HAL
open science

Contribution à une théorie physiologique et génétique de l'évolution végétale : fardeau génétique, systèmes de reproduction et évolution du taux de mutation dans les populations structurées en classes

Thomas Lesaffre

To cite this version:

Thomas Lesaffre. Contribution à une théorie physiologique et génétique de l'évolution végétale : fardeau génétique, systèmes de reproduction et évolution du taux de mutation dans les populations structurées en classes. Biologie végétale. Université de Lille, 2021. Français. NNT : 2021LILUR014 . tel-03469618

HAL Id: tel-03469618

https://theses.hal.science/tel-03469618v1

Submitted on 7 Dec 2021

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.

École doctorale n°104 : Science de la Matière, du Rayonnement et de l'Environnement

Thèse de doctorat

pour obtenir le grade de docteur délivré par

Université de Lille

Contribution à une théorie physiologique et génétique de l'évolution végétale

Fardeau génétique, systèmes de reproduction et évolution du taux de mutation dans les populations structurées en classes

---

présentée et soutenue publiquement par

Thomas LESAFFRE

le 16 Mars 2021

Composition du jury

Jacqui Shykoff, Directrice de recherche CNRS, Université Paris-Saclay	Rapportrice
Laurent Lehmann, Professeur, Université de Lausanne	Rapporteur
Christelle Fraïsse, Chargée de recherche CNRS, Université de Lille	Examinatrice
Thomas Lenormand, Directeur de recherche CNRS, Montpellier	Examinateur
Joël Cuguen, Professeur, Université de Lille	Président du jury
Sylvain Billiard, Maître de conférence, Université de Lille	Directeur de thèse

Laboratoire Évolution, Écologie et Paléontologie (EEP)

UMR CNRS 8198, F-59650 Cité Scientifique, France

On peut montrer, par des arguments d'une rigueur intellectuelle dont seuls des doctorants du laboratoire EEP sont capables, que la tarte aux pommes figure au panthéon des desserts les plus ennuyeux que la terre ait portés. Ce dessert présente une analogie surprenante avec une thèse de doctorat. En effet, nous avons dans les deux cas affaire à un objet dont la confection prend beaucoup de temps pour un résultat somme toute assez décevant.

Table des matières

I Introduction générale **2**

1 Histoires de vie **4**

1.1 Une diversité ordonnée 5

1.2 Quelques mots de physiologie végétale 7

1.2.1 Croissance 7

1.2.2 Reproduction sexuée 11

1.2.3 Longévité 12

1.2.3.1 Sénescence 13

1.2.3.2 Facteurs de mortalité extrinsèques 16

1.3 Théorie des traits histoires de vie 18

1.3.1 Avantages relatifs de l'annualité et de la pérennialité 18

1.3.2 Évolution de l'effort reproducteur 23

1.3.2.1 Évolution de la longévité par le trade-off survie/reproduction 25

1.3.2.2 Variation de l'effort reproducteur avec l'âge 29

1.3.2.3 Croissance, reproduction et survie 35

1.3.3 Évolution de l'âge à maturité 37

1.4 Conclusion 42

2 Autofécondation et dépression de consanguinité **45**

2.1 Avantages évolutifs de l'autofécondation 46
2.2 Dépression de consanguinité 48
2.2.1 Évolution du taux d'autofécondation avec une dépression de consanguinité fixée 48
2.2.2 Structure génétique de la dépression de consanguinité 51
2.2.2.1 Superdominance 52
2.2.2.2 Mutations délétères partiellement récessives 55
2.2.2.3 Importance relative des deux mécanismes 57
2.2.3 Purge des mutations délétères, interférence sélective et dépression de consanguinité 58
2.2.3.1 Purge des mutations délétères 58
2.2.3.2 Interférence sélective 59
2.3 Sur le lien entre histoire de vie et régime de reproduction 61
2.3.1 Évolution de l'autofécondation chez les espèces pérennes 61
2.3.1.1 Autofécondation antérieure et concomitante 63
2.3.1.2 Autofécondation retardée 66
2.3.2 Effet de l'autofécondation sur l'évolution de la longévité 68

3 Travail réalisé et méthodes employées dans cette thèse 72
3.1 Le travail réalisé dans cette thèse 72
3.1.1 Avant-propos 72
3.1.2 Premier chapitre : point de départ 73
3.1.3 Deuxième et troisième chapitres : génétique des populations structurées en classe et évolution du taux de mutation 75
3.1.4 Chapitre quatre : Changement d'angle 78
3.2 Le cadre analytique employé 78
3.2.1 Bases générales du cadre analytique de Barton and Turelli (1991) 79
3.2.1.1 Définition des variables 79

3.2.1.2 Du génotype au phénotype et du génotype à la fitness 82
3.2.1.3 Sélection 84
3.2.1.4 Autofécondation partielle, recombinaison et mutation 85
3.2.2 Apports méthodologiques : chevauchement de générations, lien phénotype-fitness et variables âge-dépendantes 88

II Travaux de thèse 94

1 The joint evolution of lifespan and selfing 96
1.1 Synthèse en français 96
1.1.1 Motivation 96
1.1.2 Modèle 97
1.1.3 Résultats clefs 98
1.2 Introduction 99
1.3 Methods 103
1.3.1 Outline of the model 103
1.3.2 Analytical methods 105
1.3.3 Numerical analyses and simulations methods 105
1.4 Results 106
1.4.1 Evolution of lifespan in a partially selfing population 107
1.4.2 Influence of lifespan on the evolution of self-fertilisation 109
1.4.3 Joint evolution of lifespan and self-fertilisation 112
1.5 Discussion 116
1.6 Appendices 122
1.6.1 Proportion of selfed individuals and lifetime inbreeding depression 122
1.6.2 Recursions for the evolution of reproductive effort in a partially selfing population 125

1.6.3 Recursions for the evolution of self-fertilisation in a perennial population . . . 130
1.6.4 Numerical analyses . . . 135
1.6.5 Inbreeding depression affecting fecundity . . . 138

2 On deleterious mutations in perennials : inbreeding depression, mutation load and life-history evolution 140

2.1 Synthèse en français . . . 140
2.1.1 Motivation . . . 140
2.1.2 Modèle . . . 141
2.1.3 Résultats clefs . . . 143
2.1.3.1 Effets sélectifs des mutations, nombre de mutations maintenues à l'équilibre et conséquences pour l'évolution des histoires de vie . . . 143
2.1.3.2 Dépression de consanguinité . . . 144
2.2 Introduction . . . 146
2.3 Model outline and methods . . . 149
2.4 Results . . . 155
2.4.1 Mutation-selection equilibrium . . . 156
2.4.2 Fitness effect of mutations neglecting genetic associations . . . 158
2.4.3 Average number of mutations, inbreeding depression and mutation load . . . 161
2.4.4 Consequences for life-history evolution . . . 165
2.5 Discussion . . . 166
2.6 Appendices . . . 173
2.6.1 Growth model . . . 173
2.6.2 General recursions for the effects of selection and reproduction under partial selfing . . . 175

2.6.2.1 Defining variables 175
2.6.2.2 Selection 178
2.6.2.3 Reproduction and mutation under partial selfing . . . 181
2.6.3 Analytical results 183
2.6.3.1 Lifetime fitness approach 183
2.6.3.2 Life cycle approach 190
2.6.3.3 Inbreeding depression and mutation load 195
2.6.4 Fitness landscapes 197
2.6.5 Comparison of mean phenotypic deviation for different magnitude of effect of mutations 199
2.6.6 Mutations per haploid genome, inbreeding depression and mutation load for other parameter values 201

3 Population-level consequences of inheritable somatic mutations and the evolution of mutation rates in plants 209

3.1 Synthèse en français 209
3.1.1 Motivation 209
3.1.2 Modèle 210
3.1.3 Résultats clefs 211
3.2 Introduction 212
3.3 Methods 216
3.4 Results 221
3.4.1 Evolutionarily stable mutation rate 221
3.4.2 Mutation-selection balance 226
3.5 Discussion 227
3.6 Appendices 233
3.6.1 Derivation of analytical results 233
3.6.1.1 Mutation rate evolution 233

3.6.1.2 Mutation-selection balance 244
3.6.2 Additional graphs 249

4 (W.I.P) A theoretical evaluation of indirect inbreeding depression estimators 253

4.1 Synthèse en français 253
4.1.1 Motivation 253
4.1.2 Modèle & résultats clefs 255
4.1.2.1 Estimateur de Ritland 255
4.1.2.2 Fardeau de consanguinité 256
4.2 Introduction 257
4.3 Methods 259
4.3.1 Model outline 259
4.3.1.1 Simulated population 259
4.3.1.2 Growth model 260
4.3.1.3 Genetic assumptions 260
4.3.2 Measurements 262
4.3.2.1 True inbreeding depression 262
4.3.2.2 Inbreeding coefficients 262
4.3.2.3 Ritland (1990)'s inbreeding depression estimator . . . 263
4.3.2.4 Inbreeding load estimates 264
4.4 Results & (some) discussion 265
4.4.1 Ritland's estimator 266
4.4.2 Inbreeding load 269
4.5 Appendices 273
4.5.1 Analytical prediction of juvenile inbreeding depression 273
4.5.2 Results in the absence of mutations ($U = 0$) 275

5 Conclusion **277**

5.1 Retour à la question initiale 277

5.2 Vers une une théorie physiologique et génétique de l'évolution végétale 281

5.2.1 Modéliser la croissance végétale 281

5.2.2 Questions théoriques associées 283

Première partie

Introduction générale

Chapitre 1

Histoires de vie

Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one ; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved.

- Charles Darwin, *On the Origin of Species* (1859), p.429

Dans l'ultime paragraphe de *L'Origine des espèces*, Darwin célébrait la diversité des formes de vie, citant l'émergence de ce qu'on nommait alors les animaux supérieurs comme la plus grande merveille produite par l'évolution. Il aurait aisément pu en dire autant du règne végétal. Des pins Bristlecone de Californie dont certains spécimens avoisinent les cinq mille ans d'âge (Lanner and Connor, 2001), aux éphémères fleurs du désert qui complétent leur cycle de vie en quelques jours (Friedman, 2020), pour ne citer que des cas extrêmes, les végétaux présentent en effet une très grande diversité

d'histoires de vie. C'est à cette diversité que nous allons nous intéresser dans ce premier chapitre d'introduction. Il est indispensable, dès lors, que nous définissions ce que nous entendrons par histoire de vie tout au long de ce manuscrit. De manière très large, nous appellerons histoire de vie d'un organisme l'ensemble des fécondités et chances de survie dont il sera doté au cours de son existence, à chacun des âges qu'il est susceptible d'atteindre (Partridge and Sibly, 1991; Stearns, 1992). Après avoir esquissé une façon de classifier les histoires de vie végétales, nous aborderons succinctement les mécanismes physiologiques qui les sous-tendent avant de discuter des approches théoriques traitant de leur évolution.

1.1 Une diversité ordonnée

Pour donner du sens à une collection d'éléments divers, le premier réflexe de l'esprit humain est en général de chercher à l'ordonner, à la séparer en catégories discrètes. La manière la plus simple d'ordonner les histoires de vie chez les plantes tire profit de la nature saisonnière de la vie végétale (Friedman, 2020). En fonction du nombre de mauvaises saisons qu'une espèce est capable d'endurer sans mourir, on distingue les espèces annuelles, qui naissent, croissent, se reproduisent et meurent en une unique saison favorable, des espèces pérennes qui survivent à au moins une mauvaise saison. Par ailleurs, en fonction du nombre d'épisodes de reproduction réalisés par un individu au cours de son existence, on distingue les espèces monocarpes, qui se reproduisent une seule fois, des espèces polycarpes, qui se reproduisent de manière répétée. Parmi les espèces annuelles on distingue les annuelles obligatoires, qui sont monocarpes, des annuelles facultatives, qui continuent à produire des inflorescences jusqu'à ce qu'un événement saisonnier (par exemple, une chute de température) entraîne leur mort. L'écrasante majorité des espèces pérennes sont quant à elles polycarpes, mais des espèces pérennes monocarpes existent, comme l'Agave américain (*Agave americana*),

qui se développe pendant plusieurs dizaines d'années avant d'engager un unique épisde de reproduction qui se solde par la mort de l'individu (Schaffer and Schaffer, 1977). Lorsqu'une espèce germe, endure un hiver et fleurit l'année suivante, on parle parfois d'espèce bisannuelle. Nous considérerons ici les espèces bisannuelles comme des espèces pérennes monocarpes à courte durée de vie.

Cette façon de classifier les histoires de vie végétales constitue une base utile pour ancrer les concepts au niveau verbal mais elle occulte l'essentiel de la diversité qu'elle cherche à décrire, car si les espèces annuelles représentent un ensemble relativement bien défini, les espèces dites pérennes varient en longévité par plusieurs ordres de grandeur, de plusieurs siècles pour certaines espèces ligneuses comme le Chêne pédonculé (*Quercus robur*) à quelques années pour beaucoup d'herbacées (Ehrlén and Lehtilä, 2002; Baudisch et al., 2013). De plus, ces grandes différences de longévité entre espèces pérennes sont corrélées aux variations d'un cortège d'autres traits d'histoire de vie. En effet, une espèce plus longévive est en général de plus grande taille (Aarssen et al., 2006; Marbà et al., 2007), investit une part moins importante de ses ressources dans la reproduction et devient mature sexuellement plus tardivement (Franco and Silvertown, 1996; Silvertown et al., 2001; Salguero-Gómez et al., 2016; Salguero-Gómez, 2017). Ces variations conjointes constituent ce que nous appellerons ici un syndrôme de pérennialité. Par opposition, nous pouvons également décrire un syndrôme d'annualité, qui désigne les espèces à faible longévité, au développement et à la maturation rapides, qui investissent massivement dans la reproduction (Friedman, 2020). Ces deux syndrômes, chacun poussés à leurs extrêmes, permettent de définir les bornes d'un continuum d'histoires de vies le long duquel les espèces végétales se répartissent.

1.2 Quelques mots de physiologie végétale

Afin d'étudier l'évolution des histoires de vie végétales le long de ce continuum, il est utile que nous discutions en quelques mots, fussent-ils caricaturaux, du fonctionnement physiologique des plantes. Les sections qui vont suivre abordent donc à gros traits les aspects les plus importants d'une histoire de vie, à savoir la croissance, la reproduction et la survie.

1.2.1 Croissance

Croissance primaire. Ramenée à son principe essentiel, la croissance est le processus par lequel un organisme prélève des ressources dans l'environnement pour les convertir en biomasse. Chez les végétaux, la production de biomasse est assurée par la division active de cellules indifférenciées dans des zones particulières qu'on appelle des méristèmes. Le méristème apical d'une tige produit l'ensemble des cellules qui la composent. Ce faisant, il produit des méristèmes axillaires, qui sont typiquement situés aux aisselles des bourgeons foliaires, demeurent dormants dans un premier temps et se divisent abondamment lorsqu'ils sont activés pour produire une nouvelle branche. Ils deviennent alors le méristème apical de leur branche et produisent leurs propres méristèmes axillaires (Lehtilä and Larsson, 2005; Burian et al., 2016). Par ce processus itératif de croissance, l'appareil végétatif d'une plante se construit comme un réseau de modules physiologiquement équivalents (Thomas, 2013). De nouveaux modules émergent donc continuellement des plus anciens. Les anciens méristèmes apicaux cessent leur activité au bout d'un certain temps du fait de mécanismes de mort programmée, si bien que la croissance végétale peut être interprétée comme une fuite. Tant que les méristèmes apicaux actifs parviennent à maintenir un temps d'avance sur la vague de mort programmée qui les poursuit, l'appareil végétatif persiste (Thomas et al., 2000). Le processus de croissance que nous venons de décrire porte le nom de

croissance primaire.

Croissance secondaire. Chez les espèces ligneuses, les modules morts continuent de jouer le rôle d'appareil vasculaire pour le passage de la sève entre les racines et les méristèmes apicaux actifs. Ce faisant, ils continuent à croître en épaisseur par un processus de croissance secondaire alimenté par un méristème particulier, le cambium. Ce processus est la clef des très grandes tailles atteintes par certains arbres, car une plus grande épaisseur permet de mieux supporter mécaniquement le poids de l'appareil végétatif ainsi que d'assurer une meilleure circulation de la sève (Groover et al., 2006; Baucher et al., 2007).

Allométrie et contraintes hydrauliques. La production de biomasse requiert que les nutriments prélevés dans le sol par le système racinaire et les produits de la photosynthèse soient acheminés jusqu'aux méristèmes. À mesure qu'une plante grandit, ses surfaces foliaires et racinaires augmentent et lui permettent de prélever davantage de ressources dans l'environnement. Dans le même temps, la distance à parcourir entre les racines et les méristèmes grandit, ce qui a pour effet inverse de réduire la vitesse à laquelle la plante est capable de prélever des ressources du fait de contraintes hydrauliques croissantes (Niklas and Enquist, 2001; Munné-Bosch, 2018). Il en résulte que la capacité d'une plante à consommer des ressources (son taux métabolique) croît de moins en moins vite à mesure qu'elle grandit, proportionnellement à la puissance 3/4 de sa taille totale (Encadré A - Enquist, 2002; Price et al., 2010).

Encadré A: Modèle métabolique de croissance de West et al. (2001)

La relation en puissance 3/4 qui existe entre taux métabolique et taille n'est pas une spécificité du monde végétal. En fait, elle a d'abord été observée empiriquement chez les animaux (Peters, 1983). Les arguments théoriques qui

interviennent dans sa justification chez ces derniers font appel à des notions de géométrie fractale et sont différents de ceux invoqués chez les végétaux (West et al., 1999). Indépendamment de la manière dont elle est justifiée, cette relation est la base d'une classe de modèles de croissance dits métaboliques, dont le plus simple est le modèle proposé par West et al. (2001). Ce modèle jouant un rôle de premier plan dans les chapitres deux et quatre du travail présenté dans cette thèse, il est décrit en détails dans cet encadré.

Le taux métabolique B_t d'un individu de taille G_t au temps t, c'est-à-dire la quantité totale d'énergie dont il dispose, est donné par

$$B_t = B_0 G_t^{3/4}, \tag{A.1}$$

où B_0 est une constante qui décrit le taux métabolique basal. Dans le modèle de West et al. (2001), une partie de cette énergie est d'abord consommée par le maintien des unités de biomasse existantes, et le reste est dédié à la production de nouvelles unités. Supposons que pour être maintenue, une unité de biomasse requiert toujours la même quantité d'énergie c que nous appellerons *coût de maintien*. L'énergie requise pour maintenir l'ensemble de l'organisme est alors donnée par le produit $c\ G_t$ et augmente linéairement avec la taille. Sous l'hypothèse que la production d'une unité de biomasse consomme aussi toujours la même quantité d'énergie ε, que nous appellerons *coût de production*, la croissance d'un individu peut être décrite par l'équation différentielle suivante

$$B_0 G_t^{3/4} = c\ G_t + \varepsilon \frac{\mathrm{d}G_t}{\mathrm{d}t}. \tag{A.2}$$

Une solution de cette équation est

$$G_t = \frac{e^{-\frac{c}{\varepsilon}t}}{c^4}\left(1 - B_0 e^{\frac{c}{4\varepsilon}t}\right). \tag{A.3}$$

Le comportement de cette solution en fonction du temps, et la dynamique de ressources correspondante en fonction de la taille, sont présentés dans la Figure 1.1. Après une première phase exponentielle de croissance durant laquelle les ressources requises pour le maintien augmentent moins vite que les ressources totales, la vitesse de croissance passe par un maximum et approche asymptotiquement de zéro, de sorte que le processus de croissance plafonne nécessairement à une taille maximale qui est donnée par

$$\lim_{t\to\infty} G_t = \frac{1}{c^4}. \tag{A.4}$$

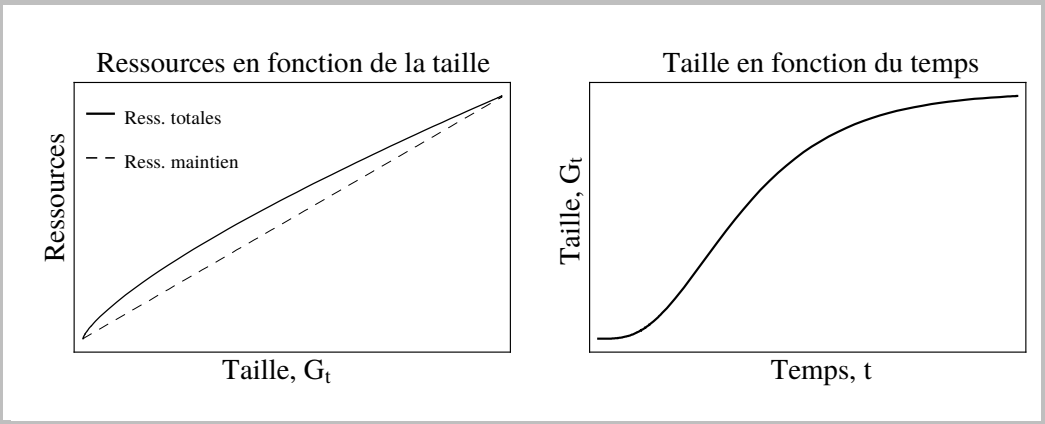

FIGURE 1.1 – Évolution de la disponibilité des ressources et de la taille d'un individu au cours du temps dans le modèle de West et al. (2001). Les valeurs de paramètres utilisées sont $B_0 = 1$, $c = \varepsilon = 0.01$.

Compétition. Outre les contraintes hydrauliques décrites plus haut, il est bien établi que la capacité d'une plante à prélever des ressources dans le milieu est fortement limitée par la compétition avec les individus proches à la fois au niveau racinaire pour l'acquisition de nutriments, et au-dessus du sol pour l'accès à la lumière (Damgaard et al., 2002; Petit and Hampe, 2006; Trinder et al., 2012; Kiær et al., 2013). Les capacités compétitives d'une plante sont très liées à sa taille. En effet, plus une plante est grande et plus elle bénéficie d'un avantage compétitif qui augmente plus que linéairement avec la taille (Nord-Larsen et al., 2006). On parle de compétition asymétrique en taille pour désigner ce phénomène, qui a pour conséquence que plus une plante est grande et plus elle a accès aux ressources (Weiner, 1990).

1.2.2 Reproduction sexuée

Pour se reproduire, une plante transforme une partie de ses méristèmes végétatifs en méristèmes floraux (Friedman, 2020). En conséquence, il existe chez les plantes une forte relation entre taille et fécondité car plus un individu est grand et plus il possède de méristèmes disponibles pour la reproduction (Harper, 1977; Klinkhamer et al., 1985; Weiner et al., 2009). Cette dernière s'opère donc au détriment de la croissance chez les végétaux, puisque les méristèmes alloués à la production de fleurs ne sont plus disponibles pour participer à la croissance végétatif (Obeso, 2002). Par ailleurs, la transition d'un méristème de l'état végétatif a l'état floral est presque toujours irréversible (Geber, 1990), si bien que dédier une fraction importante de ses méristèmes à la reproduction peut également avoir un impact sur la survie d'une plante. En effet, si l'ensemble des méristèmes d'une plante passent à l'état floral, aucun méristème ne subsistera à l'issue de la reproduction pour assurer le maintien de l'appareil végétatif, ce qui se soldera inexorablement par la mort de l'individu. Cela nous offre une manière différente d'appréhender la diversité des histoires de vie chez les plantes. Une plante monocarpe est une plante qui dédie l'ensemble de ses méristèmes à la reproduction

après une certaine période de croissance, qui peut être courte dans le cas d'une espèce annuelle, ou plus longue dans le cas d'une espèce pérenne. Une plante polycarpe conserve, elle, une fraction non-nulle de ses méristèmes à l'état végétatif afin d'alterner les phases de croissance et de reproduction (Figure 1.2 - Friedman and Rubin, 2015; Friedman, 2020).

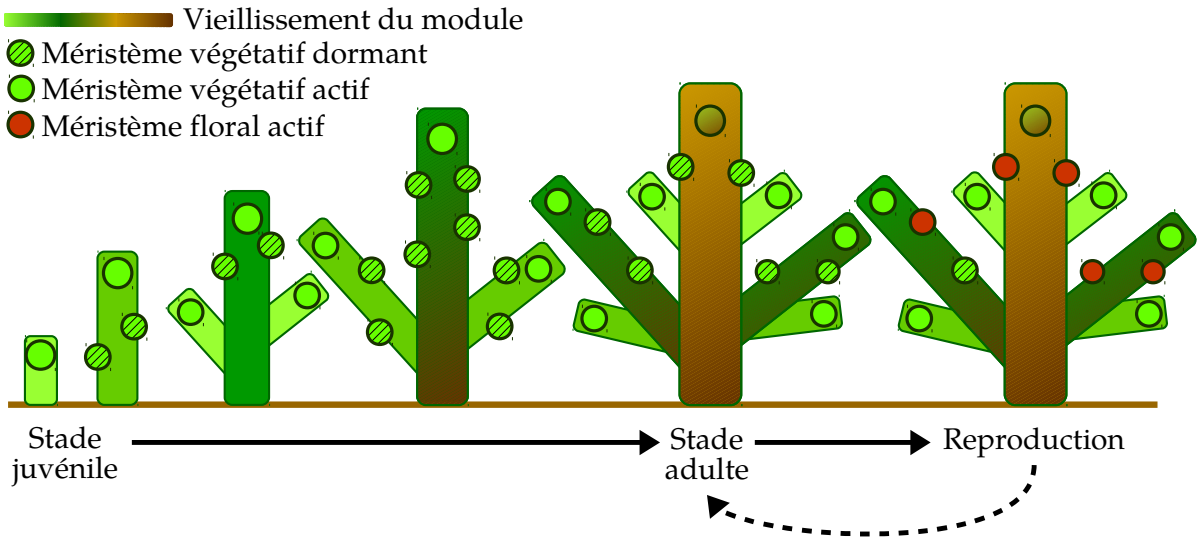

FIGURE 1.2 – Représentation schématique de la croissance végétale et de l'allocation de méristèmes à la reproduction.

1.2.3 Longévité

La notion de longévité n'est pas tout à fait triviale chez les végétaux car beaucoup d'espèces sont capables de reproduction clonale. Dans ces espèces, un individu génétique est composé de plusieurs clones physiologiquement indépendants, si bien que la mort d'un clone n'implique pas nécessairement celle de l'individu (Harper, 1977). Nous nous intéresserons ici à la longévité d'un clone dans le cas d'une espèce pérenne. Comme pour tout organisme, les facteurs de mortalité d'une plante peuvent au premier ordre être séparés entre facteurs intrinsèques et extrinsèques. Les facteurs intrinsèques désignent le flétrissement de l'appareil végétatif au cours des années, c'est-à-dire la

sénescence, et les facteurs extrinsèques l'ensemble des perturbations du milieu susceptibles de causer la mort de l'individu.

1.2.3.1 Sénescence

La vision selon laquelle aucun organisme vivant n'échappe à la sénescence, que nous définirons ici comme le déclin de la fécondité et des chances de survie d'un organisme avec l'âge, a longtemps prévalu en biologie évolutive (Munné-Bosch, 2015). Plusieurs théories ont été avancées pour expliquer l'existence de la sénescence.

Efficacité de la sélection. Puisque les chances de parvenir jusqu'à un certain âge sont d'autant plus faibles que cet âge est élevé, si bien que les individus âgés sont typiquement moins nombreux que les individus plus jeunes, l'efficacité de la sélection sur une mutation affectant la fitness devrait être d'autant plus faible qu'elle s'exprime tardivement (Williams, 1966a). Cette prédiction est supportée par le résultat théorique suivant. Supposons que le taux de croissance intrinsèque soit une bonne mesure de fitness (Encadré B), et que les chances de survie de l'âge x à $x+1$ sont données par p_x. Nous avons alors que la probabilité qu'un individu survive jusqu'à l'âge x est donnée par

$$l_x = \prod_{k=0}^{x-1} p_k, \tag{1.1}$$

si bien que nous pouvons dériver l'équation d'Euler-Lotka (Encadré B) par la probabilité de survie p_a pour obtenir

$$0 = \frac{1}{p_a} \sum_{x=a+1}^{\infty} l_x m_x e^{-rx} - \frac{\partial r}{\partial p_a} \sum_{x=0}^{\infty} m_x l_x x e^{-rx} \Leftrightarrow \frac{\partial r}{\partial p_a} = \frac{\frac{1}{p_a} \sum_{x=a+1}^{\infty} l_x m_x e^{-rx}}{\sum_{x=0}^{\infty} m_x l_x x e^{-rx}}. \tag{1.2}$$

Ainsi, comme le dénominateur est constant pour tous les âges et le numérateur décroît quand a augmente, car moins de classes d'âge sont concernées par la somme, nous pouvons en déduire que la sensibilité de la fitness à la variation de la probabilité de

survie à un âge donné décroît avec l'âge, et que la sélection est donc relâchée sur les classes d'âges plus vieilles. Le même raisonnement peut être appliqué au paramètre de fécondité m_x.

Mutations à effets tardifs et pléiotropie antagoniste. Deux mécanismes résultent de la diminution de l'efficacité de la sélection avec l'âge. D'abord, les mutations délétères s'exprimant tardivement dans la vie des organismes auront tendance à être moins bien purgées et donc à davantage s'accumuler dans la population, induisant ainsi la diminution des capacités de reproduction et de survie des organismes âgés. Ensuite, des mutations à effets pléiotropes, c'est-à-dire affectant plusieurs traits, qui auraient pour effet d'augmenter les performances au début de l'existence au détriment de la fin pourraient, suivant la balance entre leurs effets bénéfiques précoces et négatifs tardifs, être favorisées par la sélection et générer de la sénescence (Williams, 1957).

Accumulation de mutations et lignée somatique jetable. Les deux mécanismes évoqués ci-dessus se fondent sur des bases purement génétiques, aveugles aux mécanismes physiologiques qui pourraient en principe les sous-tendre. La théorie de la lignée somatique jetable (*disposable soma theory*, Kirkwood, 1977) propose une interprétation physiologique de la sénescence qui recouvre partiellement ces deux théories génétiques. Aucun mécanisme de réplication cellulaire n'est parfait, si bien que le génome subit nécessairement des dégradations au fil des divisions cellulaires (Medawar, 1952). Kirkwood propose que ces dégradations pourraient être tempérées par l'investissement par l'individu dans le maintien et la réparation de son organisme. Si investir des ressources dans le maintien et la production de la lignée germinale, c'est-à-dire dans la reproduction, est indispensable pour avoir une fitness non-nulle, l'investissement de ressources dans le maintien de la lignée somatique est, lui, moins critique car cette dernière s'éteint fréquemment du fait d'événements de mortalité extrinsèques et n'est qu'un simple véhicule pour la lignée germinale (Kirkwood, 2017). Il en résulte que, les

ressources d'un organisme étant limitées, l'investissement dans le maintien de la lignée somatique ne devrait jamais suffire à éviter son flétrissement progressif.

Ces différentes théories, dont la généralité argumente en faveur de l'universalité de la sénescence, sont cependant mises en échec par l'observation chez les plantes pérennes. En effet, aucune sénescence n'est en général observée chez ces dernières, à part chez certains arbres qui atteignent des âges très élevés (Baudisch et al., 2013). Il existe donc des mécanismes qui permettent à la majorité des espèces végétales d'échapper au flétrissement de leurs capacités de reproduction et de survie avec l'âge.

Non-séparation de la lignée germinale. Chez les plantes, il faut qu'un méristème effectue la transition vers l'état floral pour que des gamètes soient produits (section 1.2.2). Ainsi, la lignée somatique contribue à l'hérédité car c'est d'elle qu'émerge la lignée germinale (Lanfear, 2018). Elle n'est donc pas un simple véhicule pour les gènes transmis par la lignée germinale de sorte que l'idée d'une lignée somatique jetable n'est pas applicable au règne végétal. En conséquence, des mécanismes de réduction de l'accumulation de mutations au cours de la croissance ont évolué chez les végétaux. Les cellules destinées à constituer les méristèmes axillaires sont spécifiées très tôt lorsqu'elles sont produites par un méristème apical, et demeurent quasi-quiescentes jusqu'à l'activation du méristème qu'elles constituent (Burian et al., 2016). Cela a pour effet de réduire considérablement le nombre de divisions cellulaires séparant un méristème axillaire du méristème apical dont il est issu, réduisant ainsi la vitesse à laquelle les mutations s'accumulent dans l'appareil végétatif au cours de la croissance. De plus, les méristèmes sont composés de plusieurs lignées cellulaires entre lesquelles des processus de sélection peuvent avoir lieu (Klekowski, 2003). Ainsi, les plantes disposent non seulement de mécanismes permettant de limiter la production de mutations, mais aussi de mécanismes permettant d'éliminer les mutations produites au cours de la croissance.

