 x CO2-1	933	931	931	932	931	931
CO2-1 x CO3-3	971	967	967	973	968	968

Chapitre 4 : Au-delà de l'additivité, de l'importance de considérer l'effet de la dominance sur le niveau et la structure de la diversité génétique des populations

This article is under review in *Evolution*.

Présentation générale

In the different chapters of this thesis, I showed that the theoretical relationship between adaptive potential and self-fertilization is not well supported empirically, and that directional dominance seems likely to be of major importance when considering the adaptive potential of predominantly selfing species. This last point could explain the discrepancies between theory and data, since the theoretical models of quantitative traits under stabilizing selection make the strong assumption of pure additivity at the phenotypic scales (Schneemann *et al.* 2020). Even if these models are surprisingly able to fit well with biological processes (like the average dominance level of deleterious mutations, or the distribution of epistatic interactions among alleles, Martin *et al.*, 2007; Manna *et al.*, 2011), assuming a more realistic genetic architecture of quantitative traits and fitness could help to reconcile theory and data. In addition, when dominance occurs with inbreeding, the response to selection is more complex to predict, with new non-additive terms contributing to the adaptive potential of populations (Kelly, 1999). Estimations of such components remain rare in the literature (see Wolak & Keller 2014 for a review), and little is known about their contribution to the genetic variance for quantitative traits under stabilizing selection.

In this last chapter, I relaxed this assumption and explored the effect of directional dominance on a fitness-related trait under stabilizing selection for populations that differ in selfing rates. By decomposing the genetic variance into additional components specific to inbred populations, I showed that dominance components could make substantial contributions to the genetic variance of inbred populations. I also showed that ignoring these components leads to an upward bias in the predicted response to selection. Finally, I showed that when considering the effect of directional dominance, the evolutionary potential of populations remains comparable across the entire gradient in outcrossing rates, at both short- and long-term, reconciling theoretical and empirical results.

These findings suggest that other ecological or evolutionary mechanisms need to be invoked to explain the reduced evolutionary success of (partially) selfing species. All the results and their implications are discussed in the following section.

**Genetics of quantitative traits with directional dominance in partially selfing species:
implications for evolution in changing reproductive environments**

Josselin CLO ¹, Øystein OPEDAL ²

¹ AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, 34000 Montpellier, France.

² Department of Biology, Lund University, SE-223 62, Lund, Sweden.

Authors contribution : J.C. wrote the simulation code and compiled empirical data. J.C. and Ø.O. analysed the results and wrote the manuscript.

Abstract: Self-fertilization is thought to leads to reduced adaptive potential, thus leading selfing lineages down an evolutionary ‘blind alley’. The original argument is that, because selfing populations are expected to harbour less genetic diversity, their capacity for adaptation in the event of an environmental change is reduced. Though well supported theoretically, empirical support for reduced adaptability in selfing species is limited. One limitation of classical theoretical models is that they assume pure additivity of the fitness-related traits that are under stabilizing selection, despite ample evidence that quantitative traits are subject to directional dominance. Here we relax this assumption and explore the effect of directional dominance on a fitness-related trait under stabilizing selection for populations that differ in selfing rates. By decomposing the genetic variance into additional components specific to inbred populations, we show that dominance components can explain a substantial part of the genetic variance of inbred populations. We also show that ignoring these components lead to an upward bias in the predicted response to selection. Finally, we show that when considering the effect of directional dominance, the evolutionary potential of populations remains comparable across the entire gradient in outcrossing rates, reconciling theoretical and empirical results. Our findings suggest that other ecological or evolutionary mechanisms need to be invoked to explain reduced evolutionary success of (partially) selfing species.

Keywords: adaptation, directional dominance, evolvability, quantitative genetics, self-fertilization, standing genetic variation

Introduction

The contribution of non-additive genetic components to the standing genetic variance in quantitative traits is a topic of long-standing theoretical and empirical interest. The contribution of dominance variance (variance due to interactions between alleles at the same locus) has received particular attention. For Fisher (1958), dominance variance was expected to be negligible, as non-additive genetic effects are equal to zero across an infinite number of genetic backgrounds (Fisher, 1958). In agreement with this assertion, Hill *et al.* (2008) concluded from the analysis of a bi-allelic single-locus model that additive variance always explains most of the genetic variance at mutation-drift equilibrium, notably because alleles are maintained at extreme frequencies (close to 0 or 1). Wright, on the contrary, stipulated that populations are generally of finite size, therefore non-additive genetic effects cannot be ignored (Wright, 1931). Because natural populations are expected to have low effective sizes in general (Roff, 2012), dominance could be a major component of the genetic variance in natural populations. Although dominance variance does not contribute directly to the adaptive potential of diploid organisms, neglecting its contribution to the genetic variance can lead to substantial bias when estimating additive variance (Ovaskainen *et al.*, 2008). Therefore, the study of dominance effects and their role in the genetic architecture of quantitative traits is of major importance when aiming to quantify the evolvability of quantitative traits in natural populations.

Ignoring epistasis, two components contribute to the genetic variance of a quantitative trait in a randomly mating population. These are the additive variance and the dominance variance. Conveniently, the response to selection can be predicted as the product of the additive genetic variance and the selection gradient acting on a trait (Lande, 1979). This simple model has facilitated the highly successful research programme in evolutionary quantitative genetics, which has yielded countless empirical estimates of natural selection (e.g. Siepielski *et al.*, 2017) and additive genetic variance (e.g. Hansen *et al.*, 2011; Opedal, 2019). Importantly, this model

rests on the assumption that populations mate randomly. Under partial inbreeding, predicting the response to selection becomes much more complicated. In particular, while dominance variance does not contribute to selection response under random mating, this is no longer the case under partial inbreeding (Wright and Cockerham, 1985; Kelly, 1999a). When inbreeding and directional dominance occurs simultaneously, at least three additional terms contribute to the genetic variance (Cockerham and Weir, 1984). The genetic variance of a population with an average inbreeding coefficient of F is then equal to

$$\sigma^2_{G,F} = (1 + F) \sigma^2_{\text{genic}} + (1 - F) \sigma^2_{\text{DR}} + F \sigma^2_{\text{DI}} + 4F \sigma_{\text{ADI}} + F(1 - F)H^* \quad (1)$$

where σ^2_{genic} (the variance of additive effects of alleles) and σ^2_{DR} (the variance of dominance deviations of a randomly mating population) are the classical terms most often inferred in quantitative genetics. The additional terms arising under partial inbreeding include the variance of dominance deviations of a fully inbred population σ^2_{DI} , and the covariance between the additive effect of an allele and its homozygous dominance deviation σ_{ADI} . This covariance arises because the increased homozygosity in inbred populations increases the probability that an allele occurs in its homozygote state, which creates a covariance between the contribution of an allele to the breeding value and the dominance deviation of the homozygote genotype (Wolak and Keller, 2014). Note that unlike the other terms of equation (1) this term is a covariance and can thus be negative. The third additional component, H^* , is the squared per-locus inbreeding depression, summed over all loci. This last term is expected to be small if inbreeding depression is due to the cumulative effect of a large number of loci with small effects (Shaw *et al.*, 1998). While σ^2_{DR} does not affect the response to selection, the additional dominance components (notably σ^2_{DI} and σ_{ADI}) contribute to the adaptive potential, because they contribute to the covariance between parents and offspring, on which selection response depends (Wright and Cockerham, 1985; Kelly, 1999a, 1999b).

Plant mating systems are highly variable and range from functional asexuality to obligate outcrossing (Barrett, 2003). Pure mating strategies (obligate selfing or outcrossing) are uncommon, and most plant species exhibit partial inbreeding (mixed mating systems; Goodwillie *et al.*, 2005). Within mixed-mating species, outcrossing rates often vary widely among populations (Whitehead *et al.*, 2018). This variation makes plants ideal subjects for studies assessing the evolutionary importance of non-additive genetic variance components. In the light of ongoing changes in pollinator communities (including pollinator declines), a question of particular interest is how a transition from outcrossing towards partial selfing affects the immediate and long-term adaptive potential of populations. Quantitative surveys and meta-analyses have yielded somewhat conflicting results concerning the relationship between selfing rate and adaptive potential. Charlesworth and Charlesworth (1995) compiled estimates of outcrossing rates, heritabilities, and coefficients of genetic variation from published studies of plants, and detected weak overall relationships between outcrossing rates and standing genetic variation. They did, however, detect the predicted positive relationship when restricting the analysis to additive genetic coefficients of variation. More recently, Opedal *et al.* (2017) reported only minor differences in evolvabilities (mean-scaled genetic variances) of floral traits between predominantly selfing, mixed-mating, and predominantly outcrossing plants, while Clo *et al.* (2019) detected a weak positive relationship for certain traits. These results suggest that explicit consideration of non-additive components may be needed to fully understand the adaptive potential of partially selfing populations and species.

Here, we use simulation models to study the relationship between mating systems and the genetic variance, focusing in particular on the contribution of those variance components related to directional dominance interacting with inbreeding. Specifically, by introducing dominance into classical models of quantitative traits under stabilizing selection, we dissect the genetic architecture of quantitative traits at mutation-selection-drift equilibrium. We test if

incorporating all relevant components of genetic variance into a simplified model of selection response yield accurate estimates of the rate of evolution of partially selfing populations, notably compared to purely additive models. Finally, we use the insights from our model to discuss the short- and long-term effects of self-fertilization on the adaptive potential of populations.

Material and Methods

Simulation program

The model parameters, their abbreviations, and their values considered in simulations are summarized in Table 1. We consider the evolution of a quantitative trait Z in a population of size N , comprising diploid individuals reproducing through partial self-fertilization at a constant selfing rate s . The phenotypic value z of an individual is given by

$$z = g + e, (2)$$

where e is the random environmental effect drawn from a Gaussian distribution with mean 0 and variance V_E and assumed to be independent from the genetic components. The genetic effect g is controlled by L bi-allelic loci with both additive and dominance effects and is given by $g = \sum_j^L (g_j^M + g_j^P + \delta_j)$, where g_j^M and g_j^P are the additive allelic effects at locus j inherited from the maternal and paternal gamete, respectively, and δ_j is the dominance deviation (the difference between the genotypic value and the additive value) if $g_j^M \neq g_j^P$. The additive allelic effects at the j th locus are either 0 or r_{aj} , where r_{aj} is drawn from a Gaussian distribution with mean 0 and variance a^2 . The dominance deviation at the j th locus is drawn from a Gaussian distribution of mean $D.r_{aj}$ and variance d^2 . The trait Z is subject to stabilizing selection around

an optimal phenotypic value denoted Z_{opt} . The fitness value W_z of an individual with phenotype z is thus described by the Gaussian fitness function:

$$W_z = e^{-(z-Z_{opt})^2/2\omega^2}, (3)$$

where ω^2 is the width of the fitness function and represents the strength of selection. The main effect of adding directional dominance at the phenotypic scale is to change the average coefficient of dominance for a deleterious mutation h . Under a fully additive model, $h = 0.25$ (Manna *et al.*, 2011). With directional dominance, deleterious mutations with effects acting in the opposite direction of the dominance deviation have h values ranging from 0.25 to 0 (fully recessive). On the contrary, deleterious mutations with effects acting in the same direction as the dominance have h values ranging from 0.25 and 1 (fully dominant).

We initiated each simulation from a population of genetically identical individuals at the phenotypic optimum (all loci carried alleles with effect 0 and $Z_{OPT} = 0$). The life cycle can be summarized by five successive events. First, there is a phenotype-dependent choice of the first parent (selection), followed by mating-type choice (selfing versus outcrossing at rates s and $[1-s]$, respectively). In the case of outcrossing, the choice of the second parent is also phenotype dependent. Selection takes place as follows: if the ratio of the selected parent's fitness over the highest recorded fitness value in the current generation is higher than a threshold value sampled from a uniform distribution between 0 and 1, the individual is allowed to reproduce. Once the two parents are chosen, they each contribute a gamete produced by random sampling of one allele per locus. This phase is then followed by the introduction of mutations from allele 0 to allele r_{aj} and *vice versa*, the number of which is sampled from a Poisson distribution with parameter U (with $U = \mu L$, where μ is the per-locus mutation rate).

Investigated parameter values

We ran simulations for several parameter sets (Table 1). The values chosen for the mutation rate U (0.005 and 0.05) reflect the expected per-trait haploid genomic mutation rate reported in the literature (Keightley and Bataillon, 2000; Shaw *et al.*, 2002). This also reflects two mutation regimes, one where genetic associations are negligible ($U = 0.005$), and one where they are influential ($U = 0.05$, see Clo *et al.*, 2020). We chose parameter set values similar to those explored in Bürger *et al.*, (1989) and Ronce *et al.*, (2009), with the number of freely recombining loci under selection $L = 50$, and $a^2 = 0.05$ (in agreement with empirical estimates, Lynch, 1988), $V_E = 1$, $\omega^2 = 1, 9$ or 99 . We considered three levels of dominance, with the increasing allele being always dominant, $D = 0$ (non-directional dominance), $D = 0.5$ (partial dominance) and $D = 1$ (complete dominance), and $d^2 = 0.05$. Scenarios of partial and complete recessivity are not presented as they gave results similar to those for partial and complete dominance. Simulations were run for the full range of selfing rates (s varied from 0 to 1). We also considered three population sizes $N = 250, 1000$ and 10000 , which allowed us to explore the effect of drift.

Table 1. Description of model parameters, their abbreviations and range of values considered in the simulations.

<i>Parameter</i>	<i>Abbreviation</i>	<i>Value(s)</i>
Population size	N	250, 1000, 10.000
Selfing rate	S	From 0 to 1
Number of loci	L	50
Dominance level	D	0 (non-directional), 0.5 (partial dominance), 1 (complete dominance)
Strength of stabilizing selection	ω^2	1, 9, 99
Haploid phenotypic mutation rate	U	0.005 or 0.05
Mutational variance for the quantitative trait	a^2	0.05
Variance in dominance deviations	d^2	0.05

Question 1: What is the relative contribution of each genetic variance component?

Our simulations allow us, for the first time, to evaluate the contribution of each genetic parameter of equation (1) for a quantitative trait under stabilizing selection. We computed each variance component following Cockerham & Weir (1984), and we multiplied them by their

expected contribution based on the inbreeding coefficient F (equation 1, expected and observed F did not detectably differ, as reported by Glémin, 2003). Table 2 summarizes how each component of variance is calculated. We also explored the effect of dominance on the amount of variance due to genetic associations across loci (i.e. a Bulmer effect, Bulmer, 1974). The genetic covariances (σ_{LD} hereafter) are calculated for pairs of loci because pairwise associations are expected to explain most of the among-loci component (figure S1). Hence, higher-order associations appear negligible.

As a second step, we computed the proportion of within-loci variance components that is due to additivity and dominance, following the formulae of Cockerham & Weir (1984). If identity disequilibrium is null (which is the case when there is no variance in inbreeding level within the population, or when all individual inbreeding coefficients are known, Shaw *et al.*, 1998; Nietlisbach *et al.*, 2016), one can decompose the genetic variance into additive and dominance components. The additive and dominance variances of a population with a given inbreeding coefficient F are equal to

$$\sigma_{A,F}^2 = (1 + F)\sigma_{genic}^2 + 4F\sigma_{ADI} + \frac{(2F^2)}{1 + F}\sigma_{DI}^2 \quad (4)$$

and

$$\sigma_{D,F}^2 = (1 - F)\sigma_{DR}^2 + \frac{F(1 - F)}{1 + F}\sigma_{DI}^2 + F(1 - F)H^* \quad (5)$$

(Cockerham and Weir, 1984).

Table 2. Genetic variance components and their measurements. In formulas, n is the number of alleles at the k th locus, L is the number of loci that encode the quantitative trait, $p_{i,k}$ is the frequency of the i th allele at locus k , $\alpha_{i,k}$ is the additive effect of the i th allele at locus k and $\delta_{ij,k}$ is the dominance effect at locus k associated with genotype ij . Both $\alpha_{i,k}$ and $\delta_{ij,k}$ are defined from the ancestral (random mating) base population (Lynch and Walsh, 1998, chapter 4).

Variance component	Estimation
σ^2_{genic}	$2 \sum_k^L \sum_i^n p_{i,k} \alpha_{i,k}^2$
σ^2_{DR}	$\sum_k^L \sum_i^n \sum_j^n p_{i,k} p_{j,k} \delta_{ij,k}^2$
H^*	$\sum_k^L (\sum_i^n p_{i,k} \delta_{ii,k})^2$
σ^2_{DI}	$\sum_k^L (\sum_i^n p_{i,k} \delta_{ii,k}^2 - (\sum_i^n p_{i,k} \delta_{ii,k})^2)$
σ_{ADI}	$\sum_k^L \sum_i^n p_{i,k} \alpha_{i,k} \delta_{ii,k}$

Question 2: Predicting the immediate response to selection under inbreeding

In order to predict the immediate response to selection for a population with a given selfing rate, we used a simplification of the Structured Linear Model (SLM hereafter) of Kelly (1999). The SLM decomposes the population under study into cohorts sharing the number of generations of selfing that individuals have undergone. Within each of these cohorts, the model

assumes linearity between offspring and parental phenotypes. Following Kelly (1999), the mean value of a quantitative trait controlled by loci with alleles having arbitrary dominance relations can be written as

$$\bar{Z} = \mu_o - F(\mu_I - \mu_o) \quad (6.1)$$

or

$$\bar{Z} = \mu_o - FB \quad (6.2)$$

, where μ_o and μ_I are the means of the obligate outbred and obligate inbred populations, respectively, B is the directional dominance of the trait, and F the mean inbreeding coefficient of the population. Following Kelly (1999), we can write

$$\Delta\mu_o = (\sigma_{genic}^2 + F(\sigma_{genic}^2 + 2\sigma_{ADI}))S \quad (7)$$

and

$$\Delta B = (\sigma_{ADI} + F(\sigma_{ADI} + \sigma_{DI}^2))S \quad (8)$$

where S is the selection differential. In the SLM, ΔB , $\Delta\mu_o$ and S are calculated within each cohort of selfing, and the overall changes ΔB and $\Delta\mu_o$, and hence the change in the mean phenotype, are determined by a weighted average of cohort-specific values. Here, we simplified the SLM by considering the population as a single cohort of equally inbred individuals. By doing so, the expected change of the quantitative trait under study is equal to

$$\Delta Z = \left((1 + F)\sigma_{genic}^2 + F^2\sigma_{DI}^2 + (3F + F^2)\sigma_{ADI} \right) S \quad (9)$$

This quantity is very similar to the additive variance of a population with an inbreeding coefficient equal to F (equation 4) and gives similar results for the parameter sets tested (results

not shown). Nevertheless, the former result holds only under linkage equilibrium. With genetic associations, the equation becomes

$$\Delta Z = \left((1 + F)\sigma_{genic}^2 + F^2\sigma_{DI}^2 + (3F + F^2)\sigma_{ADI} + \sigma_{LD} \right) S \quad (10)$$

, where σ_{LD} represents the effect of genetic associations on the phenotypic variance. When F is equal to 0, equation 10 is equal to the breeder's equation (Walsh and Lynch, 2018), and when F is equal to 1, we find the classical model of pure-line selection (Gallais, 1990).

We first tested if incorporating all the variance components allows predicting the response to selection more accurately than is possible with a simpler “additive” model (taking into account σ_{genic}^2 and σ_{LD} only). To do so, we introduced an environmental change (coded as a shift of the phenotypic optimum) and followed the response to selection over one generation. We compared the response to selection (the observed change of the mean fitness) with predictions based on Fisher's fundamental theorem of natural selection (FTNS hereafter, Fisher, 1958). We computed the expected response to selection using the classical definition of the FTNS ($\Delta \bar{W}_F = \frac{(1+F)\sigma_{genic}^2 + \sigma_{LD}}{\bar{W}_F}$) as well as a full version derived from equation 10 ($\Delta \bar{W}_F = \frac{(1+F)\sigma_{genic}^2 + F^2\sigma_{DI}^2 + (3F + F^2)\sigma_{ADI} + \sigma_{LD}}{\bar{W}_F}$), where $\bar{W}_F$ is the mean fitness of a population with a given inbreeding coefficient F .

The assumption that all individuals have the same inbreeding coefficient is realistic for extreme selfing rates (close to either 0 or 1) but is probably overly simplistic for mixed-mating populations. Nevertheless, Kelly & Williamson (2000) compared the accuracy of this “single cohort” model with the full SLM and found that, while the full SLM is overall more accurate over the set of simulations they considered, the single-cohort model is generally a good predictor of the mean changes. The only case for which the single-cohort model can be inaccurate is when there is excessively strong inbreeding depression for fitness (Kelly and

Williamson, 2000). Such cases are not expected in our simulations, because the trait under study is already under stabilizing selection, and the inbreeding load is partially purged due to selfing (recall that the higher the selfing rate, the smaller the inbreeding load at equilibrium). In addition, the single-cohort model is probably more applicable empirically than is the full SLM.

Question 3: The effect of shifting from outcrossing to selfing on additive variance

It is generally accepted that a shift from outcrossing to selfing increases adaptive potential on the short term, because the genic variance is expected to increase by a factor $(1+F)$ due to positive correlations between additive allelic effects within loci among offspring (Wright, 1931; Lande, 1977; Peterson and Kay, 2015). However, this ignores the contribution of the other components of the genetic variance, and notably the covariance between the additive effect of an allele and its dominance deviation (σ_{ADI}) which can be substantially negative. Using equation 10 we will test if this hypothesis holds in our simulated data, and in empirical data. Only four studies have estimated all the components of genetic variance in natural plant populations (Shaw *et al.*, 1998; Kelly, 2003; Kelly and Arathi, 2003; Marriage and Kelly, 2009). These studies are restricted to two species, *Mimulus guttatus* (Kelly, 2003; Kelly and Arathi, 2003; Marriage and Kelly, 2009) and *Nemophila menziesii* (Shaw *et al.*, 1998), and a total of 22 estimates for individual traits. We used equation 10 to estimate the effect of shifting from obligate outcrossing to a given selfing rate s on adaptive potential. To compare the change in additive variance in the empirical and simulated data, we computed the ratio

$$\frac{(1 + F)\sigma_{genic}^2 + F^2\sigma_{DI}^2 + (3F + F^2)\sigma_{ADI}}{\sigma_{genic}^2} \quad (11).$$

Comparing the empirical data to simulations assuming different parameter values can also yield insights into the range of possible values for the level of dominance and strength of stabilizing selection to which traits are subject in natural populations.

Results

Question 1: What is the relative contribution of each genetic variance component?

Overall, the within-loci genetic variance (ignoring genetic associations) is dominated by the additive component for all selfing rates and all degrees of dominance (Figure 1). As expected, the contribution of dominance variance to genetic variance increases with the degree of dominance. In the case of complete dominance ($D = 1$), the dominance component explains around 25% of the genetic variance for outcrossing populations (Figure 1). These patterns are found for all simulation parameters we considered, and dominance variance never explained more than 30% of the genetic variance (Figure 1). For completely selfing populations, the contribution of dominance variance to total genetic variance is always zero (eqs. 4 and 5).

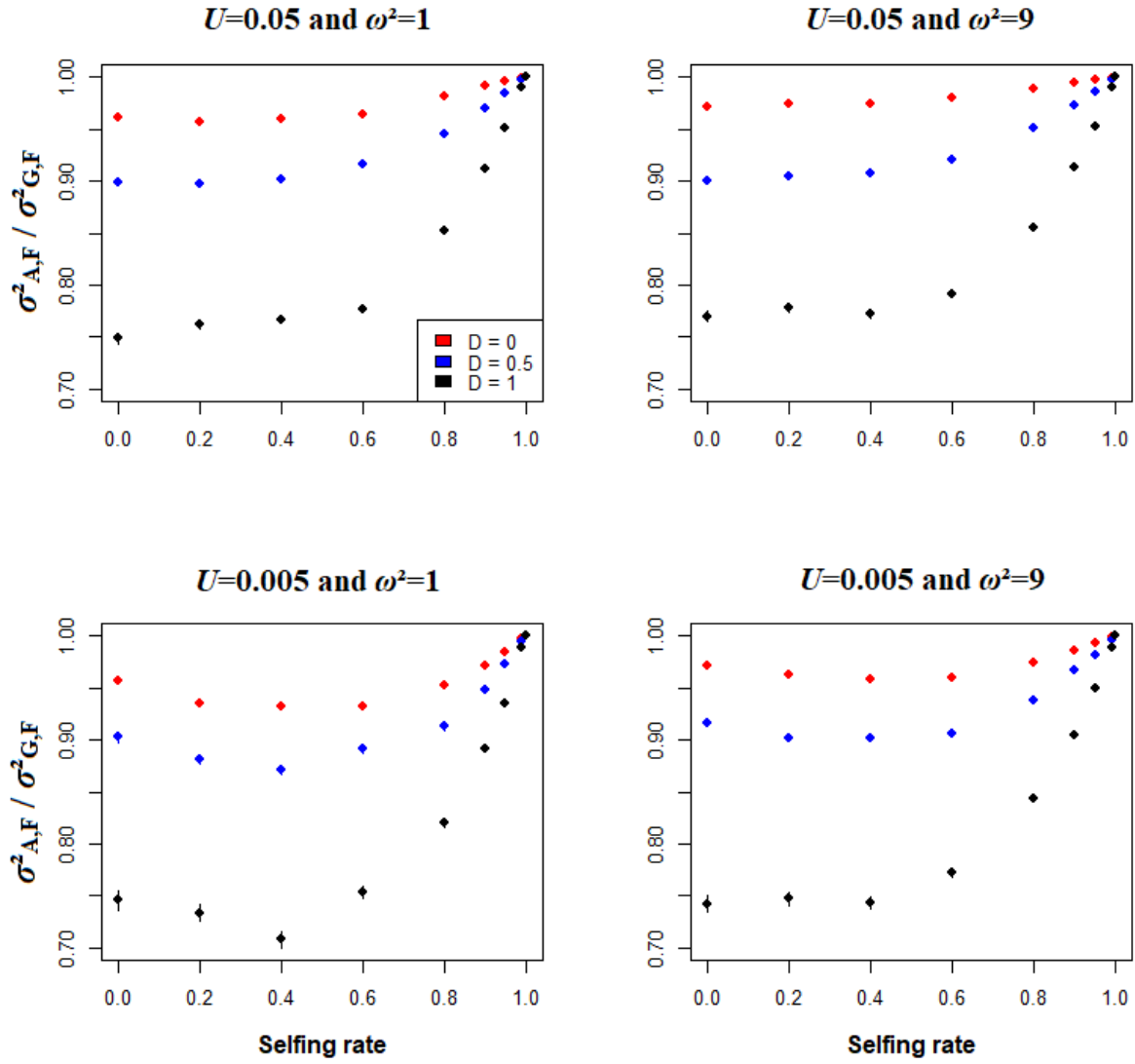


Figure 1. Effects of selfing rate and the dominance level between alleles at a given locus on the proportion of the within-loci genetic variance that is additive ($\sigma^2_{A,F}/(\sigma^2_{A,F} + \sigma^2_{D,F})$ see eqs. 4 and 5 for details). All figures are for $N=1000$ and $L=50$, while the strength of stabilizing selection and mutation rate vary among panels. The parameter D indicates the level of dominance, with 0, 0.5 and 1 indicating non-directional dominance, partial dominance and complete dominance relationships between alleles. Errors bars indicate 95% confidence intervals ($n = 100$).

The overall effect of directional dominance is to flatten the relationship between the genetic variance and the selfing rate (Figure 2A). When dominance is not directional on average ($D = 0$), the genetic variance decreases progressively with the selfing rate (Figure 2A). In the presence of directional dominance, the genetic variance is smaller in outcrossing and mixed-mating populations ($s < 0.8$), and higher in predominantly selfing populations ($s \geq 0.8$), compared to a co-dominant scenario (Figure 2A).

The genetic variance is dominated by those components contributing to the additive variance, that is the genic variance (σ^2_{genic}), the variance of homozygous dominance deviations (σ^2_{DI}), and the covariance between additive effects of alleles and their homozygous dominance deviations (σ_{ADI} , Figure 2B, D & F). The main effect of directional dominance is to increase the strength of purifying selection against deleterious mutations, leading to a loss of genic variance (σ^2_{genic}) when the degree of dominance increases (Figure 2B). The contribution of the variance of homozygous dominance deviations (σ^2_{DI}), and the covariance between additive effects and their homozygous dominance deviations (σ_{ADI}) slowly increase with the selfing rate, so that these components become important components of the genetic variance at high selfing rates ($s \geq 0.8$) and for strong dominance ($D = 1$, Figure 2D & F). Overall, the dominance variance of a random-mating population (σ^2_{DR}) and the sum over all loci of squared per-locus inbreeding depression (H^*) has minor effects on genetic variance, independently of the selfing rate and the degree of dominance (Figure 2E). The genetic associations (σ_{LD}) are important to consider only for predominantly selfing populations ($s \geq 0.8$, Figure 2G). The higher the degree of dominance, the higher the selfing rate for which genetic associations become substantial (Figure 2G). This leads to the result that the covariance between additive effects and their homozygous dominance deviations (σ_{ADI}) can be the main factor decreasing the genetic variance and thus the adaptive potential of populations (Figure 2F), notably when the mutation rate is low (Figure S2). Decreasing the mutation rate or the strength of stabilizing selection does not alter the described

patterns (Figure S2 and S3), although, when selection is weaker, positive genetic covariances between alleles at different loci can arise for mixed-mating populations (Figure S3).

The reduction of genic variance for outcrossing populations with directional dominance also lead to the surprising result that the genic variance is higher in predominantly selfing populations than in outcrossing or mixed-mating populations, independently of the mutation rate for the quantitative trait and hence the strength of genetic associations (Figure 2, S2 and S3). It is also remarkable that, even though some genetic components do not contribute to the total genetic variance due to the value of the inbreeding coefficient, it does not mean that the genetic component itself equals zero.

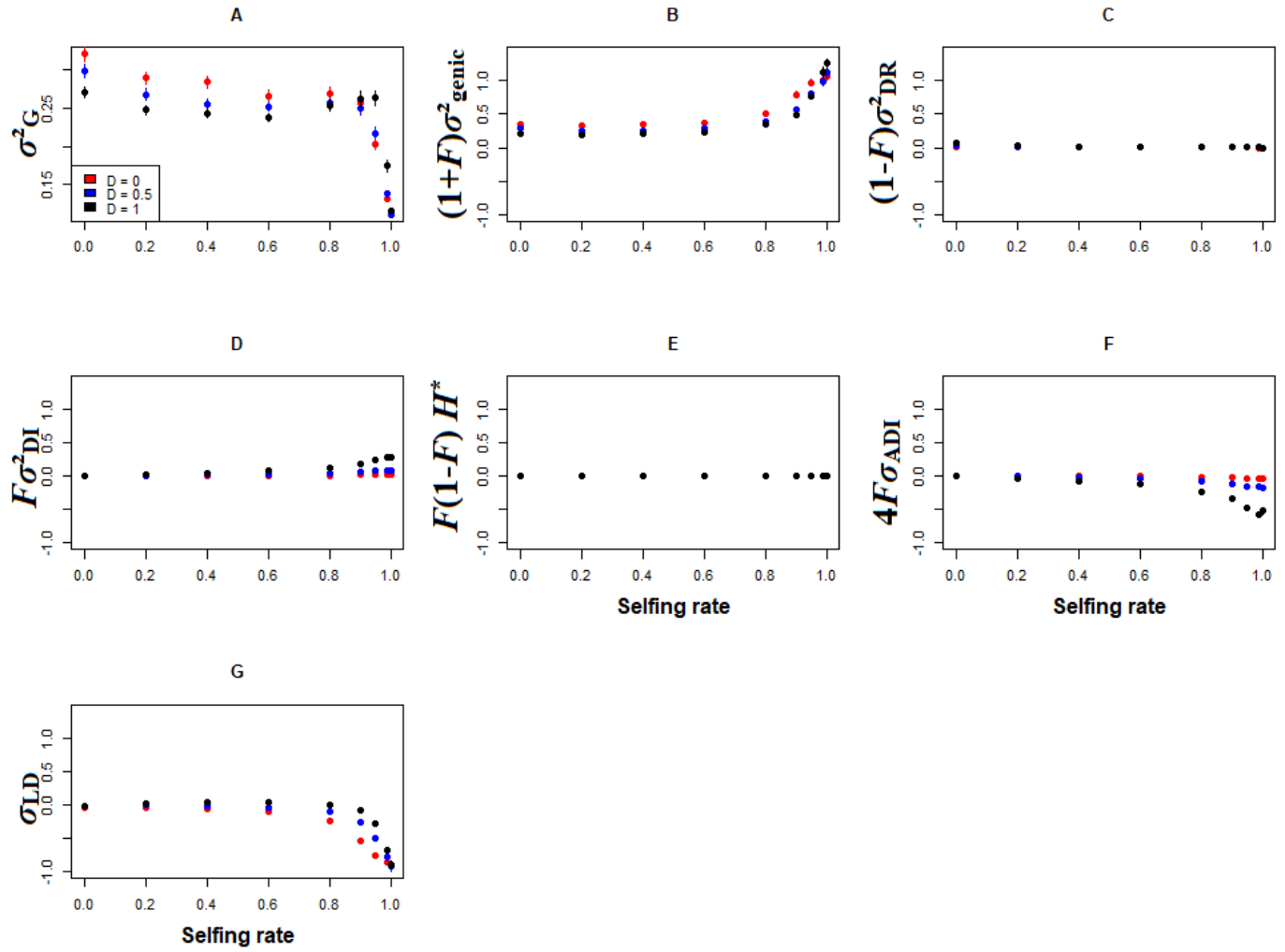


Figure 2. Genetic variance decomposition as a function of selfing rate at mutation-selection-drift equilibrium. All the genetic variance components (panels B to G) are scaled to reflect their relative contributions. **A.** Total genetic variance (σ^2_G). **B.** Genic variance (σ^2_{genic}). **C.** Dominance variance of a random mating population (σ^2_{DR}). **D.** Dominance variance of a fully inbred population (σ^2_{DI}). **E.** Sum over all loci of squared per-locus inbreeding depression (H^*). **F.** Covariance of additive effect of alleles and their homozygous dominance deviations (σ_{ADI}). **G.** Genetic covariance (or Bulmer effect, σ_{LD}). Other parameters are $N=1000$, $L=50$, $U=0.05$ and $\omega^2=1$. Errors bars indicate 95% confidence intervals ($n = 100$).

Question 2: Predicting the response to selection under inbreeding

Considering all the components of the genetic variance

Considering all relevant components of the genetic variance allows more accurate prediction of the response of a population to selection than is possible with a restricted model focused on the genic variance only (Figures 3, S4 & S5). This result holds even in the case of non-directional dominance (Figure 3, $D = 0$), notably when the rate of selfing increases (Figures 3 & S4). Both models tend to overestimate the response to selection (Figure 3, S4 & S5), but the upward bias is much greater when considering only the genic variance as a source of adaptive potential (Figures 3, S4 & S5). This effect relates largely to the contribution of the covariance of additive effect of alleles and their homozygous dominance deviations (σ_{ADI}), which is negative and thus tends to decrease the response to selection. Interestingly, the prediction of the full model can be far from the observed mean fitness change for populations that are highly but not completely selfing ($0.9 \leq s \leq 0.95$, Figures 3, 4 & S4), notably when the change in the optimum occurs in the same direction as the directional dominance (Figure 4). The higher the additive variance (and thus the less stochastic the response) and the higher the degree of dominance, the more accurate is the full model compared to the restricted model (Figures 3, S4 and S5).

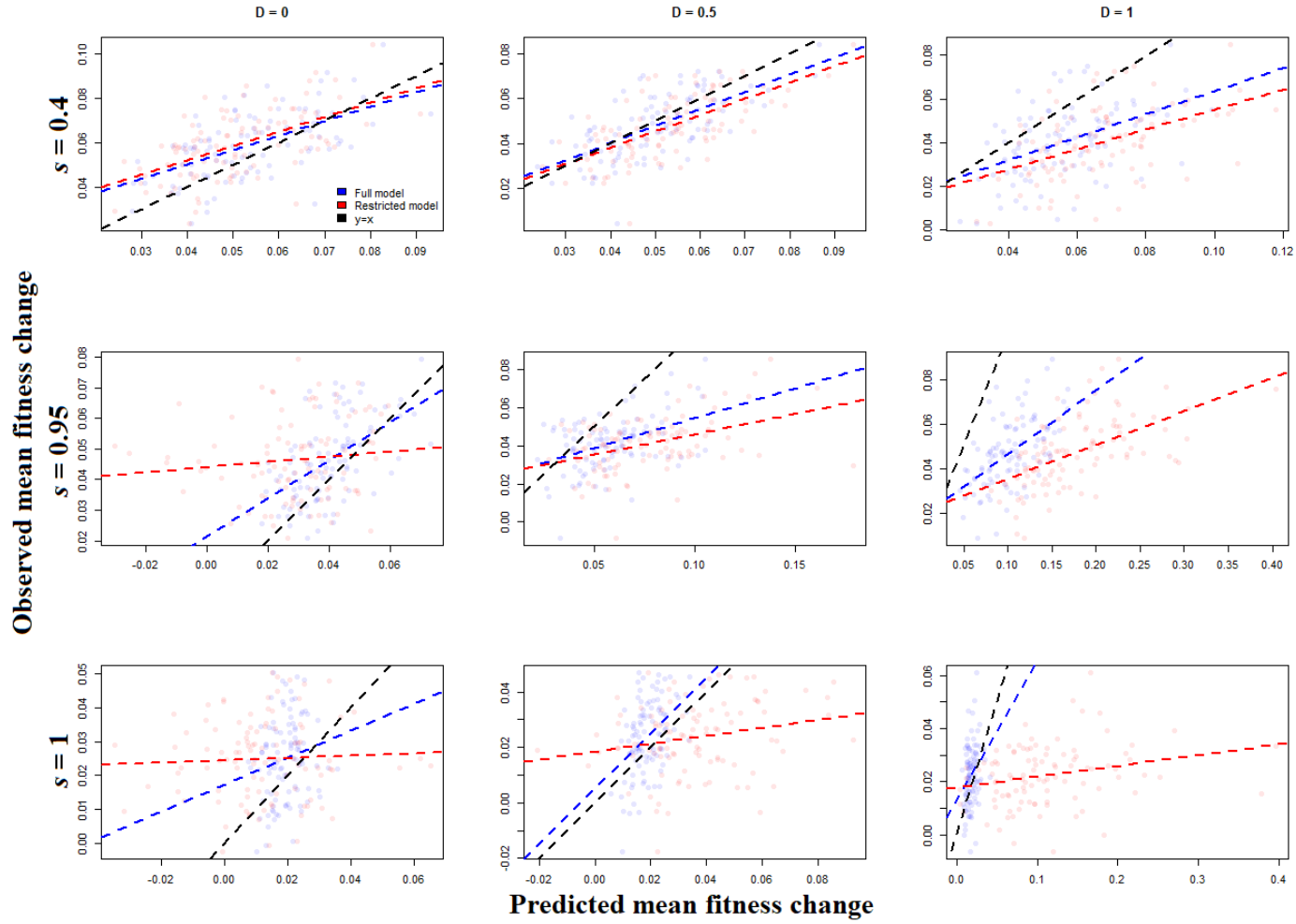


Figure 3. Observed versus predicted (through the fundamental theorem of natural selection) mean fitness change as a function of the dominance level between alleles, for three representative selfing rates ($s = 0.4$, 0.95 and 1). The blue points and line represent predictions computed by using all the components of the genetic variance that contribute to adaptive potential. The red points and line represent predictions computed by using the genic variance only. The black line represents the theoretical 1:1 relationship between observed and predicted mean fitness change. Other parameters are $N=1000$, $L=50$, $U=0.05$ and $\omega^2=1$.

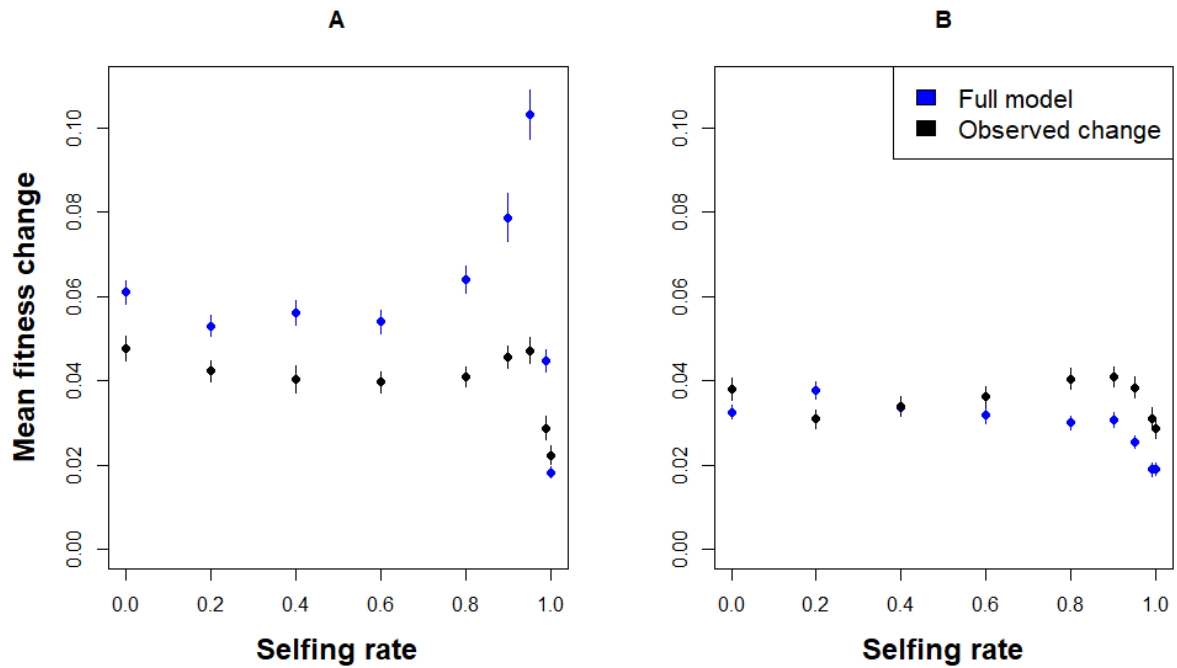


Figure 4. Observed and predicted mean fitness change after one generation of directional selection, as a function of the selfing rate and the directionality of dominance and selection. The blue points represent predictions computed by using all the components of the genetic variance that contribute to adaptive potential. The black points are the observed changes. **A.** Dominance and selection act in the same direction. **B.** Dominance and selection act in opposite directions. Other parameters are $N=1000$, $L=50$, $\omega^2 = 1$, $D = 1$ and $U=0.05$. Errors bars indicate 95% confidence intervals ($n = 100$).