Modularité. Chaque méristème donnant naissance à une branche physiologiquement équivalente à toutes les autres branches produites, un appareil végétatif peut être décrit comme une population de modules qui suit une dynamique de remplacement à mesure que certains émergent et que d'autres meurent (Thomas et al., 2000). Les espèces ligneuses, dans lesquelles les anciens modules morts servent de supports et de conduits d'acheminement des nutriments pour les modules vivants en sont la meilleure illustration (Thomas, 2013). Les modules sénescents peuvent donc mourir sans nécessairemment compromettre la survie et la fécondité de la plante dans son ensemble (Munné-Bosch, 2018), car ils sont continuellement remplacés par de nouveaux modules qui comportent moins de mutations grâce aux mécanismes de quiescence présentés plus haut. Leur élimination pourrait même constituer un avantage sur le plan physiologique, en permettant de rediriger des ressources vers les modules plus jeunes. De la même manière que la mort d'un individu n'implique pas l'extinction de la population à laquelle il appartient, la dynamique de remplacement que nous avons décrite pourrait être à l'origine de l'absence de sénescence observée chez les plantes, aussi longtemps qu'elle est capable de maintenir un rythme de remplacement suffisant (Bernard et al., 2020).

1.2.3.2 Facteurs de mortalité extrinsèques

Dormance méristématique. Du fait de leur immobilité, les plantes n'ont pas le loisir de fuir les perturbations environnementales qui les accablent, et disposent donc de mécanismes physiologiques particuliers pour y faire face. Un mécanisme déterminant pour leur survie est la capacité à maintenir des méristèmes dormants. Les méristèmes sont maintenus en dormance soit en réponse à des signaux environnementaux, on parle alors d'endodormance, soit par des signaux chimiques émis par le méristème apical, on parle alors de dominance apicale ou de paradormance (Lang, 1987). L'endodormance se met en place lorsque les conditions extérieures ne sont pas favorables à la croissance, du fait par exemple de températures trop basses ou d'une trop faible photopériode en

hiver (Vayssières et al., 2020). La capacité à entrer en endormance est un facteur déterminant de la pérennialité, car elle permet aux espèces pérennes de traverser la saison défavorable dans un état de vie ralentie, ce dont les espèces annuelles sont incapables (Friedman, 2020). Chez certaines espèces herbacées pérennes, l'appareil végétatif aérien est éliminé chaque hiver. C'est grâce aux méristèmes dormants conservés au niveau ou en-dessous du niveau du sol qu'elles sont capables d'en générer un nouveau chaque printemps.

La dominance apicale, par laquelle le méristème apical d'une tige inhibe la croissance des méristèmes axillaires proches, est un mécanisme essentiel de contrôle de l'architecture de l'appareil végétatif. En son absence, les méristèmes prolifèreraient de manière anarchique avec des conséquences désastreuses pour la viabilité de la plante dans son ensemble (Horvath et al., 2003). Elle permet aussi de maintenir une banque de méristèmes dormants, capable d'assurer la pérennité de l'appareil végétatif si le méristème apical venait à être endommagé. En effet, face à un événement catastrophique tel qu'un feu, une tempête ou l'attaque par un ravageur détruisant tout ou partie de l'appareil végétatif, ce sont ces méristèmes dormants qui seront mobilisés pour assurer la reprise de la croissance et donc la survie de l'organisme (Ott et al., 2019).

Défense. L'attaque par un herbivore peut considérablement réduire la fitness d'une plante, voire causer sa mort (Züst and Agrawal, 2017). De ce fait, beaucoup d'espèces ont acquis des mécanismes de protection contre ce type d'attaques parmi lesquels figurent la production de structures morphologiques spécialisées (e.g. épines, trichomes, Agren and Schemske, 1993) ou de composés chimiques de défense (Mithöfer and Boland, 2012). Qu'ils soient morphologiques ou chimiques, les mécanismes de protection contre l'herbivorie sont coûteux pour l'appareil végétatif, car ils consomment des métabolites primaires qui auraient pu être alloués à la croissance ou la reproduction (Herms and Mattson, 1992; Agrawal, 2011).

1.3 Théorie des traits histoires de vie

La plupart des mécanismes physiologiques exposés dans la section précédente ne sont pas pris en compte dans l'étude théorique de l'évolution des histoires de vie. La théorie des traits d'histoire de vie, qui décrit cette évolution, a en effet été développée dans l'objectif d'obtenir des résultats les plus généraux possibles, avec en tête un modèle animal comme le reflètent les ouvrages classiques de Stearns (1992) et Charlesworth (1994) par exemple, de sorte que les importantes spécificités du règne végétal n'ont pas encore trouvé leur place dans cette théorie. Dans cette section, nous allons nous intéresser aux arguments théoriques avancés pour expliquer les grandes caractéristiques des syndrômes d'annualité et de pérennialité que nous discutions plus haut, en les confrontant à chaque fois avec les éléments de physiologie que nous avons établis dans la section précédente.

1.3.1 Avantages relatifs de l'annualité et de la pérennialité

Avant d'envisager les histoires de vie comme un continuum (section 1.1), les approches théoriques se sont concentrées sur les avantages comparatifs d'une stratégie annuelle (monocarpe) par rapport à une stratégie pérenne polycarpe d'un point de vue purement démographique.

Le paradoxe de Cole. Le travail de Cole (1954) sur cette question est souvent présenté comme le déclencheur de l'effervescence intellectuelle qui a conduit dans la seconde moitié du $\mathrm{XX}^{\text{ème}}$ siècle à l'élaboration de la théorie des traits d'histoire de vie. Le modèle développé par Cole (1954) est le suivant. Supposons une population annuelle dont on note N_t l'effectif au temps t, et b_a le nombre de descendants produits par individu par pas de temps. En supposant qu'il n'y a pas de mortalité juvénile et en

l'absence de densite-dépendance, nous avons alors

$$N_{t+1} = b_a N_t \Rightarrow \lambda_a(t) = \frac{N_{t+1}}{N_t} = b_a, \tag{1.3}$$

où λ_a est le taux d'accroissement de cette population. Si nous nous intéressons maintenant à une population pérenne dans laquelle les individus sont immortels et produisent b_p descendants par pas de temps, nous avons alors

$$N_{t+1} = N_t + b_p N_t \Rightarrow \lambda_p(t) = \frac{N_{t+1}}{N_t} = b_p + 1. \tag{1.4}$$

Ainsi, une stratégie pérenne est supérieure à une stratégie annuelle dans ce modèle si et seulement si (Encadré B),

$$\lambda_a < \lambda_p \Leftrightarrow b_a < b_p + 1. \tag{1.5}$$

Différences de fécondité mises à part, il en résulte qu'adopter une stratégie pérenne est au mieux équivalent à ajouter un descendant au nombre de descendants moyen produit par une espèce annuelle, et apporte donc un avantage faible. Sachant que la majorité des espèces sont pérennes et polycarpes (Munoz et al., 2016), ce résultat semble paradoxal.

Résolution du paradoxe. Cet apparent paradoxe est cependant résolu en intégrant la mortalité dans le modèle (Charnov and Schaffer, 1973). En effet, le modèle de Cole (1954) a cela d'absurde qu'il suppose que les juvéniles parviennent tous à maturité, et que les adultes sont immortels dans le cas pérenne. Si nous supposons à présent que les juvéniles meurent avec une probabilité d_a et les adultes avec une probabilité d_p dans le cas pérennes, nous avons alors que

$$\lambda_a = b_a(1 - d_a), \tag{1.6}$$

et

$$\lambda_p = b_p(1 - d_a) + 1 - d_p. \tag{1.7}$$

Ainsi, la condition pour qu'une stratégie pérenne soit supérieure à une stratégie annuelle devient

$$\lambda_a < \lambda_p \Rightarrow b_a < b_p + \frac{1 - d_p}{1 - d_a}. \tag{1.8}$$

Cette condition dépend à présent du rapport entre la mortalité juvénile et la mortalité adulte. Le modèle de Charnov and Schaffer (1973) suggère donc que les stratégies pérennes devraient être favorisées dans les milieux où la mortalité adulte est faible et la mortalité juvénile est forte, et les stratégies annuelles lorsque la mortalité adulte est forte et la mortalité juvénile faible.

Encadré B: Stratégies évolutivement stables et mesures de fitness

Une notion essentielle dans la théorie des traits d'histoire de vie est celle de stratégie évolutivement stable (ESS ci-après, Maynard Smith and Price, 1973). Une stratégie est dite évolutivement stable si elle l'emporte compétitivement sur toutes les autres, c'est-à-dire si elle ne peut être envahie par aucune stratégie mutante lorsqu'elle est fixée dans une population. Pour établir qu'une stratégie est évolutivement stable il faut montrer qu'elle maximise la fitness, ce qui nécessite de définir une mesure de fitness adaptée au modèle considéré. L'histoire de vie des organismes qui composent une population est intrinsèquement liée à sa démographie, ainsi tout modèle d'évolution doit commencer par préciser son modèle démographique, c'est-à-dire la manière dont la reproduction et la survie des individus sont affectées par la densité de la population, afin de déterminer quelle mesure de fitness est adéquate.

Absence de densité-dépendance. Une grande partie des modèles produits en théorie des traits d'histoire de vie font l'hypothèse que la dynamique de la population n'est pas affectée par sa densité car la densité-dépendance en complique souvent considérablement l'analyse. Sous cette hypothèse, toute population grandit exponentiellement pour toujours quelle que soit sa stratégie. Une ESS est alors définie par le fait que sa fréquence dans la population tend vers 1 au fil du temps. Supposons que nous nous trouvons dans un environnement constant, et imaginons une stratégie résidente de taux de croissance r_r et d'effectif initial N_r^0 que l'on met en compétition avec une stratégie mutante de taux de croissance r_m d'effectif initial N_m^0. La fréquence de la stratégie résidente au temps t est alors donnée par

$$f_r^t = \frac{N_r^0 \exp(r_r t)}{N_r^0 \exp(r_r t) + N_m^0 \exp(r_m t)}, \tag{B.5}$$

et cette stratégie est une ESS si

$$\lim_{t \to \infty} f_r^t = 1 \;\Leftrightarrow\; r_r > r_m. \tag{B.6}$$

Le taux intrinsèque de croissance d'une stratégie est donc une bonne mesure de fitness en l'absence de densité-dépendance, et déterminer quelles sont les ESS d'un modèle revient à trouver les stratégies qui maximisent ce taux. L'histoire de vie d'un organisme est reliée à son taux intrinsèque de croissance à travers l'équation d'Euler-Lotka que nous dérivons ci-dessous. Appelons $B_t^{(x)}$ le nombre de descendants produits par des individus d'âge x au temps t et $\bar{B}_t = \sum_x B_t^{(x)}$ le nombre total de descendants produits au temps t. De plus, notons l_x la probabilité qu'un individu survive jusqu'à l'âge x et b_x la fécondité d'un individu

de cet âge. Nous avons alors que

$$\bar{B}_t = \sum_x B_t^{(x)} = \sum_x \bar{B}_{t-x} l_x m_x. \tag{B.7}$$

Supposons à présent que la population est en croissance exponentielle à un taux r, de sorte que le nombre de descendants produit au temps t est donné par

$$\bar{B}_t = \bar{B}_0 e^{rt}. \tag{B.8}$$

En injectant l'équation (B.8) dans l'équation (B.7) et après simplification, nous obtenons l'équation d'Euler-Lotka,

$$1 = \sum_x l_x b_x e^{-rx}. \tag{B.9}$$

Pour simplifier les notations, on s'intéresse parfois au taux d'accroissement $\lambda = \frac{N_{t+1}}{N_t} = e^r$ plutôt qu'au taux intrinsèque de croissance.

Populations densité-dépendantes. Lorsque la dynamique de la population dépend de sa densité en revanche, il n'est pas possible de savoir *a priori* la mesure de fitness qui sera maximisée par l'évolution. La première étape de l'analyse d'un modèle avec densité-dépendance est donc de déterminer cette mesure, ce que l'on fait au travers d'une analyse d'invasibilité (Otto and Day, 2011). L'idée est de considérer une population à l'équilibre démographique fixée pour stratégie résidente et d'y faire apparaître un mutant rare. Son effet sur la démographie est alors négligeable, et on peut s'intéresser à son taux d'accroissement dans ce cas. Une stratégie résidente est dite évolutivement stable si aucun mutant rare ne peut augmenter en fréquence dans la population.

1.3.2 Évolution de l'effort reproducteur

Même parmi les arbres, ceux qui ont produit trop de fruits dépérissent après les avoir portés, parce que le corps n'a pas reçu assez de nourriture.

- Aristote, *Génération des animaux* (330-322 av. J.-C.), Livre III, Chapitre 1.

Les modèles présentés dans le paragraphe précédent sont tellement généraux qu'ils ne prédisent pas grand chose d'un point de vue évolutif. En effet, dans ces modèles aucune distinction n'est faite entre les différents facteurs environnementaux ou stratégiques qui interviennent dans la détermination des mortalités et fécondités. De plus, aucun lien n'est fait entre les paramètres de naissance et de mort (il n'y a pas de trade-off). Enfin, l'opposition entre stratégie annuelle d'une part et pérenne polycarpe d'autre part est une approximation grossière qui ne dépeint pas toute la diversité des histoires de vie que nous décrivions en première partie. Une approche plus féconde de l'évolution des histoires de vie consiste à les étudier sous l'angle de l'allocation de ressources, c'est-à-dire sous l'hypothèse que l'histoire de vie d'un individu résulte de la manière dont il alloue les ressources dont il dispose aux différentes fonctions physiologiques qui affectent sa survie et sa fécondité au fil de son existence. Étudier l'évolution des histoires de vie revient alors à déterminer les proportions dans lesquelles un individu devrait investir dans chacune des fonctions considérées afin de maximiser sa fitness, étant donné l'environnement dans lequel il évolue (Perrin and Sibly, 1993). Les ressources dont dispose un individu étant limitées, l'allocation à une fonction physiologique se fait nécessairemment au détriment d'une ou plusieurs autres (Cody, 1966; Levins, 1968). Comme chaque fonction influence la fitness d'une manière qui lui est propre, susceptible de varier suivant la manière dont l'individu alloue ses ressources aux autres fonctions et suivant l'environnement dans lequel il évolue, des compromis évolutifs (trade-offs ci-après) émergent et contraignent l'évolution des histoires de vie

(Stearns, 1989).

Les fonctions physiologiques intégrées aux modèles peuvent varier suivant la question biologique traitée, mais on en retrouve en général trois, la reproduction, la croissance et le maintien. La reproduction regroupe les mécanismes impliqués dans la production d'organes reproducteurs, de gamètes et de descendants. La croissance désigne quant à elle l'ensemble des processus impliqués dans le développement de le la lignée somatique, c'est-à-dire dans la production de biomasse. Le maintien, enfin, regroupe l'ensemble des processus physiologiques permettant de ralentir le flétrissement de la lignée somatique dû à la sénescence. Par abus de langage, la notion de maintien est souvent utilisée de manière interchangeable avec celle de survie bien que ces deux termes recouvrent des réalités différentes (Charlesworth, 1994). La survie est une conséquence dont le maintien est une cause parmi d'autres. Cependant, les autres facteurs influençant la survie, comme la mortalité extrinsèque ou le lien entre taille et survie, étant la plupart du temps soit considérés comme constants soit négligés, le maintien est souvent l'unique facteur modélisé influençant directement la survie des individus, rendant cette confusion sémantique bénigne.

Ultimement, c'est à travers la reproduction que le maintien et la croissance interagissent avec la fitness. En conséquence, la notion d'effort reproducteur, que nous définirons ici comme la fraction des ressources dédiée à la production de descendants par un individu au cours d'une saison de reproduction (Hirshfield and Tinkle, 1975), occupe un rôle central dans notre compréhension de l'évolution des histoires de vie depuis son introduction par Williams (1966b). En effet, le modèle d'évolution de l'effort reproducteur est de loin le modèle le plus commenté de la théorie des traits d'histoire de vie (Stearns, 1992). Dans ce modèle, plus l'effort reproducteur d'un individu au cours d'une saison de reproduction est important et plus ce dernier peut espérer produire de descendants, mais plus les coûts physiologiques à payer sont élevés (Bell, 1980; Reznick, 1985; Obeso, 2002).

1.3.2.1 Évolution de la longévité par le trade-off survie/reproduction

L'évolution de l'effort reproducteur sous l'effet d'un trade-off entre survie et reproduction est l'unique hypothèse permettant de décrire l'évolution de la longévité en théorie des traits d'histoire de vie (Kirkwood and Rose, 1991; Stearns, 1992). Les effets de la structuration en âge (Gadgil and Bossert, 1970; Schaffer, 1974b), de la stochasticité environnementale (Schaffer, 1974a) ou encore de la densité-dépendance (Benton and Grant, 1999) sur cette évolution ont été explorés par de nombreux auteurs (Stearns, 1992), mais le mécanisme fondamental demeure le même : plus un individu investit de ressources dans la reproduction à l'âge x, plus ses chances de survie jusqu'à l'âge $x+1$ seront faibles. Plutôt que de passer en revue les différents modèles, leurs méthodes et leurs hypothèses, exercice utile mais long et laborieux, je prends ici le parti d'illustrer quelques-uns des résultats importants sur l'évolution de l'effort reproducteur à travers un modèle que nous allons construire pas à pas dans la suite de cette section.

Supposons une population dite résidente dont l'effectif au temps t est noté N_t, fixée pour un certain effort reproducteur $E \in [0,1]$, et dans laquelle apparaît un mutant d'effort reproducteur $E' \in [0,1]$ dont on note M_t l'effectif au temps t. Notons par ailleurs $B(\cdot)$ et $S_a(\cdot)$ les fonctions qui décrivent la fécondité et les chances de survie après la reproduction d'un individu sachant son effort reproducteur, et S_j la probabilité qu'un juvénile, une fois établi, survive jusqu'à la maturité qu'il atteint en un an. Alors, sous les hypothèses que la reproduction est asexuée et que le recrutement de juvéniles est régulé par la densité de la population adulte, l'évolution des effectifs N et M au fil du temps sont donnés par les récurrences

$$\begin{cases} N_{t+1} = S_j B(E) N_t e^{-\mu(N_t+M_t)} + S_a(E) N_t \\ M_{t+1} = S_j B(E') M_t e^{-\mu(N_t+M_t)} + S_a(E') M_t, \end{cases} \tag{1.9}$$

où μ indique l'intensité de la densité-dépendance (Ricker, 1954; Otto and Day, 2011).

Cherchons les équilibres de ce modèle. Un équilibre est atteint lorsque $N_{t+1} = N_t = \hat{N}$ et $M_{t+1} = M_t = \hat{M}$, c'est-à dire lorsque

$$\begin{cases} \hat{N} = \frac{1}{\mu} \ln\left(\frac{S_j B(E)}{1-S_a(E)}\right) - \hat{M} \\ \hat{M} = \frac{1}{\mu} \ln\left(\frac{S_j B(E')}{1-S_a(E')}\right) - \hat{N}. \end{cases} \tag{1.10}$$

Le modèle admet donc trois équilibres. Un trivial pour $\hat{N} = \hat{M} = 0$ (les deux populations sont éteintes), et deux non-triviaux, où l'un des deux types exclut l'autre. Afin de déterminer la ou les stratégies évolutivement stables de ce modèle (Encadré B), plaçons-nous dans un cas où le mutant est rare ($M_t \sim o(\varepsilon)$) de sorte que le résident est proche de son équilibre non-trivial. Dans ce cas, le taux d'accroissement du mutant est donné par

$$\lambda_m = \frac{M_{t+1}}{M_t} = \frac{S_j B(E')}{S_j B(E)}\Big(1 - S_a(E)\Big) + S_a(E'). \tag{1.11}$$

Une stratégie est dite évolutivement stable si aucun mutant rare ne peut augmenter en fréquence dans la population lorsqu'elle y est initialement fixée, c'est-à-dire lorsque $\lambda_m < 1$ pour tout $E' \neq E$. Le critère de non-invasibilité de la population est donc

$$\lambda_m < 1 \Leftrightarrow \frac{B(E')}{1 - S_a(E')} < \frac{B(E)}{1 - S_a(E)}. \tag{1.12}$$

La quantité qui apparaît dans l'équation (1.12) est le succès reproducteur total d'un individu. C'est la quantité maximisée par l'évolution dans ce modèle, c'est-à-dire une mesure de fitness (W, Encadré B). Ainsi, nous avons

$$W(E) = \frac{B(E)}{1 - S_a(E)}. \tag{1.13}$$

Afin d'obtenir une solution explicite pour l'effort reproducteur évolutivement stable,

il nous faut trouver la valeur E^* qui maximise ce critère. Pour cela, supposons que $B(E) = bE$ où b est une constante, et

$$S_a(E) = S_0(1 - E^{\varphi}), \tag{1.14}$$

où S_0 est une constante qui mesure l'intensité de la mortalité extrinsèque, et φ permet de contrôler la forme du trade-off entre survie et reproduction (Zhang, 2000). Une stratégie est évolutivement stable satisfait les conditions

$$\left.\frac{\partial W}{\partial E}\right|_{E=E^*} = 0 \text{ et } \left.\frac{\partial^2 W}{\partial E^2}\right|_{E=E^*} < 0, \tag{1.15}$$

où W est une mesure de fitness. Appliquant les conditions de l'équation (1.15) à l'équation (1.13), nous obtenons

$$E^* = \left[\frac{1 - S_0}{S_0\,(\varphi - 1)}\right]^{\frac{1}{\varphi}}, \tag{1.16}$$

qui est une stratégie évolutivement stable si et seulement si $\varphi > 1$, c'est-à-dire si le trade-off entre survie et reproduction est concave. Si le trade-off est convexe ou linéaire ($\varphi \leqslant 1$), la seule stratégie possible est un effort reproducteur maximal ($E = 1$), ce qui correspond à l'annualité dans ce modèle. De plus, lorsqu'un effort reproducteur intermédiaire est évolutivement stable, il est d'autant plus fort que la mortalité extrinsèque est élevée. Ces résultats sont classiques en théorie des traits d'histoire de vie (Schaffer, 1974b; Stearns, 1992). Plus le milieu dans lequel évolue la population est perturbé, c'est-à-dire plus les chances de mourir indépendamment de l'allocation de ressources à la survie sont élevées, plus il est avantageux d'investir dans la reproduction.

Importance de la densité-dépendance. Les résultats présentés ci-dessus ont été obtenus en supposant que la densité-dépendance affecte le recrutement des juvéniles.

Si nous avions supposé que la densité-dépendance agissait différemment dans la population, cette quantité aurait pu être différente. Par exemple, si nous supposons à présent que la densité-dépendance agit à la fois sur le recrutement des juvéniles et sur la survie des adultes, nous avons alors le système suivant,

$$\begin{cases} N_{t+1} = \Big[S_j B(E) + S_a(E)\Big] N_t e^{-\mu(N_t+M_t)} \\ M_{t+1} = \Big[S_j B(E') + S_a(E')\Big] M_t e^{-\mu(N_t+M_t)}, \end{cases} \tag{1.17}$$

qui donne à l'équilibre,

$$\begin{cases} \hat{N} = \frac{1}{\mu} \ln\Big[S_j B(E) + S_a(E)\Big] - \hat{M} \\ \hat{M} = \frac{1}{\mu} \ln\Big[S_j B(E') + S_a(E')\Big] - \hat{N}, \end{cases} \tag{1.18}$$

et donc le critère de non-invasibilité suivant,

$$\lambda_m < 1 \Leftrightarrow S_j B(E') + S_a(E') < S_j B(E) + S_a(E). \tag{1.19}$$

Ici, c'est le taux intrinsèque de croissance qui est maximisé et non plus le succès reproducteur total (Otto and Day, 2011). En donnant les mêmes formes aux fonctions $B(E)$ et $S_a(E)$, on obtient

$$W = S_j b E + S_0(1 - E^{\varphi}), \tag{1.20}$$

ce qui donne, en appliquant l'équation (1.15),

$$E^* = \left(\frac{S_j b}{S_0 \varphi}\right)^{\frac{1}{\varphi - 1}}, \tag{1.21}$$

qui est là encore une ESS si et seulement si $\varphi > 1$. L'importance de la forme du trade-off entre survie et reproduction ne change pas, mais la valeur de l'ESS dépend

maintenant non plus simplement de la mortalité extrinsèque endurée par les adultes, mais aussi des chances de survivre jusqu'à maturité des juvéniles (S_j) et du nombre de descendants moyens par unité de ressource investie dans la reproduction (b). Ce résultat illustre l'importance de préciser le modèle démographique étudié (Brommer, 2000; Mylius and Diekmann, 1995). Suivant la manière dont la taille de la population est régulée, la quantité maximisée par l'évolution peut changer et mener à des résultats quantitativement et qualitativement différents.

1.3.2.2 Variation de l'effort reproducteur avec l'âge

Dans le modèle précédent, nous avons étudié l'évolution de l'effort reproducteur en supposant qu'il est le même à tous les âges. La manière dont l'effort reproducteur varie avec l'âge est une autre question classique en théorie des traits d'histoire de vie. Pour étudier cette variation, il est nécessaire de s'intéresser à l'évolution des efforts reproducteurs âge-spécifiques pour chacun des âges qu'un individu est susceptible d'atteindre. C'est donc un problème d'une redoutable difficulté (Charlesworth, 1994), face auquel le recours aux simulations devient rapidement l'unique solution. Les premiers auteurs à s'intéresser à ce problème furent Gadgil and Bossert (1970), qui proposèrent sur la base d'exemples numériques que l'effort reproducteur devrait toujours augmenter avec l'âge car les coûts de la reproduction s'exprime sur les classes d'âge futures, qui sont de moins en moins nombreuses à mesure que l'âge augmente. Ce résultat repose sur l'hypothèse faite par Gadgil and Bossert (1970) qu'il existe un nombre fini de classes d'âges, c'est-à-dire qu'il existe un âge maximal au-delà duquel aucun individu ne survit (Schaffer, 1974b). Si l'on suppose qu'il existe une infinité de classes d'âges, mais que la probabilité d'atteindre un certain âge tend vers zéro à mesure que celui-ci augmente, la formulation de cet argument intuitif n'a plus de sens. Quelle que soit la classe d'âge considérée, il existe toujours une infinité de classes d'âges plus vieilles. De plus, Fagen (1972) puis Charlesworth and Leon (1976) ont montré que l'effort reproducteur

pouvait également diminuer avec l'âge suivant les fonctions de survie et de fécondité considérées.

Dans la suite de ce paragraphe, nous allons nous intéresser à cette question à travers le modèle que nous avons décrit au paragraphe précédent. Commençons par poser le modèle général dont nous considérerons ensuite des cas particuliers. Supposons que les populations résidente et mutante sont composées d'individus appartenant à m classes d'âges allant de 1 (juvéniles) à m, et que les effectifs de la $i^{\text{ème}}$ classe d'âge au temps t sont notés $N_t^{(i)}$ et $M_t^{(i)}$, respectivement. Les résidents et les mutants d'âge i ont un effort reproducteur E_i et E'_i, respectivement. La classe 1 est une classe de juvéniles, qui par définition ne se reproduisent pas ($E_1 = E'_1 = 0$). Posons l'hypothèse que la fécondité des individus d'âge i est pondérée par un facteur G_i, et que leurs chances de survie en fonction de leur effort reproducteur sont décrites pas la fonction $S_i(E_i)$. Les individus de la plus haute classe d'âge m continuent à survivre avec une probabilité $S_m(E_m)$ et de se reproduire, mais leurs fécondités et chances de survie ne varient plus. La probabilité en l'absence de compétition qu'un juvénile parvienne à s'implanter est une constante notée S_j, et le nombre de descendants produits par unité d'effort reproducteur est donné par la fonction $B(\cdot)$. La dynamique des populations résidente et mutante peut être décrite par le système

$$\begin{cases} N_{t+1}^{(1)} = S_j \sum_{i=2}^{m} B(E_i) G_i N_t^{(i)} e^{-\mu \sum_{i=1}^{m} \left(N_t^{(i)} + M_t^{(i)} \right)} \\ N_{t+1}^{(i)} = S_{i-1}(E_{i-1}) N_t^{(i-1)} \text{ pour } 2 \leqslant i < m, \\ N_{t+1}^{(m)} = S_{m-1}(E_{m-1}) N_t^{(m-1)} + S_m(E_m) N_t^{(m)} \\ M_{t+1}^{(1)} = S_j \sum_{i=2}^{m} B(E'_i) G_i M_t^{(i)} e^{-\mu \sum_{i=1}^{m} \left(N_t^{(i)} + M_t^{(i)} \right)} \\ M_{t+1}^{(i)} = S_{i-1}(E'_{i-1}) M_t^{(i-1)} \text{ pour } 2 \leqslant i < m, \\ M_{t+1}^{(m)} = S_{m-1}(E'_{m-1}) M_t^{(m-1)} + S_m(E'_m) M_t^{(m)}. \end{cases} \tag{1.22}$$

À l'équilibre démographique, les effectifs des classes d'âge adultes sont donnés par

$$\begin{aligned}\hat{N}^{(i)} &= \prod_{k=1}^{i-1} S_k(E_k)\hat{N}^{(1)} \text{ pour } 2 \leqslant i < m, \text{ et } N^{(m)} = \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1 - S_m(E_m)}\hat{N}^{(1)}, \\ \hat{M}^{(i)} &= \prod_{k=1}^{i-1} S_k(E'_k)\hat{M}^{(1)} \text{ pour } 2 \leqslant i < m, \text{ et } M^{(m)} = \frac{\prod_{k=1}^{m-1} S_k(E'_k)}{1 - S_m(E'_m)}\hat{M}^{(1)}.\end{aligned} \tag{1.23}$$

Injectant ces expressions dans l'équation (1.22), le nombre de juvéniles à l'équilibre dans chacune des populations est donné par

$$\begin{aligned}\hat{N}^{(1)} =& \frac{\ln\left[S_j\left(\sum_{i=2}^{m-1} B(E_i)G_i\left(\prod_{k=1}^{i-1} S_k(E_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1-S_m(E_m)}B(E_m)G_m\right)\right]}{\mu\left[1 + \sum_{i=2}^{m-1}\left(\prod_{k=1}^{i-1} S_k(E_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1-S_m(E_m)}\right]} \\ &- \frac{1 + \sum_{i=2}^{m-1}\left(\prod_{k=1}^{i-1} S_k(E'_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E'_k)}{1-S_m(E'_m)}}{1 + \sum_{i=2}^{m-1}\left(\prod_{k=1}^{i-1} S_k(E_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1-S_m(E_m)}}\hat{M}^{(1)},\end{aligned} \tag{1.24}$$

et,

$$\begin{aligned}\hat{M}^{(1)} =& \frac{\ln\left[S_j\left(\sum_{i=2}^{m-1} B(E'_i)G_i\left(\prod_{k=1}^{i-1} S_k(E'_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E'_k)}{1-S_m(E'_m)}B(E'_m)G_m\right)\right]}{\mu\left[1 + \sum_{i=2}^{m-1}\left(\prod_{k=1}^{i-1} S_k(E'_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E'_k)}{1-S_m(E'_m)}\right]} \\ &- \frac{1 + \sum_{i=2}^{m-1}\left(\prod_{k=1}^{i-1} S_k(E_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1-S_m(E_m)}}{1 + \sum_{i=2}^{m-1}\left(\prod_{k=1}^{i-1} S_k(E'_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E'_k)}{1-S_m(E'_m)}}\hat{N}^{(1)}.\end{aligned} \tag{1.25}$$

Ainsi, lorsque le mutant est rare, son taux d'accroissement λ_m est donné par

$$\lambda_m = \frac{\sum_{i=2}^{m-1} B(E'_i)G_i\left(\prod_{k=1}^{i-1} S_k(E'_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E'_k)}{1-S_m(E'_m)}B(E'_m)G_m}{\sum_{i=2}^{m-1} B(E_i)G_i\left(\prod_{k=1}^{i-1} S_k(E_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1-S_m(E_m)}B(E_m)G_m}, \tag{1.26}$$

de sorte que

$$W(\mathbf{E}) = \sum_{i=2}^{m-1} B(E_i)G_i \left(\prod_{k=1}^{i-1} S_k(E_k)\right) + \frac{\prod_{k=1}^{m-1} S_k(E_k)}{1 - S_m(E_m)} B(E_m)G_m, \tag{1.27}$$

où $\mathbf{E} = (0, E_2, ..., E_m)$ est le vecteur des E_i, est une mesure de fitness dans ce modèle. En supposant comme précédemment que

$$B(E_i) = bE_i \text{ et } S_i(E_i) = s_i\left(1 - E_i^{\varphi}\right), \tag{1.28}$$

où b et s_i sont des constantes, nous pouvons établir que l'effort reproducteur dans la classe la plus élevée (m) est une ESS si et seulement si $\varphi > 1$ et est donné par

$$E_m^* = \left(\frac{1 - s_m}{s_m(\varphi - 1)}\right)^{\frac{1}{\varphi}}. \tag{1.29}$$

Ce résultat est comparable au modèle précédent qui correspond au cas où $m = 2$ (il n'y a que des juvéniles ou des adultes avec les mêmes chances de survie et fécondités dans la population). Pour les autres classes d'âge, $2 \leqslant i \leqslant m - 1$, la même analyse donne

$$E_i^* = \left(\frac{G_i}{\varphi s_i \left[\sum_{l=i+1}^{m-1} \left(E_l G_l \prod_{k=i+1}^{l-1} S_k(E_k)\right) + E_m G_m \frac{\prod_{k=i+1}^{m-1} S_k(E_k)}{1 - S_m(E_m)}\right]}\right)^{\frac{1}{\varphi - 1}}, \tag{1.30}$$

qui est une ESS si $\varphi > 1$. L'effort reproducteur évolutivement stable à l'âge i dépend donc des efforts reproducteurs dans les classes d'âges supérieures, de $i+1$ à m. Il n'est pas possible de conclure de manière générale sur la relation entre effort reproducteur et âge à partir de ce résultat.

Considérons donc un cas particulier plus simple, dans lequel la population est com-

posée de trois types d'individus. Des juvéniles, des jeunes et des vieux adultes. À l'ESS, l'effort reproducteur des vieux adultes est donné par

$$E_3^* = \left(\frac{1 - s_3}{s_3(\varphi - 1)}\right)^{\frac{1}{\varphi}}, \tag{1.31}$$

et celui des jeunes adultes par

$$E_2^* = \left(\frac{G_2\left[1 - s_3(1 - E_3^{*\varphi})\right]}{\varphi s_2 G_3 E_3^*}\right)^{\frac{1}{\varphi - 1}}. \tag{1.32}$$

En injectant l'équation (1.31) dans l'équation (1.32), on montre que l'effort reproducteur augmente entre les jeunes et les vieux adultes ($E_2^* < E_3^*$) si et seulement si

$$\frac{s_3}{s_2} < \frac{G_3}{G_2}. \tag{1.33}$$

Ainsi, l'effort reproducteur augmente avec l'âge si la fécondité augmente davantage (ou diminue moins) que la survie chez les individus plus âgés. Ce cas particulier très simple est le seul dans lequel des résultats analytiques aisément interprétables peuvent être obtenus.

Pour $m = 4$, les efforts reproducteurs des trois classes adultes ($i = 2,\ 3,\ 4$) sont donnés par

$$\begin{aligned}
E_4^* &= \left(\frac{1 - s_4}{s_4(\varphi - 1)}\right)^{\frac{1}{\varphi}}, \\
E_3^* &= \left(\frac{G_3\left[1 - s_4(1 - E_4^{*\varphi})\right]}{\varphi s_3 G_4 E_4^*}\right)^{\frac{1}{\varphi - 1}}, \\
E_2^* &= \left(\frac{G_2\left[1 - s_4(1 - E_4^{*\varphi})\right]}{\varphi s_2 \left\{s_3\left(1 - E_3^{*\varphi}\right) G_4 E_4^* + \left[1 - s_4(1 - E_4^{*\varphi})\right] G_3 E_3^*\right\}}\right)^{\frac{1}{\varphi - 1}}.
\end{aligned} \tag{1.34}$$

La Figure 1.3 montre la manière dont ces trois efforts reproducteurs varient en fonction de la mortalité extrinsèque dans la dernière classe d'âge, tous les autres paramètres étant maintenus constants.

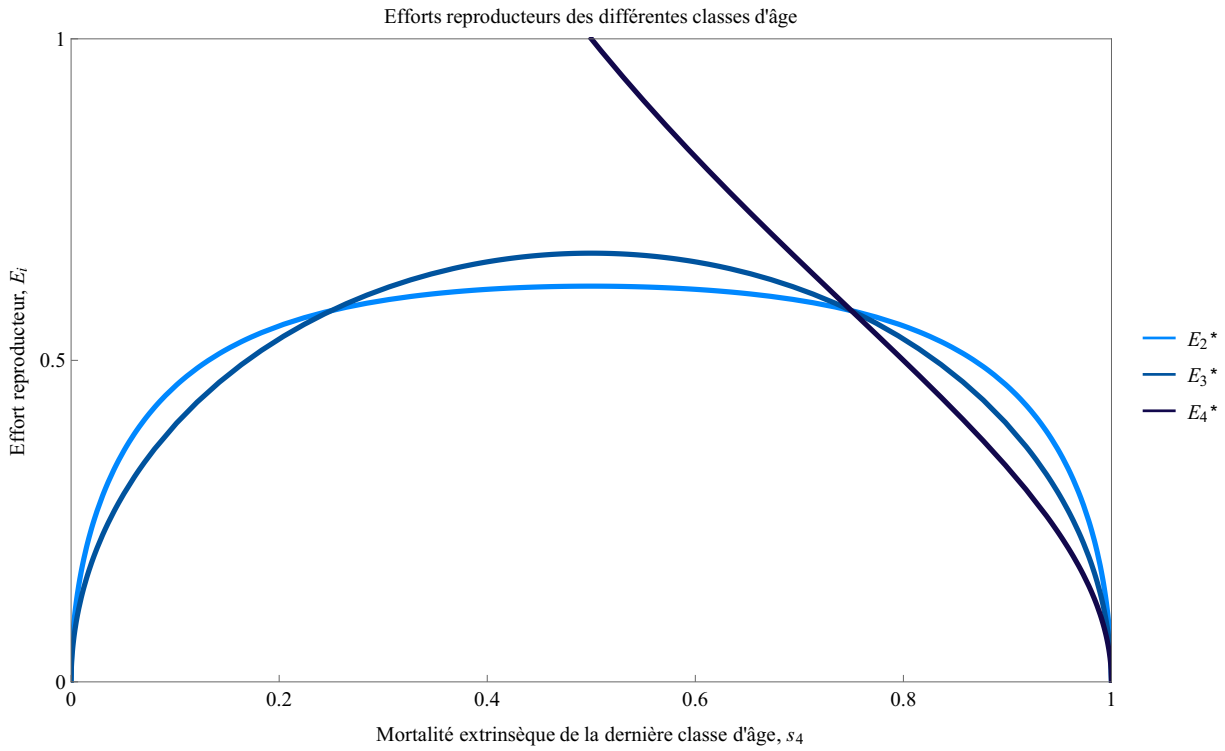

FIGURE 1.3 – Efforts reproducteurs à l'ESS pour les trois classes d'âge adultes quand $m = 4$, en fonction de la mortalité extrinsèque dans la dernière classe d'âge. Les autres paramètres sont $s_1 = s_2 = s_3 = 0.75$, $G_2 = G_3 = G_4 = 1$ et $\varphi = 2$.

Suivant la mortalité qui s'exerce sur la dernière classe, différents patrons de variation en fonction de l'âge sont observés. Quand s_4 est petit, l'effort reproducteur diminue puis augmente avec l'âge. Pour un s_4 intermédiaire, l'effort reproducteur augmente avec l'âge. Pour $s_4 = 0.75$, les trois classes adultes sont identiques en tout point, et ont le même effort reproducteur à l'ESS. Au-delà, l'effort reproducteur décroît avec l'âge. Ces résultats sont en accord avec les travaux antérieurs sur cette question (e.g. Gadgil and Bossert, 1970; Fagen, 1972; Charlesworth and Leon, 1976). Suivant la manière dont varient mortalité et fécondité en fonction de l'âge, virtuellement toutes

les relations entre effort reproducteur et âge peuvent être observer. Cela illustre que formuler des prédictions sur la manière dont l'effort reproducteur varie avec l'âge peut difficilement se faire sur la base de modèles généraux tels que ceux présentés jusqu'ici, et nécessite de plonger plus avant dans les mécanismes qui génèrent les variations de fécondité et de chances de survie en fonction de l'âge.

1.3.2.3 Croissance, reproduction et survie

L'ensemble des résultats dont nous avons discuté sur l'évolution de l'effort reproducteur reposent sur l'existence d'un trade-off entre survie et reproduction. En témoignent les écrits d'Aristote (section 1.3.2), l'idée que la reproduction influence négativement la survie est aussi ancienne que l'histoire naturelle elle-même (Jösson and Tuomi, 1994). Elle a donc un caractère intuitif en faveur duquel argumentent les nombreuses études comparatives menées au niveau interspécifique, qui montrent que l'effort reproducteur varie en général de manière inversement proportionnelle à la longévité (e.g. Harper, 1977; Primack, 1979; Charnov, 2005; Salguero-Gómez, 2017). Dans les années 1970, un fondement théorique plus solide pour cette idée a été donné par la théorie de la lignée somatique jetable que nous décrivions plus haut (Kirkwood, 1977), qui permet d'établir un lien entre allocation de ressources et sénescence. Le génome d'un organisme subit nécessairement des dégradations au fil des divisions cellulaires, que l'investissement de ressources dans le maintien permet de tempérer. Investir des ressources dans la reproduction consomme une partie des ressources nécessaires à ce maintien, accélérant ainsi la sénescence et diminuant de fait les chances de survie de l'organisme (Abrams and Ludwig, 1995; Cichoń, 1997).

Cette théorie suppose que les organismes constituent des unités physiologiques intégrées de sorte que l'ensemble de l'organisme souffre lorsqu'un organe défaille. Si cette hypothèse a un sens chez la plupart des animaux, les végétaux ne la vérifient, eux, jamais. En effet, les plantes sont des êtres constitués de modules physiologiquement

équivalents et dispensables, de sorte que la défaillance d'un module ne signifie pas la mort de l'individu entier. La théorie de Kirkwood (1977) suppose par ailleurs qu'il existe une séparation claire entre lignées germinale et somatique, permettant de considérer l'investissement dans le maintien ou dans la reproduction comme absolument distincts. La lignée germinale émergeant constamment de la lignée somatique par des mécanismes de conversion méristématique chez les plantes, cette hypothèse n'est pas vérifiée non plus par ces dernières. Ainsi, la seule théorie permettant d'établir un lien de causalité entre reproduction et longévité est mise en échec par le règne végétal. Comment, dès lors, expliquer la corrélation observée entre effort reproducteur et longévité chez les plantes ?

Nous l'évoquions dans la section 1.2.3.1, pour une plante croître est la condition de sa survie (Thomas et al., 2000). Tout méristème végétatif finit par cesser son activité et c'est aux méristèmes plus jeunes, dont ceux qu'il a produits, qu'il revient alors d'assurer la pérennité de la plante entière. Tant qu'au moins un méristème végétatif poursuit son activité, la plante survit. La reproduction entre en compétition avec ce processus, car elle implique de convertir des méristèmes végétatifs en méristèmes floraux, qui ne contribuent alors plus à la croissance de l'appareil végétatif (section 1.2.2). C'est donc à travers son effet sur la croissance que la reproduction peut indirectement influencer la survie, et non en diminuant directement l'investissement de l'organisme dans le maintien de sa lignée somatique. En l'absence de facteurs extrinsèques compromettant la survie des méristèmes végétatifs (e.g. herbivorie), l'allocation à la reproduction, mesurée par la proportion des méristèmes convertis (Bonser and Aarssen, 1996, 2006), peut être aussi élevée que l'on veut tant qu'un nombre suffisant de méristèmes végétatifs demeure pour assurer la non-extinction de la population de modules. Du point de vue de cette population, la reproduction est une source de mortalité dont l'action peut être contrée par la production suffisamment active de nouveaux modules, c'est-à-dire de nouveaux méristèmes (Watkinson and White, 1986). Les conséquences de la modula-

rité, de son interaction avec la reproduction, et en particulier de cette dynamique de population au sein même des organismes, n'ont été que peu étudiées sur le plan théorique. Par analogie, la dynamique des populations clonales peut contribuer à éclairer celle d'une population de modules. Par exemple, Gardner and Mangel (1997) ont montré que la clonalité pourrait dans une large mesure permettre à un individu génétique d'atteindre des longévités extrêmement élevées même si les individus physiologiques qui le composent meurent plus fréquemment. Cette analogie montre cependant ses limites dans le fait que les modules constitutifs d'un organisme demeurent physiologiquement liés par des dynamiques de sources-puits et par un destin commun, si la plante qu'ils constituent venait par exemple à être abattue (Galipaud and Kokko, 2020).

1.3.3 Évolution de l'âge à maturité

Dans les sections précédentes, nous avons étudié les principaux résultats obtenus en théorie des traits d'histoire de vie sur l'évolution de la longévité et sur le lien de cette dernière avec la reproduction. La seconde caractéristique importante du syndrôme de pérennialité que nous allons considérer est le lien entre âge à maturité et longévité. En effet, plus une espèce vit longtemps, plus elle a tendance à devenir mature tardivement (Salguero-Gómez et al., 2016). Comme la reproduction se fait au détriment de la croissance chez les plantes, repousser l'âge à maturité permet de croître plus rapidement, ce qui a deux avantages. D'abord, lorsque la compétition pour les ressources est intense, grandir plus vite que les autres est avantageux, car cela permet d'avoir un meilleur accès à la lumière et aux nutriments tout en privant les compétiteurs proches (Weiner, 1990; Petit and Hampe, 2006). Ensuite, une plus grande taille étant associée à une meilleure fécondité, repousser la maturité permet d'augmenter la fécondité au déclenchement de la reproduction (Kozłowski and Wiegert, 1987; Kozłowski, 1992). Cette stratégie présente toutefois un inconvénient de taille, car repousser le moment de la maturité augmente le risque de mourir avant même d'avoir pu se reproduire.

L'évolution de l'âge à maturité est donc le résultat de l'opposition entre l'effet de la reproduction sur la croissance d'une part, et le risque de mourir avant la maturité d'autre part (Stearns, 1992).

En théorie des traits d'histoire de vie, l'évolution de l'âge à maturité n'a été étudiée que par le biais de modèles d'optimisation qui négligent la compétition pour les ressources. De ce fait, elle n'a été considérée que sous l'angle de son effet sur la fécondité à travers la taille. Le modèle le plus simple portant sur cet effet fut proposé par Stearns and Crandall (1981). Dans leur modèle, la taille (et donc la fécondité) augmente linéairement avec l'âge à une vitesse différente pendant la phase juvénile et la phase adulte. Chez les plantes, si les espèces ligneuses conservent l'essentiel de leur biomasse entre deux saisons de reproduction, ce n'est pas le cas de nombre d'espèces herbacées, dont tout ou partie de l'appareil végétatif aérien est détruit chaque année (Harper, 1977). Dans ces conditions, l'influence de la croissance effectuée pendant une année sur la taille de l'individu à l'année suivante est limitée, ce qui réduit de fait l'avantage de fécondité conféré par le retardement de la maturité. L'effet de cette perte de biomasse a été considéré par Pugliese and Kozlowski (1990) dans un modèle de contrôle optimal dont l'analyse requiert des outils mathématiques compliqués. Je développe ici un modèle simple dans lequel la taille croît linéairement de manière analogue à Stearns and Crandall (1981) mais est diminuée d'une certaine proportion entre chaque saison favorable comme dans le modèle Pugliese and Kozlowski (1990). Ce modèle permet de retrouver qualitativement les résultats des différents auteurs sur l'évolution de l'âge à maturité.

Modèle de croissance. Notons q la proportion de biomasse conservée d'une saison à l'autre par un individu et g la quantité de biomasse produite en une saison de croissance en l'absence de reproduction. Avant la maturité, la taille atteinte au terme

de la $(\mathrm{i}+1)^{\text{ème}}$ saison de croissance est donnée par la suite

$$u_{i+1} = qu_i + g, \tag{1.35}$$

dont le terme général s'écrit

$$u_i = q^i \left(1 - \frac{g}{1-q}\right) + \frac{g}{1-q}. \tag{1.36}$$

Ainsi, lorsque la maturité est atteinte à l'âge a, la taille de l'individu est donnée par

$$u_a = q^a \left(1 - \frac{g}{1-q}\right) + \frac{g}{1-q}. \tag{1.37}$$

Supposons que la reproduction réduise la biomasse produite chaque année d'une proportion γ, et notons j le nombre de saisons depuis la maturité, de sorte que $j = 0$ lorsque $i = a$. Alors, la taille à l'âge j est décrite par la suite

$$v_{j+1} = qv_j + g(1-\gamma), \text{ avec comme valeur initiale } v_0 = u_a, \tag{1.38}$$

de sorte que la fécondité d'un individu d'âge i, que l'on supposera proportionnelle à sa taille, est donnée par

$$B_i(a) = \mathbf{1}_{i \geqslant a}\, b \left\{ q^{i-a} \left[q^a \left(1 - \frac{g}{1-q}\right) + \gamma \frac{g}{1-q} \right] + (1-\gamma)\frac{g}{1-q} \right\}, \tag{1.39}$$

où $\mathbf{1}_{i \geqslant a}$ est une indicatrice qui vaut zéro si $i < a$ et un sinon. Si $0 \leqslant q < 1$, la fécondité atteint au bout d'un certain temps une valeur d'équilibre donnée par

$$B_\infty = \frac{g(1-\gamma)}{1-q}. \tag{1.40}$$

Si $q = 1$, la fécondité tend linéairement vers l'infini au fil du temps, et si $q = 0$, la même fécondité $g(1-\gamma)$ est atteinte chaque année.

Dynamique de population. De la même manière que pour les modèles précédents, considérons la dynamique d'une population résidente d'âge à maturité a, et d'une population mutante d'âge à maturité a'.

$$\begin{cases} N_{t+1}^{(0)} = S \sum_{i=0}^{\infty} B_i(a) N_t^{(i)} \ \exp\left(-\mu \sum_j N_t^{(j)} + M_t^{(j)}\right) \\ N_{t+1}^{(i)} = S \ N_t^{(i-1)} \text{ pour } i \geqslant 1, \\ M_{t+1}^{(0)} = S \sum_{i=0}^{\infty} B_i(a') M_t^{(i)} \ \exp\left(-\mu \sum_j N_t^{(j)} + M_t^{(j)}\right) \\ M_{t+1}^{(i)} = S \ M_t^{(i-1)} \text{ pour } i \geqslant 1. \end{cases} \tag{1.41}$$

À l'équilibre, nous avons $\hat{N}^{(i)} = S\hat{N}^{(i-1)} = S^i \hat{N}^{(0)}$ et $\hat{M}^{(i)} = S\hat{M}^{(i-1)} = S^i \hat{M}^{(0)}$. Ainsi, nous obtenons

$$\begin{cases} \hat{N}^{(0)} = \frac{1-S}{\mu} \ln\left\{S^{a+1}\left[\left(1 - \frac{g}{1-q}\right)\frac{q^a}{1-qS} + \frac{g}{1-q}\left(\frac{\gamma}{1-qS} + \frac{1-\gamma}{1-S}\right)\right]\right\} - \hat{M}^{(0)} \\ \hat{M}^{(0)} = \frac{1-S}{\mu} \ln\left\{S^{a'+1}\left[\left(1 - \frac{g}{1-q}\right)\frac{q^{a'}}{1-qS} + \frac{g}{1-q}\left(\frac{\gamma}{1-qS} + \frac{1-\gamma}{1-S}\right)\right]\right\} - \hat{N}^{(0)}, \end{cases} \tag{1.42}$$

et, le taux d'accroissement du mutant lorsqu'il est rare est donné par

$$\lambda = \frac{S^{a'}\left[\left(1 - \frac{g}{1-q}\right)\frac{q^{a'}}{1-qS} + \frac{g}{1-q}\left(\frac{\gamma}{1-qS} + \frac{1-\gamma}{1-S}\right)\right]}{S^{a}\left[\left(1 - \frac{g}{1-q}\right)\frac{q^{a}}{1-qS} + \frac{g}{1-q}\left(\frac{\gamma}{1-qS} + \frac{1-\gamma}{1-S}\right)\right]}. \tag{1.43}$$

La quantité

$$W(a) = S^a\left[\left(1 - \frac{g}{1-q}\right)\frac{q^a}{1-qS} + \frac{g}{1-q}\left(\frac{\gamma}{1-qS} + \frac{1-\gamma}{1-S}\right)\right] \tag{1.44}$$

est donc une mesure de fitness. Si $g < 1-q$, c'est-à-dire si la croissance g n'est pas suffisante pour compenser la perte de biomasse $1-q$, la fitness décroît toujours avec

a car les termes en puissance de a deviennent toujours plus petits, et il est optimal de commencer à se reproduire le plus tôt possible. Nous considérerons que $g \leqslant 1$ pour la suite de l'analyse. Sous cette hypothèse, en appliquant les conditions décrites dans l'équation (1.15), nous obtenons que

$$a^* = \frac{1}{\ln q} \ln \left[-\frac{g\Big(1 - qS - \gamma S(1-q)\Big) \ln S}{(1-S)(1-g-q)\Big(\ln q + \ln S\Big)} \right] \tag{1.45}$$

est une ESS. Plus q est élevé, c'est-à-dire plus la proportion de biomasse conservée entre deux saisons favorables est grande, et plus l'âge à maturité à l'ESS est retardé. En effet, lorsque a tend vers zéro, nous avons

$$\lim_{q \to 0} a^* = 0, \tag{1.46}$$

et lorsque q tend vers 1,

$$\lim_{q \to 1} a^* = -\frac{g(1-S) + \Big[1 - S\Big(1 - g(1-\gamma)\Big)\Big] \ln S}{g(1-S) \ln S}. \tag{1.47}$$

La Figure 1.4a montre le comportement de l'âge à maturité à l'ESS pour une espérance de vie de $L = 10$ en fonction de la proportion de biomasse conservée q, et Figure 1.4b montre le comportement de l'âge à maturité à l'ESS en fonction de l'espérance de vie quand q tend vers 1. Dans les deux cas, différentes intensités de l'effet de la reproduction sur la croissance (γ) sont considérées. L'espérance de vie L est calculée de la manière suivante

$$L = \sum_{i=0}^{\infty} (i+1) S^i (1-S) = \frac{1}{1-S}. \tag{1.48}$$

Plus le coût de la reproduction est élevé, et plus un âge à maturité élevé est favorisé car cela permet d'aboutir à une meilleure fécondité. De la même manière, plus l'espérance

de vie est élevée, et plus les individus peuvent prendre le risque de retarder le moment de la reproduction. Par ce modèle simple, nous avons donc vu que l'âge à maturité devrait augmenter avec la longévité lorsque le milieu est stable et qu'une proportion suffisante de biomasse est conservée entre les saisons.

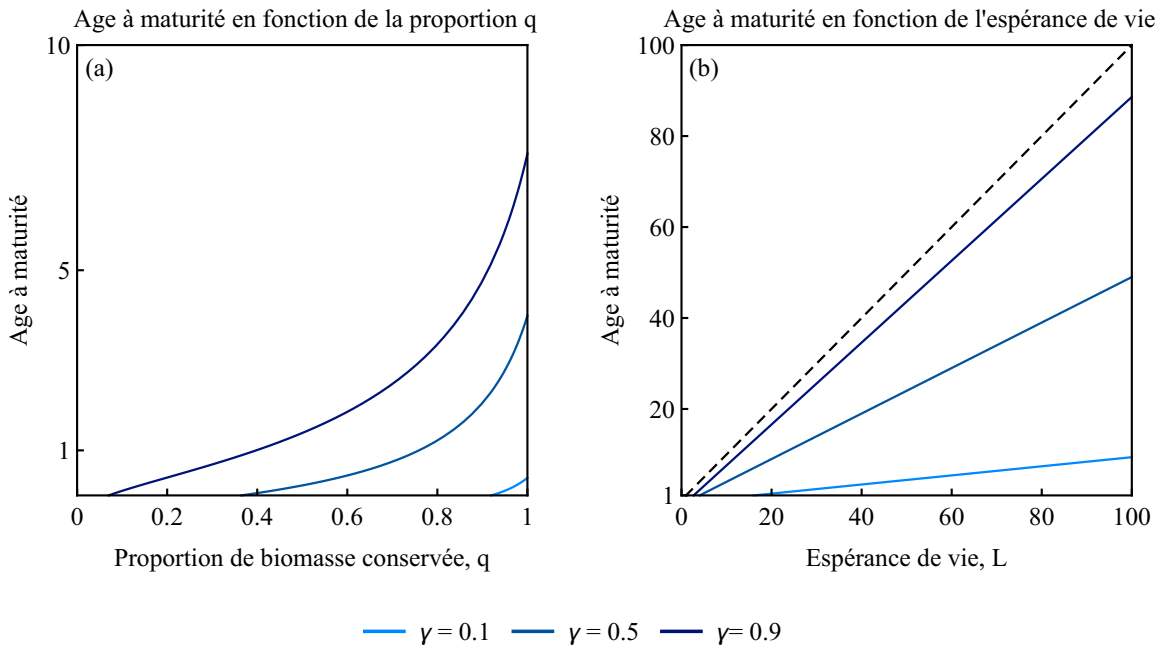

FIGURE 1.4 – (a) Age à maturité en fonction de la proportion de biomasse conservée entre saisons, q pour une espérance de vie de $L = 10$ ans, avec $g = 1$. (b) Age à maturité en fonction de l'espérance de vie lorsque $q \to 1$ et $g = 1$. L'espérance de vie est indiquée en pointillés. Pour les deux graphiques, les nuances de bleu indiquent différents niveaux de coût de la reproduction.

1.4 Conclusion

La possibilité qu'elle offrait d'expliquer une diversité aussi grande que celle des histoires de vie à partir de mécanismes aussi simples que ceux invoqués par les modèles d'allocation a valu, par son élégance, à la théorie des traits d'histoire de vie de connaître un moment de gloire lors de son émergence dans les années 1960 et pour les trente

années qui suivirent. En témoigne l'ancienneté des références théoriques citées dans ce chapitre d'introduction, l'effervescence intellectuelle qui entourait cette théorie s'est ensuite peu à peu affaiblie jusqu'à ce qu'elle tombe dans une désuétude presque complète au tournant du $\mathrm{XXI}^{\text{ème}}$ siècle. Je tiens qu'une raison de cette disgrâce vient de l'application du principe d'allocation. La théorie des traits d'histoire de vie est fondée sur ce principe, qui stipule que l'allocation de ressources à une fonction physiologique se fait nécessairement au détriment d'une ou plusieurs d'autres, car les ressources dont dispose un organisme sont limitées. Cette idée d'une logique a priori implacable se fissure rapidement à l'épreuve du réel car elle ne précise pas ce qu'est une fonction physiologique. À quel niveau se place-t-on, une enzyme dans une cascade ou la cascade elle-même ? Le phénotype cellulaire résultant, ses conséquences au niveau de l'organe ou de l'organisme entier ? La théorie des traits d'histoire de vie élimine ces questions en assimilant à la notion de fonction physiologique les conséquences phénotypiques de ces dernières. Ainsi, on parle classiquement d'allocation à la survie, à la reproduction et à la croissance dans cette théorie et l'on suppose, en vertu du principe d'allocation, que des trade-offs contraignent cette allocation. Chacun de ces phénotypes est pourtant le résultat d'une pluralité de mécanismes physiologiques plus ou moins liés les uns aux autres. Quel que soit l'organisme considéré, la croissance et le maintien partagent par exemple une dépendance primordiale à la mitose, sans laquelle ni le remplacement des cellules anciennes ni la production de biomasse supplémentaire ne pourraient s'opérer. Par ailleurs, la séparation entre lignée somatique et germinale, consubstantielle de la séparation entre maintien et reproduction, est inexistante chez les végétaux de sorte que la séparation entre croissance, maintien et reproduction paraît inadéquate chez ces derniers. Il n'est pas surprenant, dès lors, que les trade-offs supposés entre ces phénotypes, véritables pierres angulaires de la théorie des traits d'histoire de vie, soient régulièrement démentis par l'observation et ce particulièrement chez les végétaux. Une étude approfondie de la manière dont un ensemble de cellules forment un organisme

c'est-à-dire de la physiologie semble, à la lumière des éléments discutés dans ce chapitre, une bonne et peut-être l'unique manière de sortir la théorie des traits d'histoire de vie de l'impasse. Une telle étude, en permettant de redéployer le principe d'allocation sur des bases physiologiquement justifiées, pourrait en outre permettre d'ouvrir la voie à l'intégration de la génétique dans cette théorie, qui en définissant des phénotypes émergeants comme des fonctions physiologiques jette *ipso facto* un voile pudique sur l'ensemble de la relation génotype-phénotype.

Chapitre 2

Autofécondation et dépression de consanguinité

Outre la diversité d'histoires de vie que nous avons décrite en première partie, les Angiospermes, c'est-à-dire les plantes à fleurs, recélent une diversité d'organes reproducteurs qui n'est égalée par aucun autre groupe taxonomique (Barrett, 2010). Cette diversité va de paire avec une impressionnante variété de systèmes de reproduction sexuée, allant de la dioécie stricte où les fonctions mâles et femelles sont portées par des individus différents, à l'hermaphrodisme où les individus sont à la fois mâles et femelles, en passant par virtuellement toutes les combinaisons possibles de coexistence entre mâles, femelles et hermaphrodites (Barrett, 2002). Contrairement aux animaux, chez qui l'existence de sexes séparés est la règle, la dioécie est relativement rare chez les plantes à fleurs (environ 6% des espèces) et les systèmes de reproduction intermédiaires entre hermaphrodisme et dioécie le sont encore plus, si bien que les espèces hermaphrodites représentent l'écrasante majorité des Angiospermes (Renner, 2014).

Chez les espèces hermaphrodites, la coexistence de fonctions mâles et femelles sur un même individu et en général au sein des mêmes fleurs ouvre la possibilité de se reproduire sexuellement avec soi-même, c'est-à-dire de s'autoféconder (Käfer et al.,

2017). Beaucoup d'espèces possèdent des systèmes d'auto-incompatibilité leur permettant d'empêcher l'autofécondation et de se reproduire de fait presque exclusivement par allofécondation (Igic et al., 2008; Durand et al., 2020), tandis que d'autres ont acquis des spécificités morphologiques favorisant fortement l'autofécondation (Sicard and Lenhard, 2011). D'autres encore présentent des taux d'autofécondation intermédiaires et variables qui dépendent des conditions écologiques. Parmi les espèces capables de s'autoféconder, on distingue trois types d'autofécondation suivant le moment auquel elle se produit au cours de la floraison (Lloyd, 1992). Lorsque l'autofécondation se produit avant que l'allofécondation puisse avoir lieu, par exemple juste avant l'ouverture de la fleur, on parle d'autofécondation antérieure (*prior selfing*). Lorsqu'elle se produit de manière concomitante avec l'allofécondation, on parle d'autofécondation concurrente (*competing selfing*). Enfin, lorsque l'autofécondation n'est possible qu'après une période au cours de laquelle seule l'allofécondation pouvait avoir lieu, on parle d'autofécondation retardée (*delayed selfing*). Chacun de ces types a des conséquences écologiques et évolutives différentes dont nous allons discuter dans ce chapitre. L'objectif de ce dernier est d'abord d'introduire les arguments classiques qui entourent l'évolution de l'autofécondation et la structure génétique de la dépression de consanguinité, puis de discuter les résultats théoriques précédemment obtenus sur l'association entre histoires de vie et régime de reproduction.