Effects of the mating system on the immediate response to selection

Directional dominance modifies the effect of the selfing rate on the immediate response to selection (Figure 5). The co-dominant model ($D = 0$) is similar to fully additive models, with a non-linear decay of adaptive potential of populations as a function of the selfing rate. The decay is slow for mixed-mating populations ($s \leq 0.9$), and abrupt for predominantly selfing populations ($s > 0.9$). For scenarios with directional dominance ($D = 0.5$ or 1), the response to selection declines rapidly at very high selfing rates ($s \geq 0.99$), while the relationship is more U-shaped for more moderate selfing rates (Figure 5). The initial decline in response to selection for intermediate rates of inbreeding followed by an increase for highly selfing populations ($0.6 \leq s \leq 0.95$) means that predominantly selfing populations may exhibit adaptive potential comparable to that for outcrossing populations (Figure 5), notably for higher degrees of dominance ($D = 1$). These patterns remain true whatever the strength of selection, and the mutation rate (Figures S7 & S8).

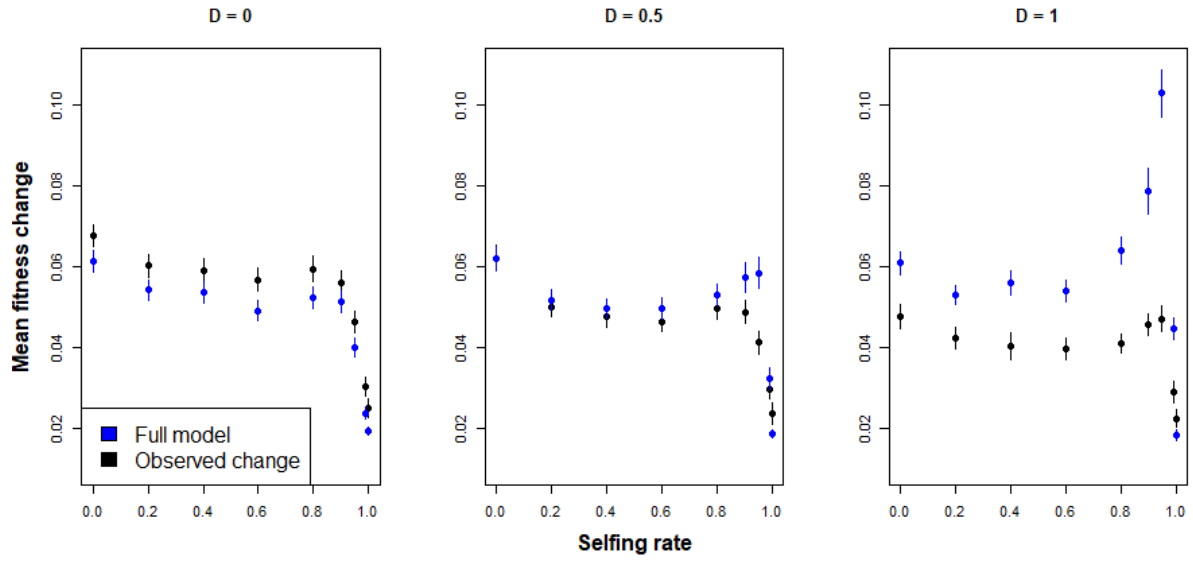


Figure 5. Observed and predicted mean fitness change after one generation of directional selection, as a function of the selfing rate and the degree of dominance. The blue points represent predictions computed by using all the components of the genetic variance that contribute to adaptive potential. The black points are the observed changes. Other parameters are $N=1000$, $L=50$, $\omega^2 = 1$ and $U=0.05$. Errors bars indicate 95% confidence intervals ($n = 100$).

Question 3: The effect of shifting from outcrossing to selfing on additive variance

The empirical estimates of all the genetic-variance components for floral and fitness traits compiled from the literature suggest that considering all the components is critical to understanding the adaptive potential of historically outcrossing populations in the event of a shift towards selfing (Figure 6). In the full model, switching from obligate outcrossing to complete selfing increases the adaptive potential by a factor of 3, compared to a factor of 2 for the additive model (Figure 6). The simulated data are in agreement with this pattern when genetic diversity is low (i.e. for low mutation rates and/or strong stabilizing selection, Figures 6, S6 & S7). The rate of increase in adaptive potential with increasing selfing rate depends on the degree of directional dominance and is stronger for complete dominance (Figure 6). When genetic diversity is higher (for high mutation rate and/or weak stabilizing selection), on the other hand, the increase in adaptive potential with selfing is weaker under the full model than under the additive model (Figure S8). Interestingly, several of our simulation sets correspond well to the pattern in the empirical data (Figure 6, S6, S7), especially the case $N = 1000$, $U = 0.005$, $L = 50$, $D = 1$ and $\omega^2 = 9$ (Figure 6). These observations suggest that the studied traits may tend to be subject to strong directional dominance ($0.5 \leq D \leq 1$) and strong stabilizing selection ($1 \leq \omega^2 \leq 9$).

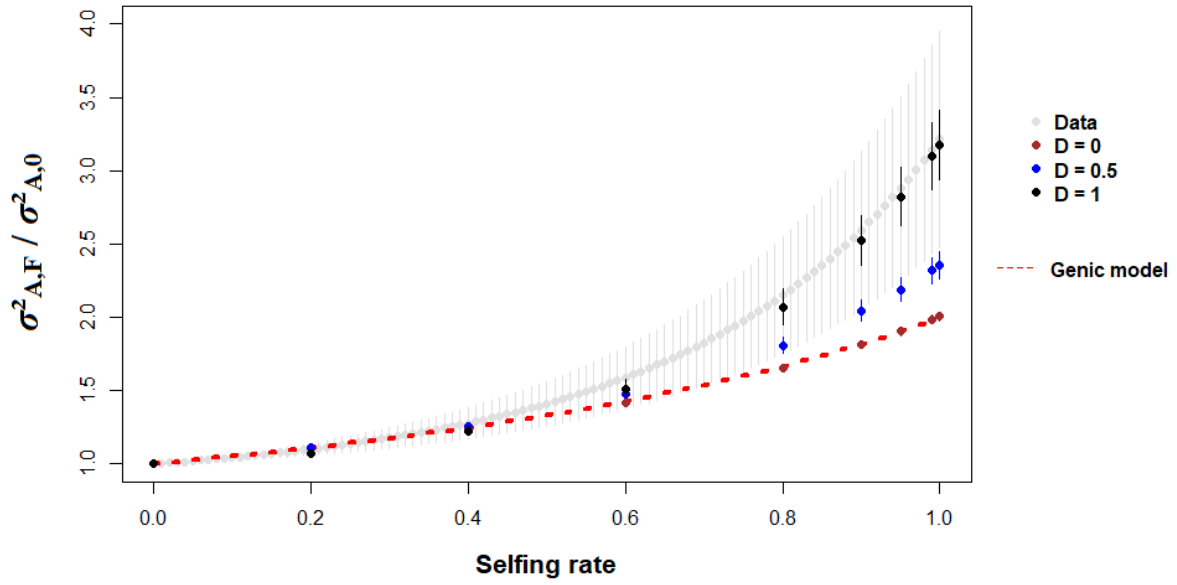


Figure 6. Evolution of the additive variance (equation 10) for a randomly mating population ($s = 0$) in the event of a switch to a given selfing rate s . Grey points represent average values for empirical estimates (± 1 SE, see eq. 11 for details), whereas red, blue and black points represent simulation results for $D = 0, 0.5$ and 1 , respectively. The genic model represents the classical prediction of an increase of the adaptive potential of a factor $(1+F)$. Other parameters values are $N = 1000$, $U = 0.005$, $L = 50$, and $\omega^2 = 9$. Errors bars indicate 95% confidence intervals for simulated values ($n = 100$), and standard errors for empirical values ($n = 22$).

Discussion

Shifts in mating systems can have important implications for trait evolvability and hence the potential of populations to adapt to shifting adaptive optima. Here, we have explored the effect of directional dominance on a quantitative trait under stabilizing selection. We have shown that the additive variance is always the largest component of the genetic variance, but the relative contribution of each component of the genetic variance depends on the degree of dominance and the selfing rate of the population under study. For inbred populations, even if the contribution of those dominance terms specific to inbred populations can be moderate, incorporating them into models of response to selection greatly improves the predictive power. We also showed that a simplification of existing models of selection response in partially selfing populations do not lead to dramatic reductions in predictive power. The overall effect of directional dominance is to flatten the effect of mating system on the immediate response to selection, notably by reducing the standing genetic variation of randomly mating populations. Finally, based on empirical estimates of all relevant components of genetic variance in partially inbred species, we have shown that the transition from obligate outcrossing to selfing is probably more advantageous in terms of adaptive potential than previously expected, suggesting that selfing may evolve under a wider range of conditions than traditionally thought. The comparison of empirical estimates of variance components to our simulation results allowed us to formulate hypotheses about the strength of stabilizing selection typically acting on fitness-related traits.

The decomposition of the genetic variance under directional dominance

Our simulations demonstrate that most of the genetic variance is explained by the additive variance (as defined by Cockerham and Weir, 1984). For outcrossing populations, where only

the genic and classical dominance variance (σ^2_{DR}) are relevant to selection response, we found that dominance variance never explains more than 30% of the genetic variance. This result extends the theoretical predictions of Hill *et al.* (2008), namely that dominance variance contributes only a small fraction of the genetic variance and that adding selection on the focal trait does not alter qualitatively these results. Empirical results are more variable, and dominance variance can explain from 0 to almost 100% of the genetic variance (see Wolak and Keller, 2014 & Class and Brommer, 2020 for brief reviews). Estimates from natural populations are rare and hence prevent generalization about the contribution of dominance to genetic variance, but most estimates suggest a dominance contribution of less than 50% (Mullin *et al.*, 1992; Abney *et al.*, 2000; Merilä *et al.*, 2001; Waldmann *et al.*, 2008; e Silva *et al.*, 2010; Class and Brommer, 2020), in agreement with our simulation results.

When considering all relevant components of the genetic variance, we found that the genic variance, the variance of dominance deviations in a fully inbred populations, and the covariance between additive effects of alleles and their homozygous dominance deviations are important components to consider if we want to understand the adaptive potential of populations. The contribution of these terms to the genetic variance is important mainly for highly selfing populations for which the inbreeding coefficient is high. This result agrees with the few empirical data available in the literature (Hoeschele and Vollema, 1993; Shaw *et al.*, 1998; Edwards and Lamkey, 2002; Kelly, 2003; Kelly and Arathi, 2003; Marriage and Kelly, 2009). Again, more empirical results are needed to gain insight into the evolutionary importance of these components. The scarcity of such data in the literature almost certainly relates to the complexity of their inference, which necessitates either complicated crossing designs (Shaw *et al.*, 1998), or the use of fully inbred lines (Kelly and Arathi, 2003). Exploiting predominantly selfing species is probably the easiest way forward towards greater empirical insight into these processes (Gay *et al.*, 2013), because their populations are naturally organized into fully inbred

lines (Siol *et al.*, 2008; Jullien *et al.*, 2019). By considering the selfed progenies of these natural lines as well as the rare hybrids produced by residual allogamy between the repeated multi-locus genotypes, it is possible to estimate all the components of the genetic variance (Kelly and Arathi, 2003).

Predicting the response to selection in partially selfing populations

To evaluate the importance of non-additive variance components in selection response, we tested if a simplification of the structured linear model (Kelly, 1999a) integrated into the fundamental theorem of natural selection (FTNS) can be a good predictor of the evolution of population mean fitness. Although the additional genetic variance components relevant to inbred populations (σ^2_{DI} , σ_{ADI} and H^*) contribute only moderately to the genetic variance of the population, the single-cohort model we used proved always to be more accurate than a fully additive model, even when dominance is not directional on average, in agreement with a previous investigation by Kelly and Williamson (2000). The additive model generally overestimates the response to selection compared to the full model, notably because it ignores the covariance between the additive effects of alleles and their homozygous dominance deviations, which can be substantially negative as previously pointed out by Shaw *et al.* (1998). The single-cohort model is expected to be inaccurate when inbreeding depression is strong (Kelly and Williamson, 2000), which was not the case in our simulations because the quantitative trait under study is already under selection, or when the selected individuals are not a random sample of the inbreeding classes (Kelly, 1999a; Walsh and Lynch, 2018). We found that this latter problem can greatly bias the predictions derived from the simplified model for highly but not fully selfing populations when the dominance deviations and selection act in the same direction. This discrepancy is notably due to the fact that for such selfing rates

populations are organized into inbred lines (Lande and Porcher, 2015; Abu Awad and Roze, 2018), and hybridization events between those lines creates F_1 hybrids in which directional dominance is fully expressed. If the dominance and selection act in the same direction, it will split the population into two cohorts, one with those F_1 that were selected against before the change, and another cohort of inbred lines that are less fit than their hybrids (Clo *et al.*, 2020). In such cases, the inbreeding coefficient of the selected individuals is much lower than expected for such selfing rates, explaining the difference between predicted and observed changes (Kelly, 1999a). If dominance and selection act in different directions, F_1 hybrids will be less fit than inbred lines in the new environment, and the selected individuals will have an inbreeding coefficient closer to that expected under such a high selfing rate, explaining why our simplification is more accurate in this case (Figure 4). Because the response to selection after an environmental change can be fuelled by the remobilization of cryptic diversity stored in negative linkage disequilibrium by residual allogamy in predominantly selfing species (Clo *et al.*, 2020), we expect that the alignment between dominance and selection will have a major effect on the short-term adaptive potential of inbred populations. Notably, if hybrids are selected against because dominance generates recombinant individuals far from the new phenotypic optimum, the remobilization of cryptic diversity by residual allogamy will not occur, and response to selection will be based on inbred lines, limiting the adaptive potential of selfing populations (Noël *et al.*, 2017; Clo *et al.*, 2020).

Consequences of a mating system shift for short- and long-term adaptive potential

Based on a fully additive model, a shift from outcrossing to selfing increases the adaptive potential by a factor $1 + F$ (Wright, 1931). By considering a more realistic genetic architecture of quantitative traits, and by compiling empirical estimates of all genetic components of mixed-

mating plants, we have shown that the gain in adaptive potential can be greater than in the additive model, notably when tending toward high selfing rates ($s \geq 0.8$). This strong beneficial effect can contribute to explain why transitions from outcrossing to selfing are so frequent (Wright *et al.*, 2013). The good fit between our simulation results and the empirical data allowed us to hypothesize that the floral traits studied in *N. menziesii* (Shaw *et al.*, 1998) and *M. guttatus* (Kelly, 2003; Kelly and Arathi, 2003; Marriage and Kelly, 2009) tend to be under strong stabilizing selection ($1 \leq \omega^2 \leq 9$). This insight must be taken as provisional, however, because other parameter sets not considered here could conceivably fit the empirical data equally well. Estimates of the strength of stabilizing selection acting on fitness-related traits are extremely rare in the literature, and our results stress the importance of further investigations.

One important insight from our simulations is that, on the long term, the reduction of genic variance due to purifying selection and directional dominance in outcrossing and mixed-mating populations lead to a limited effect of the selfing rate on the immediate response to selection. This is in agreement with the overall weak relationship between the selfing rate and the adaptive potential of populations detected in existing meta-analyses (Charlesworth and Charlesworth, 1995; Geber and Griffen, 2003; Ashman and Majetic, 2006; Opedal *et al.*, 2017; Clo *et al.*, 2019). While these studies have assessed relationships between mating systems and standing genetic variance, we are not aware of any experiments comparing the realized response to selection for a trait in populations that truly differ in their mating regimes (Holeski and Kelly, 2006 & Noël *et al.*, 2017 artificially manipulated the selfing rates of their model organisms). Such experiments are necessary to assess the realized effect of mating systems on adaptive potential because no integrative measures of the additive variance of populations as a function of their selfing rate are available in the literature. The best way to estimate the “realized” evolvability (additive variance) of a trait remains to impose directional selection, and control for the selection differential (as in Noël *et al.*, 2017). In addition, this may be the most promising

way forward for estimating additive variance in mixed-mating and predominantly selfing populations, notably if the exact value of each component of the genetic variance is not desired.

Conclusions and implications for evolution in changing reproductive environments

Our results offer some insights into the consequences of a change in the reproductive environment, such as a decline in the reliability of pollination leading to a shift towards a greater rate of selfing. Based on our simulations, if a reduction in pollination reliability concurs with a shift in the phenotypic optimum, the ability of populations to track this shift is not compromised by an increased rate of selfing. This suggests that reduced adaptive potential is unlikely to make predominantly selfing an evolutionary ‘dead end’ (Stebbins, 1957), at least on the short term. Indeed, if the mutation rate on the phenotypic trait is sufficiently high so that genetic associations contribute to genetic variance, the release of such diversity could allow predominantly selfing lineages to evolve faster than outcrossing or mixed-mating lineages (Clo *et al.*, 2020). While some support for Stebbins’ classic hypothesis has come from phylogenetic studies showing increased diversification rates of some self-incompatible taxa (e.g. Goldberg *et al.*, 2010), it is not well understood whether similar trends exist for mixed-mating lineages with continuous variation in outcrossing rates (Wright *et al.*, 2013). Our finding that evolutionary potential remains comparable across the entire gradient in outcrossing rates suggest that other ecological or evolutionary mechanisms need to be invoked to explain reduced evolutionary success of (partially) selfing species.

Acknowledgement:

We thank J.K. Kelly for providing information about the interpretation of the quantitative traits measured in *Mimulus*. We thank C. Pelabon, J. Ronfort, J. Sztepanacz and E.V. Tambarussi for helpful comments on the manuscript. We also thank the evolvability working group at the Centre for Advanced Study in Oslo for hosting us during parts of the work. This work has been conducted with the help of the data and calculation centre South Green of the CIRAD-UMR AGAP.

Bibliography:

- Abney M, McPeck MS, Ober C (2000). Estimation of variance components of quantitative traits in inbred populations. *The American Journal of Human Genetics* **66**: 629–650.
- Abu Awad D, Roze D (2018). Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution* **72**: 751–769.
- Ashman TL, Majetic CJ (2006). Genetic constraints on floral evolution: a review and evaluation of patterns. *Heredity* **96**: 343–352.
- Barrett SC (2003). Mating strategies in flowering plants: the outcrossing–selfing paradigm and beyond. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **358**: 991–1004.
- Bulmer MG (1974). Linkage disequilibrium and genetic variability. *Genetics Research* **23**: 281–289.
- Bürger R, Wagner GP, Stettinger F (1989). HOW MUCH HERITABLE VARIATION CAN BE MAINTAINED IN FINITE POPULATIONS BY MUTATION–SELECTION BALANCE? *Evolution* **43**: 1748–1766.
- Charlesworth D, Charlesworth B (1995). Quantitative genetics in plants: the effect of the breeding system on genetic variability. *Evolution* **49**: 911–920.
- Class B, Brommer JE (2020). Can dominance genetic variance be ignored in evolutionary quantitative genetic analyses of wild populations? *Evolution*.
- Clo J, Gay L, Ronfort J (2019). How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species. *Evolution*.
- Clo J, Ronfort J, Abu Awad D (2020). Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. *Journal of Evolutionary Biology* **33**: 1203–1215.

- Cockerham CC, Weir BS (1984). Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics*: 157–164.
- Edwards JW, Lamkey KR (2002). Quantitative genetics of inbreeding in a synthetic maize population. *Crop Science* **42**: 1094–1104.
- Fisher RA (1958). *The genetical theory of natural selection*, 2nd ed. Dover Publication, New York.
- Gallais A (1990). *Théorie de la sélection en amélioration des plantes*. Elsevier Mason SAS.
- Gay L, Siol M, Ronfort J (2013). Pedigree-free estimates of heritability in the wild: promising prospects for selfing populations. *PLoS one* **8**: e66983.
- Geber MA, Griffen LR (2003). Inheritance and natural selection on functional traits. *International Journal of Plant Sciences* **164**: S21–S42.
- Glémin S (2003). How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57**: 2678–2687.
- Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, Igić B (2010). Species Selection Maintains Self-Incompatibility. *Science* **330**: 493.
- Goodwillie C, Kalisz S, Eckert CG (2005). The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annu Rev Ecol Evol Syst* **36**: 47–79.
- Hansen TF, Pélabon C, Houle D (2011). Heritability is not evolvability. *Evolutionary Biology* **38**: 258.
- Hill WG, Goddard ME, Visscher PM (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* **4**: e1000008.
- Hoeschele I, Vollema AR (1993). Estimation of variance components with dominance and inbreeding in dairy cattle. *Journal of Animal Breeding and Genetics* **110**: 93–104.
- Holeski LM, Kelly JK (2006). Mating system and the evolution of quantitative traits: an experimental study of *Mimulus guttatus*. *Evolution* **60**: 711–723.
- Jullien M, Navascués M, Ronfort J, Loridon K, Gay L (2019). Structure of multilocus genetic diversity in predominantly selfing populations. *Heredity* **123**: 176–191.
- Keightley PD, Bataillon TM (2000). Multigeneration Maximum-Likelihood Analysis Applied to Mutation-Accumulation Experiments in *Caenorhabditis elegans*. *Genetics* **154**: 1193.
- Kelly JK (1999a). Response to selection in partially self-fertilizing populations. I. Selection on a single trait. *Evolution* **53**: 336–349.
- Kelly JK (1999b). Response to selection in partially self-fertilizing populations. II. Selection on multiple traits. *Evolution* **53**: 350–357.
- Kelly JK (2003). Deleterious mutations and the genetic variance of male fitness components in *Mimulus guttatus*. *Genetics* **164**: 1071–1085.

- Kelly JK, Arathi HS (2003). Inbreeding and the genetic variance in floral traits of *Mimulus guttatus*. *Heredity* **90**: 77.
- Kelly JK, Williamson S (2000). Predicting response to selection on a quantitative trait: a comparison between models for mixed-mating populations. *Journal of theoretical biology* **207**: 37–56.
- Lande R (1977). The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* **86**: 485–498.
- Lande R (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution*: 402–416.
- Lande R, Porcher E (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* **200**: 891–906.
- Lynch M (1988). The rate of polygenic mutation. *Genetics Research* **51**: 137–148.
- Lynch M, Walsh B (1998). *Genetics and analysis of quantitative traits*. Sinauer Sunderland, MA.
- Manna F, Martin G, Lenormand T (2011). Fitness Landscapes: An Alternative Theory for the Dominance of Mutation. *Genetics* **189**: 923.
- Marriage TN, Kelly JK (2009). Inbreeding depression in an asexual population of *Mimulus guttatus*. *Journal of Evolutionary Biology* **22**: 2320–2331.
- Merilä J, Kruuk LEB, Sheldon BC (2001). Natural selection on the genetical component of variance in body condition in a wild bird population. *Journal of Evolutionary Biology* **14**: 918–929.
- Mullin TJ, Morgenstern EK, Park YS, Fowler DP (1992). Genetic parameters from a clonally replicated test of black spruce (*Picea mariana*). *Canadian Journal of Forest Research* **22**: 24–36.
- Nietlisbach P, Keller LF, Postma E (2016). Genetic variance components and heritability of multiallelic heterozygosity under inbreeding. *Heredity* **116**: 1–11.
- Noël E, Jarne P, Glémin S, MacKenzie A, Segard A, Sarda V, *et al.* (2017). Experimental evidence for the negative effects of self-fertilization on the adaptive potential of populations. *Current Biology* **27**: 237–242.
- Opedal ØH (2019). Evolutionary Potential of Herkogamy. *eLS*: 1–8.
- Opedal ØH, Bolstad GH, Hansen TF, Armbruster WS, Pélabon C (2017). The evolvability of herkogamy: Quantifying the evolutionary potential of a composite trait. *Evolution* **71**: 1572–1586.
- Ovaskainen O, Cano JM, Merilä J (2008). A Bayesian framework for comparative quantitative genetics. *Proceedings of the Royal Society B: Biological Sciences* **275**: 669–678.
- Peterson ML, Kay KM (2015). Mating system plasticity promotes persistence and adaptation of colonizing populations of hermaphroditic angiosperms. *The American Naturalist* **185**: 28–43.
- Roff DA (2012). *Evolutionary quantitative genetics*. Springer Science & Business Media.
- Ronce O, Shaw FH, Rousset F, Shaw RG (2009). IS INBREEDING DEPRESSION LOWER IN MALADAPTED POPULATIONS? A QUANTITATIVE GENETICS MODEL. *Evolution* **63**: 1807–1819.

- Shaw RG, Byers DL, Shaw FH (1998). Genetic components of variation in *Nemophila menziesii* undergoing inbreeding: morphology and flowering time. *Genetics* **150**: 1649–1661.
- Shaw FH, Geyer CJ, Shaw RG (2002). A COMPREHENSIVE MODEL OF MUTATIONS AFFECTING FITNESS AND INFERENCES FOR *ARABIDOPSIS THALIANA*. *Evolution* **56**: 453–463.
- Siepielski AM, Morrissey MB, Buoro M, Carlson SM, Caruso CM, Clegg SM, *et al.* (2017). Precipitation drives global variation in natural selection. *Science* **355**: 959–962.
- e Silva JC, Hardner C, Potts BM (2010). Genetic variation and parental performance under inbreeding for growth in *Eucalyptus globulus*. *Annals of Forest Science* **67**: 606–606.
- Siol M, Prosperi JM, Bonnin I, Ronfort J (2008). How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in *Medicago truncatula*. *Heredity* **100**: 517.
- Stebbins GL (1957). Self fertilization and population variability in the higher plants. *The American Naturalist* **91**: 337–354.
- Waldmann P, Hallander J, Hoti F, Sillanpää MJ (2008). Efficient Markov chain Monte Carlo implementation of Bayesian analysis of additive and dominance genetic variances in noninbred pedigrees. *Genetics* **179**: 1101–1112.
- Walsh B, Lynch M (2018). *Evolution and selection of quantitative traits*. Oxford University Press.
- Whitehead MR, Lanfear R, Mitchell RJ, Karron JD (2018). Plant mating systems often vary widely among populations. *Frontiers in Ecology and Evolution* **6**: 38.
- Wolak ME, Keller LF (2014). Dominance genetic variance and inbreeding in natural populations. *Quantitative genetics in the wild*: 104–127.
- Wright S (1931). Evolution in Mendelian populations. *Genetics* **16**: 97.
- Wright AJ, Cockerham CC (1985). Selection with partial selfing. I. Mass selection. *Genetics* **109**: 585–597.
- Wright SI, Kalisz S, Slotte T (2013). Evolutionary consequences of self-fertilization in plants. *Proc R Soc B* **280**: 20130133.

Matériel supplémentaire

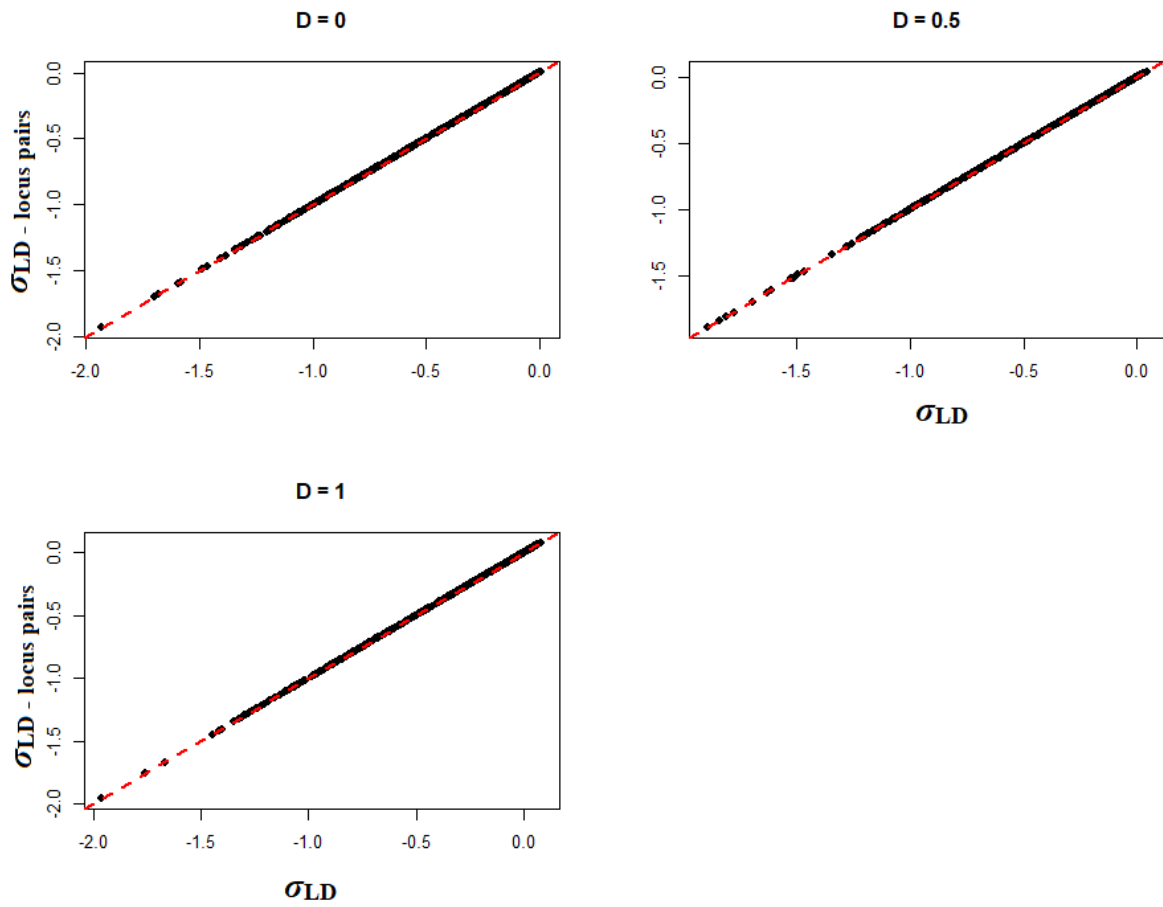


Figure S1. Comparison between the overall genetic covariances (the difference between the observed genetic variance and the within-locus components) and the calculation of the genetic covariances by locus pairs, for the three level of dominance. The red line is the 1:1 relationship. For all the panels, there is a perfect 1:1 relationship ($y = x$, $R^2 = 1$). Other parameters are $N=1000$, $L=50$, $U=0.05$ and $\omega^2=1$.

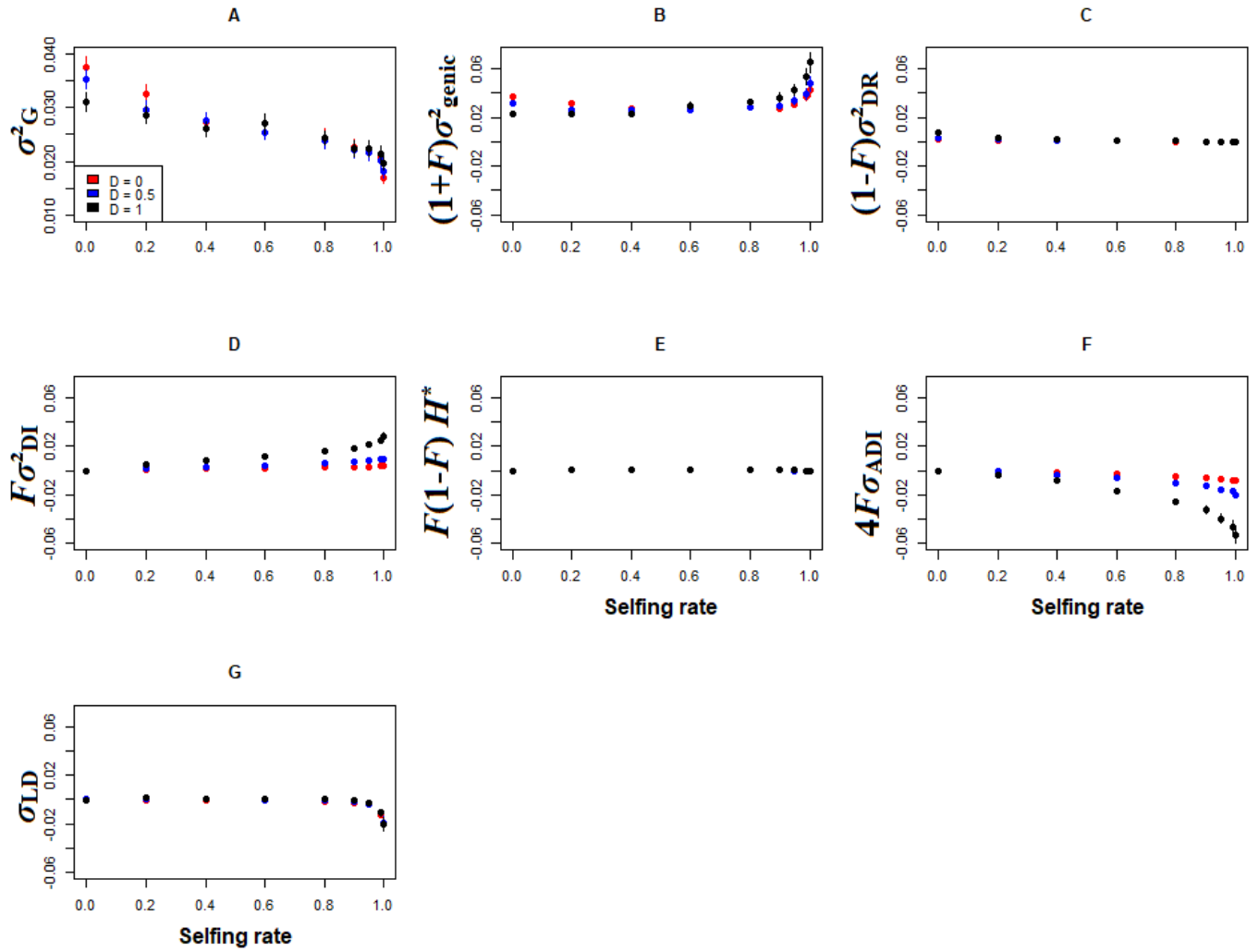


Figure S2. Genetic variance and its decomposition (contribution of each component) at mutation-selection-drift equilibrium, as a function of the selfing rate. All the components of the genetic variance (panels B to G) have the same scaled in order to compare their contributions to the genetic variance. **A.** Total genetic variance (σ^2_G). **B.** Genic variance (σ^2_{genic}). **C.** Dominance variance of a random mating population (σ^2_{DR}). **D.** Dominance variance of a fully inbred population (σ^2_{DI}). **E.** Sum over all loci of squared per-locus inbreeding depression (H^*). **F.** Covariance of additive effect of alleles and their autozygous dominance deviations (σ_{ADI}). **G.** Genetic covariance (or Bulmer effect, σ_{LD}). Other parameters are $N=1000$, $L=50$, $U=0.005$ and $\omega^2=1$. Errors bars stand for 95% confidence intervals ($n = 100$).

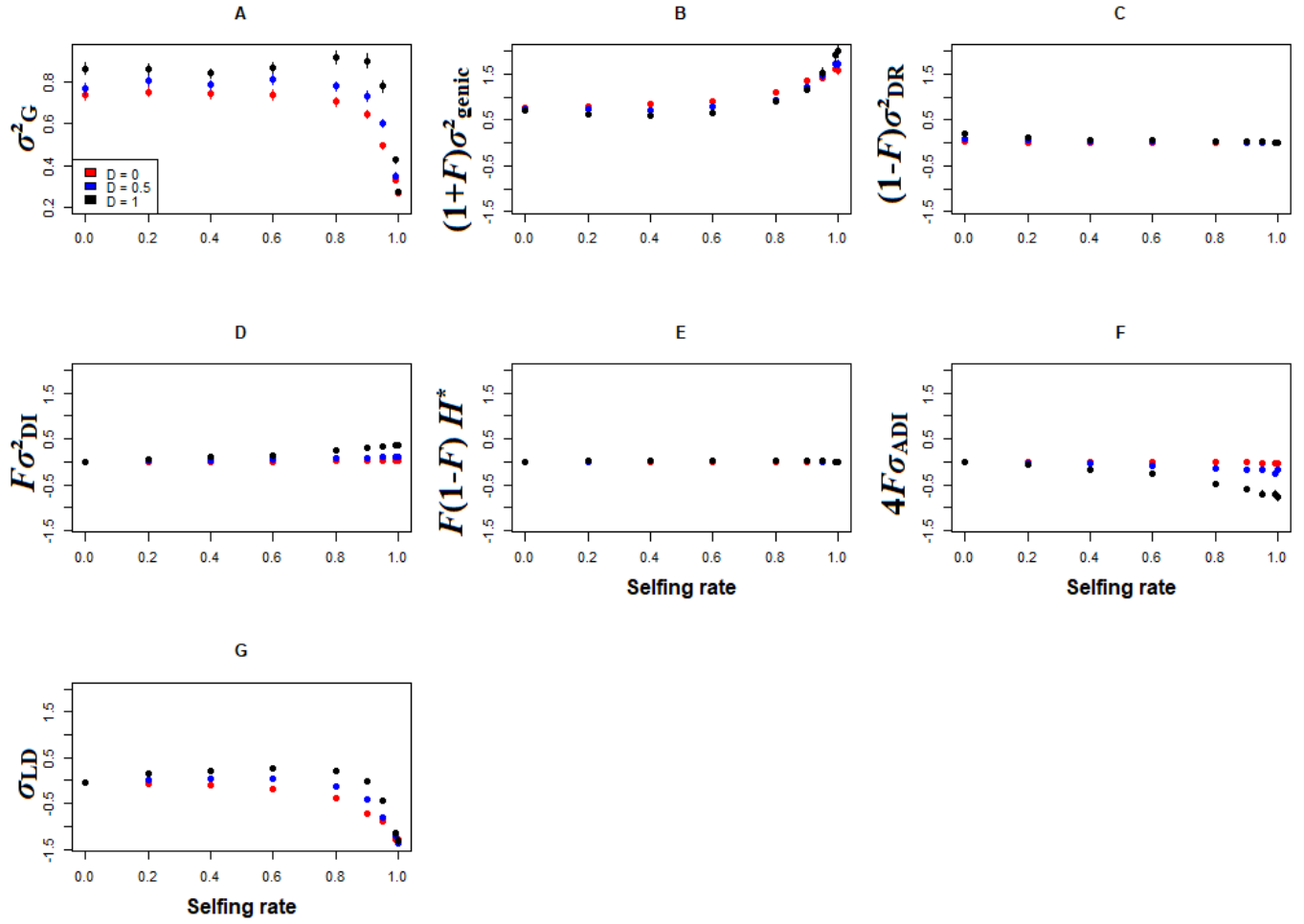


Figure S3. Genetic variance and its decomposition (contribution of each component) at mutation-selection-drift equilibrium, as a function of the selfing rate. All the components of the genetic variance (panels B to G) have the same scaled in order to compare their contributions to the genetic variance. **A.** Total genetic variance (σ^2_G). **B.** Genic variance (σ^2_{genic}). **C.** Dominance variance of a random mating population (σ^2_{DR}). **D.** Dominance variance of a fully inbred population (σ^2_{DI}). **E.** Sum over all loci of squared per-locus inbreeding depression (H^*). **F.** Covariance of additive effect of alleles and their autozygous dominance deviations (σ_{ADI}). **G.** Genetic covariance (or Bulmer effect, σ_{LD}). Other parameters are $N=1000$, $L=50$, $U=0.05$ and $\omega^2=9$. Errors bars stand for 95% confidence intervals ($n = 100$).

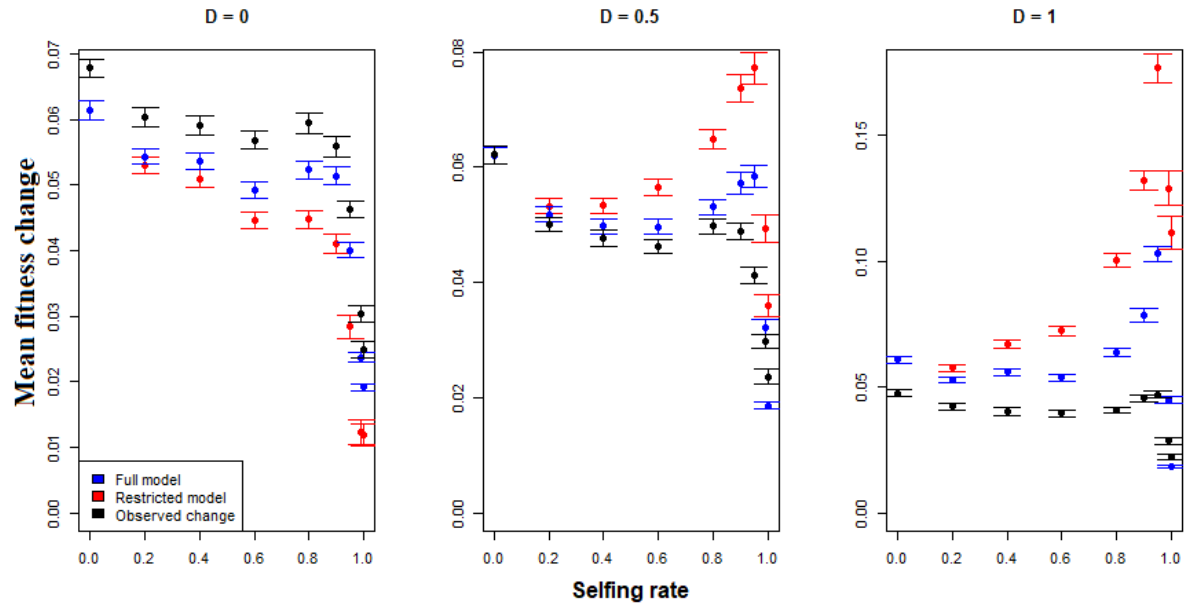


Figure S4. Observed and predicted adaptive potential of populations as a function of the selfing rate and the degree of dominance. The blue points represent predictions computed by using all the components of the genetic variance that contribute to adaptive potential. The red points represent predictions computed by using the genic variance only. The black points are the observed response to selection. Other parameters are $N=1000$, $L=50$, $U=0.05$ and $\omega^2=1$. Errors bars indicate 95% confidence intervals ($n = 100$).