2.1 Avantages évolutifs de l'autofécondation

Avantage automatique de Fisher. D'un point de vue génétique, un variant qui augmente le taux d'autofécondation permet à son porteur d'être à la fois mère et père d'une plus grande fraction de sa descendance, ce qui résulte en un avantage automatique de transmission dont l'existence a été montrée par Fisher (1941) par le modèle suivant. Supposons une population annuelle, dans laquelle ségrége un locus

biallélique modifiant le taux d'autofécondation de son porteur. À ce locus, l'allèle 0 code pour un taux d'autofécondation nul, de sorte qu'un individu homozygote pour cet allèle (00) a un taux d'autofécondation $\alpha_{00} = 0$. L'allèle 1 code pour un taux d'autofécondation a, tel que $\alpha_{11} = a$. Les deux allèles sont codominants ($\alpha_{01} = \alpha_{10} = \frac{a}{2}$). On suppose que l'autofécondation ne réduit pas la contribution des individus au nuage de pollen disponible pour l'allofécondation, et que l'allofécondation se fait en panmixie totale. De plus, on suppose que l'autofécondation a lieu avant ou en même temps que l'allofécondation. Les deux génotypes hétérozygotes (01 et 10) étant équivalents, nous pouvons les rassembler dans la même classe et désigner les trois types d'individus par le nombre i de copies de l'allèle 1 qu'ils portent ($00 \to 0$, 01 et $10 \to 1$, $11 \to 2$). Notons $P_t^{(i)}$ la fréquence de la classe i au temps t et $q_t = \frac{1}{2}P_t^{(1)} + P_t^{(2)}$ la fréquence de l'allèle 1 au temps t. En appliquant les règles de ségrégation Mendelienne, la dynamique des fréquences génotypiques au sein de la population peut alors être décrite par le système suivant

$$\begin{cases} P_{t+1}^{(0)} = P_t^{(0)}(1-q_t) + P_t^{(1)}\left[\frac{a}{2}\frac{1}{4} + \left(1-\frac{a}{2}\right)\frac{1-q_t}{2}\right] \\ P_{t+1}^{(1)} = P_t^{(0)}q_t + P_t^{(1)}\left[\frac{a}{2}\frac{1}{2} + \left(1-\frac{a}{2}\right)\frac{1}{2}\right] + P_t^{(2)}(1-a)(1-q_t) \\ P_{t+1}^{(2)} = P_t^{(1)}\left(\frac{a}{2}\frac{1}{4} + \left(1-\frac{a}{2}\right)\frac{q_t}{2}\right) + P_t^{(2)}\left[a + (1-a)q_t\right]. \end{cases} \tag{2.1}$$

Supposons que l'allèle 1 est rare au temps t, de sorte que $P_t^{(2)} \approx 0$, $P_t^{(1)} \approx \varepsilon$ et $P_t^{(0)} \approx 1 - \varepsilon$. Nous avons alors que

$$q_t = \frac{1}{2}P_t^{(1)} + P_t^{(2)} \approx \frac{\varepsilon}{2} \Leftrightarrow \varepsilon = 2q_t. \tag{2.2}$$

Au premier ordre en ε, nous obtenons que

$$q_{t+1} = \frac{1}{2}P_{t+1}^{(1)} + P_{t+1}^{(2)} \approx q_t\left(1 + \frac{a}{4}\right). \tag{2.3}$$

Ainsi, un allèle augmentant le taux d'autofécondation augmentera toujours en fréquence dans la population lorsqu'il est rare du fait de l'avantage automatique de transmission dont il dispose.

Assurance reproductive. Le second avantage de l'autofécondation est d'ordre écologique. En l'absence de pollinisateurs ou de partenaires compatibles, un individu incapable de s'autoféconder ne pourra pas se reproduire et verra sa fitness réduite à zéro, là où un individu autofécondant pourra tout de même produire des descendants. L'autofécondation agit donc comme un mécanisme d'assurance reproductive, dont l'importance varie suivant les conditions écologiques dans lesquelles se trouve la population (Eckert et al., 2006).

2.2 Dépression de consanguinité

Les espèces fortement autofécondantes sont cependant minoritaires parmi les plantes à fleurs (Goodwillie et al., 2005), ce qui suggère l'existence de facteurs contrebalançant les avantages évolutifs de l'autofécondation. Au niveau microévolutif, le principal facteur empêchant l'évolution de l'autofécondation est la dépression de consanguinité, c'est-à-dire la réduction de fitness des individus issus d'autofécondation par rapport aux individus issus d'allofécondation.

2.2.1 Évolution du taux d'autofécondation avec une dépression de consanguinité fixée

Mettons dans un premier temps de côté les mécanismes génétiques qui génèrent la dépression de consanguinité, et supposons que cette dernière est une constante, notée δ. Dans cette section, nous supposerons que la dépression de consanguinité réduit les chances de survie à maturité d'un ovule autofécondé d'une proportion δ. Introduisons

également la constante c, qui mesure l'intensité de l'effort de pollinisation (à $c = 1$, les pollinisateurs sont présents en abondance et tous les ovules non-autofécondés seront allofécondés, et à $c = 0$ il n'y a pas de pollinisateur) ; et la constante κ, définie telle que la proportion de pollen consommée par l'autofécondation est κa où a est le taux d'autofécondation, et $1 - \kappa a$ mesure la contribution de l'individu au nuage de pollen disponible pour l'allofécondation. Nous pouvons alors étendre le modèle décrit dans la section 2.1 en y introduisant les effets de la dépression de consanguinité, de l'intensité de la pollinisation et de la proportion de pollen consommée par l'autofécondation. Supposer qu'une fraction $1 - c$ des ovules disponibles pour l'allofécondation ne seront pas fécondées revient à supposer que cette fraction des ovules est bel et bien fécondée, mais meurt immédiatemment après. Ainsi, nous pouvons découper notre modèle en deux phases, une phase d'appariement des gamètes, presque analogue à celle que nous avions décrite en section 2.1, et une phase de sélection au cours de laquelle les individus issus d'autofécondation subissent la dépression de consanguinité, et l'effet de l'effort de pollinisation est incorporé parmi les allofécondés. Afin de pouvoir incorporer les nouveaux éléments apportés dans cette section, il est nécessaire que nous distinguions les individus autofécondés des individus allofécondés. La fréquence des individus portant i copies de l'allèle 1 parmi les individus issus d'autofécondation et d'allofécondation au temps t sont notées $P_t^{(i,\mathrm{autof})}$ et $P_t^{(i,\mathrm{allof})}$, respectivement. De plus, nous noterons $P_t^{(i)} = P_t^{(i,\mathrm{autof})} + P_t^{(i,\mathrm{allof})}$ la fréquence totale des individus portant i copies au temps

t. Nous avons alors le système suivant

$$\begin{cases} P_{t+1}^{(0,\text{autof},R)} = \frac{1}{4}\frac{a}{2}P_t^{(1)} \\ P_{t+1}^{(0,\text{allof},R)} = P_t^{(0)}\hat{p}_t + \frac{a}{2}P_t^{(1)}\frac{\hat{p}_t}{2} \\ P_{t+1}^{(1,\text{autof},R)} = \frac{1}{2}\frac{a}{2}P_t^{(1)} \\ P_{t+1}^{(1,\text{allof},R)} = P_t^{(0)}(1-\hat{p}_t) + \frac{1}{2}\frac{a}{2}P_t^{(1)} + (1-a)\,P_t^{(2)}\hat{p}_t \\ P_{t+1}^{(2,\text{autof},R)} = \frac{1}{4}\frac{a}{2}P_t^{(1)} + aP_t^{(2)} \\ P_{t+1}^{(2,\text{allof},R)} = \frac{a}{2}P_t^{(1)}\frac{1-\hat{p}_t}{2} + (1-a)P_t^{(2)}(1-\hat{p}_t), \end{cases} \tag{2.4}$$

où le R en exposant désigne le fait que nous nous trouvons après l'appariement des gamètes et avant la phase de sélection, et $\hat{p}_t$ est la fréquence de l'allèle 0 dans le nuage pollinique au temps t, donnée par

$$\hat{p}_t = \frac{P_t^{(0)} + \frac{1}{2}\left(1-\frac{a}{2}\kappa\right)P_t^{(1)}}{P_t^{(0)} + \left(1-\frac{a}{2}\kappa\right)P_t^{(1)} + (1-a\kappa)P_t^{(2)}}. \tag{2.5}$$

Pendant la phase de sélection, les fitness des individus issus d'autofécondation sont données par $w_{\text{autof}} = 1-\delta$, et celles des individus issus d'allofécondation par $w_{\text{allof}} = c$. La fitness moyenne est donc donnée par

$$\bar{w} = w_{\text{autof}}\sum_i P_t^{(i,\text{autof},R)} + w_{\text{allof}}\sum_i P_t^{(i,\text{allof},R)}. \tag{2.6}$$

Sous l'hypothèse que l'allèle 1 est rare, nous pouvons de nouveau supposer que $P_t^{(2)} \approx 0$, $P_t^{(1)} \approx \varepsilon$ et $P_t^{(0)} \approx 1-\varepsilon$, et donc que $\varepsilon = 2q_t$ (avec $q_t = 1-p_t$, équation 2.2). La fréquence de l'allèle 0 dans le pollen est alors donnée par

$$\hat{p}_t = 1 - \frac{\varepsilon}{2}\left(1-\frac{a}{2}\kappa\right), \tag{2.7}$$

au premier ordre en ε. Le changement de fréquence de l'allèle 0 entre t et $t+1$ au premier ordre en ε, Δp_t, est donc

$$\Delta p_t = \frac{aq_t}{4c}\Big(c(1+\kappa) - 2(1-\delta)\Big). \tag{2.8}$$

Le premier terme du produit est toujours du signe de a et peut donc être ignoré. L'autofécondation est favorisée lorsque $\Delta p_t < 0$, c'est-à-dire

$$\delta < 1 - \frac{c}{2}(1+\kappa). \tag{2.9}$$

En-dessous de ce seuil, l'autofécondation est favorisée. Au-dessus, l'allofécondation est maintenue. Lorsque la pollinisation est très efficace ($c = 1$) et la quantité de pollen consommée par l'autofécondation est négligeable ($\kappa = 0$), seule l'opposition entre l'avantage automatique de Fisher et la dépression de consanguinité gouverne l'évolution de l'autofécondation, et la condition devient $\delta < \frac{1}{2}$ (Lloyd, 1979). Plus la pollinisation est mauvaise, plus l'autofécondation est favorisée du fait de l'assurance reproductive qu'elle procure. Plus la quantité de pollen consommée par l'autofécondation est grande, et moins l'autofécondation est favorisée. Dans cette section, nous avons supposé que l'autofécondation est réalisée avant ou de manière concomitante avec l'allofécondation (*prior* ou *competing selfing*). L'autofécondation retardée a quant à elle tendance à être plus largement favorisée (Lloyd, 1992), car elle n'interfère pas avec l'allofécondation puisqu'elle ne se produit que lorsqu'il reste des ovules à féconder à l'issue de cette dernière, c'est-à-dire quand $c < 1$.

2.2.2 Structure génétique de la dépression de consanguinité

Les effets délétères de l'autofécondation sont connus depuis Darwin (1876), mais ce n'est qu'avec la redécouverte des lois de ségrégation de Mendel que les mécanismes

générant la dépression de consanguinité purent être étudiés. Partant du constat que le principal effet de la consanguinité était d'augmenter l'homozygotie des individus concernés, deux hypothèses génétiques ont été proposées pour expliquer le déclin de leur fitness par rapport aux individus non-consanguins (Charlesworth and Charlesworth, 1987). Dans cette section nous allons reprendre quelques calculs classiques afin d'illustrer ces deux hypothèses, puis nous discuterons leurs importances relatives.

2.2.2.1 Superdominance

La première hypothèse suppose l'existence de loci auxquels les hétérozygotes ont une fitness supérieure à celle des homozygotes, c'est-à-dire une situation de superdominance. Considérons le modèle de superdominance le plus simple possible, tel que celui présenté dans le chapitre 6 (p. 228) du livre de Hartl and Clark (1997). Supposons une population diploïde de très grande taille constante, à générations non-chevauchantes. À chaque pas de temps, les individus produits au pas de temps précédent se développent et produisent des gamètes proportionnellement à leur fitness, puis se reproduisent par autofécondation à un taux α et par allofécondation panmictique pour le reste, et meurent. Dans cette population, étudions un locus auquel ségrégent deux allèles, 0 et 1. Lorsqu'ils sont à l'état homozygote, les allèles n'ont pas d'effet particulier sur la fitness et nous avons $w_{00} = w_{11} = 1$. Lorsqu'ils sont présents à l'état hétérozygote en revanche, les individus ont une fitness augmentée d'un facteur s ($w_{01} = w_{10} = 1 + s$, Figure 2.1).

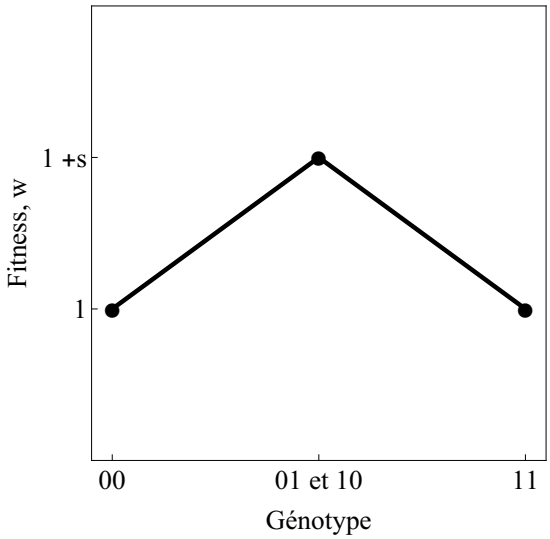

FIGURE 2.1 – Fitness en fonction du génotype dans une situation de superdominance à un seul locus biallélique.

De la même manière que précédemment, notons $P_t^{(i)}$ la fréquence des individus porteurs de i copies de l'allèle 1. À l'issue de la reproduction, les fréquences $P_t^{(i)}$ sont données par

$$\begin{cases} P_t^{(0)} = {p_t}^2 + p_t q_t F \\ P_t^{(1)} = 2 p_t q_t (1-F) \\ P_t^{(2)} = {q_t}^2 + p_t q_t F, \end{cases} \tag{2.10}$$

où F est le coefficient de consanguinité de Wright (1922), qui mesure la déviation de la panmixie due à la consanguinité, p_t et q_t sont les fréquences des allèles 0 et 1 au temps t, respectivement, avec $p_t = 1 - q_t$.

La fitness moyenne est alors donnée par

$$\bar{w} = P_t^{(0)} + P_t^{(2)} + (1+s) P_t^{(1)} = 1 + 2 s p_t q_t (1-F), \tag{2.11}$$

et les fréquences génotypiques après sélection sont

$$\begin{cases} P_t^{(0,\mathrm{sel})} = \frac{{p_t}^2 + p_t q_t F}{1 + 2 s p_t q_t (1-F)} \\ P_t^{(1,\mathrm{sel})} = (1+s) \frac{2 p_t q_t (1-F)}{1 + 2 s p_t q_t (1-F)} \\ P_t^{(2,\mathrm{sel})} = \frac{{q_t}^2 + p_t q_t F}{1 + 2 s p_t q_t (1-F)}. \end{cases} \tag{2.12}$$

Ainsi, la fréquence de l'allèle 0 après sélection est donnée par

$$p_t^{\mathrm{sel}} = \frac{{p_t}^2 + p_t q_t F}{1 + 2 s p_t q_t (1-F)} + \frac{1}{2}(1+s) \frac{2 p_t q_t (1-F)}{1 + 2 s p_t q_t (1-F)}. \tag{2.13}$$

La reproduction ne modifiant pas les fréquences alléliques (pas de mutation), nous avons que $p_t^{\text{sel}} = p_{t+1}$. Ainsi, à l'équilibre, c'est-à-dire lorsque $p_t = p_{t+1} = p^*$, nous obtenons

$$p^* = \frac{1}{2}. \tag{2.14}$$

Les allèles 0 et 1 sont maintenus équifréquents dans la population sous l'effet de l'avantage des hétérozygotes qui génère une sélection balancée. Les fréquences génotypiques au sein des individus issus d'autofécondation juste après la reproduction sont données par

$$\begin{cases} P_t^{(0,\text{autof})} = P_t^{(0)} + \frac{1}{4}P_t^{(1)} \\ P_t^{(1,\text{autof})} = \frac{1}{2}P_t^{(1)} \\ P_t^{(2,\text{autof})} = P_t^{(2)} + \frac{1}{4}P_t^{(1)}, \end{cases} \tag{2.15}$$

de sorte que la fitness moyenne de ces derniers à l'équilibre est donnée par

$$w_{\text{autof}} = 1 + \frac{s}{4}(1 - F). \tag{2.16}$$

Quant aux individus issus d'allofécondation, nous avons

$$w_{\text{allof}} = (p^*)^2 + (1 - p^*)^2 + 2p^*(1 - p^*)(1 + s) = 1 + \frac{s}{2}. \tag{2.17}$$

Ainsi, la dépression de consanguinité générée par la superdominance à un locus est donnée par

$$\delta = 1 - \frac{w_{\text{autof}}}{w_{\text{allof}}} = \frac{s(1 + F)}{2(2 + s)}. \tag{2.18}$$

Elle est positive dès lors que $s > 0$, c'est-à-dire dès lors qu'il y a superdominance. Le mécanisme qui la génère est le suivant. L'autofécondation génère plus d'individus homozygotes que l'allofécondation, ce qui a pour effet de diminuer la fitness moyenne des individus issus d'autofécondation car les homozygotes ont une fitness plus faible

que les hétérozygotes.

2.2.2.2 Mutations délétères partiellement récessives

La seconde hypothèse génétique proposée pour expliquer l'existence de la dépression de consanguinité fait intervenir des mutations délétères partiellement récessives. Ces mutations sont constamment produites du fait des erreurs commises par la machinerie de réplication de l'ADN, et éliminées de la population par sélection. Sous les mêmes hypothèses démographiques que dans la section précédente, intéressons-nous maintenant à la dépression de consanguinité générée par de telles mutations dans un modèle simple, tel que celui présenté par Lande and Schemske (1985). Supposons un locus biallélique auquel l'allèle 0 est un allèle sain qui n'a pas d'effet sur la fitness ($w_{00} = 1$), et l'allèle 1 est un mutant d'effet délétère s ($w_{11} = 1 - s$). Chez les hétérozygotes, l'allèle 1 s'exprime proportionnellement à son coefficient de dominance h ($w_{01} = w_{10} = 1 - sh$, Figure 2.2). L'allèle 0 mute vers l'allèle 1 au moment de la méiose à un taux u. Ainsi, le modèle est composé de trois phases : mutation, appariement des gamètes et sélection. Plaçons nous entre l'appariement des gamètes et la sélection. À ce stade, les fréquences des différents génotypes sont les mêmes que celles données par l'équation (2.10) dans le modèle précédent.

De ce fait, la fitness moyenne est donnée par

$$\bar{w} = P_t^{(0)} + P_t^{(1)}(1 - sh) + P_t^{(2)}(1 - s) = 1 - sq_t\Big[1 - p_t(1 - 2h)(1 - F)\Big]. \quad (2.19)$$

Après sélection, nous obtenons la fréquence allélique suivante pour l'allèle 1,

$$q_t^{\text{sel}} = (1-s)\frac{q_t^2 + p_t q_t F}{1 - sq_t\Big[1 - p_t(1 - 2h)(1 - F)\Big]} + \frac{1}{2}(1-sh)\frac{2p_t q_t(1 - F)}{1 - sq_t\Big[1 - p_t(1 - 2h)(1 - F)\Big]}. \quad (2.20)$$

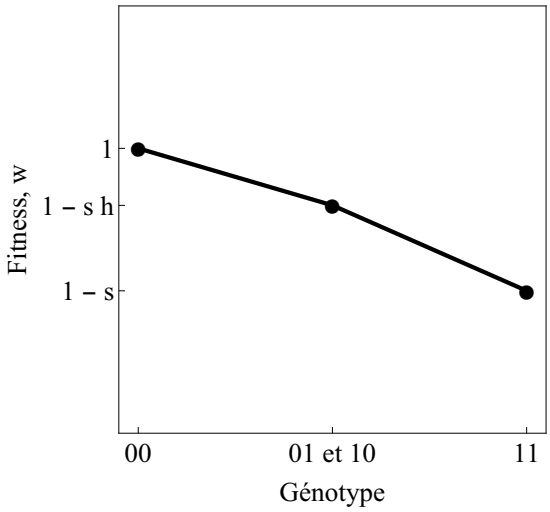

FIGURE 2.2 – Fitness en fonction du génotype à un locus auquel ségrége une mutation délétère partiellement récessive.

La mutation diminue ensuite la fréquence de l'allèle 0 au profit de l'allèle 1 d'une proportion $1-u$, de sorte que nous avons que

$$q_{t+1} = q_t^{\text{sel}} + u(1 - q_t^{\text{sel}}). \tag{2.21}$$

Sous l'hypothèse que l'allèle 1 est rare et que le taux de mutation est faible (q_t et u d'ordre ε), la variation en fréquence de l'allèle délétère, $\Delta q_t = q_{t+1} - q_t$ au premier ordre en ε est donnée par

$$\Delta q_t = u - s\left[h + F(1-h)\right] q_t, \tag{2.22}$$

et la fréquence à l'équilibre mutation-sélection de l'allèle délétère est

$$q^* = \frac{u}{s\left[h + F(1-h)\right]}. \tag{2.23}$$

Calculons la dépression de consanguinité qui résulte de sa présence. Les fréquences génotypiques au sein des individus issus d'autofécondation juste après la reproduction sont données les mêmes que celles données pour le modèle précédent dans l'équation (2.15). Nous avons donc

$$w_{\text{autof}} = 1 - sq^*\left[1 - \frac{p^*}{2}(1-2h)(1-F)\right]. \tag{2.24}$$

De plus, la fitness moyenne des individus issus d'allofécondation est donnée par

$$w_{\text{allof}} = 1 - sq^*\Big[1 - p^*(1-2h)\Big], \tag{2.25}$$

de sorte que la dépression de consanguinité à l'équilibre mutation-sélection, sous l'hypothèse que l'allèle 1 est rare, est donnée par

$$\delta = 1 - \frac{w_{\text{autof}}}{w_{\text{allof}}} = s(1-2h)\frac{1+F}{2}q^* = \frac{u(1+F)(1-2h)}{2\left[h+F(1-h)\right]}. \tag{2.26}$$

Là encore, une dépression de consanguinité positive est observée pourvu que les mutations délétères soient récessives ($h < \frac{1}{2}$). Le mécanisme à l'oeuvre est le suivant. Un individu issu d'autofécondation a plus de chances d'être homozygote pour l'allèle 0 ou pour l'allèle 1 qu'un individu issu d'allofécondation, et la fitness moyenne des homozygotes,

$$w_{\text{hom}} = \frac{1+1-s}{2} = 1 - \frac{s}{2}, \tag{2.27}$$

est plus faible que celle des hétérozygotes ($w_{\text{het}} = 1 - sh$) pourvu que l'allèle 1 soit partiellement récessif ($h < \frac{1}{2}$), de sorte que la fitness moyenne des autofécondés est plus faible que celle des allofécondés.

2.2.2.3 Importance relative des deux mécanismes

Les calculs présentés dans les deux sections précédentes montrent que la superdominance et le maintien de mutations délétères récessives à l'équilibre mutation-sélection sont deux mécanismes théoriquement capables de générer de la dépression de consanguinité lorsqu'un seul locus est considéré (Charlesworth and Charlesworth, 1987). Cependant, un génome est composé d'un grand nombre de loci. La question se pose donc de l'importance relative des deux mécanismes. Parmi les nouvelles mutations non-neutres, l'écrasante majorité sont délétères (Eyre-Walker and Keightley, 2007), et une relation de proportionnalité existe entre l'intensité de leur effet délétère et le niveau de dominance des mutations : plus une mutation est délétère, et plus elle est récessive (Manna et al., 2011). Le processus par lequel l'ensemble de la diversité génétique est engendrée, c'est-à-dire la mutation, produit donc constamment des mutations délétères

récessives capables de générer de la dépression de consanguinité. La superdominance requiert quant à elle des situations génétiques plus compliquées pour se produire car elle nécessite l'existence d'un avantage sélectif à l'hétérozygotie, c'est-à-dire à maintenir une diversité génétique intra-organisme. Bien que des exemples célèbres de telles situations existent, comme celui du lien entre l'allèle provoquant la Drépanocytose et la Malaria, qui lorsqu'il est présent à l'état homozygote induit une grave maladie du sang, mais procure à l'état hétérozygote une résistance accrue contre la Malaria sans causer de symptômes aigus (Aidoo et al., 2002), la superdominance demeure un phénomène assez rare et qui se manifeste surtout dans la résistance face aux pathogènes et parasites (Gemmell and Slate, 2006). À l'inverse, l'existence de quantités parfois importantes de mutations délétères récessives dans les génomes est bien documentée (Crow, 1999; Charlesworth and Willis, 2009; Brown and Kelly, 2020). Les données empiriques suggèrent donc que le maintien de mutations délétères récessives dans les populations est le principal mécanisme à l'origine de la dépression de consanguinité (Charlesworth and Charlesworth, 1987).

2.2.3 Purge des mutations délétères, interférence sélective et dépression de consanguinité

2.2.3.1 Purge des mutations délétères

L'autofécondation, en augmentant l'homozygotie des individus qui en sont issus, expose plus efficacement les mutations récessives à la sélection. Ces dernières ont donc tendance à diminuer en fréquence dans la population lorsque le taux d'autofécondation augmente (équations 2.23 et 2.26). L'efficacité de ce phénomène dit de purge des mutations délétères dépend en premier lieu du niveau de dominance (h) et l'intensité de l'effet sélectif des mutations (s). De manière générale, plus une mutation a un effet fort et moins elle est récessive (*i.e.* plus s et h sont élevés), mieux elle est éliminée

par la sélection car elle lui est plus exposée. L'effet de l'autofécondation sur cette élimination est quant à lui d'autant plus fort que les mutations sont récessives (*i.e.* que h est petit), car son effet révélateur prend alors davantage d'importance (Charlesworth, 1990). Il en résulte que le niveau de dominance influence significativement la magnitude de la dépression de consanguinité car il quantifie les écarts de fitness entre homozygotes et hétérozygotes. Quant à l'effet de l'intensité de la sélection sur l'élimination des mutations délétères, son impact sur la dépression de consanguinité est faible voire négligeable, car si moins de mutations sont maintenues à l'équilibre mutation-sélection lorsqu'elles sont plus délétères, leur plus grand effet sur la fitness compense leur raréfaction (Bataillon and Kirkpatrick, 2000).

2.2.3.2 Interférence sélective

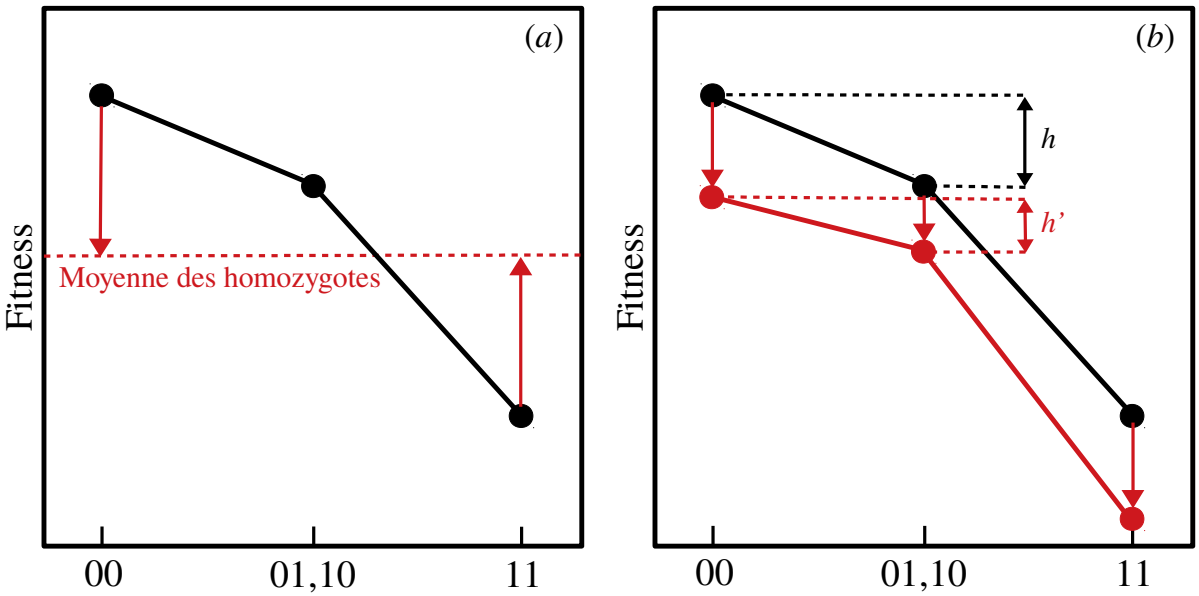

FIGURE 2.3 – (a) Fitness moyennes des individus homozygotes et hétérozygotes en présence de mutations délétères récessives. (b) Effet des associations génétiques sur la dominance efficace.

Tous les effets discutés précédemment se produisent même lorsque les loci ségrégent

indépendamment les uns des autres. Cependant, des associations génétiques entre loci sont générées dès lors que la reproduction n'est pas parfaitement panmictique ou que la population est de taille finie. Ces associations sont détruites par la recombinaison, et résultent de plusieurs effets dont l'intensité relative dépend de la structure de croisement, de l'intensité de la sélection et de la récessivité des mutations (Roze, 2015). Le point clef est que les individus homozygotes à un locus du fait de la consanguinité ont également tendance à être homozygotes à d'autres loci. Supposons des loci bialléliques auxquels ségrégent des mutations délétères récessives. Dans ce cas les homozygotes (incluant les homozygotes pour l'allèle sain) ont une fitness en moyenne plus faible que les hétérozygotes (Figure 2.3a), ce qui crée une surreprésentation d'hétérozygotes dans la population et diminue l'efficacité de la sélection contre les mutations délétères. Par ailleurs, la fitness d'un individu homozygote à un locus a tendance à être réduite d'un plus gros facteur que celle d'un individu hétérozygote par la présence de mutations aux autres loci, puisque les homozygotes à un locus ont tendance à être aussi homozygotes aux autres. Cela a pour effet de réduire le niveau de dominance efficace de la mutation délétère à ce locus (*i.e.* de la rendre plus récessive, Figure 2.3b) en modifiant les écarts de fitness entre homozygotes et hétérozygotes à ce locus. Enfin, la sélection contre les mutations délétères récessives étant moins efficace chez les hétérozygotes, elles sont surreprésentées chez ces derniers, ce qui réduit encore l'efficacité de la sélection. Ces trois effets induisent une diminution de l'efficacité de la sélection que l'on appelle interférence sélective. Cette diminution est d'autant plus marquée que les mutations sont récessives, et conduit au maintien de davantage de mutations et donc d'une dépression de consanguinité plus élevée. Dans le cas particulier de mutation totalement ou presque totalement récessives ($h \approx 0$) avec un taux de mutation suffisamment élevé, l'interférence sélective décrite plus haut devient si forte que les mutations délétères ne parviennent plus à être éliminées de la population en-dessous d'un certain seuil de taux d'autofécondation (Lande et al., 1994; Roze, 2015).

2.3 Sur le lien entre histoire de vie et régime de reproduction

Dans le premier chapitre de cette introduction, nous nous sommes penchés sur l'évolution des histoires de vie au travers du syndrôme de pérennialité et de la théorie des traits d'histoire de vie. Dans le deuxième chapitre, nous nous sommes intéressés à l'évolution de l'autofécondation et à son interaction avec la dépression de consanguinité. À l'image de l'essentiel de la littérature produite sur cette question, nous nous sommes cantonnés jusqu'ici à l'étude de populations annuelles, dans lesquelles les générations sont non-chevauchantes. Une forte association existe pourtant entre système de reproduction et pérennialité chez les Angiospermes, la plupart des plantes autofécondantes étant annuelles et la grande majorité des espèces pérennes étant allofécondantes (Figure 2.4, Barrett and Harder, 1996; Munoz et al., 2016). Cette troisième section a pour but de résumer les arguments théoriques qui ont été formulés pour tenter d'expliquer l'existence de cette association.