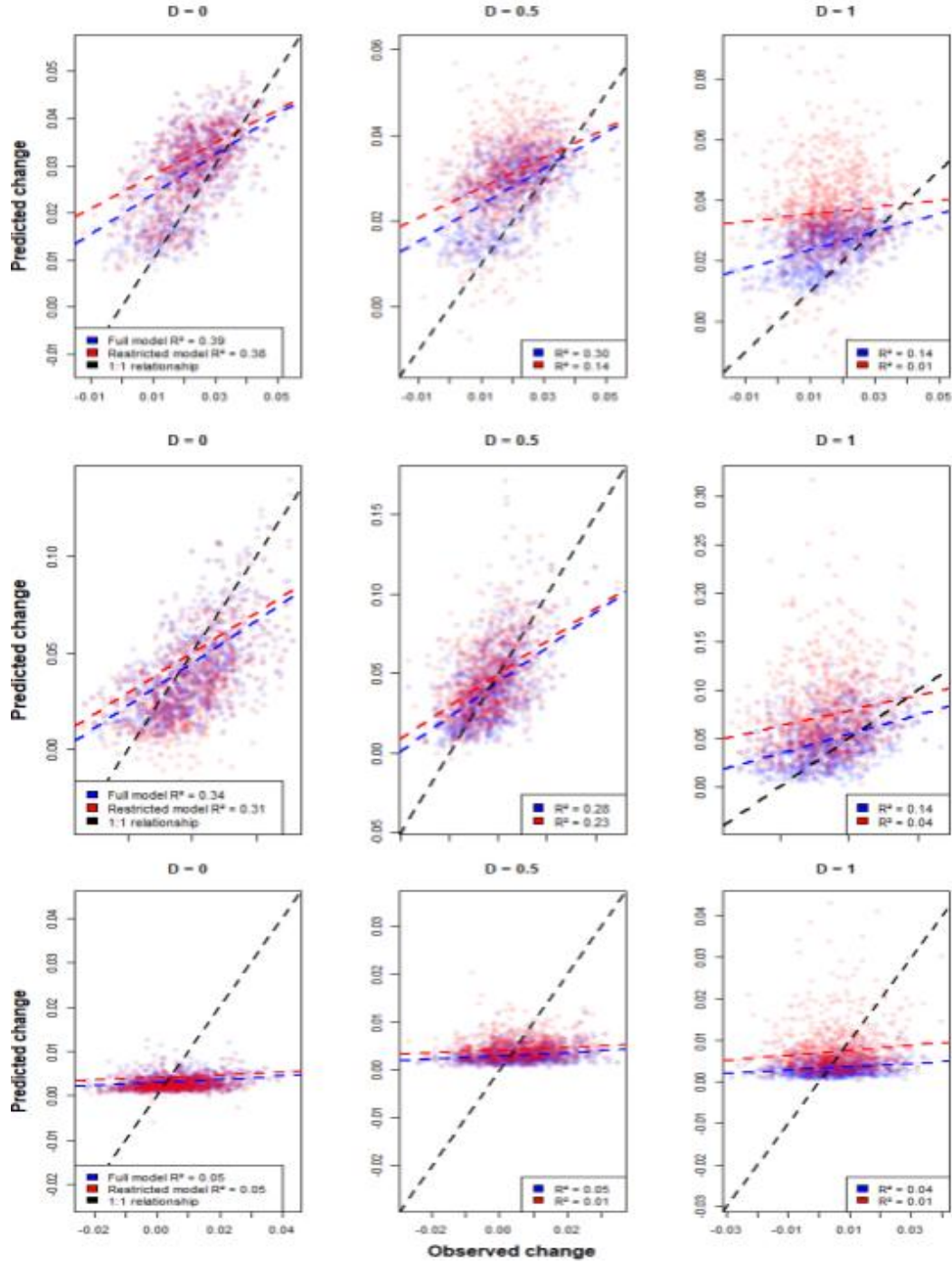


Figure S5. Observed versus predicted (through FTNS) mean fitness change, in function of the dominance level between alleles, for all the simulated selfing rates. Blue points and line represents predictions computed by using all the components of the genetic variance that contribute to adaptive potential. Red points and line represents predictions computed by only using genic variance. The black line represents the theoretical 1:1 relationship between observed and predicted mean fitness change. Other parameters are $N=1000, L=50, U=0.05$ and $\omega^2=9$ for the top three panels, $N=250, L=50, U=0.05$ and $\omega^2=1$ for the middle three panels, $N=1000, L=50, U=0.005$ and $\omega^2=1$ for the bottom three panels.

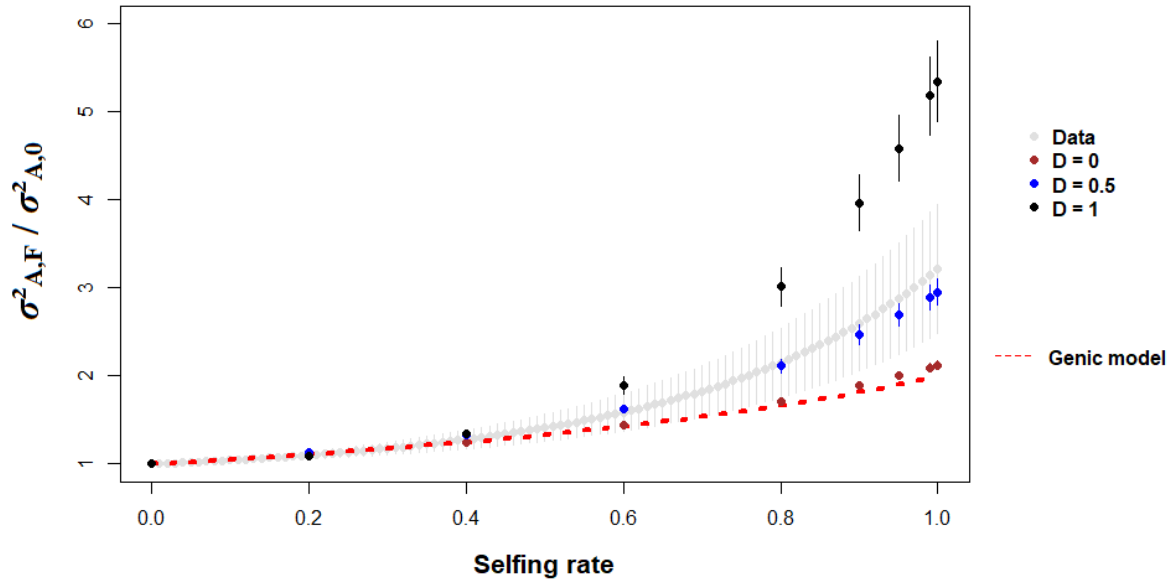


Figure S6. Evolution of the additive variance (equation X) for a randomly mating population ($s = 0$) that switches its selfing rate to a given selfing rate s . Grey points represent values for empirical estimates, whereas red, blue and black points represent simulation values for respectively $D = 0, 0.5$ and 1 . The genic model represents the classical predictions of an increase of the adaptive potential of a factor $(1+F)$. Other parameters values are $N = 1000$, $U = 0.005$, $L = 50$, and $\omega^2 = 1$. Error bars indicate 95% confidence intervals for simulated values ($n = 100$), and standard errors for empirical values ($n = 22$).

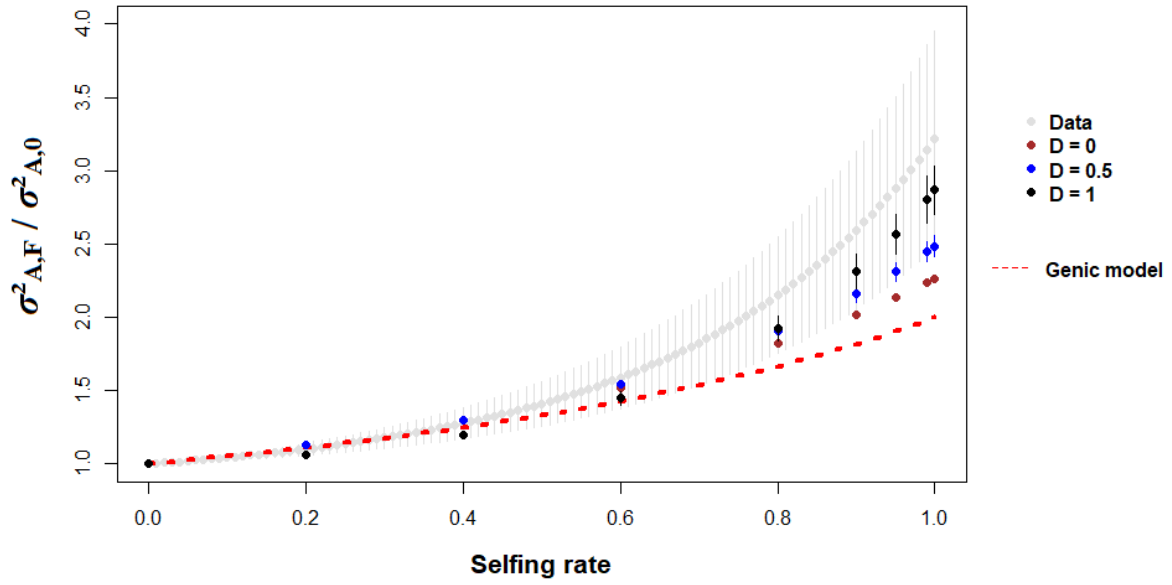


Figure S7. Evolution of the additive variance (equation X) for a randomly mating population ($s = 0$) that switches its selfing rate to a given selfing rate s . Grey points represent values for empirical estimates, whereas red, blue and black points represent simulation values for respectively $D = 0, 0.5$ and 1 . The genic model represents the classical predictions of an increase of the adaptive potential of a factor $(1+F)$. Other parameters values are $N = 1000$, $U = 0.05$, $L = 50$, and $\omega^2 = 1$. Error bars indicate 95% confidence intervals for simulated values ($n = 100$), and standard errors for empirical values ($n = 22$).

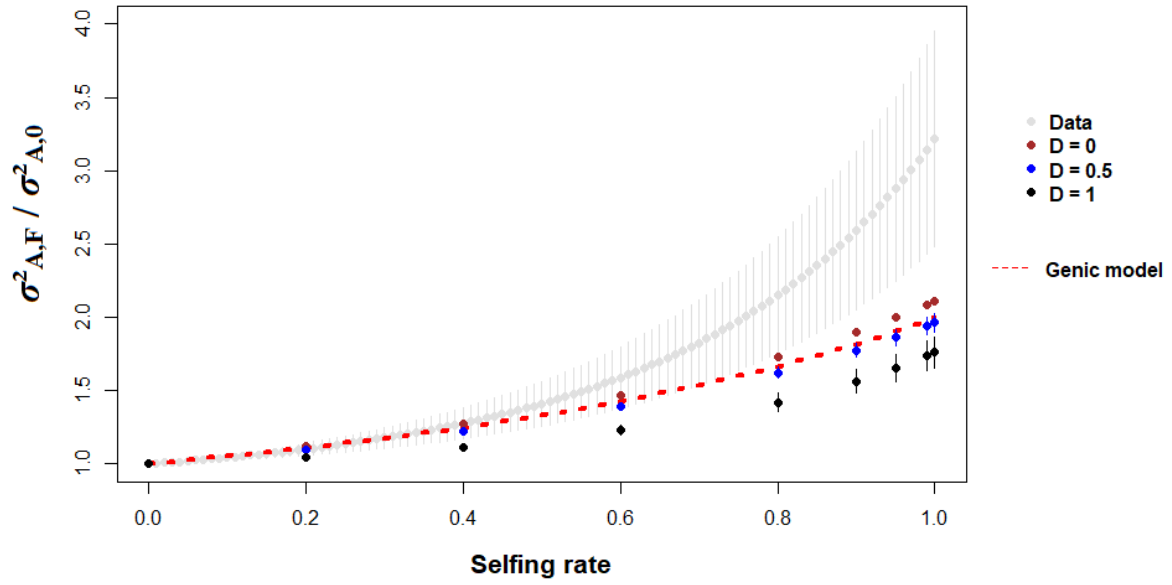


Figure S8. Evolution of the additive variance (equation X) for a randomly mating population ($s = 0$) that switches its selfing rate to a given selfing rate s . Grey points represent values for empirical estimates, whereas red, blue and black points represent simulation values for respectively $D = 0, 0.5$ and 1 . The genic model represents the classical predictions of an increase of the adaptive potential of a factor $(1+F)$. Other parameters values are $N = 1000$, $U = 0.05$, $L = 50$, and $\omega^2 = 9$. Error bars indicate 95% confidence intervals for simulated values ($n = 100$), and standard errors for empirical values ($n = 22$).

Discussion générale et perspectives

Résumé des principaux résultats

Au cours de cette thèse, j'ai montré que l'effet de l'autofécondation sur le niveau de variance génétique des populations est plus complexe que ce que pouvaient laisser penser les modèles théoriques. A l'aide d'une méta-analyse menée à large échelle taxonomique (Chapitre 1), j'ai montré que si la variance génétique des populations tend à diminuer avec le taux d'autofécondation, le système de reproduction n'explique que très partiellement le niveau de variance génétique au sein des populations (entre 2 et 20%). Ceci confirme les résultats d'autres méta-analyses menées sur des jeux de données plus modestes. Au cours des chapitres 2 et 4, nous avons vu que prendre en compte les associations génétiques et l'effet de la dominance directionnelle permet expliquer l'apparente contradiction entre les attendus théoriques et les données empiriques. Premièrement, les associations génétiques qui se créent dans les populations consanguines ont pour effet de diminuer l'efficacité de la sélection sur les allèles délétères. Ces derniers augmentent ainsi en fréquence dans la population, ce qui alimente la variance génétique des populations autogames (chapitre 2, figure 2). Deuxièmement, en considérant une architecture génétique plus réaliste des caractères quantitatifs, j'ai montré que la variance génétique attendue dans une population allogame est moindre lorsque l'on considère l'effet de la dominance directionnelle sur le caractère quantitatif sous sélection (Chapitre 4, figure 2B). Ainsi, les modèles additifs sous-estimeraient l'impact de la sélection sur la variance des populations allogames. En tenant compte de ces deux effets, les modèles théoriques prédisent que la variance génétique est peu ou pas affectée par le taux d'autofécondation des populations (chapitre 4).

Même si malgré ces effets, le niveau de variance génétique est plus faible en populations autogames par rapports aux allogames ou à régime mixte, nous avons vu que la capacité de répondre à un changement environnemental n'était pas forcément réduite. En effet, dans le chapitre 2 nous avons montré que si le taux de mutation affectant les gènes déterminant des caractères quantitatifs est élevé, alors les populations autogames sont en mesure de stocker de la diversité *via* les associations génétiques entre locus. En cas de sélection directionnelle (par exemple suite à un changement environnemental), cette diversité (partie cachée de la variance additive à l'équilibre mutation-sélection-dérive) peut être mobilisée suite à un événement d'allogamie entre lignées pures ayant fixé des combinaisons alléliques différentes. Nous avons vu que dans ces conditions, les populations autogames peuvent répondre aussi bien au changement environnement que leurs homologues allogames ou à régimes mixtes. Si le taux de

mutation est faible, alors la sélection purificatrice, plus efficace chez les populations autogames (car les mutations délétères sont récessives en espérance), permet à elle seule de maintenir la population proche de l'optimum phénotypique. Dans ces conditions les associations sont rares, et la variance génétique réduite dans les populations autogames. Celles-ci sont alors moins à même de répondre à un changement environnemental, par rapport aux populations plus allogames.

Néanmoins, la remobilisation de la variance via l'allofécondation nécessite que les descendants de ces croisements ne soient pas contre-sélectionnés. Or, nous avons montré, chez *Medicago truncatula*, que les évènements d'hybridation entre des lignées collectées dans les populations naturelles sont en espérance délétères (Chapitre 3). Nos analyses suggèrent que la dépression de croisement serait due à la relation de dominance qui existe entre les différents allèles portés par les lignées pour les caractères sous sélection, éloignant les hybrides de l'optimum local. Ceci crée un mécanisme d'isolement reproducteur impliquant la dominance directionnelle qui maintient le potentiel évolutif des populations autogames. Il est délétère dans un environnement stable et permet l'évitement des évènements de recombinaison entre lignées (à l'échelle locale ou globale). Il peut néanmoins être avantageux lors d'un changement d'optimum en permettant le relargage de la diversité génétique stockée via les associations génétiques (comme confirmé par les simulations du chapitre 4, figure 4), cela dépend de si la direction de la dominance permet aux individus F_1 de s'approcher ou non du nouvel optimum phénotypique.

Ces différents résultats, pris dans leur globalité, permettent de relativiser l'argument disant que l'autogamie est un cul de sac évolutif car cet argument repose sur l'hypothèse que les populations autogames auraient moins de diversité génétique que les populations plus recombinantes. Nos résultats suggèrent ainsi que d'autres mécanismes, génétiques ou écologiques doivent être avancés pour expliquer le moindre succès évolutif des espèces autogames.

Discussion : Repenser l'hypothèse du cul-de-sac évolutif

Plusieurs de mes résultats permettent de remettre en perspective l'hypothèse du cul-de-sac évolutif pour les espèces autogames. Dans cette discussion, je reprendrai les principales hypothèses sur lesquelles repose la théorie du cul de sac évolutif, et les discuterai au vu des résultats de ma thèse et de la littérature sur le sujet. En particulier, l'hypothèse du cul-de-sac évolutif repose sur deux observations majeures : (1) l'unidirectionnalité des transitions de l'allogamie vers l'autogamie, et (2) le taux de diversification négatif des espèces autogames (Takebayashi and Morrell, 2001; Igic and Busch, 2013).

Les raisons des transitions unidirectionnelles du système de reproduction

La transition de l'autofécondation vers l'allofécondation est un processus théoriquement peu probable, ce qui alimente l'hypothèse du cul-de-sac évolutif énoncée par Stebbins (1957). Aucune étude empirique n'a mis en évidence sans équivoque une transition de l'autofécondation vers l'allofécondation (voir Takebayashi and Morrell 2001 pour une revue). Plusieurs hypothèses génétiques ont été proposées pour expliquer pourquoi la réversion de l'autofécondation vers l'allofécondation est compliquée et dans quelles situations théoriques elle est possible. Dans un premier temps, je synthétiserai ces différents résultats, puis j'étendrai ces résultats à partir des résultats de ma thèse.

Le rôle des mutations délétères dans la transition de l'autogamie vers l'allogamie

Dans ce paragraphe, nous ne présenterons que des hypothèses génétiques permettant théoriquement la transition de l'autofécondation vers l'allofécondation. En effet, lorsque la transition originale de l'allofécondation vers l'autofécondation est due à l'assurance de reproduction associée à l'autogamie, dans un contexte de déclin des pollinisateurs par exemple, alors la réversion vers l'allogamie est peu probable, voire impossible, puisque les pollinisateurs sont une composante centrale de la valeur sélective des individus se reproduisant par croisement (Busch and Delph, 2012; Igic and Busch, 2013).

D'un point de vue génétique, pour que la transition du système autogame vers le système allogame soit possible, il faut que la hausse du taux de recombinaison soit accompagnée d'une hausse de la valeur sélective de la population. Si le niveau de dépression de consanguinité est supérieur à 0.5 dans une population autogame, et qu'un modificateur du taux d'autofécondation augmentant le taux d'allogamie apparaît dans la population, alors la transition de l'autofécondation vers l'allofécondation est théoriquement possible (car la dépression de consanguinité l'emporte sur l'avantage automatique de l'autofécondation, Lande and Schemske, 1985). La prise en compte de l'effet de purge des mutations délétères lié à l'autofécondation rend cependant ce scénario très improbable, la purge réduisant la dépression de consanguinité en deçà du seuil de 0.5 (Abu Awad & Roze, 2018). On peut noter que cet effet de purge est soutenu empiriquement (Winn *et al.* 2011).

Toutefois, du fait de leurs plus faibles tailles efficaces et de leurs taux de recombinaison réduit, les populations autogames sont plus à même de fixer des mutations délétères (Muller, 1964). Ces mutations vont participer à l'émergence d'hétérosis lors d'évènements d'hybridation entre différentes populations d'une même espèce ayant accumulé ces mutations à des locus différents (Crow, 1948; Lynch, 1991; Glémin, 2003). Il a alors été proposé que la mise en contact secondaire de populations complètement autogames ayant fixé des mutations délétères pourrait être le terrain parfait pour la transition de l'autofécondation vers l'allofécondation (Busch and Delph, 2012; Harkness *et al.*, 2019). Cependant, les modèles théoriques ont montré que si l'hétérosis pouvait permettre à un modificateur du taux d'autofécondation d'augmenter en fréquence, sa probabilité de se fixer, et donc d'établir un système allogame, restait faible (Harkness *et al.*, 2019). En effet, l'apparition du morphe allofécondant va permettre d'augmenter la recombinaison et de créer des haplotypes moins impactés par les mutations délétères. Ainsi, lorsque l'allèle autogame ancestral se retrouve dans un haplotype purgé, il va permettre de transmettre cette combinaison allélique favorable en limitant la recombinaison, et va donc finir par éliminer l'allèle d'allofécondation (Harkness *et al.*, 2019). En conséquence, la plus petite valeur d'hétérosis permettant une probabilité de fixation du modificateur du taux d'autofécondation supérieure aux attendus neutres est de 2812%, ce qui rend cet événement hautement improbable. De plus, il subsiste généralement un résidu d'allogamie dans les populations autogames (Kamran-Disfani and Agrawal, 2014), il est donc fortement probable que l'allogamie résiduelle suffise, lors de contacts secondaires de populations autogames, à provoquer les mêmes effets d'hétérosis que dans ces modèles qui ne considèrent que la probabilité de transition de 100% à 0% d'autofécondation, et que cela diminue encore la

probabilité de fixation d'un modificateur du taux d'autofécondation. Enfin, ces modèles ne considèrent que l'effet de mutations inconditionnellement délétères sur la valeur sélective des populations, qui favorisent fortement un allèle allofécondant puisque ce dernier va permettre de masquer l'effet de ces mutations.