2.3.1 Évolution de l'autofécondation chez les espéces pérennes

Peu de théoriciens se sont penchés sur l'association entre longévité et système de reproduction chez les Angiospermes. Elle a d'abord été envisagée sous l'angle de l'effet de la longévité sur l'évolution de l'autofécondation, avec pour présupposé que les systèmes de reproduction évoluent plus rapidement que les histoires de vie chez les Angiospermes. Le premier argument verbal a été formulé par Stebbins (1950), qui a proposé que l'assurance reproductive procurée par l'autofécondation pourrait être moins avantageuse chez les espèces pérennes pour qui ne pas se reproduire au cours d'une année a des conséquences moins dramatiques sur la fitness que pour une espèce annuelle, qui n'a elle qu'une seule opportunité pour cela. Plus tard, Lloyd (1992) a proposé un autre mécanisme basé sur la notion de *seed discounting*, que nous tradui-

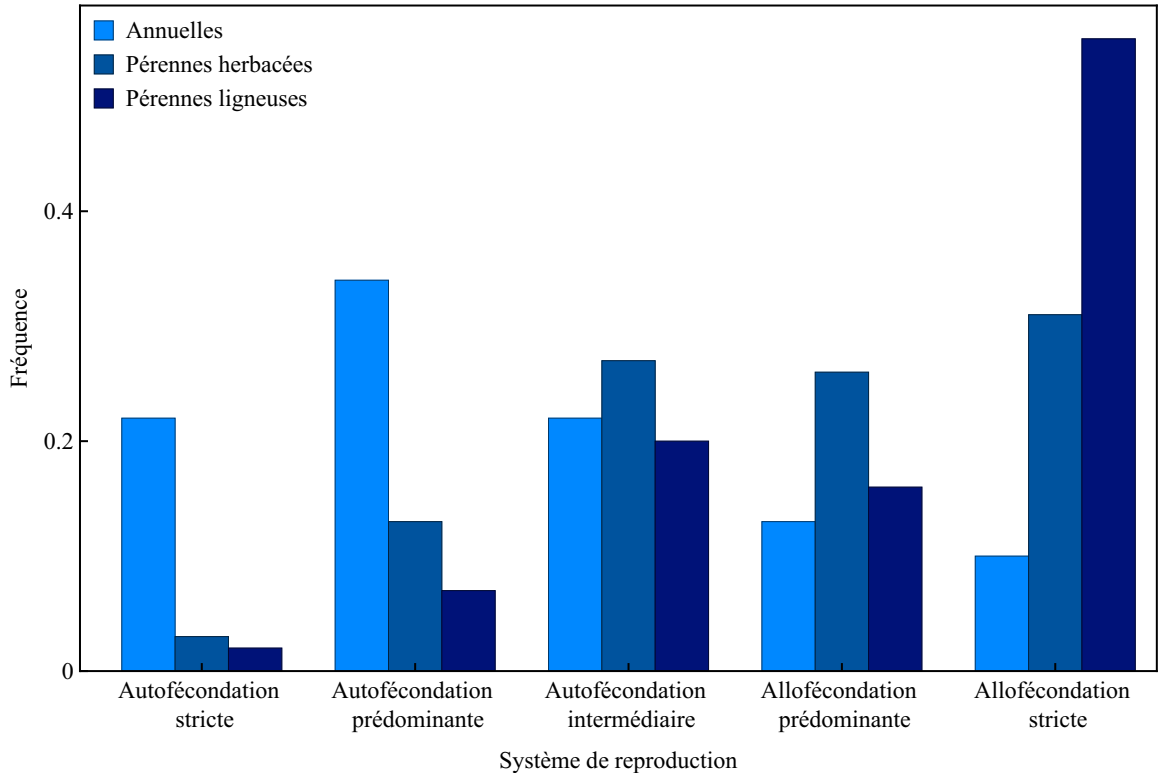

FIGURE 2.4 – Fréquence des systèmes de reproduction répartis en 5 classes suivant la proportion d'autofécondation réalisée pour trois catégories d'histoire de vie, les annuelles, les pérennes herbacées et les ligneuses. Les espèces herbacées ont une longévité moyenne significativement plus faible que les ligneuses (Ehrlén and Lehtilä, 2002). Données de $n = 1996$ espèces tirées de Munoz et al. (2016).

rons ici par perte de graines. L'idée de base est la suivante. Quel que soit le degré de pérennialité de l'espèce considérée, si l'autofécondation se produit avant ou en même temps que l'allofécondation, alors des graines peuvent être autofécondées alors qu'elles auraient pu être allofécondées, menant ainsi à la production d'une descendance moins viable et donc à une perte de graines au cours de la saison. On parle de perte de graines intrasaisonnière (*within-season seed discounting*). Chez les espèces annuelles, cela produit un coût de l'autofécondation lorsque cette dernière se produit avant ou pendant l'allofécondation mais pas lorsqu'elle se produit de manière retardée, puisque toutes les graines qui pouvaient être allofécondées l'auront été dans ce cas. Chez les espèces

pérennes, Lloyd (1992) propose qu'il pourrait exister une seconde forme de perte de graines dite intersaisonnière (*between-seasons seed discounting*), si l'autofécondation induit la fertilisation d'ovules et donc la consommation de ressources que l'individu aurait eu meilleur compte à investir dans sa survie, sa croissance ou sa fécondité future. Dans ce cas, l'autofécondation pourrait générer un coût additionnel chez les espèces pérennes indépendamment du moment où elle se produit. Morgan et al. (1997) a cherché à rassembler ces arguments verbaux au sein d'un modèle théorique que nous allons étudier dans la suite de cette section. Le lecteur familier avec le travail de Morgan et al. (1997) constatera quelques écarts méthodologiques dans la manière dont les résultats sont obtenus et présentés dans cette section. Les conclusions biologiques demeurent inaffectées par ces changements.

2.3.1.1 Autofécondation antérieure et concomitante

Supposons une population de taille constante $\bar{N}$, où les individus deviennent adultes dès la première année de leur existence et survivent d'une année sur l'autre avec une probabilité qui dépend de la mortalité extrinsèque, quantifiée par la constante $S \in [0, 1]$, et de la dépression de consanguinité qui s'exprime sur la survie des adultes avec une magnitude δ_a lorsque l'individu considéré est issu d'autofécondation. Ainsi, un individu adulte survit entre deux années avec une probabilité S s'il est issu d'allofécondation, et avec une probabilité $S(1 - \delta_a)$ s'il est issu d'autofécondation. Il résulte de ces hypothèses qu'une fraction $1 - S(1 - \theta\delta_a)$ des adultes meurt chaque année, où θ est la proportion d'individus autofécondés parmi les adultes, laissant ainsi sa place aux juvéniles produits au cours de l'année. Supposons par ailleurs que tous les individus produisent le même nombre d'ovules, et que les graines produites sont en large excès par rapport aux places disponibles dans la population, si bien que la population conserve toujours une taille $\bar{N}$. Les juvéniles sont recrutés avec une probabilité proportionnelle à leur fitness. Cette dernière est de 1 si le juvénile est issu d'allofécondation, et $1 - \delta_j$,

où δ_j mesure l'intensité de la dépression de consanguinité chez les juvéniles, sinon.

Dans cette population, supposons d'abord que l'autofécondation se produit à un taux α avant ou pendant l'allofécondation. La proportion d'individus autofécondés parmi les adultes varie alors suivant la récurrence

$$\theta_{t+1} = S(1-\theta_t\delta_a)\frac{(1-\delta)\theta_t}{(1-\delta_a)\theta_t + 1 - \theta_t} + [1 - S(1-\theta_t\delta_a)]\frac{\alpha(1-\delta_j)}{\alpha(1-\delta_j) + 1 - \alpha}, \quad (2.28)$$

et à l'équilibre démographique, nous avons

$$\theta^*(\alpha) = \frac{\alpha(1-\delta_j)(1-S)}{S(1-\alpha)\delta_a + (1-S)(1-\alpha\delta_j)}. \quad (2.29)$$

Introduisons un mutant de taux d'autofécondation $\alpha' \neq \alpha$, et cherchons à déterminer son taux d'accroissement λ lorsqu'il est rare. Supposons que le résident et le mutant soient présents au temps t en effectifs N_t et N'_t, respectivement, avec $N'_t \ll N_t$ et de sorte que $N'_t + N_t = \bar{N}$. Les effectifs du résident et du mutant à $t+1$ sont alors donnés par

$$N_{t+1} = \{N_t\left[1 - S\left(1-\theta_t\delta_a\right)\right] + N'_t\left[1 - S\left(1-\theta'_t\delta_a\right)\right]\}\frac{B(\alpha)N_t}{B(\alpha)N_t + B(\alpha')N'_t} + S\left(1-\theta_t\delta_a\right)N_t, \quad (2.30)$$

et,

$$N'_{t+1} = \{N_t\left[1 - S\left(1-\theta_t\delta_a\right)\right] + N'_t\left[1 - S\left(1-\theta'_t\delta_a\right)\right]\}\frac{B(\alpha')N'_t}{B(\alpha)N_t + B(\alpha')N'_t} + S\left(1-\theta'_t\delta_a\right)N'_t. \quad (2.31)$$

où la fonction $B(\alpha)$ indique la quantité de graines viables produites par un individu ayant un taux d'autofécondation α et contient l'effet de la dépression de consanguinité affectant les juvéniles, et θ'_t désigne la proportion d'autofécondés parmi les mutants. Comme nous avons $N'_t + N_t = N'_{t+1} + N_{t+1} = \bar{N}$, nous pouvons diviser les deux équations par $\bar{N}$ pour obtenir des récurrences sur les fréquences des deux types dans

la population totale. En supposant ensuite que la fréquence du mutant est d'ordre ε, on obtient au premier ordre en ε que

$$P'_{t+1} = [1 - S\,(1 - \theta_t \delta_a)]\,\frac{B(\alpha')}{B(\alpha)} P'_t + S\,(1 - \theta'_t \delta_a)\,P'_t, \tag{2.32}$$

où P'_t est la fréquence du mutant dans la population, ce qui donne

$$\lambda_t = \frac{P'_{t+1}}{P'_t} = [1 - S(1 - \theta_t \delta_a)]\,\frac{B(\alpha')}{B(\alpha)} + S(1 - \theta'_t \delta_a). \tag{2.33}$$

Morgan et al. (1997) donnent alors la forme suivante à la fonction $B(\alpha)$,

$$B(\alpha) = n\left[(1 - \delta_j)\alpha + \frac{1 - \alpha + 1 - \bar{\alpha}}{2}\right], \tag{2.34}$$

où $\bar{\alpha}$ est le taux d'autofécondation moyen dans la population et n est le nombre d'ovules produits par individu, et recherchent, en se restreignant au cas de l'invasion d'une population résidente allofécondante ($\alpha = 0$) par un mutant complètement autofécondant ($\alpha' = 1$), les cas où l'autofécondation est favorisée, c'est-à-dire les cas où

$$\left.\frac{\partial \lambda}{\partial \alpha'}\right|_{\alpha=0,\alpha'=1} > 0. \tag{2.35}$$

En injectant le nombre d'autofécondés à l'équilibre démographique parmi les mutants et les résidents, on obtient alors la condition suivante

$$\left(\frac{1}{2} - \delta_j\right)(1 - \delta_j) > \frac{S}{(1 - S)^2}\delta_a[1 - S(1 - \delta_a)]. \tag{2.36}$$

Dans le cas annuel, $S = 0$, elle se réduit au résultat classique $\delta_j < \frac{1}{2}$. Lorsque $S > 0$, l'évolution de l'autofécondation devient de plus en plus difficile à mesure que la pérennialité augmente, comme le montre la Figure 2.5.

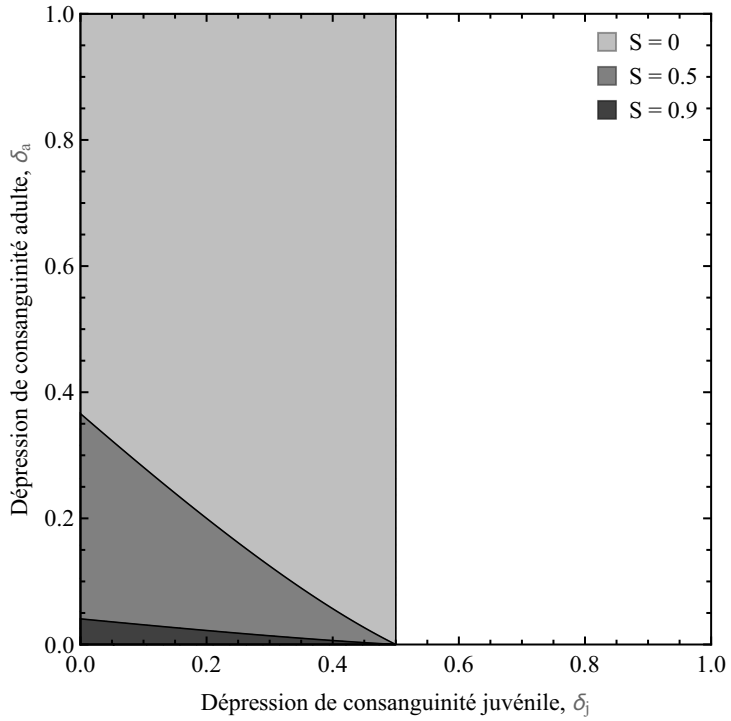

FIGURE 2.5 – Condition pour l'évolution de l'autofécondation antérieure ou concomitante pour diverses niveaux de pérennialité en fonction des dépression de consanguinité juvénile et adulte. Les zones grisées représentent les valeurs de dépression de consanguinité pour lesquelles l'autofécondation envahit la population. Plus la zone est sombre et plus l'espérance de vie est élevée.

2.3.1.2 Autofécondation retardée

Lorsque l'autofécondation se produit avant ou pendant l'allofécondation, Morgan et al. (1997) font l'hypothèse qu'elle ne permet pas d'augmenter la quantité totale de graines produites. Lorsqu'elle se produit après l'allofécondation en revanche, elle peut permettre de féconder les ovules qui ne l'ont pas été lorsque l'allofécondation était possible. C'est dans ce modèle que les auteurs incorporent la perte de graine intersaisonnière (*between-seasons seed discounting*) proposé par Lloyd (1992). Pour cela, ils supposent que tous les individus produisent le même nombre d'ovules o, et S, la probabilité de survie, n'est plus une constante mais une fonction de la proportion

d'ovules fécondés. Nous considérerons la fonction suivante,

$$S(\alpha) = s\left[1 - k\left(c + \alpha(1-c)\right)^{\gamma}\right], \tag{2.37}$$

où c désigne l'efficacité de la pollinisation et s l'intensité de la mortalité extrinsèque. Plus c est petit et plus grande est la proportion d'ovules disponibles pour l'autofécondation, qui ne se produit qu'après l'allofécondation. Cette fonction diffère légérement de celle utilisée par Morgan et al. (1997) par la constante s qui est absente de leur travail. L'ajout d'une telle constante permet de considérer l'évolution de l'autofécondation retardée en fonction de la longévité et en l'absence de perte de graines intersaisonnière. En incorporant la nouvelle fonction $S(\alpha)$ dans l'expression de la proportion d'adultes issus d'autofécondation $\theta^*(\alpha)$, et en réalisant la même analyse que dans la section précédente, nous obtenons qu'un mutant complétement autofécondant peut envahir une population strictement allofécondante si la condition suivante est satisfaite

$$o(1-c)(1-\delta_j)(1-s) - s(1-k)\delta_a \frac{k + \delta_a(1-k)}{k(1-\delta_j)} - sk(1-c)\left[\gamma(1-\delta_a) - oc^{\gamma}(1-\delta_j)\right] > 0. \tag{2.38}$$

Cette condition est toujours satisfaite et l'autofécondation retardée est toujours favorisée dans le cas annuel ($s = 0$, Fig. 2.6a). Chez les espèces pérennes, l'autofécondation est d'autant moins favorisée que la longévité augmente (Figure 2.6a). Dans leur article, Morgan et al. (1997) ont proposé que cela pouvait en partie être dû à l'effet de l'autofécondation sur la perte de graines intersaisonnière. En utilisant une fonction de survie légérement plus générale que la leur, nous montrons qu'à longévité maximale constante l'augmentation de l'intensité de cette perte augmente l'espace de paramètres dans lequel l'autofécondation retardée peut être favorisée (Figure 2.6b). Ce résultat s'explique par la baisse de longévité moyenne qui résulte de la perte de graines intersaisonnière chez les allofécondés aussi bien que les autofécondés, qui réduit de fait l'impact de la dépression de consanguinité adulte. L'argument de Lloyd (1992) ne permet donc

pas, sous cette forme, de favoriser le maintien de l'allofécondation chez les espèces pérennes.

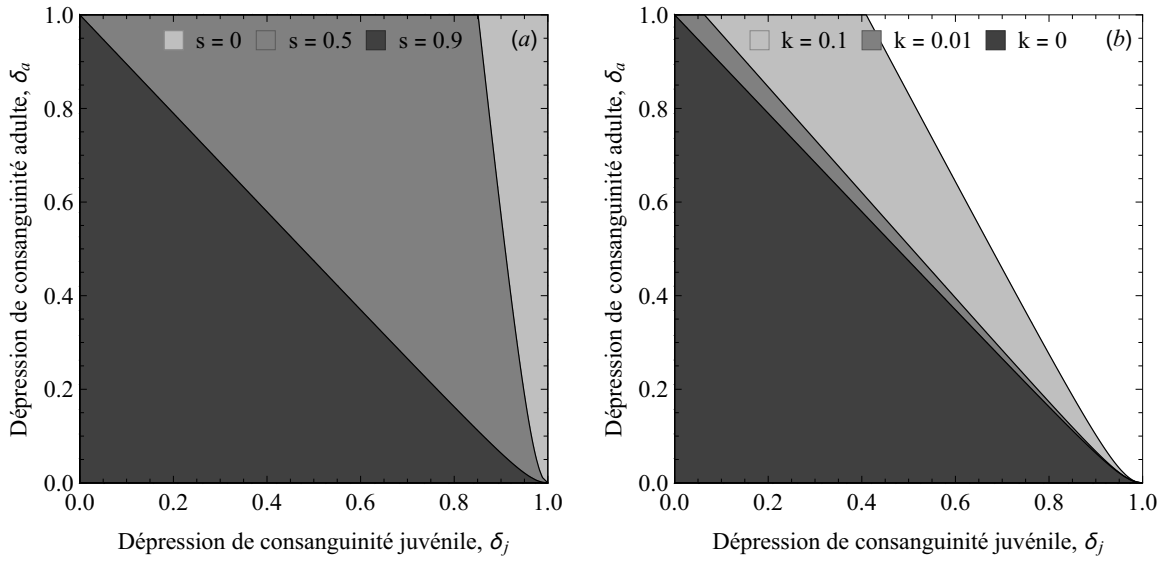

FIGURE 2.6 – Conditions pour l'évolution de l'autofécondation retardée, (a) pour diverses niveaux de pérennialité avec $k = 0$ et (b) pour différents niveaux d'intensité de l'effet de la maturation de graines sur la survie en fonction des dépression de consanguinité juvénile et adulte, avec $s = 0.9$. Les zones grisées représentent les valeurs de dépression de consanguinité pour lesquelles l'autofécondation envahit la population. Les paramètres utilisés sont $\gamma = 0.4$, $c = 0.25$ et $o = 100$.

2.3.2 Effet de l'autofécondation sur l'évolution de la longévité

L'effet de l'autofécondation sur l'évolution de la longévité n'avait à ma connaissance été discuté par aucun auteur avant que Zhang (2000) produise un modèle sur cette question. Pour cela, il s'est intéressé à la coévolution de l'effort reproducteur sous un trade-off entre survie et reproduction et de l'allocation au sexe (c'est-à-dire à la production de pollen, Brunet, 1992). Supposons qu'un individu investisse une fraction E de ses ressources à la reproduction, et que ces ressources soient ensuite réparties entre la fonction mâle, dans une proportion r, et la fonction femelle dans une proportion

$1-r$. Ainsi, un tel individu investit $1-E$ ressources dans la survie, $M = Er$ ressources dans la production de pollen et $F = E(1-r)$ ressources dans la production d'ovules. Supposons que les individus atteignent la maturité en un an et se reproduise par autofécondation au taux α. Les juvéniles parviennent à maturité avec une probabilité s_j ou $s_j(1-\delta_j)$ si ils sont issus d'allo- ou d'autofécondation, respectivement. Aucune dépression de consanguinité n'agit chez les adultes. Zhang (2000) considère alors la mesure de fitness suivante pour un mutant rare d'effort reproducteur et d'allocation au sexe E' et r',

$$W = 2S_a(E') + s_j f(F')[2\alpha(1-\delta_j) + 1 - \alpha] + s_j(1-\alpha)f(F)\frac{m(M')}{m(M)}, \tag{2.39}$$

où $S_a(E)$ est la fonction qui décrit la survie des individus adultes en fonction de leur effort reproducteur, $f(F)$ et $m(M)$ celles qui décrivent la production d'ovules et de pollen, respectivement, en fonction de l'effort reproducteur et de l'allocation au sexe. Le terme associé à la survie adulte est augmenté d'un facteur 2 parce qu'un survivant contribue 2 copies de son génome au pas de temps suivant (Zhang, 2000). Cette mesure de fitness admet un extremum là où sa dérivée s'annule, c'est-à-dire

$$\begin{cases} 2\frac{\partial S_a(E')}{\partial E'} + s_j\left[\Big(1+\alpha(1-2\delta_j)\Big)\frac{\partial f(E',r')}{\partial E'} + (1-\alpha)\frac{f(E,r)}{m(E,r)}\frac{\partial m(E',r')}{\partial E'}\right] = 0 \\ s_j\left[\Big(1+\alpha(1-2\delta_j)\Big)\frac{\partial f(E',r')}{\partial r'} + (1-\alpha)\frac{f(E,r)}{m(E,r)}\frac{\partial m(E',r')}{\partial r'}\right] = 0. \end{cases} \tag{2.40}$$

Les fonctions $m(M)$ et $f(F)$ peuvent être vues comme des fonctions composées de E et r car $M = Er$ et $F = E(1-r)$. Pour toute fonction composée de deux fonctions dérivables $(g \circ f)(x)$, la règle de dérivation en chaîne donne

$$\frac{\partial(g \circ f)(x)}{\partial x} = \frac{\partial(g \circ f)(x)}{\partial f(x)}\frac{\partial f(x)}{\partial x}. \tag{2.41}$$

En appliquant cette règle aux fonctions $m(M) = (m \circ M)(E, r)$ et $f(F) = (f \circ F)(E, r)$, on obtient les relations suivantes

$$\frac{\partial f}{\partial E} = \frac{r-1}{E}\frac{\partial f}{\partial r} \quad \text{et} \quad \frac{\partial m}{\partial E} = \frac{r}{E}\frac{\partial m}{\partial r}, \tag{2.42}$$

où $f = f(E, r)$ et $m = m(E, r)$. Ces relations peuvent être injectées dans les équations (2.40) pour obtenir

$$\begin{cases} \frac{\partial S_a(E')}{\partial E'} + s_j \frac{1+\alpha(1-2\delta_j)}{2(1-r')} \frac{\partial f(E',r')}{\partial E'} = 0 \\ \frac{\partial f(E',r')}{\partial r'} + \frac{1-\alpha}{1+\alpha(1-2\delta_j)} \frac{f(E,r)}{m(E,r)} \frac{\partial m(E',r')}{\partial r'} = 0. \end{cases} \tag{2.43}$$

Ainsi, sous l'hypothèse que la production d'ovules augmente linéairement avec l'allocation de ressources, $f(E, r) = f_{max}E(1 - r)$, on a

$$\frac{\partial f(E', r')}{\partial E'} = f_{max}(1 - r'). \tag{2.44}$$

En supposant la forme suivante pour la probabilité de survie en fonction de l'effort reproducteur, $S_a(E) = s_a(1 - E^{\varphi})$, l'effort reproducteur à l'extremum, E^*, est donné par

$$E^* = \left\{ \frac{s_j f_{max} \left[1 + \alpha(1 - 2\delta_j)\right]}{2\varphi s_a} \right\}^{\frac{1}{\varphi - 1}}. \tag{2.45}$$

Pour déterminer si E^* est une ESS, il faut considérer comment la fitness varie lorsqu'on se déplace dans chacune des directions dans lesquelles le couple (E, r) peut se déplacer. Considérer les dérivées secondes pour chaque variable ne suffit pas car cela ne représente que deux axes suivant lesquels le couple peut varier sur l'infinité possibles. Reprenons donc l'investissement dans les voies mâle et femelle, $M = Er$ et $F = E(1-r)$, comme variables et intéressons-nous à la matrice Hessienne de la fitness, $H(W)$, c'est-à-dire

la matrice des dérivées secondes qui est ici

$$H(W) = \begin{pmatrix} \frac{\partial^2 W}{\partial M^2} & \frac{\partial^2 W}{\partial M \partial F} \\ \frac{\partial^2 W}{\partial M \partial F} & \frac{\partial^2 W}{\partial F^2} \end{pmatrix}. \tag{2.46}$$

Les caractéristiques de cette matrice permettent de déterminer si les extremums d'une fonction sont des maximum ou des minimum. Comme nous sommes en dimension deux, nous pouvons étudier le signe du déterminant de H, D_H,

$$D_H = h_{11}h_{22} - h_{12}h_{21} = 2s_j(1-\alpha)\frac{f(F)}{m(M)}\frac{\partial^2 m}{\partial M^2}\frac{\partial^2 S_a}{\partial E^2}. \tag{2.47}$$

Étant donnée la forme que nous avons supposée pour $S_a(E)$, sa dérivée seconde est négative, et le signe de D_H dépend de la dérivée seconde m. Comme Zhang (2000), supposons que $m(M) = M^n$ est une fonction puissance avec $n < 1$. Dans ce cas, $D_H > 0$ et E^* est une ESS si la trace T_H de la matrice H,

$$T_H = \frac{\partial^2 W}{\partial M^2} + \frac{\partial^2 W}{\partial F^2}, \tag{2.48}$$

est négative. Comme nous avons supposé que la fonction $f(F)$ est linéaire, la fitness est également linéaire en F et sa dérivée seconde par rapport à F est nulle. De ce fait, E^* est une ESS pourvu que $n < 1$. Zhang (2000) montre donc que l'évolution de l'effort reproducteur, et donc de la longévité, est influencée par le taux d'autofécondation d'une manière différente suivant la valeur de la dépression de consanguinité juvénile δ_j. Lorsque la dépression de consanguinité juvénile est inférieure à $1/2$, E^* augmente avec le taux d'autofécondation, et inversement.

Chapitre 3

Travail réalisé et méthodes employées dans cette thèse

3.1 Le travail réalisé dans cette thèse

3.1.1 Avant-propos

Au début de ma thèse, je ne suis pas parti d'un projet prévu pour les trois années à venir. J'ai avancé de question en question, de proche en proche, à tâtons, avec l'intuition que la cohérence générale de mon travail se constaterait *a posteriori*. Il y a trois ans j'ai poussé une porte et je n'ai cessé, depuis, de suivre la piste qu'elle cachait. Le conditionnement que nous impose la mise en compétition permanente - et à mon avis délétère - des chercheurs et des chercheuses entre eux m'a d'abord poussé à chercher une façon d'introduire mes travaux qui les englobe et les justifie dans leur ensemble, comme s'ils étaient le résultat d'un plan de bataille minutieusement établi à l'avance et diligemment exécuté. Rien n'est pourtant plus éloigné de la vérité. Au cours de ma thèse, je me suis planté, beaucoup, j'ai pensé, un peu, et j'ai progressé, sans doute. Me retournant sur ces trois années de thèse, j'ai repris avec un certain embarras

les travaux que j'avais réalisé en début de thèse. Il y a tant de choses que j'aurais fait différemment, si j'avais su. Il m'est alors apparu que la véritable cohérence de mon travail est résolument celle d'un parcours, qu'il est inutile de chercher à maquiller en plan de bataille.

À mes yeux, ce manuscrit de thèse est donc un point d'étape sur ce parcours dont il n'est, je l'espère, pas la fin. C'est dans cet esprit que je l'ai rédigé, et le premier chapitre de l'introduction en est l'illustration. J'ai écrit ce dernier au terme d'une réflexion autarcique à la fois forcée et permise par le confinement. Il dépasse par de nombreux aspects son rôle introductif, d'abord parce qu'il aborde nombre de questions dont mon travail de thèse ne traite pas, mais surtout parce que je me suis servi de sa rédaction pour préciser mes connaissances auparavant limitées sur la théorie des traits d'histoire de vie et pour essayer de comprendre le fonctionnement physiologique des végétaux, cherchant ce faisant à repérer les liens qui unissent ces deux champs scientifiques. Davantage que mon travail de thèse, ce chapitre introduit le projet de recherche que j'ai envie de mener dans les années futures. Ce projet, dont je préciserai les contours en conclusion de ce manuscrit, n'est toutefois pas sans rapport avec mon travail de thèse. En cohérence avec les précédentes lignes, il en est même à mon avis une poursuite naturelle. En guise d'introduction à ce travail de thèse, je raconte ici en quelques phrases le parcours dont il résulte. L'idée n'est pas de résumer le contenu des différents chapitres, chacun étant à cet effet muni d'une synthèse en français, mais bien de présenter comment ils s'articulent les uns aux autres.

3.1.2 Premier chapitre : point de départ

Dans le chapitre deux de l'introduction, nous avons vu que l'existence de dépression de consanguinité affectant la survie des adultes réduit drastiquement la zone de paramètres dans laquelle l'autofécondation peut évoluer chez les espèces pérennes, et ce d'autant plus que la longévité augmente. Pour montrer cela, Morgan et al. (1997)

se sont toutefois cantonnés au cas d'un mutant complétement autofécondant tentant d'envahir une population totalement allofécondante, hypothèse incompatible avec l'importante proportion d'espèces à taux d'autofécondation intermédiaires empiriquement observée (Goodwillie et al., 2005). De plus, l'utilisation d'un modèle phénotypique gomme la nature intrinsèquement fréquence-dépendante de l'évolution des régimes de reproduction (section 2.1). Zhang (2000) montre quant à lui que l'autofécondation peut parfois influencer l'évolution de la longévité en l'absence de dépression de consanguinité adulte et sous des hypothèses strictes sur le lien entre allocation de ressources et survie, production d'ovules et de pollen. Il utilise pour cela le taux de croissance intrinsèque de la population comme mesure de fitness sans préciser la démographie de la population étudiée, ce qui revient au mieux à dire que la population est en croissance exponentielle indéfinie (Encadré B), et multiplie arbitrairement le terme associé avec la survie adulte par deux alors que rien ne justifie ce choix. En effet, un mutant rare est presque toujours à l'état hétérozygote. Ainsi, l'évolution de l'autofécondation peut être influencée par la longévité et l'évolution de la longévité peut être influencée par le régime de reproduction, mais les modèles qui le montrent souffrent de plusieurs limites. À titre personnel, je n'étais convaincu par les conclusions d'aucun de ces deux articles. La motivation initiale du travail réalisé dans cette thèse était de se placer à la croisée de ces travaux en s'intéressant à la coévolution de ces deux traits plutôt qu'à leurs évolutions séparées.

À cet effet, j'ai construit dans le premier chapitre un modèle d'évolution conjointe du taux d'autofécondation et de l'effort reproducteur, dans lequel l'évolution de la longévité se produisait donc à travers un trade-off entre survie et reproduction. Nous avons là un bon exemple de ce qu'une thèse est un parcours. Le premier chapitre de mes travaux de thèse se fonde sur une hypothèse que je me suis employé à déconstruire dans l'introduction générale. Nous y reviendrons. Les résultats de ce chapitre montrent l'importance clef de la dépression de consanguinité affectant la survie des adultes entre

les saisons de reproduction pour la coévolution entre autofécondation et longévité. En effet, l'existence d'une telle dépression de consanguinité, par son effet répété sur la survie des individus autofécondés, a deux conséquences. D'abord, parce qu'elle réduit la longévité des autofécondés et donc leur succès reproducteur total, elle induit une augmentation marquée de la dépression de consanguinité totale, c'est-à-dire mesurée à l'échelle de l'existence complète des individus (*lifetime inbreeding depression*) lorsque la longévité maximale considérée augmente. Cela a pour effet de réduire drastiquement la zone de paramètres dans laquelle l'autofécondation peut évoluer chez les espèces plus longévives, en accord avec les résultats de Morgan et al. (1997). Seules l'autofécondation complète ou l'allofécondation stricte peuvent être évolutivement stable dans ce modèle. C'est lorsque la dépression de consanguinité totale n'est pas suffisamment élevée pour empêcher l'évolution de l'autofécondation que s'exprime la seconde conséquence de la dépression de consanguinité affectant la survie adulte, qui induit alors une augmentation de l'effort reproducteur évolutivement stable et donc une baisse de la longévité. À travers l'action de la dépression de consanguinité sur la survie des adultes, le modèle présenté dans le chapitre un suggère donc que l'autofécondation devrait évoluer plus difficilement chez les espèces pérennes, et induire une réduction de la longévité chez ces dernières si elle parvient à évoluer.