Les transitions de l'autogamie vers l'allogamie freinées par la dépression de croisement

Si des croisements entre populations autogames ayant fixé du fardeau de mutation à différents locus peuvent générer de l'hétérosis, on s'attend aussi à ce que ces dernières soient différenciées génétiquement (Wright *et al.*, 2013) et aient également fixé des incompatibilités épistatiques (Dobzhansky, 1937; Lynch, 1991), des locus sous-dominants (Schierup and Christiansen, 1996) ou des réarrangements chromosomiques (Lande, 1985). Ces effets pourraient limiter ou annuler la hausse de valeur sélective due à l'hétérosis, et donc limiter encore un peu plus les conditions dans lesquelles un modificateur du taux d'autofécondation pourrait permettre une transition de l'autogamie vers l'allogamie. De plus, même sans fixation des incompatibilités mentionnées précédemment, les différentes populations peuvent également être adaptées à leurs environnement, en fixant différentes combinaisons alléliques, à l'échelle intra- et inter- populations. Or nous avons vu dans le chapitre 3 que la dominance directionnelle pouvait jouer un rôle dans l'isolement reproducteur entre populations (ou même lignées) autogames. Ainsi, la différenciation génétique sur des caractères liés à la valeur sélective des individus et soumis à la dominance directionnelle pourrait également limiter l'effet d'hétérosis pour l'allèle modificateur du système de reproduction. En effet, la sélection du modificateur va dépendre de la valeur sélective des recombinants dans lesquels il va se retrouver (F_1), et dans lesquels la dominance directionnelle des allèles est pleinement exprimée. Ceci peut conduire à la (sub-) létalité des F_1 , notamment lorsque la différenciation génétique est importante (Figure 1).

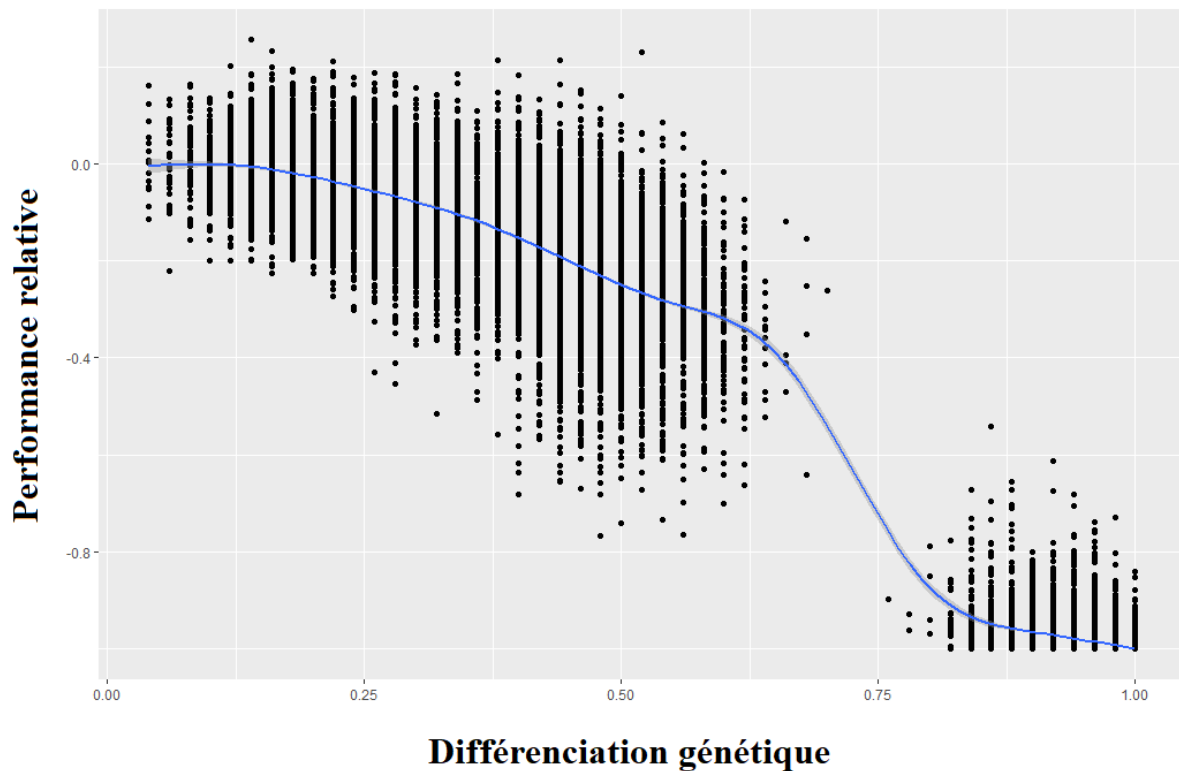


Figure 1. Performance relative (comme définie au chapitre 3) des hybrides F_1 issus de croisements théoriques entre lignées pures simulées via le modèle du chapitre 2, en fonction de la différenciation génétique entre les deux lignées utilisées pour générer les hybrides. La dominance directionnelle est introduite en piochant la valeur de l'hétérozygote dans une loi uniforme bornée par les valeurs du parent moyen et de la plus grande breeding value des deux homozygotes. Les paramètres de simulation sont $N = 250$, $L = 50$, $\omega^2 = 1$, $s = 0.95$. Le taux de mutation varie, permettant de générer des lignées pures ayant différents degrés de différenciation génétique pour le caractère quantitatif sous sélection stabilisante. Josselin Clo, données non-publiées.

Les coûts potentiels d'une transition de l'autogamie vers l'allogamie

Une autre hypothèse, peu exploitée, pouvant expliquer pourquoi l'on n'observe pas de transitions de l'autogamie vers l'allogamie, est que les avantages associés à l'autofécondation sont tellement forts qu'une transition vers l'allofécondation serait délétère. A court terme, une transition de l'autofécondation vers l'allofécondation aurait pour effet de libérer toute la variance cachée dans les associations génétiques. Le niveau de variance étant négativement

corrélé à la valeur sélective en milieu stable (Abu Awad and Roze 2018), ceci aurait pour premier effet de diminuer la viabilité de la population. De plus, il serait nécessaire de faire évoluer un certain nombre de traits floraux (taille de fleur, coloration, production de nectar, production de pollen...) permettant d'attirer des pollinisateurs, et des mécanismes d'évitement de l'autofécondation, qui peuvent être coûteux à mettre en place, notamment si l'autofécondation a généré des compromis évolutifs avec d'autres traits de valeur sélective (Slotte *et al.* 2012, Wright *et al.* 2013). Sur le long terme, une transition vers l'allofécondation est également synonyme de purge moins efficace des mutations délétères (Lande and Porcher 2015, Abu Awad & Roze, 2018).

Une manière de regarder si une transition de l'autogamie vers l'allogamie est possible, et de quantifier son coût sur la valeur sélective des populations serait de tester expérimentalement si l'on peut sélectionner ou non pour un plus grand taux d'allogamie dans des populations hautement autogames (Igic and Busch, 2013). L'évolution du système de reproduction peut être extrêmement rapide (moins de 10 générations, Bodbyl Roels and Kelly, 2011), la sélection sur des traits liés au système de reproduction pourrait permettre de tester les différentes hypothèses présentées ici (possibilité ou non de sélectionner pour plus d'allogamie, les coûts associés à une transition vers l'allofécondation...). L'herkogamie pourrait être le trait idéal, puisqu'il est fortement relié au taux d'autofécondation des populations, et présente généralement un niveau de diversité génétique important dans les populations (Herlihy and Eckert, 2002, 2007; Opedal *et al.*, 2017; Opedal, 2018, 2019).

Il apparaît ainsi qu'une transition de l'autofécondation vers l'allofécondation est hautement improbable, en grande partie parce que l'allogamie résiduelle permet d'imiter les avantages qui seraient associés à une évolution vers de forts taux d'allofécondation, et que les coûts potentiels associés à une telle transition pourraient tout simplement empêcher un changement de système de reproduction. Ainsi, les avantages associés à un système hautement (mais pas complètement) autogame pourraient suffire à expliquer l'uni-directionnalité des transitions de système de reproduction. Nous allons maintenant tenter d'expliquer l'observation de taux d'extinction plus fort des espèces autogames par rapport aux allogames.

Le taux de diversification négatif des populations autogames

La 2^{ème} observation qui soutient l'hypothèse du cul-de-sac évolutif des espèces autogames est le taux de diversification négatif des espèces auto-compatibles, qui serait une conséquence de leur fort taux d'extinction (Goldberg *et al.*, 2010). Ici, je reprendrai les principales hypothèses précédemment mentionnées pour expliquer le taux d'extinction plus fort chez les espèces autogames, en montrant qu'elles peuvent, pour la plupart, être remises en question théoriquement et/ou empiriquement. Je proposerai ensuite des pistes alternatives d'analyse.

Les rôles des mutations dans la probabilité d'adaptation et l'extinction des populations

Comme vu en introduction, l'autofécondation peut modifier la probabilité de fixer une mutation avantageuse permettant l'adaptation à un changement environnemental, et les différents modèles se sont focalisés sur la probabilité de fixation d'une ou plusieurs mutations, en fonction de leur niveau de dominance (Glémin and Ronfort, 2013; Hartfield and Glémin, 2014, 2016). Parmi les principaux résultats, on peut rappeler qu'en régime autogame, la probabilité de fixer deux mutations bénéfiques ségrégeant simultanément dans la population est plus faible qu'en régime allogame (Hartfield and Glémin, 2016), et que la probabilité d'entraîner une mutation délétère pendant la phase de balayage sélectif est également plus forte en population autofécondante (Hartfield and Glémin, 2014). Ces modèles suggèrent donc que la probabilité de s'adapter *via* des nouvelles mutations est plus importante pour les populations allogames que pour les populations autogames. Néanmoins, utiliser la probabilité de fixation d'une ou deux mutations comme proxy de la probabilité d'adaptation d'une population est trompeur, puisque les caractères adaptatifs sont probablement, dans la majorité des cas, polygéniques. Plusieurs mutations bénéfiques peuvent ségréger au même locus lorsque plusieurs combinaisons haplotypiques tendent à augmenter en fréquence lors de la phase d'adaptation. En modifiant le modèle de simulations développé dans le chapitre 4, j'ai pu observer que lorsque plusieurs nouvelles mutations sont susceptibles de participer à l'adaptation à un nouvel environnement, l'autofécondation s'adapte toujours au moins aussi bien que les populations allogames (Figure 2). Dans ce modèle, les populations sont uniformes et n'ont pas de diversité génétique, et la dynamique évolutive est uniquement due à l'augmentation de

fréquence et/ou la fixation de nouvelles mutations avantageuses. Il apparaît que l'autofécondation répond toujours au moins aussi bien que les populations plus recombinales (Figure 2). Lorsque les mutations sont récessives, la dynamique est plus rapide en régime autogame, comme attendu dans les modèles mono-locus (Figure 2, Glémin and Ronfort, 2013). Lorsque les mutations sont co-dominantes ou dominantes, alors les dynamiques d'adaptation sont similaires entre les différents systèmes de reproduction (figure 2), alors que les attendus mono- et bi-locus suggéraient que l'adaptation était facilitée dans des régimes plus allogames (Glémin and Ronfort, 2013; Hartfield and Glémin, 2014, 2016). Cela pourrait venir du fait que l'autofécondation permet une augmentation en fréquence plus rapide des mutations avantageuses et le maintien de combinaisons haplotypiques favorables, tout en maintenant une purge efficace des mutations délétères.

Enfin, nous avons vu en introduction que les mutations n'avaient pas toujours un effet bénéfique sur l'adaptation des populations. En effet, à cause de leur plus forte probabilité de fixer des mutations délétères, les populations autogames sont plus à même de s'éteindre via des phénomènes de fonte mutationnelle (Abu Awad *et al.*, 2014; Abu Awad and Billiard, 2017). Si empiriquement un tel processus est soutenu, (Escobar *et al.*, 2010; Willi, 2013), l'extinction est alors attendue tellement rapide que les espèces ne laissent aucune trace phylogénétique (pas de spéciations), les rendant alors indétectables dans les études utilisant les espèces actuelles (Abu Awad *et al.*, 2014; Abu Awad and Billiard, 2017).

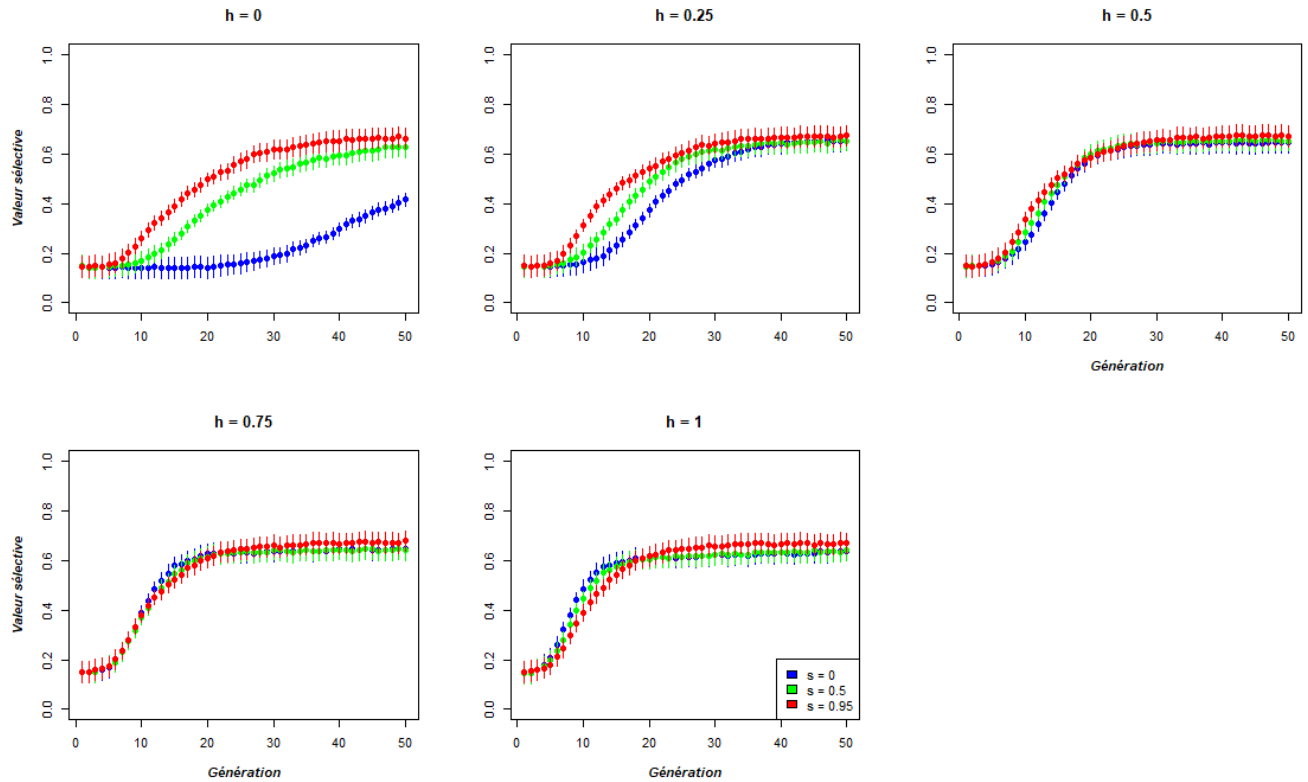


Figure 2. Dynamique d’adaptation à partir de nouvelles mutations, d’une population sans variance génétique pré-existante, en fonction du taux d’autofécondation et du coefficient de dominance des mutations bénéfiques (la dominance varie de la récessivité totale à la dominance totale). Le modèle utilisé est celui du chapitre 4, avec les conditions de changement d’optimum du Chapitre 2 (tous les individus sont uniformes et ont une valeur phénotypique de 0, et le nouvel optimum est égal à 2.5). Les autres paramètres sont $N = 250$, $U = 0.05$, $L = 50$, $\omega^2 = 1$ et $n = 100$ réplicats par taux d’autofécondation. Les barres d’erreurs représentent l’intervalle de confiance à 95%. Josselin Clo, données non-publiées.

L’adaptation des populations autogames à partir de la variation génétique pré-existante

Le principal argument avancé pour expliquer la plus grande probabilité d’extinction des populations autogames est leur manque de diversité génétique (Stebbins, 1957; Takebayashi and Morrell, 2001; Igic and Busch, 2013), limitant ainsi leur capacité à répondre à des changements environnementaux. Au cours de cette thèse nous avons vu que les données empiriques ne montrent pas un signal fort dans ce sens et que cette prédiction théorique n’est plus valable sous certaines conditions. Ici je reprendrai les principales preuves empiriques qui

réfutent cette hypothèse, et discuterai leurs limites, puis je proposerai des manières plus intégratives de mesurer le potentiel évolutif des populations autofécondates, et enfin, je soulignerai l'importance des associations génétiques dans le potentiel évolutif des populations consanguines.

L'autofécondation n'affecte pas ou peu le niveau de variance génique en populations naturelles

Le faible niveau de variance génétique en population autofécondante est bien soutenu théoriquement (Charlesworth and Charlesworth, 1995; Lande and Porcher, 2015; Abu Awad and Roze, 2018). Néanmoins, les données empiriques peinent à confirmer ces attendus. Notre méta-analyse à large échelle montre que la diversité additive de caractères quantitatifs semble être peu affectée par le système de reproduction (chapitre 1). Ce résultat est en adéquation avec les précédentes méta-analyses portant sur le même sujet (Charlesworth and Charlesworth, 1995; Geber and Griffen, 2003; Ashman and Majetic, 2006; Opedal *et al.*, 2017) et soutenu par les modèles théoriques lorsqu'ils sont actualisés pour tenir compte des associations génétiques entres mutations délétères (chapitre 2) et de la dominance directionnelle (chapitre 4). Ce dernier chapitre a également permis de rappeler que le potentiel évolutif des espèces autogames ne peut pas se résumer à la variance génique des caractères quantitatifs (Cockerham and Weir, 1984; Wright and Cockerham, 1985; Kelly, 1999). Il n'existe pas de manière triviale d'estimer le potentiel évolutif des espèces autofécondantes, notamment pour des taux d'autofécondations intermédiaires (Kelly and Williamson, 2000). La meilleure manière de quantifier la variance additive « réalisée » reste de réaliser des expériences de sélection directionnelle sur des caractères quantitatifs. A ma connaissance, seule deux études de ce genre ont été réalisées, je vais les résumer dans la section suivante.

Les expériences de sélection artificielles, et l'effet de l'autofécondation sur la variance additive réalisée

Deux expériences ont regardé l'effet de l'autofécondation sur le niveau de réponse à une sélection artificielle. Holesky & Kelly (2006) sur *Mimulus guttatus*, et Noël *et al.* (2017) sur *Physa acuta*, ont montré que le système de reproduction influençait la vitesse de réponse à la

sélection directionnelle. Un résultat surprenant dans ces deux études est que les populations autogames répondent généralement aussi bien à la sélection que les allogames, au moins sur le court terme. Dans leur expérience, Noël *et al.* (2017) ont imposé une sélection par troncation sur un trait morphologique, la circularité de la coquille chez la physe, en sélectionnant les 25% d'individus ayant la plus grande circularité de coquille. Ils ont tout d'abord créé deux lignées à partir d'une population ancestrale, une lignée 100% allogame (35 générations d'allofécondation), et une lignée 100% autogame (29 générations d'autofécondation). Les principaux résultats de cette étude sont que (1) les groupes autogames répondent plus rapidement durant les premières générations de sélection (Figure 3), mais que (2) cet avantage disparaît sur le long terme (après trois générations) (Figure 3). Holesky & Kelly (2006) ont trouvé des résultats similaires chez *Mimulus guttatus*. À partir d'un pool de lignées pures, les auteurs ont créé trois groupes de populations, des populations allogames (4 réplicats), à régimes mixtes ($s \approx 50\%$, 10 réplicats) et autogames (6 réplicats), et ont imposé de la sélection directionnelle sur deux traits négativement corrélés : la largeur de la corolle et la date de floraison. Pour la taille de la corolle, sur les quatre générations de sélection, plus la population est consanguine, plus la réponse à la sélection est forte et variable. Il ne semble pas y avoir de déclin de réponse à la sélection pour les populations consanguines, mais l'expérience dure moins longtemps que celle de Noël *et al.* (2017).

De manière très intéressante, ces expériences fournissent des données empiriques qui supportent les modèles théoriques, en démontrant l'avantage à court terme et le désavantage à long terme de l'autofécondation. Cependant, elles renseignent peu sur ce que serait l'évolution à court et long terme d'une véritable population autogame. En effet, les deux espèces étudiées sont préférentiellement allogames (Holeski & Kelly, 2006; Escobar *et al.*, 2008), et le système de reproduction est imposé « artificiellement », sans qu'il corresponde à un quelconque optimum phénotypique. Or, imposer un régime autogame à une population précédemment allogame a pour conséquence d'augmenter la réponse immédiate à la sélection, au moins d'un point de vue purement additif (Wright, 1931). En effet, la variance additive est doublée dans le régime consanguin, puisque la variance génique contribue d'un facteur $(1 + f)$, contre un facteur 1 en population allogame. Lorsque toutes les composantes de la variance génétique sont prises en compte et contribuent significativement, aucune prédiction de l'effet d'un passage de l'allofécondation à l'autofécondation sur le potentiel évolutif n'est triviale (Kelly, 1999a, Chapitre 4). Néanmoins, d'après la décomposition de la variance génétique faite par Kelly & Arathi (2003) sur les traits floraux du même pool de lignées qu'ont utilisé Holesky & Kelly

(2006), nous pouvons prédire que la variance additive (au sens de Cockerham & Weir, 1984) de la largeur de corolle devrait être deux fois plus importante dans les lignées purement autogames par rapport aux allogames. Deuxièmement, imposer un régime purement autogames est probablement peu réaliste au vu du résidu d'allogamie qui est toujours observé en populations naturelles d'espèces majoritairement autogames, et impose une forte sélection sur les lignées et une baisse de la variance génétique sans doute beaucoup plus drastique que ce qui serait observé avec de l'allogamie résiduelle (David *et al.*, 1993; Chelo *et al.*, 2017).

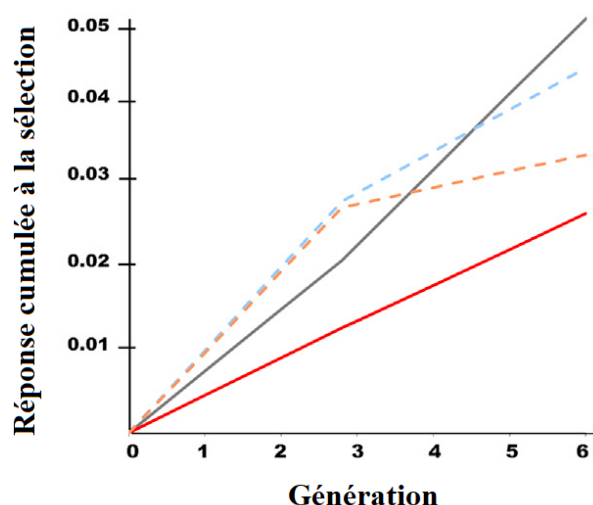


Figure 3. Réponse cumulée à la sélection par troncation au cours du temps, chez *Physa acuta*. Les courbes rouges représentent les groupes issus de la lignée 100% autogame, les bleues les groupes issus de la lignée 100% allogame. Les courbes pleines représentent les groupes étant maintenus à un régime 100% allogame pendant la phase de sélection par troncation, alors que les courbes pointillées représentent les groupes maintenus à un régime 100% autogame. Figure adaptée de Noël *et al.* (2017).

Enfin, en quelques générations d'autofécondation imposée, on ne s'attend pas à ce que les populations aient accumulé de la diversité génétique en fixant des associations génétiques. Or ces associations, qui pourraient limiter la réponse initiale à la sélection, pourraient la dynamiser sur le moyen terme. Ces trois effets expérimentaux fragilisent les conclusions que l'on peut tirer des résultats de ces deux études. Il serait pertinent de reproduire ces expériences d'évolution expérimentales mais sur différentes populations ayant des taux d'autofécondations qui diffèrent réellement, et non artificiellement. Des espèces comme *Arabidopsis thaliana* (Tedder *et*

al., 2011) ou *Eichhornia paniculata* (Barrett and Husband, 1990) semblent idéales pour traiter ce genre de questions, car elles présentent une forte variabilité de taux d'autofécondation en populations naturelles.

L'accumulation de variance via les associations génétiques, et les contraintes à sa remobilisation

L'hypothèse de potentiel évolutif appauvri chez les autogames ne prend pas en compte les associations entre locus. Nous avons vu en effet que la variance stockée dans les associations génétiques peut être remobilisée via les événements d'allogamie résiduelle, permettant un niveau d'adaptation similaire voire supérieur des espèces autogames par rapport aux allogames lors de la réponse des populations à un changement environnemental. Il est compliqué de quantifier la part de variance cachée par les associations génétiques. La manière la plus simple est de faire des croisements entre lignées pures, et de regarder la variabilité phénotypique au sein des individus F_2 . Des études ont reporté que les populations autogames sont plus à même de produire des descendance F_2 sortant de la gamme des valeurs parentales (Rieseberg *et al.*, 1999, 2003), suggérant que les associations génétiques sont une composante importante de la variabilité génétique des espèces autofécondantes.

Néanmoins un tel processus nécessite que les hybrides, sur lesquels la remobilisation de la diversité se fait, ne soient pas inconditionnellement contre-sélectionnés. J'ai mis en lumière, chez *Medicago truncatula*, que les événements d'hybridation sont en moyenne délétères à la génération F_1 , et que la baisse de valeur sélective est due en partie au moins à la dominance directionnelle. Cette dernière limite donc les événements d'hybridation lorsque le milieu est stable. Cependant, lors de changements environnementaux, les croisements pourraient être bénéfiques en termes de valeur sélective du fait de cette architecture génétique. Il se pourrait que ce mécanisme basé sur la dominance directionnelle soit répandu chez les populations autogames. En effet, un isolement reproducteur dû à la dominance directionnelle constitue un mécanisme particulièrement efficace pour maintenir le potentiel évolutif des populations autogames: il permet d'éviter les événements de recombinaison entre lignées (à l'échelle locale ou globale) et le relargage de la diversité génétique stockée via les associations génétiques, qui serait défavorable dans un environnement stable, et permet (voire favorise) l'expression de la variabilité cachée lors d'un changement d'optimum. Oakley *et al.* (2015) ont montré que

l'hétérosis et la dépression d'allofécondation observés dans leurs croisements de lignées d'*Arabidopsis thaliana* étaient dus à la dominance. De plus, une ré-analyse des données de Gimond *et al.* (2013) et de Roumet & Widmer (données non publiées), a permis d'établir la corrélation décrite dans le chapitre 3 entre la similarité parentale et la probabilité que les hybrides sortent de la gamme parentale (Figure 4), suggérant un rôle de la dominance directionnelle dans les performances des hybrides pour ces deux études également.

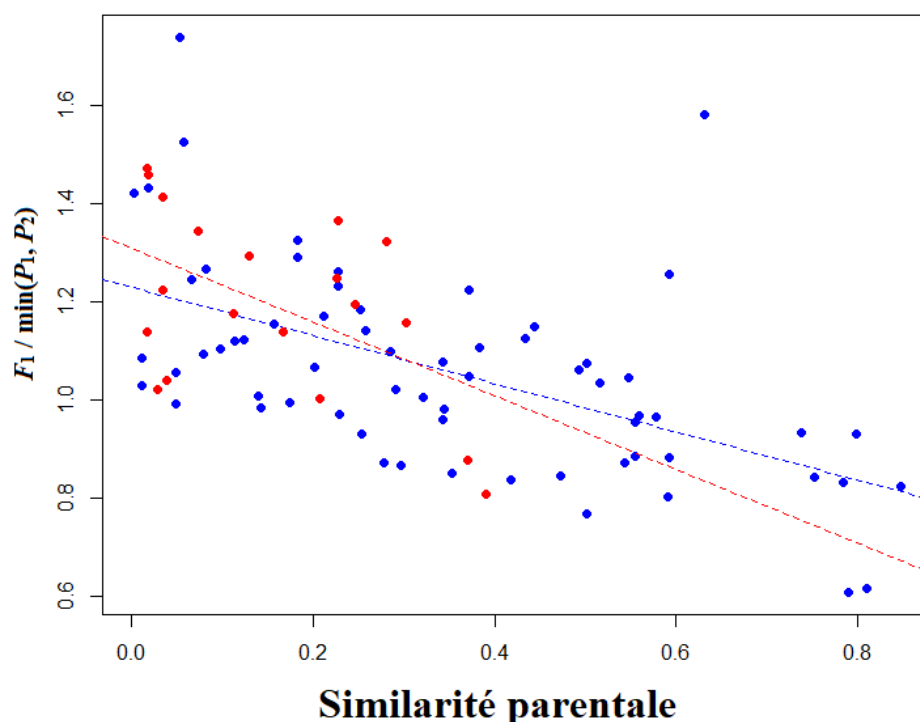


Figure 4. Relation entre la similarité parentale et le ratio $F_1/\max(P_1, P_2)$ (se référer au chapitre 3 pour plus de détails sur les notations), pour les espèces *Arabidopsis thaliana* (points bleus, $n = 72$, données de Roumet & Widmer, non publiées) et *Caenorhabditis elegans* (points rouges, $n = 18$, données de Gimond *et al.* 2013). Josselin Clo, données non-publiées.

Pour conclure, les différents résultats de mon travail de thèse ont montré que les hypothèses classiques avancées pour expliquer le taux d'extinction accru des espèces autogames ne sont pas nécessairement valables lorsque l'on considère une architecture génétique plus réaliste des caractères adaptatifs. Dans ces conditions, le niveau de variance génétique disponible ne semble pas être limitant pour l'adaptation à partir de la variation

génétique pré-existante, et une analyse rapide de l'adaptation à partir de nouvelles mutations suggère que celle-ci est aussi efficace en régime autogame par rapport aux régimes plus allofécondants. Il est alors nécessaire de proposer d'autres mécanismes pour expliquer les données empiriques démontrant un taux de diversification négatif des espèces autogames, ce que nous allons tenter de faire dans la section suivante.

Hypothèses alternatives : extinctions globales vs locales, et l'hypothèse du suicide évolutif

Il a été montré que le système de reproduction a un effet sur l'aire de distribution des espèces (Lowry and Lester, 2006; Grossenbacher *et al.*, 2015). Les espèces autogames ont généralement des aires de distribution plus larges que les espèces allogames, avec notamment une capacité à coloniser de plus hautes latitudes (Grossenbacher *et al.*, 2015, Figure 5). De plus, au sein de certaines espèces, il a été montré que les populations autogames sont plutôt présentes en bordure d'aire de distribution (Pujol *et al.*, 2009; Griffin and Willi, 2014; Paggi *et al.*, 2015). Or, les limites d'aires de distribution des espèces et les hautes latitudes sont généralement des habitats plus rudes que le cœur de l'aire, avec des extrêmes de températures, des saisonnalités plus marqués, et moins de diversité biotique qu'au cœur des aires de distribution (Mittelbach *et al.*, 2007). L'assurance de reproduction associée à l'autofécondation joue sans doute un rôle prépondérant dans ces patrons de distribution, et on peut penser que les phases de colonisations de ces régions ont probablement été accompagnées de forts goulets d'étranglement. De plus, l'adaptation aux conditions locales est nécessaire pour que la population puisse s'installer durablement dans ces conditions extrêmes (Pannell, 2015). Du fait de ces effets démographiques et sélectifs, on s'attend à ce que le niveau de diversité génétique en limites d'aire de distribution soit inférieur à celui observé au cœur de ces aires (Martinez-Padilla *et al.*, 2017). Finalement, cette baisse de diversité et de taille efficace des populations pourrait limiter la capacité de ces populations à répondre à des changements environnementaux (Bürger *et al.*, 1989; Abu Awad and Roze, 2018), et/ou favorisent l'accumulation du fardeau de mutation (Perrier *et al.*, 2020). Ainsi, une manière d'expliquer le plus fort taux d'extinction des populations autogames pourrait être un taux d'extinction local très prononcé en limite d'aire de distribution, où l'on retrouve uniquement les taxons autogames, alors que le taux d'extinction pourrait être similaire en cœur d'aire à celui des espèces plus allogames (Barrett *et al.*, 2014).

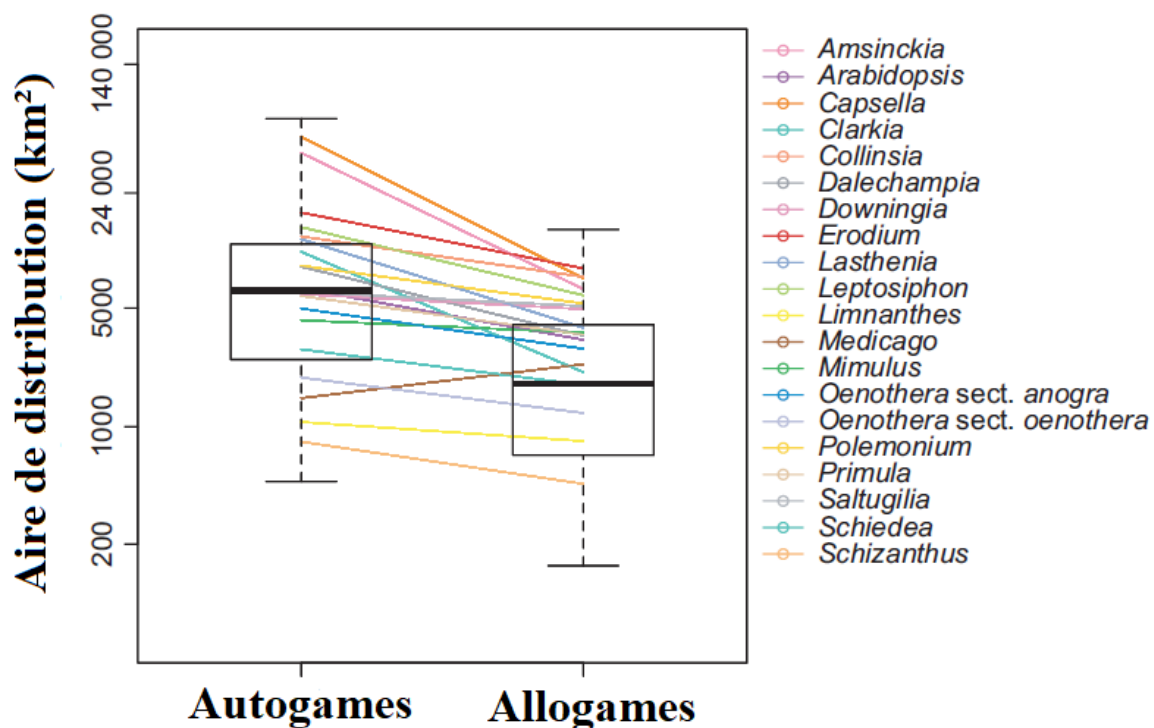


Figure 5. Comparaison de la taille des aires de distributions (en km²) entre espèces auto- et allofécondantes pour 20 différents clades d'espèces tempérées et tropicales. Figure modifiée de Grossenbacher *et al.* (2015).

Une autre piste intéressante à creuser serait de compléter les estimations des taux d'extinction des espèces en considérant les espèces à régimes mixtes et autogames. Jusqu'ici les modèles phylogénétiques ont en effet fait la distinction entre espèces auto-incompatible et auto-compatible (Goldberg *et al.*, 2010), notamment parce que les méthodes permettant d'inférer les différents paramètres du modèle (taux d'extinctions, spéciations notamment) à partir d'un arbre phylogénétique ne gèrent qu'un caractère à deux états (Maddison *et al.*, 2007). Or toutes les espèces auto-compatibles n'ont pas nécessairement le même taux d'extinction, qui pourrait par exemple être plus fort chez les espèces à régime mixte. En effet, la transition de l'allogamie vers l'autogamie, bien que pouvant permettre le sauvetage évolutif des populations, notamment lors de déclin de pollinisateurs (Busch and Delph, 2012), peut aussi représenter un suicide évolutif (Cheptou, 2004, 2019). En effet, les conséquences démographiques associées à la baisse de valeur sélective dues à la dépression de consanguinité peuvent entraîner un taux de croissance des populations inférieur à 1, sous le seuil de viabilité, et donc une extinction des populations (Cheptou, 2004). C'est notamment le cas lorsque le gain de valeur sélective associé

à l'assurance de reproduction est modéré, et ne compense pas l'expression de la dépression de consanguinité (Cheptou, 2019). Abu Awad & Billard (2017) ont également montré que lors d'une phase d'évolution de l'allofécondation vers l'autofécondation, l'accumulation de mutations délétères due à la baisse des tailles efficace et démographique des populations peut entraîner un processus de fonte mutationnelle. Ainsi, il apparaît que la transition de l'allogamie vers l'autogamie n'est pas sans risque pour les populations, et on pourrait supposer que le taux d'extinction plus fort pour les espèces auto-compatibles soit lié aux conséquences génétiques et démographiques qu'entraîne l'évolution vers l'autofécondation. Une telle analyse phylogénétique permettrait également de débattre de la stabilité des espèces à régime mixte (Winn *et al.*, 2011; Igic and Busch, 2013).

Perspectives

A travers cette discussion, j'ai présenté plusieurs arguments qui relativisent l'hypothèse du cul-de-sac évolutif. En particulier, les résultats de ma thèse suggèrent que l'autofécondation ne serait pas si désavantageuse sur le long terme, au moins lorsqu'elle est pratiquée dans les mêmes conditions écologiques que l'allogamie. Dans ce cadre, une piste de recherche prometteuse serait d'explorer plus finement les hypothèses alternatives qui pourraient conduire à ce patron de taux d'extinction plus élevé chez les espèces auto-incompatibles, en différenciant les taux d'extinction locaux et globaux, ou en affinant l'estimation des taux d'extinction des espèces auto-compatibles, en différenciant les espèces à régimes mixte et autogame. Cela demanderait néanmoins des analyses phylogénétiques poussées. En effet, le nombre d'états des modèles d'inférences phylogénétique (correspondant par exemple au nombre de classes d'autofécondation que l'on mettrait dans le modèle) détermine le nombre de paramètres à estimer, et même pour un modèle à trois états, des inférences précises nécessiteraient des jeux de données conséquents (voir Igic and Busch, 2013 pour une discussion sur le sujet), qui ne sont pour le moment pas disponibles. Des familles d'angiospermes avec un grand nombre d'espèces modèles, comme les brassicacées ou les poacées, pour lesquelles les systèmes de reproduction sont bien documentés et des génomes de référence sont disponibles, pourraient permettre d'explorer ces questions dans les années à venir. D'autres pistes plus simples à creuser seraient les tests de réversion du système de reproduction, et une étude expérimentale du potentiel évolutif des espèces en fonction de leur système de reproduction, qui sont des

questions certes expérimentalement lourdes, mais qui ne demandent pas de nouvelles méthodes à développer. Ainsi, l'étude de l'évolution de l'autofécondation et de ses conséquences sur le potentiel évolutif des populations ne semble toujours pas être un cul-de-sac pour la carrière des futures générations d'évolutionnistes.

Bibliographie

- Abu Awad D, Billiard S (2017). The double edged sword: The demographic consequences of the evolution of self-fertilization. *Evolution* **71**: 1178–1190.
- Abu Awad D, Gallina S, Bonamy C, Billiard S (2014). The Interaction between Selection, Demography and Selfing and How It Affects Population Viability. *PLOS ONE* **9**: e86125.
- Abu Awad D, Roze D (2018). Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution* **72**: 751–769.
- Ashman TL, Majetic CJ (2006). Genetic constraints on floral evolution: a review and evaluation of patterns. *Heredity* **96**: 343–352.
- Barrett SCH, Arunkumar R, Wright SI (2014). The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**: 20130344.
- Barrett SC, Husband BC (1990). Variation in outcrossing rates in *Eichhornia paniculata*: the role of demographic and reproductive factors. *Plant Species Biology* **5**: 41–55.
- Bodbyl Roels SA, Kelly JK (2011). Rapid evolution caused by pollinator loss in *Mimulus guttatus*. *Evolution: International Journal of Organic Evolution* **65**: 2541–2552.
- Bull JJ, Charnov EL (1985). On irreversible evolution. *Evolution* **39**: 1149–1155.
- Bürger R, Wagner GP, Stettinger F (1989). How much heritable variation can be maintained in finite populations by mutation–selection balance? *Evolution* **43**: 1748–1766.
- Busch JW, Delph LF (2012). The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Annals of botany* **109**: 553–562.
- Charlesworth D, Charlesworth B (1995). Quantitative genetics in plants: the effect of the breeding system on genetic variability. *Evolution* **49**: 911–920.

- Chelo, I.M., Afonso, B., Carvalho, S., Theologidis, I., Goy, C., Pino-Querido, A., *et al.* 2017. Partial selfing eliminates inbreeding depression while maintaining genetic diversity. *bioRxiv* 210005.
- Cheptou P-O (2004). Allee effect and self-fertilization in hermaphrodites: reproductive assurance in demographically stable populations. *Evolution* **58**: 2613–2621.
- Cheptou P-O (2019). Does the evolution of self-fertilization rescue populations or increase the risk of extinction? *Annals of Botany* **123**: 337–345.
- Cockerham CC, Weir BS (1984). Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics*: 157–164.
- Crow JF (1948). Alternative hypotheses of hybrid vigor. *Genetics* **33**: 477.
- David, J.L., Savy, Y. & Brabant, P. 1993. Outcrossing and selfing evolution in populations under directional selection. *Heredity* **71**: 642–651. Nature Publishing Group.
- Dobzhansky T (1937). *Genetics and the origin of species*. NY: Columbia Univ. Press. 364 p.
- Escobar JS, Cenci A, Bolognini J, Haudry A, Laurent S, David J, *et al.* (2010). An integrative test of the dead-end hypothesis of selfing evolution in Triticeae (Poaceae). *Evolution: International Journal of Organic Evolution* **64**: 2855–2872.
- Escobar, J.S., Nicot, A. & David, P. 2008. The Different Sources of Variation in Inbreeding Depression, Heterosis and Outbreeding Depression in a Metapopulation of *Physa acuta*. *Genetics* **180**: 1593.
- Fisher RA (1941). AVERAGE EXCESS AND AVERAGE EFFECT OF A GENE SUBSTITUTION. *Annals of Eugenics* **11**: 53–63.
- Geber MA, Griffen LR (2003). Inheritance and natural selection on functional traits. *International Journal of Plant Sciences* **164**: S21–S42.
- Gimond C, Jovelin R, Han S, Ferrari C, Cutter AD, Braendle C (2013). OUTBREEDING DEPRESSION WITH LOW GENETIC VARIATION IN SELFING CAENORHABDITIS NEMATODES. *Evolution* **67**: 3087–3101.
- Glémin S (2003). How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57**: 2678–2687.
- Glémin S, Ronfort J (2013). Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evolution: International Journal of Organic Evolution* **67**: 225–240.
- Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, Igić B (2010). Species Selection Maintains Self-Incompatibility. *Science* **330**: 493.