3.1.3 Deuxième et troisième chapitres : génétique des populations structurées en classe et évolution du taux de mutation

Les résultats obtenus dans le premier chapitre sont en accord avec le patron empiriquement observé d'association entre autofécondation et annualité (Figure 2.4). Ils se fondent néanmoins sur l'hypothèse que la dépression de consanguinité est fixée par une constante. Autrement dit, les bases génétiques de la dépression de consan-

guinité sont négligées. Partant de l'hypothèse que la dépression de consanguinité est principalement générée par des mutations délétères récessives maintenues à l'équilibre mutation-sélection dans la population, j'ai donc commencé par étudier la dynamique de mutations diminuant la survie des individus entre deux saisons de reproduction. Loin de générer l'augmentation de dépression de consanguinité prédite par les modèles à dépression de consanguinité fixée, de telles mutations, parce qu'elles sont soumises à une sélection plus intense dans les espèces plus longévives, sont éliminées plus efficacement de la population et aboutissent à une diminution de la dépression de consanguinité avec la longévité à l'équilibre mutation-sélection. Ce résultat jette le discrédit sur ceux obtenus au premier chapitre, car le mécanisme sur lequel ils se fondent ne fonctionne plus sous l'hypothèse la mieux supportée quant aux bases génétiques de la dépression de consanguinité (Charlesworth and Willis, 2009). Empiriquement, on observe pourtant que la dépression de consanguinité augmente avec la pérennialité chez les Angiospermes (Petit and Hampe, 2006; Duminil et al., 2009; Angeloni et al., 2011). La question des mécanismes qui sous-tendent cette augmentation et l'association entre annualité et autofécondation, du même coup, reste donc sans réponse et fait l'objet des chapitres deux et trois qui constituent le coeur de mon travail de thèse.

Deuxième chapitre. Dans le deuxième chapitre, j'ai exploré l'hypothèse selon laquelle l'augmentation de la dépression de consanguinité chez les espèces pérennes pourrait être due à des mutations affectant des traits dont les effets sur la fitness varient avec la pérennialité, les mutations affectant la survie n'étant après tout qu'un cas particulier de cette idée plus générale. Pour cela, partant du constat que la fécondité d'une plante est proportionnelle à sa taille, comme nous l'évoquions dans le premier chapitre de l'introduction, et que la taille est l'un des attributs les plus variables à la fois au niveau intra- et interspécifique au sein du règne végétal, j'ai étudié la dynamique de mutations délétères affectant la croissance des individus. L'étude de telles

mutations nécessitant de modéliser la croissance, j'ai cherché à fonder mon approche sur des hypothèses mécanistes plutôt que phénoménologiques en tirant parti du modèle allométrique de West et al. (2001), présentés dans l'encadré A.

Troisième chapitre. Dans le chapitre un de l'introduction générale, nous avons vu que la lignée germinale émerge constamment de la lignée somatique chez les plantes. À partir de ce constat, Scofield and Schultz (2006) ont proposé que l'augmentation de la dépression de consanguinité chez les espèces plus longévives pourrait résulter de leur plus grande propension à accumuler des mutations somatiques héritables pendant leur croissance puisqu'elles atteignent, du moins en ce qui concerne les espèces ligneuses, des tailles beaucoup plus élevées que les espèces annuelles. Cette hypothèse est la seule autre explication proposée dans la littérature pour expliquer verbalement l'augmentation de la dépression de consanguinité. Elle n'a cependant jamais reçu de traitement théorique et suppose du reste que la taille des individus, c'est-à-dire le nombre de mitoses qui se sont produites en leur sein, est la seule chose qui diffère entre espèces pérennes et annuelles. Un déterminant majeur de l'importance du mécanisme décrit par Scofield and Schultz (2006) est le taux auquel les mutations se produisent pendant la croissance. En effet, si les espèces de plus grande taille présentent davantage d'opportunités d'accumuler des mutations délétères pendant leur croissance, il se pourrait que des mécanismes permettant de réduire leur taux de mutation somatique aient évolué. Dans ce troisième chapitre, l'objectif était de construire le modèle le plus simple possible permettant de décrire des individus chimériques qui accumulent des mutations héritables à mesure qu'ils croissent, afin d'étudier l'évolution du taux de mutation dans de tels organismes ainsi que les conséquences à l'échelle de la population de l'accumulation de mutations somatiques héritables en lien avec la longévité.

3.1.4 Chapitre quatre : Changement d'angle

Les chapitres précédents traitent de la dépression de consanguinité d'un point de vue purement théorique, prenant pour acquis les patrons révélés par les études empiriques. Dans ce quatrième chapitre, nous adoptons une posture différente en nous intéressant aux méthodes indirectes utilisées pour estimer la dépression de consanguinité. Ces méthodes sont les plus utilisées chez les espèces pérennes car mesurer directement la dépression de consanguinité implique d'avoir accès au succès reproducteur total des individus, ce qui est faisable pour une espèce qui complète son cycle de vie en une saison, mais nettemment moins envisageable pour une espèce ligneuse. Nous identifions deux grands types d'approches, celles basées sur le choix d'un proxy de fitness (Nietlisbach et al., 2019), et celles basées sur les variations de l'homozygotie parmi les individus (Ritland, 1990). Nous modifions le modèle étudié dans le chapitre deux pour y introduire des marqueurs neutres afin de mesurer l'homozygotie des individus et appliquons les méthodes identifiées à la population simulée afin d'en évaluer la qualité des prédictions produites par rapport à la dépression de consanguinité réellement à l'oeuvre dans la population.

3.2 Le cadre analytique employé

Tout au long de ma thèse, je me suis servi d'un cadre analytique de génétique des populations multilocus introduit par Barton and Turelli (1991) et généralisé par Kirkpatrick et al. (2002) pour étudier les questions auxquelles je me suis intéressé. Pour cela, j'ai dû l'adapter afin au cas des populations structurées en classes. Dans cette section, je donne les bases générales de ce cadre analytique, ou du moins des outils méthodologiques dont j'ai fait usage en son sein, puis discute brièvement la manière dont j'ai procédé à cette adaptation. L'objectif de cette section est simplement de donner une intuition de ces adaptations, le détail de leur fonctionnement étant donné

dans les chapitres concernés.

3.2.1 Bases générales du cadre analytique de Barton and Turelli (1991)

Par souci de concision, nous nous placerons dans le cas simple d'une population de taille constante N que nous supposerons grande, composée d'individus hermaphrodites et diploïdes. Dans cette population, nous nous intéresserons à la dynamique génétique suivie par des loci autosomaux et bialléliques qui affectent un phénotype lié à la fitness de leur porteur.

3.2.1.1 Définition des variables

Au lieu d'étudier les variations des fréquences des différents génotypes dans la population, le cadre analytique de Barton and Turelli (1991) étudie le changement en fréquence des différents allèles présents et des associations génétiques qui existent entre les différents positions dans le génome, c'est-à-dire de la mesure dans laquelle la manière dont les allèles ségrégent à ces positions s'écarte de la panmixie, comme par exemple l'excès en homozygotes à un locus. Cette approche ne réduit pas en elle-même la complexité des modèles étudiés, puisqu'on se retrouve toujours avec le même nombre de variables à considérer que si l'on s'en tenait à l'étude des fréquences génotypiques, mais elle permet dans de nombreux cas d'obtenir des approximations élégantes auxquelles aucune autre méthode ne parvient. Pour cela, elle repose sur la définition de variables indicatrices qui permettent d'encoder le génotype multilocus des individus.

Variables indicatrices. Considérons deux loci i et j, où ségrégent deux allèles 0 et 1. Les variables $X_{i,\varnothing}$ et $X_{\varnothing,i}$ (resp. $X_{j,\varnothing}$ et $X_{\varnothing,j}$) sont définies comme les variables indicatrices de la présence de l'allèle 1 au locus i (resp. j) sur les chromosomes hérités

paternellement et maternellement, respectivement, c'est-à-dire qu'elles valent 1 lorsque l'allèle 1 est présent et 0 sinon. Nous avons alors que

$$\mathbb{E}\left[X_{i,\varnothing}\right] = \mathbb{E}\left[X_{\varnothing,i}\right] = 1p_i + 0(1-p_i) = p_i, \tag{3.1}$$

où p_i est la fréquence de l'allèle 1 au locus i.

Variables centrées. À partir de ces variables indicatrices, nous pouvons définir les variables centrées notées $\zeta_{i,\varnothing}$ et $\zeta_{\varnothing,i}$ (resp. $\zeta_{j,\varnothing}$ et $\zeta_{\varnothing,j}$) telles que

$$\zeta_{i,\varnothing} = X_{i,\varnothing} - p_i \text{ et } \zeta_{\varnothing,i} = X_{\varnothing,i} - p_i. \tag{3.2}$$

Ces variables sont dites centrées car leur espérance est nulle.

Associations génétiques. Les espérances des produits de ces variables peuvent quant à elles être non-nulles et permettent de mesurer les associations génétiques entre positions du génome. Par exemple, l'espérance du produit $\zeta_{i,\varnothing} \times \zeta_{\varnothing,i}$ est donnée par

$$\mathbb{E}\left[\zeta_{i,\varnothing} \times \zeta_{\varnothing,i}\right] = \mathbb{E}\left[X_{i,\varnothing} \times X_{\varnothing,i}\right] - p_i^2, \tag{3.3}$$

elle mesure l'association génétique entre les allèles hérités paternellement et maternellement au locus i, c'est-à-dire l'excès en homozygotes à ce locus. En effet, dans l'équation (3.3), l'espérance $\mathbb{E}\left[X_{i,\varnothing} \times X_{\varnothing,i}\right]$ mesure la fréquence observée des homozygotes au locus i dans la population et p_i^2 donne leur fréquence attendue sous l'hypothèse de panmixie, c'est-à-dire lorsque les allèles ségrégent indépendamment. De manière générale, l'espérance du produit

$$\zeta_{\mathbb{U},\mathbb{V}} = \prod_{u\in\mathbb{U}} \zeta_{u,\varnothing} \times \prod_{v\in\mathbb{V}} \zeta_{\varnothing,v} \tag{3.4}$$

où $\mathbb{U}$ et $\mathbb{V}$ sont des ensembles de positions génomiques situées sur les chromosomes hérités paternellement et maternellement, respectivement, permet de mesurer l'association génétique entre les allèles ségrégeant à ces ensembles de positions. On note

$$D_{\mathbb{U},\mathbb{V}} = \mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}}\right] \tag{3.5}$$

cette association. Ainsi, l'association $D_{i,i}$ mesure l'excès en homozygotes au locus i évoqué plus haut, et l'association $D_{ij,\varnothing}$ le déséquilibre de liaison entre les locus i et j sur le chromosome d'origine paternelle. Lorsque toutes les associations sont nulles, la population est panmictique. Une notation abrégée pour la moyenne des associations symétriques (par exemple, $D_{ij,\varnothing}$ et $D_{\varnothing,ij}$), est

$$\tilde{D}_{ij} = \frac{D_{ij,\varnothing} + D_{\varnothing,ij}}{2}. \tag{3.6}$$

Cette notation est particulièrement utile pour alléger les expressions lorsqu'aucun effet parental, comme cela pourrait être le cas pour les loci situés sur des chromosomes sexuels par exemple, n'est inclus dans le modèle.

Régles de simplification. Des indices répétés apparaissent souvent dans les calculs impliquant les associations génétiques (par exemple $D_{ii,i}$ où l'allèle porté au chromosome paternel est présent deux fois). Ces répétitions peuvent être éliminées de la manière suivante. Prenons l'association $D_{\mathbb{U}ii,\mathbb{V}}$ où les deux indices i sont les deux seuls indices répétés. Alors, nous avons que

$$D_{\mathbb{U}ii,\mathbb{V}} = \mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}} \times \zeta_{ii,\varnothing}\right] = \mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}} \times (X_{i,\varnothing} - p_i)^2\right]. \tag{3.7}$$

En remarquant que $X_{i,\varnothing}^2 = X_{i,\varnothing}$ car $X_{i,\varnothing}$ ne peut valoir que 0 ou 1, nous obtenons après simplification que

$$D_{\mathbb{U}ii,\mathbb{V}} = (1-2p_i)D_{\mathbb{U}i,\mathbb{V}} + p_i q_i D_{\mathbb{U},\mathbb{V}} \tag{3.8}$$

avec $q_i = 1 - p_i$. Lorsque plusieurs indices répétés sont présents, l'équation (3.8) est appliquée séquentiellement jusqu'à ce qu'il ne reste plus d'indices répétés.

3.2.1.2 Du génotype au phénotype et du génotype à la fitness

Phénotype. Le phénotype qui résulte du génotype d'un individu peut être encodé à partir de ces variables. Prenons un trait quantitatif quelconque z, et supposons qu'il est déterminé par un unique locus biallélique. Quatre génotypes sont possibles, les deux homozygotes (11 et 00) et les deux hétérozygotes (10 et 01). Notons $z_{x_1x_2}$ le phénotype qui résulte du génotype x_1x_2. Nous pouvons alors écrire que le phénotype z d'un individu pris au hasard dans la population est donné par

$$z = z_{00}(1-X_{i,\varnothing})(1-X_{\varnothing,i}) + z_{10}X_{i,\varnothing}(1-X_{\varnothing,i}) + z_{01}(1-X_{i,\varnothing})X_{\varnothing,i} + z_{11}X_{i,\varnothing}X_{\varnothing,i}, \tag{3.9}$$

étant donné qu'une seule de ces combinaisons de variables indicatrices peut être non-nulle à la fois. En injectant les variables centrées dans cette expression et en s'intéressant à sa moyenne, on obtient ensuite que

$$\begin{aligned} z = \bar{z} &+ \left[z_{10} - z_{00} + p_i\left(z_{11} + z_{00} - z_{10} - z_{01}\right)\right]\zeta_{i,\varnothing} \\ &+ \left[z_{01} - z_{00} + p_i\left(z_{11} + z_{00} - z_{10} - z_{01}\right)\right]\zeta_{\varnothing,i} \\ &+ \left(z_{11} + z_{00} - z_{01} - z_{10}\right)\left(\zeta_{i,i} - D_{i,i}\right). \end{aligned} \tag{3.10}$$

En remarquant que $D_{i,\varnothing} = D_{\varnothing,i} = 0$, cette équation peut être réécrite de la manière suivante,

$$z = \bar{z} + \sum_{\mathbb{X}\subseteq\mathbb{G}} \Delta z_{\mathbb{X}} \left(\zeta_{\mathbb{X}} - D_{\mathbb{X}}\right), \tag{3.11}$$

où $\mathbb{G}$ est l'ensemble des positions du génome, $\mathbb{X}$ est un sous-ensemble de $\mathbb{G}$ qui peut contenir à la fois des positions situées sur le chromosome paternel et sur le chromosome maternel, et $\Delta z_{\mathbb{X}}$ est un coefficient qui varie suivant les hypothèses du modèle. Le raisonnement présenté ci-dessus pour un locus se généralise sans difficulté supplémentaire (autre que la longueur des calculs) à un nombre arbitraire de loci, de sorte que l'équation (3.11) exprime la manière dont le phénotype est relié au génotype de manière générale dans le cadre de Barton and Turelli (1991).

Fitness. Le trait z est relié à la fitness d'un individu au cours d'une phase de sélection donnée par une fonction W qu'on supposera non-négative pour tout z et dérivable,

$$W(z) = W\left(\bar{z} + \sum_{\mathbb{X}\subseteq\mathbb{G}} \Delta z_{\mathbb{X}} \left(\zeta_{\mathbb{X}} - D_{\mathbb{X}}\right)\right). \tag{3.12}$$

En supposant que les coefficients $\Delta z_{\mathbb{X}}$ peuvent tous être écrits comme des multiples d'une constante $\nu > 0$, la fonction W peut alors être réexprimée comme une série de Taylor

$$W(z) = \sum_{n=0}^{\infty} \frac{W^{(n)}(\bar{z})}{n!} \left[\sum_{\mathbb{X}\subseteq\mathbb{G}} \Delta z_{\mathbb{X}} \left(\zeta_{\mathbb{X}} - D_{\mathbb{X}}\right)\right]^n, \tag{3.13}$$

où $W^{(n)}$ est la $n^{\text{ième}}$ dérivée de la fonction W en fonction de z. Ainsi, la fitness relative d'un individu, c'est-à-dire sa fitness renormalisée par la fitness moyenne parmi les individus soumis à sélection, est donnée par la série

$$\frac{W}{\overline{W}} = 1 + \frac{1}{W(\bar{z})} \sum_{n=1}^{\infty} \frac{W^{(n)}(\bar{z})}{n!} \left[\sum_{\mathbb{X}\subseteq\mathbb{G}} \Delta z_{\mathbb{X}} \left(\zeta_{\mathbb{X}} - D_{\mathbb{X}}\right)\right]^n. \tag{3.14}$$

Sous l'hypothèse que la sélection est faible, les termes d'ordre $n > 2$ voire $n > 1$ peuvent être négligés pour parvenir à une expression très simplifiée de la fitness relative d'un individu.

3.2.1.3 Sélection

À partir de l'équation (3.14), il est possible de quantifier les effets de la sélection sur les fréquences alléliques et les associations génétiques. On procède pour cela en deux temps.

Associations génétiques prenant la population avant sélection comme référence. Commençons par exprimer l'association génétique considérée *après* sélection en prenant pour référence les fréquences alléliques et associations *avant* sélection ($D^{\bullet}_{\mathbb{U}}$) en utilisant l'équation suivante

$$D^{\bullet}_{\mathbb{U}} = \mathbb{E}\left[\zeta_{\mathbb{U}} \frac{W}{\overline{W}}\right]. \tag{3.15}$$

Au premier ordre, c'est-à-dire en excluant les termes $n > 1$ de l'équation (3.14), cela donne

$$D^{\bullet}_{\mathbb{U}} = D_{\mathbb{U}} + \frac{W^{(1)}(\bar{z})}{W(\bar{z})} \sum_{\mathbb{X} \subseteq \mathbb{G}} \Delta z_{\mathbb{X}} \left(D_{\mathbb{U} \cup \mathbb{X}} - D_{\mathbb{U}} D_{\mathbb{X}}\right). \tag{3.16}$$

Cette expression pour les associations génétiques *après* sélection prenant pour référence la population *avant* sélection peut être utilisée pour obtenir le changement en fréquence des allèles du fait de la sélection ($\Delta^s p_i$). En effet, l'association $D^{\bullet}_{i,\varnothing}$ est donnée par

$$D^{\bullet}_{i,\varnothing} = \mathbb{E}\left[\zeta_{i,\varnothing} \frac{W}{\overline{W}}\right] = \mathbb{E}\left[X_{i,\varnothing} \frac{W}{\overline{W}}\right] - p_i = \Delta^s p_i, \tag{3.17}$$

car $\mathbb{E}\left[\frac{W}{\overline{W}}\right] = 1$ par définition.

Mise à jour des valeurs de référence. En remarquant que la fréquence allélique *après* sélection p'_i est donnée par $p'_i = p_i + \Delta^s p_i$, nous pouvons écrire les variables centrées *après* sélection prenant pour référence la population *après* sélection en fonction des variables centrées prenant pour référence la population *avant* sélection de la manière suivante,

$$\zeta'_{i,\varnothing} = X'_{i,\varnothing} - p_i - \Delta^s p_i, \text{ et } \zeta'_{\varnothing,i} = X'_{\varnothing,i} - p_i - \Delta^s p_i, \tag{3.18}$$

où les X' sont mesurés *après* sélection. Nous voyons alors apparaître des termes en $X' - p$ qui correspondent aux variables centrées mesurées *après* sélection en prenant pour référence les fréquences alléliques *avant* sélection, ce qui permet d'exprimer les associations génétiques *après* sélection en fonction des associations génétiques *après* sélection mais prenant en référence les fréquences alléliques *avant* sélection et des changements de fréquences alléliques,

$$D'_{\mathbb{U}} = D^{\bullet}_{\mathbb{U}} + \sum_{\mathbb{X} \subseteq \mathbb{U}} D^{\bullet}_{\mathbb{U} \setminus \mathbb{X}} \prod_{i \in \mathbb{X}} \left(-\Delta^s p_i\right). \tag{3.19}$$

3.2.1.4 Autofécondation partielle, recombinaison et mutation

Autofécondation et recombinaison. Toute structure de croisement, parce qu'elle écarte la population de la panmixie, est susceptible de générer des associations génétiques intra- et interlocus. Au cours de ma thèse, je n'ai considéré qu'une structure de croisement, l'autofécondation partielle, dans laquelle les individus se reproduisent par autofécondation à un taux α et par allofécondation (*random mating*) pour le reste. L'autofécondation a pour conséquence d'augmenter l'homozygotie des individus aux différents loci considérés, et de créer ce faisant des associations génétiques entre ces derniers. Par exemple, un individu homozygote a un locus aura une chance plus élevée d'être homozygote à d'autres loci avoisinants. À l'inverse, la recombinaison a pour

effet de détruire les associations génétiques existantes. Aucun de ces deux processus n'a d'effet sur les fréquences alléliques, sauf à considérer un locus modifieur du taux d'autofécondation ou que la recombinaison provoque l'apparition de mutations, par exemple. L'expression générale de leur effet conjoint sur l'association $D_{\mathbb{A},\mathbb{B}}$ où $\mathbb{A}$ et $\mathbb{B}$ contiennent les positions héritées paternellement et maternellement, respectivement, est donnée par

$$\begin{aligned} D^r_{\mathbb{A},\mathbb{B}} &= \sum_{\substack{\mathbb{S},\mathbb{T} \\ \mathbb{S}\cup\mathbb{T}=\mathbb{A}}} \sum_{\substack{\mathbb{U},\mathbb{V} \\ \mathbb{U}\cup\mathbb{V}=\mathbb{B}}} t_{\mathbb{S},\mathbb{T}} t_{\mathbb{U},\mathbb{V}} \left\{ (1-\alpha)\mathbb{E}\left[\zeta_{\mathbb{S},\mathbb{T}}\right]\mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}}\right] + \alpha\ \mathbb{E}\left[\frac{\zeta_{\mathbb{S}\cup\mathbb{U},\mathbb{T}\cup\mathbb{V}} + \zeta_{\mathbb{S}\cup\mathbb{V},\mathbb{T}\cup\mathbb{U}}}{2}\right]\right\} \\ &= \sum_{\substack{\mathbb{S},\mathbb{T} \\ \mathbb{S}\cup\mathbb{T}=\mathbb{A}}} \sum_{\substack{\mathbb{U},\mathbb{V} \\ \mathbb{U}\cup\mathbb{V}=\mathbb{B}}} t_{\mathbb{S},\mathbb{T}} t_{\mathbb{U},\mathbb{V}} \left\{ (1-\alpha) D_{\mathbb{S},\mathbb{T}} D_{\mathbb{U},\mathbb{V}} + \alpha \frac{D_{\mathbb{S}\cup\mathbb{U},\mathbb{T}\cup\mathbb{V}} + D_{\mathbb{S}\cup\mathbb{V},\mathbb{T}\cup\mathbb{U}}}{2}\right\} \end{aligned} \tag{3.20}$$

où les ensembles $\mathbb{S}$ et $\mathbb{T}$ et $\mathbb{U}$ et $\mathbb{V}$ sont respectivement des partitions de $\mathbb{A}$ et de $\mathbb{B}$, et $t_{\mathbb{U},\mathbb{V}}$ est la probabilité que les positions comprises dans $\mathbb{U}$ proviennent d'un des chromosomes du parent, et celles contenues dans $\mathbb{V}$ de l'autre. L'équation (3.20) paraît compliquée du fait de la lourdeur des notations, mais se révèle la plupart du temps assez simple d'utilisation car les associations génétiques entre plus de deux ou trois loci sont souvent négligeables lorsque la sélection est faible et la recombinaison est suffisamment forte. La Figure 3.1 présente le raisonnement qui permet d'aboutir aux coefficients t dans le cas de l'association $D_{ijk,ijk}$ entre trois loci notés i, j, k et placés dans cet ordre sur les chromosomes. Ici, la recombinaison peut mener à quatre cas de figure différents suivant les crossing-overs qui se produisent et dans chaque cas de figure, chaque chromosome a une probabilité de $\frac{1}{2}$ d'être transmis. Il suffit ensuite de combiner les différentes possibilités. Par hypothèse, l'allofécondation se produit entre individus non-apparentés. De ce fait, les associations d'allèles héritées par les descendants par la voie maternelle et paternelle sont indépendantes et sont obtenues en multipliant les associations génétiques héritées de chaque parent entre elles. L'autofécondation implique en revanche une forte dépendance entre les chromosomes hérités, puisqu'ils

proviennent du même individu. Cette dépendance est prise en compte en s'intéressant à l'espérance des produits et non aux produits des espérances.

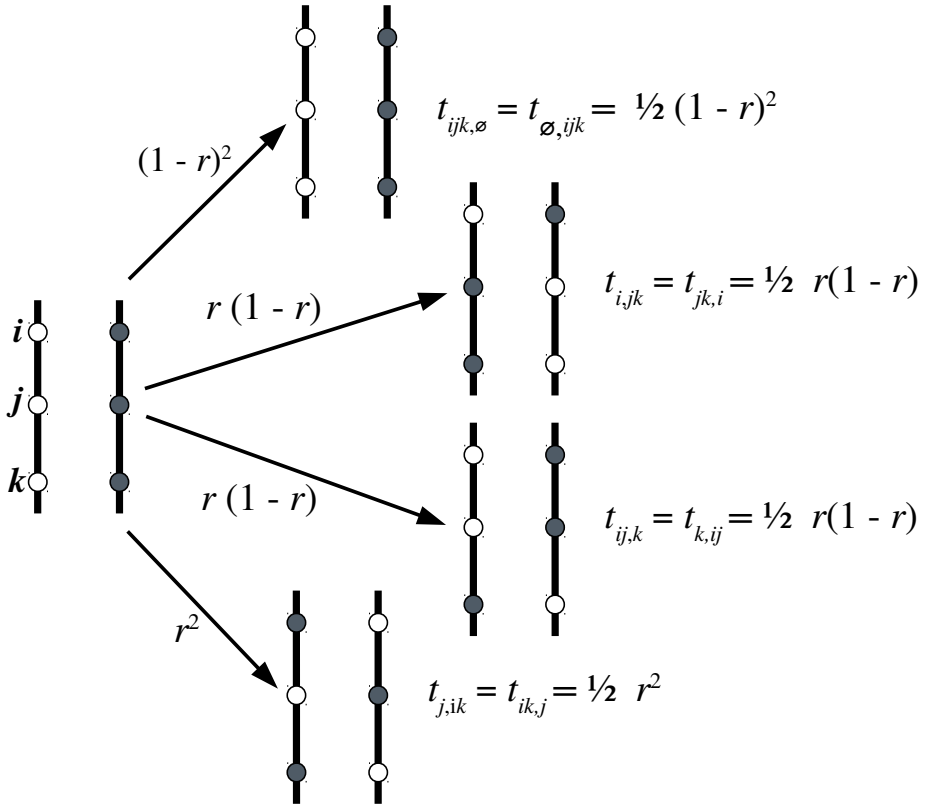

FIGURE 3.1 – Combinaisons possibles des allèles d'origine paternelle (en blanc) et maternelle (en gris) à l'issue d'une phase de recombinaison où l'on suppose que les taux de recombinaison entre les loci sont constants et égaux ($r_{ij} = r_{jk} = r$), et les probabilités de transmission $t_{\mathbb{U},\mathbb{V}}$ associées dans un cas à trois loci.

Mutation. La mutation a pour effet de changer l'identité d'un allèle et affecte de ce fait les fréquences alléliques et les associations génétiques. Supposons qu'à chaque locus, l'allèle 0 mute vers l'allèle 1 à un taux u, et l'allèle 1 mute vers l'allèle 0 à un taux v. La variable indicatrice $X_{i,\varnothing}$ (resp. $X_{\varnothing,i}$), qui décrit le génotype des individus au locus i sur le chromosome paternel (resp. maternel), est affectée de la manière

suivante

$$X_{i,\varnothing}^m = X_{i,\varnothing}(1-v) + u(1-X_{i,\varnothing}) = X_{i,\varnothing}\,(1-(u+v)) + u. \tag{3.21}$$

Si l'individu est de génotype 1 avant mutation (de sorte que $X_{i,\varnothing} = 1$), il demeure dans cet état avec probabilité $1-v$. Si il est de génotype 0 en revanche ($X_{i,\varnothing} = 0 \Leftrightarrow 1 - X_{i,\varnothing} = 1$), il mute avec probabilité u. Après mutation, la fréquence de l'allèle 1 à cette position est donnée par

$$p_i^m = \mathbb{E}\left[X_{i,\varnothing}^m\right] = p_{i,\varnothing}\Big(1-(u+v)\Big) + u. \tag{3.22}$$

Ainsi, la variable centrée associée à cette position après mutation est donnée par

$$\zeta_{i,\varnothing}^m = X_{i,\varnothing}^m - p_{i,\varnothing}^m = \zeta_{i,\varnothing}\Big(1-(u+v)\Big). \tag{3.23}$$

Notant $\mu = u+v$, nous obtenons donc que l'effet de la mutation sur les produits de variables centrées, c'est-à-dire sur les associations génétiques, est donné par

$$D_{\mathbb{U}}^m = \mathbb{E}\left[(1-\mu)^{|\mathbb{U}|}\zeta_{\mathbb{U}}\right] = (1-\mu)^{|\mathbb{U}|}D_{\mathbb{U}}, \tag{3.24}$$

où $|\mathbb{U}|$ désigne le cardinal de l'ensemble $\mathbb{U}$, c'est-à-dire le nombre de positions distinctes qu'il contient.

3.2.2 Apports méthodologiques : chevauchement de générations, lien phénotype-fitness et variables âge-dépendantes

Avant les travaux présentés dans cette thèse, le cadre théorique présenté dans la section précédente n'avait, à ma connaissance, jamais été employé pour étudier des populations structurées en âges où les générations sont chevauchantes. Au fil des

chapitres, j'ai apporté trois éléments méthodologiques nouveaux que j'évoque succinctement dans cette section. Le détail de ces méthodes est donné dans les appendices de chaque chapitre.

Chevauchement de générations. Au cours de ma thèse, j'ai introduit le chevauchement des générations de la manière la plus simple possible en supposant que les juvéniles ne peuvent être recrutés qu'en remplacement des adultes décédés au pas de temps précédent, afin que la population demeure de taille constante. Dans ces conditions, il est possible de séparer les processus qui se produisent parmi les juvéniles de ceux se produisant parmi les adultes, c'est-à-dire de s'intéresser à ces deux sous-populations séparément pour les rassembler ensuite. De manière générale, si on note $D^j_{\mathbb{U}}$ et $D^a_{\mathbb{U}}$ les associations génétiques résultant des différentes phases du modèle (e.g. sélection, reproduction, mutation...) parmi les juvéniles et les adultes, respectivement, cette variable mesurée dans la population complète à la fin du pas de temps ($D'_{\mathbb{U}}$)est donnée par

$$D'_{\mathbb{U}} = (1 - \bar{S})D^j_{\mathbb{U}} + \bar{S}D^a_{\mathbb{U}}, \tag{3.25}$$

où $\bar{S}$ est la probabilité de survie moyenne parmi les adultes. Cette probabilité peut dépendre de la dynamique génétique de la population, c'est-à-dire du génotype des individus aux loci modélisés. Dans un modèle à deux loci bialléliques et sous l'hypothèse que la sélection est faible, comme c'est le cas dans le premier chapitre, cette dépendance est faible et peut être négligée pour obtenir des approximations analytiques valables. Dans un modèle où les loci sont plus nombreux, comme c'est le cas dans le chapitre deux où une infinité de loci affectant la survie est considérée, des méthodes numériques doivent être employées.