- Griffin PC, Willi Y (2014). Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. *Ecology Letters* **17**: 484–490.
- Grossenbacher D, Briscoe Runquist R, Goldberg EE, Brandvain Y (2015). Geographic range size is predicted by plant mating system. *Ecology letters* **18**: 706–713.
- Harkness A, Brandvain Y, Goldberg EE (2019). The evolutionary response of mating system to heterosis. *Journal of Evolutionary Biology* **32**: 476–490.
- Hartfield M, Glémin S (2014). Hitchhiking of Deleterious Alleles and the Cost of Adaptation in Partially Selfing Species. *Genetics* **196**: 281.
- Hartfield M, Glémin S (2016). Limits to Adaptation in Partially Selfing Species. *Genetics* **203**: 959.
- Herlihy CR, Eckert CG (2002). Genetic cost of reproductive assurance in a self-fertilizing plant. *Nature* **416**: 320–323.
- Herlihy CR, Eckert CG (2007). Evolutionary analysis of a key floral trait in *Aquilegia canadensis* (Ranunculaceae): genetic variation in herkogamy and its effect on the mating system. *Evolution: International Journal of Organic Evolution* **61**: 1661–1674.
- Holeski LM, Kelly JK (2006). Mating system and the evolution of quantitative traits: an experimental study of *Mimulus guttatus*. *Evolution* **60**: 711–723.
- Igic B, Busch JW (2013). Is self-fertilization an evolutionary dead end? *New Phytologist* **198**: 386–397.
- Kamran-Disfani A, Agrawal AF (2014). Selfing, adaptation and background selection in finite populations. *Journal of evolutionary biology* **27**: 1360–1371.
- Kelly JK (1999). Response to selection in partially self-fertilizing populations. I. Selection on a single trait. *Evolution* **53**: 336–349.
- Kelly JK, Arathi HS (2003). Inbreeding and the genetic variance in floral traits of *Mimulus guttatus*. *Heredity* **90**: 77.
- Kelly JK, Williamson S (2000). Predicting response to selection on a quantitative trait: a comparison between models for mixed-mating populations. *Journal of theoretical biology* **207**: 37–56.
- Lande R (1985). The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity* **54**: 323–332.
- Lande R, Porcher E (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* **200**: 891–906.

- Lande R, Schemske DW (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**: 24–40.
- Lowry E, Lester SE (2006). The biogeography of plant reproduction: potential determinants of species' range sizes. *Journal of Biogeography* **33**: 1975–1982.
- Lynch M (1991). The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45**: 622–629.
- Maddison WP, Midford PE, Otto SP (2007). Estimating a binary character's effect on speciation and extinction. *Systematic biology* **56**: 701–710.
- Martínez-Padilla, J., Estrada, A., Early, R., & García-González, F. (2017). Evolvability meets biogeography: evolutionary potential decreases at high and low environmental favourability. *Proceedings of the Royal Society B: Biological Sciences*, 284(1856), 20170516.
- Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10**: 315–331.
- Muller HJ (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1**: 2–9.
- Noël E, Jarne P, Glémin S, MacKenzie A, Segard A, Sarda V, *et al.* (2017). Experimental evidence for the negative effects of self-fertilization on the adaptive potential of populations. *Current Biology* **27**: 237–242.
- Oakley CG, Ågren J, Schemske DW (2015). Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. *Heredity* **115**: 73.
- Opedal ØH (2018). Herkogamy, a Principal Functional Trait of Plant Reproductive Biology. *International Journal of Plant Sciences* **179**: 677–687.
- Opedal ØH (2019). Evolutionary Potential of Herkogamy. *eLS*: 1–8.
- Opedal ØH, Bolstad GH, Hansen TF, Armbruster WS, Pélabon C (2017). The evolvability of herkogamy: Quantifying the evolutionary potential of a composite trait. *Evolution* **71**: 1572–1586.
- Paggi GM, Palma-Silva C, Bodanese-Zanettini MH, Lexer C, Bered F (2015). Limited pollen flow and high selfing rates toward geographic range limit in an Atlantic forest bromeliad. *Flora-Morphology, Distribution, Functional Ecology of Plants* **211**: 1–10.
- Pannell JR (2015). Evolution of the mating system in colonizing plants. *Molecular Ecology* **24**: 2018–2037.

- Perrier A, Sánchez-Castro D, Willi Y (2020). Expressed mutational load increases toward the edge of a species' geographic range. *Evolution*.
- Pujol B, Zhou S-R, Vilas JS, Pannell JR (2009). Reduced inbreeding depression after species range expansion. *Proceedings of the National Academy of Sciences* **106**: 15379–15383.
- Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity*, 83(4), 363-372.
- Rieseberg, L. H., Widmer, A., Arntz, A. M., & Burke, B. (2003). The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1434), 1141-1147.
- Schierup MH, Christiansen FB (1996). Inbreeding depression and outbreeding depression in plants. *Heredity* **77**: 461–468.
- Slotte, T., Hazzouri, K.M., Stern, D., Andolfatto, P. & Wright, S.I. 2012. Genetic architecture and adaptive significance of the selfing syndrome in *Capsella*. *Evol. Int. J. Org. Evol.* **66**: 1360–1374. Wiley Online Library.
- Stebbins GL (1957). Self fertilization and population variability in the higher plants. *The American Naturalist* **91**: 337–354.
- Takebayashi N, Morrell PL (2001). Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany* **88**: 1143–1150.
- Tedder A, Ansell SW, Lao X, Vogel JC, Mable BK (2011). Sporophytic self-incompatibility genes and mating system variation in *Arabis alpina*. *Annals of Botany* **108**: 699–713.
- Willi Y (2013). MUTATIONAL MELTDOWN IN SELFING *ARABIDOPSIS LYRATA*. *Evolution* **67**: 806–815.
- Winn AA, Elle E, Kalisz S, Cheptou P-O, Eckert CG, Goodwillie C, *et al.* (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution: International Journal of Organic Evolution* **65**: 3339–3359.
- Wright AJ, Cockerham CC (1985). Selection with partial selfing. I. Mass selection. *Genetics* **109**: 585–597.
- Wright SI, Kalisz S, Slotte T (2013). Evolutionary consequences of self-fertilization in plants. *Proc R Soc B* **280**: 20130133.

Annexe 1 - Digest: How mutational bias could explain the maintenance of sex

This article is published in *Evolution*.

Clo, J. (2018). Digest: How mutational bias could explain the maintenance of sex. *Evolution*, 72(9), 1970-1971.



Digest: How mutational bias could explain the maintenance of sex*

Josselin Clo^{1,2}

¹Unité Mixte de Recherche AGAP, Montpellier SupAgro, 2 place Pierre Viala, 34060 Montpellier Cedex 02, France

²E-mail: josselin.clo@supagro.fr

Received July 19, 2018

Accepted August 2, 2018

How does mutational bias affect the fitness of populations under different reproductive strategies? Vanhoenacker et al. (2018) found that mutational bias can greatly reduce the mean fitness of asexual populations, offering a new hypothesis for the maintenance of sex.

Understanding the prevalence of sexual reproduction among eukaryotes has been dubbed the “queen of problems” in evolutionary biology (Bell 1982). Sex comes with a cost—it is quite resource-consuming considering the production of males, expensive reproductive attributes, and behaviors such as locating or courting mates. Its ubiquity despite these biological limitations is thus puzzling. Though several hypotheses have been proposed to explain the maintenance of sex, empirical evidence is inconclusive (see Neiman et al. 2018 for a review). More experimental approaches are necessary to validate or refute current hypotheses and determine the interactions between different mechanisms involved in the maintenance or loss of sex.

Recently, theoretical works have focused on the effect of mating systems on the genetic architecture of adaptive traits and their consequences for mean population fitness. When modeling a quantitative trait under stabilizing selection, predominantly, and completely asexual populations should be favored, as the reduced (or nonexistent) recombination would aid in maintaining favorable genetic associations with compensatory epistatic effects on fitness (Fig. 1). Increasing sexual reproduction increases the recombination rate and contributes to breaking up the beneficial combinations, leading to less-well-adapted populations (Roze and Blanckaert 2014). However, these classical models make the assumption that the effects of mutations on fitness are symmetrically distributed around zero, increasing the probability of compensatory mutations, an unrealistic hypothesis (Charlesworth 2013).

In this issue, Vanhoenacker et al. (2018) developed mathematical and simulation models to explore the evolution of sex in the presence of mutational bias (an increased probability of new mutations moving the phenotype away from the optimum, introducing a component of directional selection; Fig. 1). Previously considered in a context where reproduction was only sexual, this approach is the first to bring to light the potential of mutational bias to affect a species’ reproductive strategy. Their results predict that, in regimes where mutations have weak to moderate effects on fitness, introducing a mutational bias greatly reduces the mean fitness of asexual populations compared to sexual ones. This is mainly caused by the combined effects of the mutational bias and drift that together contribute to moving the mean phenotype away from the optimum. The strength of the mutational bias on fitness can be influenced by several parameters, such as the number of loci under selection, population size, pleiotropic effects of loci on traits, and the dimensionality of the fitness landscape. Furthermore, their individual-based simulations confirm that in the presence of mutational bias, the rate of sex increases, whereas in its absence, asexuality is the more successful reproductive strategy.

Vanhoenacker et al.’s (2018) work offers an alternative explanation for the maintenance of sex. As the authors have mentioned, despite the links between theoretical parameters and experimental results not always being obvious, there is empirical evidence for mutational bias (Lyman et al. 1996). Even though the strength of this bias cannot be quantified, this work gives a new direction to explore in works aiming to understand the maintenance of sex. More empirical works are now needed to evaluate levels

*This article corresponds to Vanhoenacker, E., L. Sandell, and D. Roze. 2018. Stabilizing selection, mutational bias, and the evolution of sex. *Evolution*. <https://doi.org/10.1111/evo.13547>.

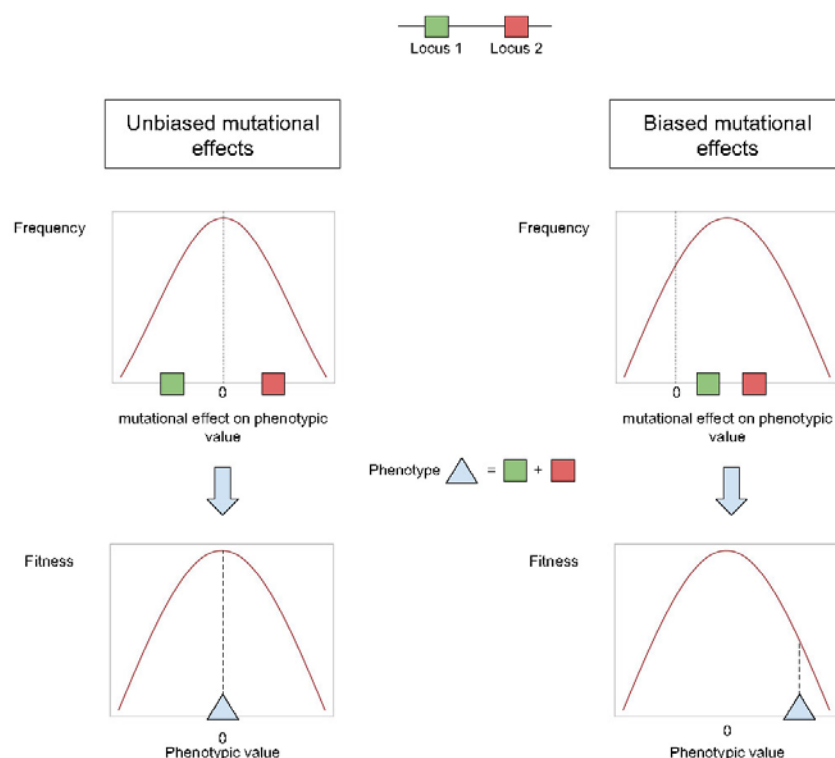


Figure 1. Schema presenting the effect of the distribution of mutational effects on compensatory associations between mutations. Under the hypothesis of unbiased mutational effects, a mutation has the same probability of increasing or decreasing the phenotypic value. Considering a two-locus system, given the value of the green mutation on one locus, the red mutation on the other has a probability tending toward 0.5 of being beneficial (compensating the effect of the green mutation, even if the red mutation could overshoot the optimum and lead to a more strongly deleterious phenotype). Under this scenario, asexuality is favored, as it allows the maintenance of compensatory associations (linkage disequilibrium) between mutations. If there is mutational bias, mutations tend to move the phenotype in a given direction (not toward 0), and hence away from the optimum, introducing a component of directional selection. In this case, given the value of the green mutation, the red mutation has a probability that is less than 0.5 of being advantageous. This favors sexual reproduction, as this type of deleterious interaction will be better selected against due to higher recombination rates.

of mutational bias, and once key aspects of mutational bias are better understood and identified (which traits, the strength of the bias, etc.), experiments on model species with variable levels of sex (such as *Townsendia hookeri*, a plant with sexual and asexual populations) will allow future research to test their predictions.

Roze, D., and A. Blanckaert. 2014. Epistasis, pleiotropy, and the mutation load in sexual and asexual populations. *Evolution* 68:137–149.

Vanhoenacker, E., L. Sandell, and D. Roze. 2018. Stabilizing selection, mutational bias and the evolution of sex. *Evolution* 72:1740–1758.

Associate Editor: K. Moore
Handling Editor: Mohamed A. F. Noor

LITERATURE CITED

- Bell, G. 1982. *The masterpiece of nature*. Croon Helm, London, U.K.
- Charlesworth, B. 2013. Stabilizing selection, purifying selection and mutational bias in finite populations. *Genetics* 194:955–971.
- Lyman, R. F., F. Lawrence, S. V. Nuzhdin, and T. F. C. Mackay. 1996. Effects of single *P* element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* 143:277–292.
- Neiman, M., P. G. Meirmans, T. Schwander, and S. Meirmans. 2018. Sex in the wild: how and why field-based studies contribute to solving the problem of sex. *Evolution* 72:1194–1203.

SUBMIT A DIGEST

Digests are short (~500 word), news articles about selected original research included in the journal, written by students or postdocs. These digests are published online and linked to their corresponding original research articles. For instructions on Digests preparation and submission, please visit the following link: <https://sites.duke.edu/evodigests/>.

Annexe 2 - Very local gene dispersal in the pseudometallicolous herb *Noccaea* *caerulescens* in the south of France

This article is under revision (first submission in *Heredity*).

Mathilde Mousset^{1,2,5}, **Josselin Clo**^{1,2}, Lea Prats¹, Juliette Pouzadoux¹, Stefano Donega³, David Carbonnel¹, Elodie Flaven¹, Christophe Petit¹, Patrice David⁴, François Rousset¹, Ophélie Roncé¹, Agnès Mignot¹

¹ Institut des Sciences de l'Evolution, Université de Montpellier, CNRS, IRD, EPHE, CC 065, Place Eugène Bataillon, 34095 Montpellier cedex 05, France

² Montpellier SupAgro, 2 place Pierre Viala, 34060 MONTPELLIER Cedex 02, France

³ International Center for Genetic Engineering and Biotechnology, AREA Science Park Padriciano 9934149 Trieste, Italy

⁴ Centre d'Ecologie Fonctionnelle et Evolutive, CEFE-UMR 5175, Campus CNRS, 1919 Route de Mende, 34293 Montpellier cedex, France

Abstract: Higher selfing rates have been predicted to evolve in polluted sites to prevent maladapted gene flow from non-polluted areas. Despite initial empirical support for this prediction, recent studies of variation in mating systems in pseudometallophytes have revealed patterns not consistent with this expectation. Whether gene flow constrains the evolution of local adaptation, and whether variation in mating system is primarily driven by its consequences for local adaptation has therefore been questioned. Here we explored the fine scale genetic structure and variation in selfing rates of the pseudometallophyte *Noccaea caerulescens* within abandoned mining sites, to address these questions. In particular, we estimated the spatial scale of gene flow within one site, and tested whether variation of local plant density alone could explain variation in selfing rates, within two former mines, using two methods to estimate selfing rates. We found no effect of density on the selfing rates of *Noccaea caerulescens* in either population, but observed that gene dispersal occurs at a very small scale (<2m). These results suggest that local adaptation to varying soil conditions is unlikely to be constrained by gene flow, while the causes of variation in selfing rates within and across sites remain elusive.

Annexe 3 - How mating systems influence evolvability

This book chapter is a part of a book named “Evolvability“. Thomas F. Hansen and Christophe Pelabon lead the redaction of the book.

How mating systems influence evolvability

Jacqueline Sztepanacz¹, Øystein H. Opedal², **Josselin Clo**³

¹ Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Canada

² Departement of Biology, Lund University, SE-223 62, Lund, Sweden.

³ AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, 34000 Montpellier, France.

L'autofécondation est souvent décrite comme un cul-de-sac évolutif. Du fait de leurs tailles efficaces plus petites, augmentant la force de la dérive, une purge plus efficace des mutations délétères, et des associations génétiques prédominantes, les populations autogames sont supposées avoir moins de diversité génétique, limitant ainsi leur potentiel évolutif. Néanmoins, peu de données empiriques supportent ces attendus théoriques, et la dynamique des populations autogames et leur potentiel adaptatif restent mal compris. L'objectif de ma thèse est de mieux comprendre la dynamique évolutive des populations d'espèces autogames et notamment le rôle potentiel de l'allogamie résiduelle dans cette dynamique et la réponse adaptative des populations aux contraintes environnementales. Pour y parvenir, je combine des approches théoriques et empiriques, en utilisant l'espèce préférentiellement autogame *Medicago truncatula*. Dans un premier temps, j'ai montré, via une méta-analyse à l'échelle des angiospermes, que si les populations autogames pouvaient avoir moins de variance génétique additive que les populations plus allofécondantes, le système de reproduction n'explique que très partiellement le niveau de diversité au sein des populations. Dans un deuxième temps, j'ai utilisé la modélisation afin de montrer que même si le niveau de variance génétique des populations autogames était inférieur, la remobilisation par allogamie résiduelle de la variance génétique stockée dans les associations entre gènes permet un niveau d'adaptation similaire à celui des populations allogames. J'ai ensuite montré que la différenciation génétique accumulée entre lignées autogames aux échelles intra- et inter-populationnelles, est en moyenne délétère lors d'évènements d'hybridation. Ceci est dû à un effet négatif de la dominance, mais ce mécanisme d'isolement n'est pas un frein à la potentiel remobilisation de la variance cachée chez *Medicago truncatula*. Finalement, j'ai montré que rendre plus réalistes les modèles théoriques de traits quantitatifs sous sélection stabilisante en introduisant de la dominance directionnelle à l'échelle phénotypique permet de réconcilier la théorie avec l'observation d'un faible effet du système de reproduction sur le potentiel évolutif des populations.

Mots clés : autofécondation, variance génétique, dominance, résidu d'allofécondation, dépression de croisement.

Self-fertilization is often described as an evolutionary dead end. Due to their smaller effective population size, increasing the strength of genetic drift, the more efficient purge of deleterious mutations, and predominant genetic associations, selfing populations are expected to harbour less genetic diversity, limiting their adaptive potential. However, there is little empirical data supporting these theoretical expectations, and the dynamics and adaptive potential of selfing populations remain poorly understood. The aim of this thesis is to better understand the evolutionary dynamics of selfing populations, and notably the possible role of residual allogamy in this dynamics and in populations' response to environmental constraints. To do so, I combined theoretical and empirical approaches, using the predominantly selfing species *Medicago truncatula*. First, I performed a meta-analysis in angiosperm species and showed that if selfing populations have less genetic diversity than their outcrossing counterparts can, the mating system is a poor predictor of the amount of genetic diversity within populations. I then used models to show that even if the amount of genetic variance is smaller in selfing populations, the remobilization by residual allogamy of the genetic variance stored in genetic associations allows a similar level of adaptation as in outcrossing populations. I then showed that the genetic differentiation accumulated between inbred lines, at within- and among-population scales, is deleterious, on average, during hybridization events. This is due to negative dominance effects, but this mechanism of reproductive isolation does not hinder the remobilization of the hidden genetic variance in *Medicago truncatula*. Finally, we showed that making theoretical models of quantitative traits under stabilizing selection more realistic, by adding directional dominance at the phenotypic scale, could reconcile theory with empirical data that report a weak effect of the mating system on the adaptive potential of populations.

Keywords: Self-fertilization, genetic variance, dominance, residual allogamy, outbreeding depression