Lien phénotype-fitness. De manière plus générale, dès lors que de nombreux loci affectent un phénotype lié non-linéairement à la fitness, c'est-à-dire que

$$\frac{\partial W^2(z)}{\partial z^2} \neq 0, \tag{3.26}$$

il est nécessaire, pour modéliser correctement la dynamique génétique se produisant à un locus, de connaître l'état de la population aux autres loci car cet état détermine la position de l'individu sur le paysage de fitness et donc l'intensité voire la direction de la sélection qui s'exerce audit locus. Dans le chapitre deux de mon travail de thèse, je tiens compte de cet effet en résolvant numériquement le modèle pour le nombre de mutations par haplotype à l'équilibre mutation-sélection et la valeur moyenne du phénotype affecté par les mutations.

Variables âge-dépendantes. Le chapitre trois, qui porte sur les conséquences populationnelles de l'accumulation de mutations somatiques héritables à une infinité de loci bialléliques, apporte une nouvelle difficulté dans le fait que contrairement aux modèles précédents les individus modélisés sont chimériques, constitués de modules d'âges différents dans lesquels les mutations somatiques ont donc eu plus ou moins longtemps pour s'accumuler. En outre, étant donné qu'un locus modifieur du taux de mutation est considéré, la vitesse à laquelle ces mutations s'accumule dépend du génotype de l'individu. Ainsi, la fitness d'un module d'âge n au sein d'un individu n'est pas la même que celle d'un module d'âge $n+1$. Pour palier à ce problème, j'introduis dans le chapitre trois des variables indicatrices âge-dépendantes. Par exemple, la variable $X^n_{i,\varnothing}$ décrit le génotype des modules d'âge n sur le chromosome hérité paternellement au $i^{\text{ème}}$ locus. Dans ce chapitre, les différents modules ne diffèrent que par l'accumulation de mutations somatiques. De ce fait, les variables décrivant le génotype d'un module d'âge $n+1$ peuvent être déduites de celles décrivant les modules d'âge n par une suite

arithmético-géométrique, donnée en l'occurrence par

$$X_{i,\varnothing}^{n+1} = X_{i,\varnothing}^{n} + u(1 - X_{i,\varnothing}^{n}), \tag{3.27}$$

où u est le taux de mutation de l'individu considéré, et dont le terme général est

$$X_{i,\varnothing}^{n} = 1 - (1-u)^n (1 - X_{i,\varnothing}^{0}), \tag{3.28}$$

où $X_{i,\varnothing}^{0}$ est la variable qui décrit le génotype des modules d'âge zéro, c'est-à-dire les graines. À l'aide de ces variables, il est possible de quantifier les changements de fréquence et des associations génétiques à chaque locus parmi les modules d'âge n puis d'obtenir les changements moyens en faisant la moyenne des changements dans chaque classe d'âge, ce qui nécessite de connaître la fréquence de chacune et permet d'obtenir les approximations présentées dans le chapitre trois.

Deuxième partie

Travaux de thèse

Chapitre 1

The joint evolution of lifespan and selfing

Auteurs : Thomas Lesaffre & Sylvain Billiard

Statut de publication : Le travail présenté dans ce chapitre a été recommandé par *Peer Community In Evolutionary Biology*. La recommandation par Thomas Bataillon peut être trouvée ici : lien. Il est à présent publié dans *Journal of Evolutionary Biology* (lien).

1.1 Synthèse en français

1.1.1 Motivation

Il existe une forte association entre annualité et autofécondation prédominante d'une part, et pérennialité et allofécondation d'autre part chez les Angiospermes (Figure 2.4, Stebbins, 1950; Barrett and Harder, 1996; Munoz et al., 2016). La littérature suggère que l'autofécondation et la longévité peuvent influencer l'évolution l'une de l'autre, au travers des arguments présentés dans la section 2.3 du chapitre deux de

l'introduction. Aucune étude ne s'est cependant penchée sur l'évolution conjointe de ces deux traits. L'objectif de ce premier chapitre est d'étudier cette coévolution.

1.1.2 Modèle

Pour cela, un modèle de génétique des populations dans lequel l'effort reproducteur et le taux d'autofécondation évoluent conjointement est considéré. L'utilisation d'un modèle de génétique des populations plutôt que d'un modèle phénotypique plus classique est initialement basée sur une préférence personnelle mais permet par ailleurs de supposer que chaque trait est déterminé par un locus modifieur, ce qui permet aux individus d'adopter n'importe quelle stratégie d'effort reproducteur et d'autofécondation, contrairement au modèle de Morgan et al. (1997) dans lequel l'évolution de l'autofécondation était restreinte à l'allofécondation ou l'autofécondation stricte. Les individus survivent entre les saisons de floraison avec une probabilité qui dépend de la mortalité extrinsèque, de leur effort reproducteur (*i.e.* un trade-off est supposé entre survie et reproduction) et de la dépression de consanguinité si ils sont issus d'autofécondation. Les juvéniles sont recrutés en remplacement des individus morts au pas de temps précédent de sorte que la population conserve une taille constante que l'on suppose grande. En plus de la dépression de consanguinité qui agit sur la survie à l'âge adulte, les individus issus d'autofécondation souffrent de dépression de consanguinité au stade juvénile, ce qui réduit la probabilité qu'ils soient recrutés, comme dans le modèle de Morgan et al. (1997). Des résultats analytiques sont d'abord obtenus pour l'évolution de chaque trait séparément, considérant l'autre comme fixé, ce qui nous permet de formuler des prédictions sur leur évolution conjointe. Ces prédictions sont ensuite testées à l'aide de simulations individu-centrées.

1.1.3 Résultats clefs

Nous montrons que l'évolution de chaque trait séparément permet de décrire précisément les stratégies évolutivement stables qui résultent de la coévolution entre autofécondation et effort reproducteur. Suivant la magnitude de la dépression de consanguinité juvénile et adulte et de la mortalité extrinsèque, trois situations sont possibles. Soit l'allofécondation stricte est toujours favorisée et l'effort reproducteur se stabilise à un niveau qui dépend des conditions de mortalité extrinsèque subies par la population, soit l'autofécondation stricte évolue toujours, auquel cas l'effort reproducteur atteint un niveau plus élevé que dans une espèce strictement allofécondante à l'ESS du fait de la dépression de consanguinité adulte qui réduit la longévité des individus. Enfin, une situation dans laquelle deux équilibres stables existent peut être observée. Dans ce cas, soit la population évolue vers l'allofécondation stricte et se stabilise à l'effort reproducteur correspondant, soit elle évolue vers l'autofécondation complète et un effort reproducteur plus élevé est maintenu à l'ESS. L'équilibre atteint par la population dépend alors des conditions initiales de son évolution. Ces résultats montrent l'importance clef de la dépression de consanguinité agissant sur la survie adulte pour la coévolution de la longévité et de l'autofécondation. En son absence, ces deux traits évoluent de manière indépendante et l'évolution de l'autofécondation est aussi probable dans une espèce pérenne que dans une espèce annuelle. Lorsqu'elle est présente en revanche cette dépression de consanguinité induit, par son effet répété sur la survie des individus, une augmentation rapide de la dépression de consanguinité endurée par les individus consanguins à l'échelle de leur existence complète à mesure que la longévité maximale de l'espèce augmente, ce qui réduit considérablement les possibilités d'évolution de l'autofécondation chez les espèces plus longévives, en accord qualitatif avec les résultats de Morgan et al. (1997) même si le seuil pour la zone dans laquelle l'évolution de l'autofécondation est possible décroît moins vite avec la longévité dans notre modèle. Ce résultat s'explique par le fait que nous laissons l'autofécondation

évoluer par petits incréments (sélection faible) alors que Morgan et al. (1997) considérent l'invasion d'une population allofécondante ($\alpha = 0$) par un mutant totalement autofécondant ($\alpha' = 1$). Lorsque la dépression de consanguinité n'est pas suffisamment élevée pour empêcher l'évolution de l'autofécondation complète, cette dernière envahit la population et la dépression de consanguinité adulte a pour effet d'augmenter l'effort reproducteur à l'ESS, c'est-à-dire de réduire la longévité des individus. Les résultats que nous obtenons dans ce modèle sont en accord avec l'association empiriquement observée de l'autofécondation avec l'annualité et de la pérennialité avec l'allofécondation.

1.2 Introduction

In Angiosperms, strong associations exist between mating systems and other life-history traits, such as dispersal (Auld and Rubio de Casas, 2012), allocation to male *vs.* female functions (Brunet, 1992), or lifespan. Despite significant empirical work (Barrett and Harder, 1996; Scofield and Schultz, 2006; Munoz et al., 2016), the latter association has received limited attention from the theoretical standpoint. Indeed, while it has long been recognised that most self-fertilising species are short-lived and most predominant or obligate outcrossers are long-lived (Stebbins, 1950; Barrett and Harder, 1996; Duminil et al., 2009; Munoz et al., 2016), relatively few theoretical arguments have been advanced to explain this association.

One of the most prevalent argument to explain the evolution of self-fertilisation is reproductive assurance : contrary to an outcrosser, a selfer is able to reproduce in the absence of pollinators or compatible mating partners, which grants it an advantage when pollination is inefficient (Eckert et al., 2006). Stebbins (1950) proposed that this advantage may be weaker in perennials compared to annuals, since perennials get more than one try at reproducing, and therefore may be less impacted by poor pollination

during a given mating season. Later, Lloyd (1992) suggested that self-fertilisation, because it ensures reproduction, could also cause the consumption of resources that could have been more advantageously allocated to post-breeding survival or future outcrossed reproduction in perennials, yielding an additional cost to self-fertilisation is such species (between-seasons seed discounting). Morgan et al. (1997) investigated the validity of these arguments by developing a phenotypic model and concluded that the association between annuality and selfing is more likely to be due to between-seasons seed discounting, rather than reproductive assurance. They also showed the importance of the repeated effect of adult inbreeding depression for the maintenance of outcrossing in perennials. Empirical evidence also suggests that inbreeding depression is overall higher in perennials than in annuals (Angeloni et al., 2011), which could constitute an additionnal barrier to the evolution of self-fertilisation in perennials.

These arguments focus on the consequences of perenniality for the evolution of self-fertilisation, considering lifespan as a fixed characteristic. Yet, lifespans evolve in nature (Stearns, 1992), and joint evolutionary shifts of lifespan and mating system have been documented. Indeed, the transition to self-fertilisation is often associated with significant morphological changes, such as vegetative size and flower size reduction (the selfing syndrome, Sicard and Lenhard, 2011), and lifespan shortening compared to outcrossing relatives (Ehrlén and Lehtilä, 2002). For example, *Arabidopsis thaliana* is a highly selfing annual that recently differentiated from its self-incompatible and perennial relatives *A. lyrata* and *A. halleri* (Clauss and Mitchell-Olds, 2006). Furthermore, studying pairs of sister species across 15 families, Grossenbacher et al. (2015) found numerous joint shifts towards selfing and annuality from outcrossing, perennial ancestors in genera such as *Mimulus* or *Medicago*, and very few shifts to longer lifespans in association with selfing. In fact, the only such shifts they found were observed in the *Oenothera* genus, where segregation and recombination are suppressed when reproducing by self-fertilisation, which implies that selfing individuals are effectively

reproducing clonally (Johnson et al., 2009).

These examples show that joint shifts of mating system and lifespan almost always occur in the same direction, from outcrossing more long-lived ancestors towards self-fertilising derived species with a shorter lifespan. In such situations, lifespan shortening could have allowed for the evolution of self-fertilisation. Alternatively, the transition to self-fertilisation could have induced evolution towards a shorter lifespan. This possibility has seldom been investigated. Indeed, classical studies of the evolution of lifespan predict that it should be fine tuned to best fit the extrinsic mortality conditions experienced by the considered population, through optimal allocation of resources to reproduction, growth or survival (e.g. Schaffer, 1974a; Cichoń, 1997) and senescence (e.g. Silvertown et al., 2001), but rarely consider the evolution of lifespan in interaction with other traits. From a theoretical standpoint, the only study that, to our knowledge, has investigated the influence of the mating system on lifespan evolution is that of Zhang (2000), who developed a phenotypic model for the joint evolution of reproductive effort and sex allocation in partially self-fertilising hermaphrodites. Assuming a survival *vs.* reproduction trade-off (Stearns, 1992), they reached the conclusion that reproductive effort increases (and lifespan decreases) when the selfing rate increases through greater allocation to the female function, provided that inbreeding depression is weak ($\delta < \frac{1}{2}$), female reproduction is very costly, and juvenile survivorship is constrained within a narrow range of values. Importantly however, Zhang (2000) assumed inbreeding depression to only affect the survival of juveniles to maturity, although inbreeding depression typically occurs over all stages of the life cycle (Husband and Schemske, 1996).

In summary, on the one hand, the influence of lifespan on the evolution of the mating system has been studied, considering lifespan as a fixed characteristic (Morgan et al., 1997). On the other hand, the potential influence of the mating system on the evolution of lifespan has only been scarcely investigated, assuming no inbreeding

depression occured in adults (Zhang, 2000). Finally, the question of the joint evolution of lifespan and mating system has never been tackled.

In this paper, we build modifier models (Kirkpatrick et al., 2002) to investigate the joint evolution of lifespan and selfing in a population with overlapping generations, including inbreeding depression affecting various steps in individuals' life cycle as fixed parameters. Following previous authors, we model the evolution of lifespan through that of reproductive effort, assuming a survival *vs* reproduction trade-off for which we consider various shapes (*i.e.* convex, linear or concave Stearns, 1992). We incorporate extrinsic mortality as a constant parameter (Schaffer, 1974a; Zhang, 2000). We first study the evolution of each trait separately, taking the other as fixed, and incorporate inbreeding depression affecting juvenile and adult survival. In each case, we obtain accurate analytical approximations for the evolutionary stable strategies. We show that self-fertilisation is expected to favour evolution towards shorter lifespans when inbreeding depression affects adult survival. Conversely, we show that the range of inbreeding depression under which selfing can evolve rapidly shrinks as lifespan increases, in agreement with previous work (Morgan et al., 1997). Then, using individual-centered simulations along with our previous analytical approximations, we study the joint evolution of lifespan and selfing. We study the effects of inbreeding depression affecting various steps in the life cycle, and discuss how extrinsic mortality conditions are expected to affect evolutionary associations. In particular, we show that selfers may sometimes remain short-lived even in a very stable habitat, as a strategy to avoid the deleterious effects of inbreeding.

1.3 Methods

1.3.1 Outline of the model

Life cycle and demographic assumptions. We consider a very large population made of partially selfing hermaphrodites, which are assumed to be diploid. We assume the population stays at carrying capacity. This implies that juveniles may only settle in replacement of recently deceased adults (Figure 1.1). Once settled, juveniles reach maturity before the next mating event.

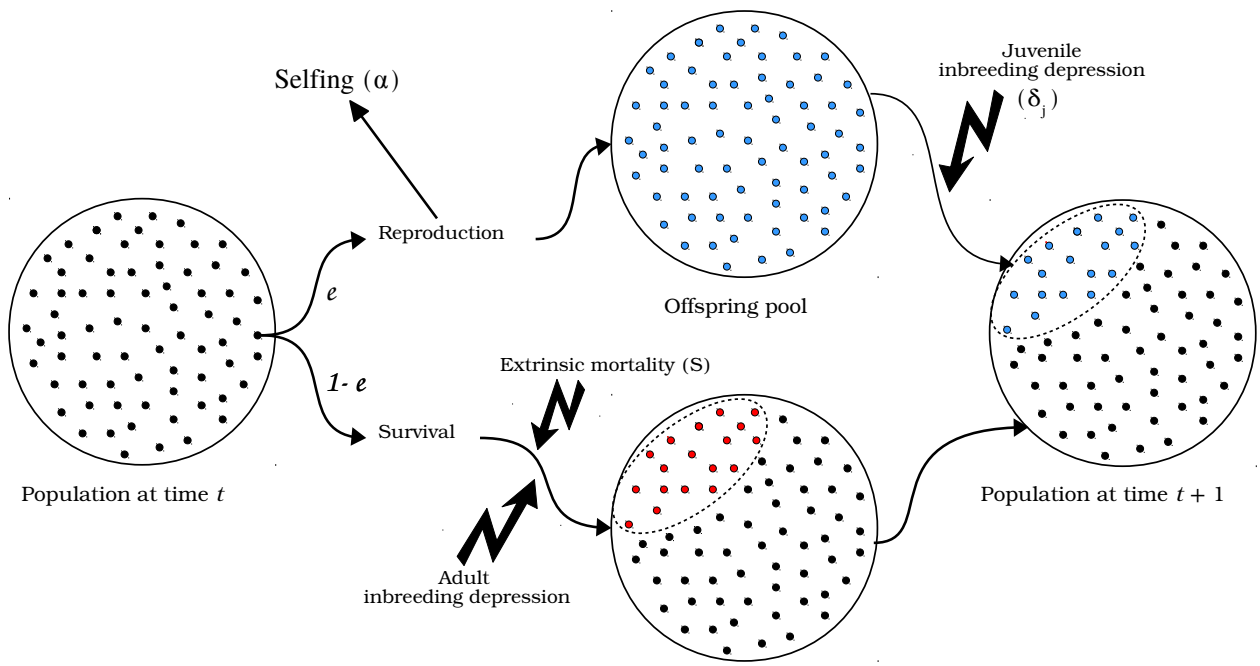

FIGURE 1.1 – Schematic representation of the life cycle and demography assumed in the model. Established individuals allocate a fraction e of their resources to reproduction, and the remaining $1 - e$ to survival. Juveniles replace deceased adults. Red dots depict deceased adults, and blue dots depict juveniles.

We assume that adults keep the same fecundity and survival probability throughout their lives (*i.e.* no age-specific effects). Established individuals allocate a fraction e of

their resources to reproduction, and the remaining fraction $1 - e$ to post-breeding survival. Consequently, sexually mature individuals have a certain probability of survival between mating events (say, flowering seasons) and generations may overlap : the more resources an individual allocates to reproduction, the larger its reproductive output, but the lower its chances of survival until the next mating event. During each mating event, individuals reproduce by self-fertilisation in a proportion α, and by random mating otherwise. Selfed offspring suffer from inbreeding depression (Charlesworth and Charlesworth, 1987) differently depending on the considered stage. As juveniles, they suffer from juvenile inbreeding depression, denoted δ_j, which decreases their probability of survival to maturity. If they reach maturity, they suffer from adult inbreeding depression, denoted δ_a, which decreases their survival probability between mating events. Denoting $\mathcal{S}_o(e)$ the survival probability between two mating events of an outcrossed individual as a function of its reproductive effort, that of a selfed individual, $\mathcal{S}_s(e)$, is therefore given by $\mathcal{S}_s(e) = \mathcal{S}_o(e) \times (1 - \delta_a)$. In simulations, we also considered the case where selfed adults suffer from inbreeding depression on fecundity, which diminishes selfed individuals' contribution to the gamete pool by a proportion δ_f (Appendix V).

Whether lifetime inbreeding depression, that is the decrease in lifetime fitness of selfed individuals relative to the outcrossed, varies with life expectancy depends on the life stages we assume inbreeding depression to affect. When inbreeding depression affects juvenile survival or fecundity, lifetime inbreeding depression is unaffected by life expectancy. On the contrary, when inbreeding depression affects adult survival, lifetime inbreeding depression increases with life expectancy (Appendix I). Indeed, in the latter case, selfed individuals have less opportunities to reproduce, while in the former, selfed individuals have the same life expectancy as the outcrossed.

Genetic assumptions. We assume that individuals' selfing rate and reproductive effort are each entirely determined by a single biallelic modifier locus. In each case, we consider the population to be initially fixed with one allele (the resident) and introduce a rare mutant allele which has a small effect on its bearer's phenotype, that is, we assume weak selection. We then follow the change in frequency of this mutant when it is rare, and look for situations where no mutant can increase in frequency, that is, Evolutionarily Stable Strategies (ESS, Maynard Smith and Price, 1973).

1.3.2 Analytical methods

For each model, we obtained analytical predictions for the evolutionary stable strategies, using the theoretical framework introduced by Barton and Turelli (1991) and generalized by Kirkpatrick et al. (2002). Only a summary of the results is given in the main text, detailed recursions can be found in Appendix I, II and III.

1.3.3 Numerical analyses and simulations methods

All programs used in the present study are available on GitHub : link

Numerical analyses. The analytical results with approximations we obtain when studying the evolution of reproductive effort and selfing separately are compared with that of exact numerical analyses. For each model, the exact recursions, that is tracking genotypic frequencies (rather than allelic frequencies and genetic associations) with no approximations, are run for 10^6 generations. A rare mutant with a small effect ($p_m = 10^{-4}$, $\varepsilon = 0.01$) is introduced at $t = 0$, and the resulting frequency of the mutant is compared to its initial one. When the mutant increases in frequency, the mutant allele is taken as resident, and recursions are run again introducing a new mutant, until mutants no longer increase in frequency or the analysis hits a bound (*i.e.* 0 or 1). As these analyses were only conducted for validation purpose, outputs

are presented in Appendix IV. Our results showed that the approximations we obtain are very close to the numerically obtained ESS, when considering the evolution of reproductive effort and selfing separately.

Individual-centered simulations. To study the joint evolution of lifespan and selfing, we performed individual-centered C++ simulations, incorporating inbreeding depression affecting juvenile (δ_j) and adult survival (δ_a). In Appendix V, we also consider the influence of inbreeding depression affecting fecundity (δ_f). In simulations, individuals follow the same life cycle as described above (Fig. 1.1), and their selfing rate and reproductive effort are each determined by one modifier locus, which are allowed to mutate in both directions, following a uniform distribution in $[\alpha_0 - d, \alpha_0 + d]$ and $[e_0 - d, e_0 + d]$ where α_0 and e_0 are the parent's selfing rate and reproductive effort, respectively, with $d = 10^{-1}$ and a mutation rate $U_m = 10^{-2}$. Free recombination ($r = \frac{1}{2}$) is assumed between the two loci.

1.4 Results

Throughout the following sections, we will need to track the proportion of selfed individuals in the population (Θ), which stays close to its equilibrium value in the absence of mutants (Θ^*) when mutants at modifier loci are rare and only weakly deviate from the resident strategy. As shown in Appendix I, this equilibrium value is given by

$$\Theta^* = \frac{\alpha\,(1-\delta_j)\,(1-\mathcal{S}_o(e))}{\alpha\,(1-\delta_j)\,(1-\mathcal{S}_o(e)) + (1-\alpha)\,\delta_a \mathcal{S}_o(e)}. \tag{1.1}$$

Equation (1.1) is a decreasing function of $\mathcal{S}_o(e)$, that is the equilibrium proportion of selfed individuals in the population is decreased by overlapping generations. We show in Appendix I that this is due to the repeated effect of adult inbreeding depression on post-breeding survival.

1.4.1 Evolution of lifespan in a partially selfing population

In this section, we analyse a model for the evolution of lifespan under partial selfing through the evolution of reproductive effort, considering the selfing rate α as a parameter, and assuming inbreeding depression affects juvenile and adult survival. We assume that the reproductive effort of a given individual is entirely determined by its genotype at a single biallelic modifier locus. Alleles M and m, which we assume to be codominant, code for a reproductive effort $e = e_0$ and $e = e_0 + \varepsilon$ ($\varepsilon \ll 1$), respectively. Furthermore, following Zhang (2000), we assume that the survival probability between two mating events of an outcrossed individual as a function of its reproductive effort e, is given by

$$\mathcal{S}_o(e) = S\left(1 - e^x\right), \tag{1.2}$$

where S is the maximal survival probability, that is a measure of extrinsic mortality, and x controls the shape of the survival *vs* reproduction trade-off. We use this function form because it is flexible and allows for the consideration of a variety of trade-off shapes. The detailed recursions are given in Appendix II. In brief, we follow the variation of three variables, namely, the frequency of allele m (p_m), which is assumed to be a very rare mutant with a small effect ($\varepsilon \ll 1$), the deviation in homozygosity at the modifier locus as compared to the panmictic expectation ($D_{m,m}$), and the proportion of selfed individuals in the population (Θ), which is assumed to remain close to its equilibrium value Θ^*. We look for Evolutionarily Stable Strategies (ESS, Maynard Smith and Price, 1973), that is situations where no mutant allele m may invade the population and replace the resident allele M.

Allelic frequencies change. To leading order in ε, we can express the change in frequency of allele m between two timesteps (Δp_m) as

$$\Delta p_m = \frac{\varepsilon}{2e_0}\Big(p_m\,(1-p_m) + D_{m,m}\Big)V + o(\varepsilon^2), \tag{1.3}$$

with $V = 1 - S\,(1-\Theta\delta_a)\,(1+(x-1)\,{e_0}^x)$. Using a separation of timescales approximation (Kirkpatrick et al., 2002), that is assuming the deviation in homozygosity at the modifier rapidly equilibrates in comparison with allelic frequencies, we obtain a quasi-equilibrium value by solving $\Delta D_{m,m} = 0$ for $D_{m,m}$ (where $\Delta D_{m,m}$ is the change in excess in homozygotes at the modifier between two timesteps). Since $D_{m,m} \geqslant 0$ provided that $\varepsilon \ll 1$, the first term in Equation (1.3) is always of the sign of ε, and only V matters for the determination of the equilibrium.

Evolutionarily Stable Strategy. In order to determine the equilibrium reproductive effort e^*, we plug Equation (1.1) into V, and solve $V = 0$ for e_0, which yields

$$e^* = \left[\frac{\mathcal{B} + \sqrt{(-\mathcal{B})^2 + 4(1-S)(x-1)(1-\delta_a)[1-SS(1-\delta_a)](1-\alpha\delta_j)^2}}{2S(x-1)(1-\delta_a)(1-\alpha\delta_j)}\right]^{\frac{1}{x}}, \tag{1.4}$$

with $\mathcal{B} = (1-S)(2-x)(1-\alpha\delta_j) - \delta_a\,[(1-S(2-x))(1-\alpha\delta_j) - x\alpha(1-\delta_j)]$.

For this equilibrium to be stable, that is for Equation (1.4) to be an ESS, it is required that

$$\frac{\partial V}{\partial e} = x(1-x)e^{x-1}S(1-\Theta\delta_a) < 0. \tag{1.5}$$

One can see that when the survival *vs.* reproduction trade-off is convex or linear ($x \leqslant 1$), condition (1.5) is not fulfilled, which implies that $\Delta p_m \geqslant 0$. Thus, alleles increasing their bearer's reproductive effort are always favoured, and only annuality ($e^* = 1$) can evolve. On the other hand, when the trade-off is concave ($x > 1$), $\frac{\partial V}{\partial e} < 0$, Equation (1.4) is an ESS and perennial strategies can be maintained. This result is classical, and

has been reached by numerous authors before (e.g. Schaffer, 1974a; Bell, 1980; Zhang, 2000). In addition, Equation (1.4) is a decreasing function of S. That is, higher extrinsic mortality favours higher reproductive efforts : allocating resources to survival becomes less and less advantageous as extrinsic mortality increases (*i.e.* as S decreases), because individuals are more likely to die independently of the resources they spend on survival. Hence, our model also predicts the classical observation that annuals are more likely to evolve in more disturbed habitats, where extrinsic mortality is higher.

Assuming $x > 1$, that is a concave trade-off, differentiating Equation (1.4) with respect to α yields

$$\frac{\partial e^*}{\partial \alpha} = \frac{\delta_a \left(1 - \delta_j\right) e^*}{\left(1 - \alpha\delta_j\right) \sqrt{(-\mathcal{B})^2 + 4(1-P)(x-1)(1-\delta_a)[1 - S(1-\delta_a)](1-\alpha\delta_j)^2}} \geqslant 0. \tag{1.6}$$

Equation (1.6) shows that the ESS reproductive effort is an increasing function of the selfing rate α, as long as some selfed juveniles survive to maturity ($\delta_j \neq 1$), and adults suffer from inbreeding depression ($\delta_a \neq 0$). In other words, through its impact on adult survival, self-fertilisation favours evolution towards higher reproductive efforts, and hence shorter lifespans. This effect becomes stronger as adult inbreeding depression increases, because allocating resources to survival becomes less beneficial in selfed adults, and as juvenile inbreeding depression decreases, because it reduces the proportion of selfed individuals entering the adult population. Therefore, we show that self-fertilisation favours evolution of shorter lifespans.

1.4.2 Influence of lifespan on the evolution of self-fertilisation

In this section, we study the influence of lifespan on the evolution of self-fertilisation assuming inbreeding depression only affects juvenile and adult survival, taking reproductive effort, *i.e.* lifespan, as a fixed parameter. We assume that the selfing rate of a given parent is entirely determined by a single biallelic locus. Alleles M and m, which

are assumed to be codominant, code for selfing rates α_0 and $\alpha_0 + a$, respectively. Allele m is assumed to be a rare mutant with a weak effect ($a \ll 1$). In this section, we do not need to make any assumption regarding the shape of the survival *vs* reproduction trade-off. Hence, the general $\mathcal{S}_o(e)$ notation will be used. Recursions are given in Appendix III. As in the previous section, we follow three variables : the frequency of the mutant allele (p_m), the deviation in homozygosity at the modifier locus as compared to the panmictic expectation ($D_{m,m}$), and the proportion of selfed individuals in the population (Θ), which is assumed to remain close to its equilibrium value (Θ^*, Equation (1.1)) when the mutant is rare. We look for the conditions under which allele m invades the population.

Allelic frequencies change. Plugging Equation (1.1) into the allelic frequencies change (Δp_m) yields, to leading order a,

$$\Delta p_m = -\frac{a\left(1-\alpha_0\delta_j\right)D_{m,m}\left(1-\mathcal{S}_o(e)\right)}{2\left[\mathcal{S}_o(e)\left(1-\alpha_0\right)\delta_a + \left(1-\mathcal{S}_o(e)\right)\left(1-\alpha_0\delta_j\right)\right]^2} \times T + o\left(a^2\right) \tag{1.7}$$

with $T = \left[1-\mathcal{S}_o(e)\left(1+\delta_a\right)\right]\left[\mathcal{S}_o(e)\delta_a - \left(1-\mathcal{S}_o(e)\right)\left(1-2\delta_j\right)\right]$.

Using the same arguments as in the previous section, we show in Appendix III that for all α, $D_{m,m} > 0$. Hence, the first term in Equation (1.7) is always negative, and only T matters for the determination of the equilibrium. Since T does not depend on the selfing rate, there are only two possible situations : either $T > 0$ and full outcrossing ($\alpha = 0$) is favoured, or $T < 0$ and full selfing evolves ($\alpha = 1$), similar to the findings of previous authors (e.g. Lande and Schemske, 1985). Solving $T < 0$ for δ_j, we have

$$T < 0 \;\Leftrightarrow \delta_j < \frac{1-\mathcal{S}_o(e)\left(1+\delta_a\right)}{2\left(1-\mathcal{S}_o(e)\right)}, \tag{1.8}$$

which simplifies to $\delta_j < \frac{1}{2}$ in the annual case ($\mathcal{S}_o(e) = 0$). Otherwise, this threshold is a decreasing function of $\mathcal{S}_o(e)$ when $\delta_a > 0$. Hence, we find that the range of condi-

tions under which self-fertilisation can evolve in a population decreases when lifespan increases, provided that inbreeding depression affects adult survival, in agreement with previous results (Morgan et al., 1997). This result implies that even very weak adult inbreeding depression is sufficient to prevent evolution of self-fertilisation in long-lived species. Equation (1.8) was validated using numerical analyses, which were found to be in very good agreement with the analytical prediction (Appendix IV).

1.4.3 Joint evolution of lifespan and self-fertilisation

In this final section, we study the joint evolution of lifespan and selfing. Figure 1.2 highlights the different situations that arise when considering the joint evolution of lifespan and selfing. The threshold reproductive effort above which selfing can evolve ($\bar{e}$) can be obtained by setting $\mathcal{S}_o(e) = S\,(1 - e^x)$ in Equation (1.8), and solving $T < 0$ for e. It is given by

$$\bar{e} = \left(\frac{S(1+\delta_a) + 2\delta_j(1-S) - 1}{S(1+\delta_a - 2\delta_j)} \right)^{\frac{1}{x}}. \tag{1.9}$$

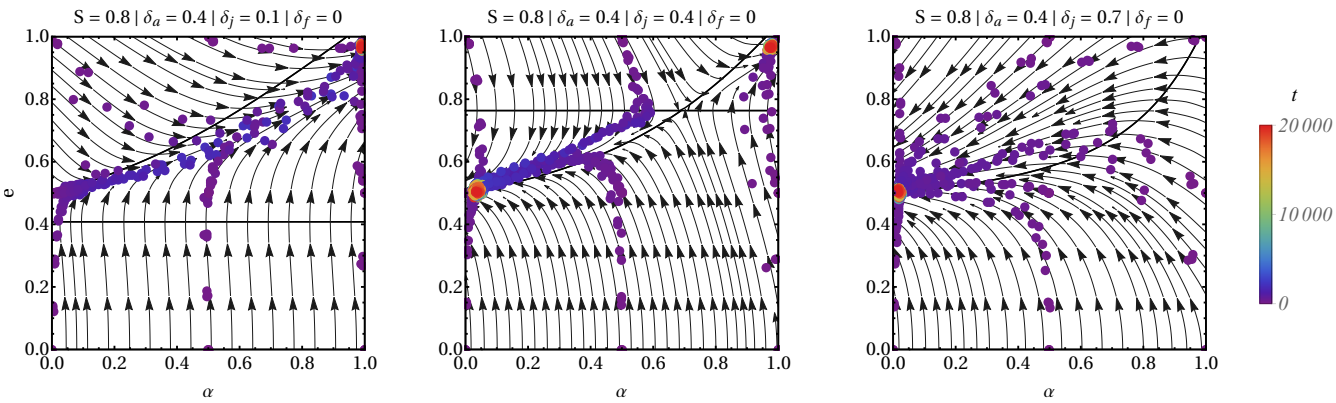

FIGURE 1.2 – Phase diagrams highlighting the three kinds of behaviour that can arise from the coevolutionary dynamics of lifespan and selfing. Reproductive effort (e) is plotted against selfing rate (α). Solid lines depict isoclines, and arrows indicate how the joint evolution behaves. Points depict simulation results, and are colored with respect to time.

Overall, the transition to selfing is always associated with increased reproductive effort, that is reduced lifespan. Depending on the inbreeding depression and extrinsic mortality conditions, one of three things can happen : (1) selfing is always favored

(Fig. 1.2a), and reproductive effort evolves to

$$e^*_{self} = \left(\frac{1 - S(1 - \delta_a)}{S(1 - \delta_a)(x - 1)} \right)^{\frac{1}{x}}, \tag{1.10}$$

(2) outcrossing is always maintained (Fig. 1.2c), and reproductive effort stabilises at

$$e^*_{out} = \left(\frac{1 - S}{S(x - 1)} \right)^{\frac{1}{x}}, \tag{1.11}$$

(3) either of the two stable equilibria ($\alpha = 0$ and $\alpha = 1$) can be reached, depending on initial conditions (Fig. 1.2b). This occurs when adult inbreeding depression is high enough, provided that $\delta_j < \frac{1}{2}$ (Fig. 1.3). Selfers then stabilise at a high reproductive effort, which allows them to escape the deleterious effects of adult inbreeding depression. In other words, in situations where extrinsic mortality conditions would allow for the evolution of longer lifespans, selfers could remain short-lived solely owing to the deleterious effects of inbreeding.

Figure 1.3 summarizes the evolutionary stable outcomes of the joint evolution of lifespan and selfing in various inbreeding depression and extrinsic mortality conditions. Simulation results are in good agreement with analytical predictions, although small deviations are observed in Figure 1.2 and 1.3 for high and low selfing rates and reproductive efforts. These deviations are a consequence of the value constraints applied to the traits to keep them within bounds. Indeed, as the trait value (*i.e.* selfing rate or reproductive effort) approaches a bound of its definition domain (that is, 0 or 1 on the $[0, 1]$ interval), it will tend to mutate slightly away from it, as all mutations exceeding the bound will be cut back to the bound value (for example, a mutation at the reproductive effort modifier coding for a reproductive effort $e > 1$ will be cut to $e = 1$).

In Appendix V, we also consider the effect of inbreeding depression affecting fe-

cundity (δ_f). We show that δ_f or δ_j have the same effect : they reduce the range of conditions under which the full selfing strategy can evolve. The mechanisms underlying these similar behaviours are however slightly different. Indeed, while in both cases they are caused by a reduction of the participation of selfed individuals to reproduction, δ_j decreases the proportion of selfed individuals in the population. On the other hand, δ_f reduces the contribution of selfed adults to reproduction by reducing their contribution to the gamete pool.

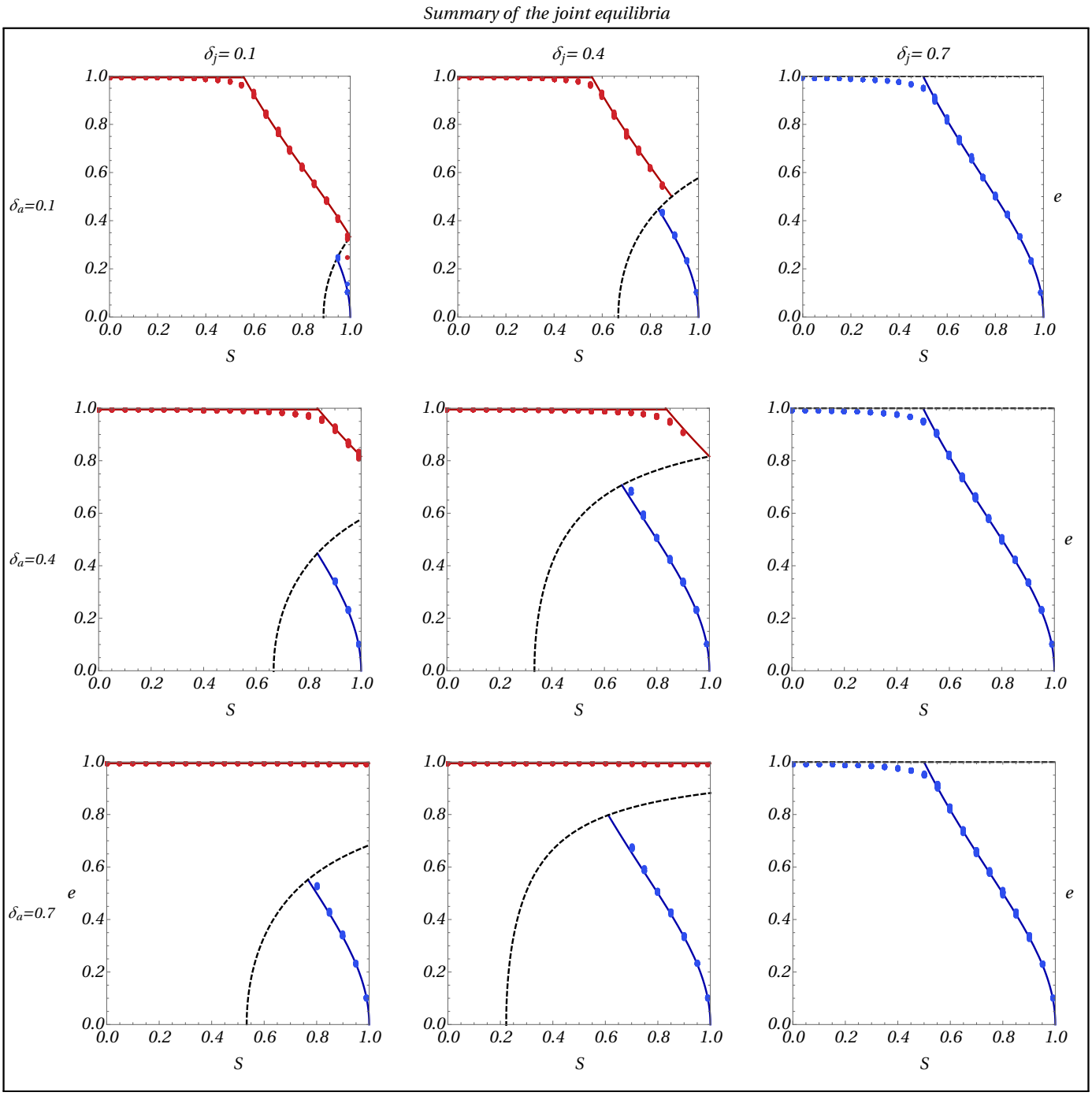

FIGURE 1.3 – Joint evolutionary equilibria. All results presented here assume $x = 2$. Y-axis is the ESS reproductive effort (e^*), and X-axis is extrinsic mortality (S). The lines correspond to analytical results, while dots depict simulations results. The red ones corresponds to the equilibrium reproductive effort when $\alpha = 1$ is reached, while the blue line corresponds to the equilibrium reproductive effort when $\alpha = 0$ is reached. The dashed black line represents the threshold lifetime inbreeding depression under which outcrossing can be maintained.

1.5 Discussion

We built modifier models to investigate the joint evolution of lifespan and mating system. To the best of our knowledge, this is the first time modifier models are used to study the evolution of life-history traits in a population with overlapping generations.

Although mixed mating is common in flowering plants (Munoz et al., 2016), we found only two stable mating systems : full outcrossing or full selfing, similar to previous authors (e.g. Lande and Schemske, 1985; Charlesworth et al., 1990). This discrepancy between theoretical predictions and empirical observations has given rise to multiple theoretical investigations. All of which proposed complications of earlier models which allow mixed mating to be maintained as an evolutionarily stable strategy (Goodwillie et al., 2005; Barrett and Harder, 2017), but go beyond the scope of the present paper. The role of lifespan in the maintenance of mixed mating may however represent a track worth following.

In agreement with Morgan et al. (1997), we found that increasing lifespan considerably reduces the range of conditions under which self-fertilisation can evolve when inbreeding depression affects adult survival. This effect is due to the fact that life expectancy (and therefore lifetime reproductive sucess, that is fitness in our model) scales exponentially with survival probability. Indeed, when survival probability in the population is higher, its reduction among selfed individuals due to inbreeding depression is more consequential for fitness than it is in more short-lived species, which leads to higher lifetime inbreeding depression and therefore greater selection against selfing. Furthermore, we found inbreeding depression on juvenile survival and fecundity (Appendix V) to affect the coevolutionary dynamics similarly, as they both reduced the range of conditions under which the full selfing strategy can be reached. This is because decreasing the contribution of selfed parents to the production of offspring by reducing the amount of gametes they produce (fecundity), or diminishing their frequency among the parents (juvenile survival) is equivalent in our model. Conversely, we

showed that self-fertilisation is expected to cause evolution towards shorter lifespans due to inbreeding depression affecting adult survival, even under the low inbreeding depression conditions required for selfing to evolve in long-lived species. In addition, we showed that self-fertilising species could remain short-lived even in a very stable habitat, as a strategy to avoid the deleterious effects of inbreeding. Finally, we showed that long-lived selfers can only emerge under very weak adult inbreeding depression and should be very rare. This is corroborated by the work of Munoz et al. (2016), who found that only 2% of woody perennials are obligatory selfers.

The only previous study of the influence of the mating system on the evolution of lifespan is that of Zhang (2000). While our main conclusions are in line with theirs, they also differ in a number of ways. Zhang (2000) concluded that self-fertilisation could induce evolution of shorter lifespans through increased allocation to female reproduction, without any role for adult inbreeding depression. Zhang (2000)'s results and ours are not mutually exclusive, since they rely on different mechanisms. However, we expect our prediction to apply more generally. Indeed, while our predictions only require inbreeding depression to affect adult survival, which seems reasonable, as inbreeding depression commonly occurs over all stages of life (Husband and Schemske, 1996), Zhang (2000)'s prediction requires that female reproduction is sufficiently costly, that inbreeding depression remains low and that juvenile survivorship is constrained within a narrow range of values. Importantly, Zhang (2000)'s approach and ours also differ widely in the methods used, as we assume that juveniles are produced in large excess with respect to the available resources, so that the population reachs its carrying capacity before each flowering seasons and the population size is kept constant, whereas Zhang (2000) modeled adult survival as an additive term (see Equation 3 in their study), which is only valid in the case of an exponentially growing population, where juvenile recruitment is not limited by resources availability (Stearns, 1992; Kozłowski, 1993). Small founding populations, which cannot produce enough offspring to saturate their

habitat over the course of one mating season, may exhibit such rapid growth. However, no population grows forever as resources sooner or later become limiting. Since non-annual species, and particularly long-lived species such as trees, tend to inhabit more stable environments (Schaffer, 1974a; Petit and Hampe, 2006), we expect relatively stable populations to be more widespread than exponentially growing populations, which is a necessarily transient state, and therefore would also expect our approach to hold greater generality than that of Zhang (2000).

Overall, both theoretical and empirical results indicate that predominantly selfing species are almost always short-lived, while predominantly outcrossing ones are mostly long-lived, but the causative mechanisms underlying this correlation still require theoretical and empirical work to be unravelled. One role of models such as the present study is to provide putative ecological and genetic conditions under which a given association may emerge. Given these results, we propose that lifespan could act as a confounding factor when considering the joint evolution of self-fertilisation with other traits. For example, lifespan shortening following the transition to self-fertilisation could affect other traits, thereby contributing to the emergence of the selfing syndrome, rather than self-fertilisation *per se* alone. Indeed, as a shorter lifespan implies that individuals have to complete their life cycle more rapidly, they could grow smaller flowers regardless of an adaptation to more efficient selfing (Sicard and Lenhard, 2011). Accounting for lifespan could also shed new light on some long-standing evolutionary questions related to the evolution of mating systems. For instance, lifespan could influence the joint evolution of dispersal and self-fertilisation in two ways. First, increased lifespan induces higher local relatedness (Duputié and Massol, 2013), and perennials tend to occupy more saturated, competitive environments. Hence, kin competition should be greater in more long-lived species, and long-distance dispersal could be favoured as a mean to avoid it. Since perennials outcross more than annuals, this could generate an indirect association between long-distance dispersal and outcrossing through lifespan.

Second, the repeated effect of adult inbreeding depression considerably increases the detrimental effects of mating among relatives in perennials. Long-distance dispersal and outcrossing could thus be favoured jointly in such species as an inbreeding avoidance strategy (Auld and Rubio de Casas, 2012).

Throughout this work, we assumed fixed inbreeding depression. Inbreeding depression is generally thought to be caused by recessive deleterious mutations segregating at low frequencies in populations (Charlesworth and Willis, 2009). The population genetics of such mutations in populations with overlapping generations are poorly understood theoretically. Evidence from large-scale meta-analyses suggest that the measured magnitude of inbreeding depression increases as species' lifespan increases (Duminil et al., 2009; Angeloni et al., 2011). However, it is unclear whether this pattern is due to mating system differences between long-lived and short-lived species (Munoz et al., 2016), or to lifespan. Additionally, empirical studies of inbreeding depression in perennials are rather scarce, and rarely span over several years, let alone individuals' entire lives (Husband and Schemske, 1996). Consequently, the quantities measured in such studies are not likely to depict inbreeding depression in its classic definition, that is the lifetime fitness decrease of selfed individuals as compared to outcrossed ones (Charlesworth and Charlesworth, 1987), but rather its magnitude at a given timestep or stage. Therefore, estimates in short-lived and long-lived species may not be readily comparable, and higher inbreeding depression levels reported in perennials could in part reflect measurement biases (Angeloni et al., 2011). Lifespan may also interact with the mutation load in non-trivial ways. On the one hand, longer lifespan may increase selection against deleterious mutations, because of the increased number of opportunities for selection to occur, thereby leading to lower levels of inbreeding depression through better purging (Morgan, 2001). On the other hand, more long-lived species may endure significantly more mitotic mutations throughout their lives owing to their overall larger stature, which could result in an increase in inbreeding depression as plants do not have

a separated germline (Scofield and Schultz, 2006). Furthermore, based on theoretical arguments (Abu Awad et al., 2016), it was proposed that perennials may experience reduced purging of mutations affecting juvenile fitness, as some of these mutations would remain as neutral mutations among adults, and be recurrently reintroduced in the population through reproduction.

Importantly, these various predictions stem from vastly different approaches, and remain mostly verbal. In particular, the rare models that considered populations with overlapping generations vary considerably in terms of the life stage they assume deleterious mutations to affect. Indeed, while some authors assume they only act on juvenile survival or gametes production (Abu Awad et al., 2016), others assume they affect their bearers' survival throughout their lives (Morgan, 2001). As for somatic mutations (Scofield and Schultz, 2006), although it is clear such mutations occur and can be inherited in large-statured plants such as oaks (*Quercus robur*, Plomion et al., 2018), the population-level consequences of this mechanism were never investigated theoretically. Besides, no study has yet considered the interaction between separate loads affecting different life stages, although these situations are likely to occur since differential purging has been reported between life stages (Angeloni et al., 2011). Finally, every theoretical study so far has assumed individuals' fecundities and survivorship to not depend upon their age, which is a strong simplifying assumption (Franco and Silvertown, 1996; Petit and Hampe, 2006). In the future, significant insight for the evolution of mating systems and life-histories is to be gained through investigations of the dynamics of mutation loads in perennials.

Acknowledgements

This preprint has been reviewed and recommended by Peer Community In Evolutionary Biology (https ://dx.doi.org/10.24072/pci.evolbiol.100070). We thank Thomas

Bataillon, along with two anonymous reviewers, for taking the time to consider our manuscript, and ultimately recommending our work. Their comments greatly improved the quality of the manuscript. We also thank Sylvain Glémin, Florence Débarre, Diala Abu Awad, Vincent Castric, Louis Mahé and Denis Roze for helpful discussions and comments on the manuscript. This work was funded by the European Research Council (NOVEL project, grant #648321). The authors also thank the Région Hauts-de-France, and the Ministère de l'Enseignement Supérieur et de la Recherche (CPER Climibio), and the European Fund for Regional Economic Development for their financial support.

1.6 Appendices

1.6.1 Proportion of selfed individuals and lifetime inbreeding depression

In this section, we detail the derivation of the equilibrium proportion of selfed individuals in the population, and show that its decrease when longevity increases is attributable to stronger selection against inbred individuals.

Proportion of selfed individuals. Following reproduction, adult individuals survive with probability $\mathcal{S}_o(e)$ if they were produced by outcrossing, and $(1-\delta_a)\,\mathcal{S}_o(e)$ if they were selfed. Therefore, denoting Θ the proportion of selfed individuals in the population at a given timestep, the proportion of adult individuals surviving between two mating events ($\bar{\mathcal{S}}$) is given by

$$\bar{\mathcal{S}} = \mathcal{S}_o(e)\left[\Theta\left(1-\delta_a\right)+1-\Theta\right], \tag{A1}$$

and the proportion of selfed individuals among them, Θ^s is

$$\Theta^s = \frac{\Theta\left(1-\delta_a\right)}{\Theta\left(1-\delta_a\right)+1-\Theta}. \tag{A2}$$

Among the juveniles settling in the population at a given mating event, the proportion of selfed individuals (Θ^j) given the selfing rate α can be expressed as

$$\Theta^j = \frac{\alpha(1-\delta_j)}{\alpha(1-\delta_j)+1-\alpha}. \tag{A3}$$

Hence, the proportion of selfed individuals at the next timestep (Θ') is

$$\Theta' = \bar{\mathcal{S}}\;\Theta^s + \left(1-\bar{\mathcal{S}}\right)\Theta^j, \tag{A4}$$

and, solving $\Delta\Theta = \Theta' - \Theta = 0$, we obtain the equilibrium proportion of selfed individuals in the population, Θ^* :

$$\Theta^* = \frac{\alpha\,(1-\delta_j)\,(1-\mathcal{S}_o(e))}{\mathcal{S}_o(e)\,(1-\alpha)\,\delta_a + (1-\mathcal{S}_o(e))\,(1-\alpha\delta_j)}. \tag{A5}$$

Lifetime inbreeding depression. In our model, inbreeding depression only affects survival, in juveniles and in adults. Thus, the decrease in fitness selfed individuals suffer from in comparison with the outcrossed throughout their lives, and not only between two mating events (*i.e.* lifetime inbreeding depression) can be expressed in terms of life expectancies. The life expectancies for selfed ($\mathcal{L}_s$) and outcrossed ($\mathcal{L}_o$) individuals, respectively, are given by

$$\mathcal{L}_s = \frac{1-\delta_j}{1-\mathcal{S}_o(e)\,(1-\delta_a)} \quad \text{and,} \quad \mathcal{L}_o = \frac{1}{1-\mathcal{S}_o(e)}, \tag{A6}$$

and lifetime inbreeding depression can be expressed as

$$\bar{\delta} = 1 - \frac{\mathcal{L}_s}{\mathcal{L}_o} = 1 - \frac{(1-\mathcal{S}_o(e))\,(1-\delta_j)}{1-\mathcal{S}_o(e)\,(1-\delta_a)}. \tag{A7}$$

Equation (A7) is an increasing function of $\mathcal{S}_o(e)$, that is selection against selfed individuals grows stronger as longevity increases. This is due to the repeated effect of adult inbreeding depression on post-breeding survival.In the annual case ($\mathcal{S}_o(e) = 0$), Equation (A7) reduces to $\bar{\delta} = \delta_j$. Interestingly, Equation (A7) relates to Θ^* (Equation (A5)) in the following way :

$$\Theta^* = \frac{\alpha\left(1-\bar{\delta}\right)}{\alpha\left(1-\bar{\delta}\right) + 1 - \alpha}. \tag{A8}$$

Therefore, the equilibrium proportion of selfed individuals in the population decreases when generations overlap (that is, $\mathcal{S}_o(e)$) increases, due to stronger selection against

selfed individuals owing to the repeated effect of adult inbreeding depression on adult survival.

1.6.2 Recursions for the evolution of reproductive effort in a partially selfing population

In this section, we detail the mathematical derivations leading to the ESS reproductive effort expression we obtain. The selfing rate is fixed and reproductive effort is controlled by a single modifier locus. To describe the changes happening in the population between two timesteps, we use the theoretical framework first described by Barton and Turelli (1991) and generalized by Kirkpatrick et al. (2002). We follow the variation of three variables, namely, the frequency (p_m) of the mutant allele, which is assumed to be a rare with a small effect (ϵ) on its bearer's reproductive effort, the excess in homozygosity at the modifier locus ($D_{m,m}$), and the proportion of selfed individuals in the population (Θ).

We define the indicator variables X_m and $\overset{*}{X}_m$, corresponding to the two allelic positions of the modifier locus located on the paternally and maternally inherited chromosomes, respectively. These variables can take two values, 1 if the mutant allele (m) is present at the considered position, and 0 otherwise. Since there is no maternal or paternal effect on the expression of the alleles at the modifier, we have :

$$\mathbb{E}[X_m] = \mathbb{E}[\overset{*}{X}_m] = p_m. \tag{A9}$$

Using these indicator variables, we may define the centered variables ζ_m and $\overset{*}{\zeta}_m$ as :

$$\zeta_m = X_m - p_m \text{ and, } \overset{*}{\zeta}_m = \overset{*}{X}_m - p_m \text{ so that } \mathbb{E}\left[\zeta_m\right] = \mathbb{E}\left[\overset{*}{\zeta}_m\right] = 0. \tag{A10}$$

These variables allow us to define the excess in homozygotes in the population at the modifier as compared to the panmictic expectation ($D_{m,m}$) as :

$$D_{m,m} = \mathbb{E}[\zeta_{m,m}] = \mathbb{E}[X_m \overset{*}{X}_m] - p_m^2, \tag{A11}$$

where $\zeta_{m,m} = \zeta_m \times \overset{*}{\zeta}_m$.

Reproduction Using indicator variables, the reproductive effort of a given individual, e, can be expressed as follows

$$e = e_0 + \frac{\varepsilon}{2}\left(X_m + \overset{*}{X}_m\right) = e_0 + \varepsilon\, p_m + \frac{\varepsilon}{2}\left(\zeta_m + \overset{*}{\zeta}_m\right). \tag{A12}$$

The average reproductive effort is thus

$$\bar{e} = \mathbb{E}[e] = e_0 + \varepsilon\, p_m, \tag{A13}$$

which yields, to leading order in ε, the relative contribution of a given individual to reproduction ($\tilde{e}$) during a given mating event

$$\tilde{e} = \frac{e}{\bar{e}} \approx 1 + \hat{\varepsilon}_j\left(\zeta_m + \overset{*}{\zeta}_m\right), \text{ with } \hat{\varepsilon}_j = \frac{\varepsilon}{2e_0}. \tag{A14}$$

Frequency of the mutant. Using Equation (A14), the frequency of the mutant allele (p_m^j) among juveniles following reproduction is

$$p_m^j = \mathbb{E}\left[\tilde{e}\ \frac{X_m + \overset{*}{X}_m}{2}\right] = \frac{1}{2}\mathbb{E}\left[\tilde{e}\ \left(\zeta_m + \overset{*}{\zeta}_m + 2p_m\right)\right] = p_m + \hat{\varepsilon}_j\left(p_m q_m + D_{m,m}\right), \tag{A15}$$

with $q_m = 1 - p_m$, and the change in frequency of the mutant among juveniles is therefore

$$\Delta^j p_m = \hat{\varepsilon}_j\left(p_m q_m + D_{m,m}\right). \tag{A16}$$

Excess in homozygotes. The change in homozygosity at the modifier in juveniles can be divided in two phases, selection and syngamy. The excess in homozygotes

at the modifier following selection, $D^{j\bullet}_{m,m}$ is given by

$$D^{j\bullet}_{m,m} = \mathbb{E}\left[\left(X^j_m - p^j_m\right)\left(\overset{*}{X}{}^j_m - p^j_m\right)\right], \tag{A17}$$

where X^j_m and $\overset{*}{X}{}^j_m$ are indicator variables defined among juveniles following selection (so that $\mathbb{E}\left[X^j_m\right] = \mathbb{E}\left[\overset{*}{X}{}^j_m\right] = p^j_m$). Noting that $p^j_m = p_m + \Delta^j p_m$, and expressing Equation (A17) in terms of ζ-variables, we have

$$\begin{aligned} D^{j\bullet}_{m,m} &= \mathbb{E}\left[\tilde{e}\ \zeta_{m,m} - \Delta^j p_m\left(\tilde{e}\ \zeta_m\right) - \Delta^j p_m\left(\tilde{e}\ \overset{*}{\zeta}_m\right) + \left(\Delta^j p_m\right)^2\right] \\ &= \mathbb{E}\left[\tilde{e}\ \zeta_{m,m}\right] + \left(\Delta^j p_m\right)^2 \approx D_{m,m}\left(1 + 2\hat{\varepsilon}_j\right) \end{aligned} \tag{A18}$$

to leading order in ε and assuming p_m is small. During syngamy, homozygosity is generated by inbreeding. In our model, we assumed partial selfing. Therefore, the excess in homozygotes among juveniles, after accounting for juvenile inbreeding depression, is given by

$$D^j_{m,m} = \Theta^j\,\mathbb{E}\left[\frac{1}{4}{\zeta^j_m}^2 + \frac{1}{2}\zeta^j_{m,m} + \frac{1}{4}{\overset{*}{\zeta}{}^j_m}^2\right] = \frac{\Theta^j}{2}\left(D^{j\bullet}_{m,m} + p^j_m q^j_m\right) \approx \frac{\Theta^j}{2}\left(D_{m,m}\left(1 + 3\hat{\varepsilon}_j\right) + p_m\left(1 + \hat{\varepsilon}_j\right)\right), \tag{A19}$$

where $\zeta^j_m = X^j_m - p^j_m$ (resp. $\overset{*}{\zeta}{}^j_m = \overset{*}{X}{}^j_m - p^j_m$) is the centered variable defined for the paternally (resp. maternally) inherited chromosome among juveniles following selection, that is taking the frequency of the mutant among juveniles following selection (p^j_m) as reference value (Kirkpatrick et al., 2002).

Survival We denote by $\mathcal{S}$ the probability of survival of a parent taken at random in the parental population. Using our indicator variables, and given the proportion Θ of

selfed individuals in the population, the probability of survival is

$$\mathcal{S} = \Theta\ (1-\delta_a)\,\mathcal{S}_o(e) + (1-\Theta)\mathcal{S}_o(e) = S\left[1 - \left(e_0 + \frac{\varepsilon}{2}\left(X_m + \overset{*}{X}_m\right)\right)^x\right]\left(\Theta\,(1-\delta_a) + 1 - \Theta\right). \tag{A20}$$

Substituting indicator variables with ζ-variables and assuming that ϵ and p_m are small, we obtain, to leading order,

$$\mathcal{S} = S\,(1-\Theta\ \delta_a)\left[1 - {e_0}^x - \frac{x\ \varepsilon\ {e_0}^{x-1}}{2}\left(\zeta_m + \overset{*}{\zeta}_m\right)\right] + o\left(\varepsilon^2\right). \tag{A21}$$

Thus, the mean survival probability is

$$\bar{S} = \mathbb{E}\left[\mathcal{S}\right] \approx S\,(1-\Theta\ \delta_a)\left[1 - {e_0}^x\right], \tag{A22}$$

and the relative survival probability ($\tilde{\mathcal{S}}$) simplifies to

$$\tilde{\mathcal{S}} = \frac{\mathcal{S}}{\bar{\mathcal{S}}} \approx 1 - \hat{\varepsilon}_s\left(\zeta_m + \overset{*}{\zeta}_m\right), \text{ with } \hat{\varepsilon}_s = \frac{x\ \varepsilon\ {e_0}^{x-1}}{2\,(1-{e_0}^x)}. \tag{A23}$$

Frequency of the mutant. Using Equation (A23), the frequency of the mutant (p_m^s) among the surviving parents is

$$p_m^s = \mathbb{E}\left[\tilde{\mathcal{S}}\ \frac{X_m + \overset{*}{X}_m}{2}\right] \approx p_m - \hat{\varepsilon}_s\left(p_m q_m + D_{m,m}\right), \tag{A24}$$

and the variation in frequency of the mutant owing to selection among parents is therefore given by

$$\Delta^s p_m = -\hat{\varepsilon}_s\left(p_m q_m + D_{m,m}\right) \tag{A25}$$

Excess in homozygotes. Since only selection acts at this stage, the excess in homozygotes at the modifier locus among the parents ($D^s_{m,m}$) is given by

$$D^s_{m,m} = \mathbb{E}\left[(X^s_m - p^s_m)(\overset{*}{X}{}^s_m - p^s_m)\right]. \tag{A26}$$

Noting that $p^s_m = p_m + \Delta^s p_m$, and expanding Equation (A26), we have

$$\begin{aligned} D^s_{m,m} &= \mathbb{E}\left[\tilde{\mathcal{S}}\ \zeta_{m,m} - \Delta^s p_m \left(\tilde{\mathcal{S}}\ \zeta_m\right) - \Delta^s p_m \left(\tilde{\mathcal{S}}\ \overset{*}{\zeta}_m\right) + (\Delta^s p_m)^2\right] \\ &= \mathbb{E}\left[\tilde{\mathcal{S}}\ \zeta_{m,m}\right] + (\Delta^s p_m)^2 \approx D_{m,m}\left(1 - 2\hat{\varepsilon}_s\right). \end{aligned} \tag{A27}$$

Next timestep From our previously derived expressions, the frequency of the mutant (p_m'), the excess in homozygotes ($D'_{m,m}$) and the proportion of selfed individuals in the population (Θ') in the next timestep is

$$\begin{cases} p_m' = \bar{\mathcal{S}}\ p^s_m + \left(1 - \bar{\mathcal{S}}\right)\ p^j_m \\ D'_{m,m} = \bar{\mathcal{S}}\ D^s_{m,m} + \left(1 - \bar{\mathcal{S}}\right) D^j_{m,m} \\ \Theta' = \bar{\mathcal{S}}\ \Theta^s + \left(1 - \bar{\mathcal{S}}\right)\Theta^j. \end{cases}$$

In order to determine the *Evolutionary Stable Strategy* (ESS, Maynard Smith and Price, 1973) for the population given the parameters of the model, one has to determine the value of e_0 for which no mutant allele (m) may invade the population and replace the resident (M). To do so, we will consider the change in frequency of the mutant allele over one timestep (Δp_m), which can be expressed as

$$\Delta p_m \approx \left[\left(1 - \bar{\mathcal{S}}\right)\hat{\varepsilon}_j - \bar{\mathcal{S}}\ \hat{\varepsilon}_s\right]\left(p_m q_m + D_{m,m}\right) = \hat{\varepsilon}_j\left(p_m q_m + D_{m,m}\right)V, \tag{A28}$$

with $V = 1 - S\left(1 - \Theta\delta_a\right)\left(1 + (x-1)\, {e_0}^x\right)$.

Separation of timescales approximation When selection is weak (ε is small), we may assume that the excess in homozygotes in the population reaches a value close to equilibrium much faster than the allelic frequencies. It is obtained by solving

$$D'_{m,m} - D_{m,m} = 0$$

for $D_{m,m}$. Thus, assuming the mutant is rare (p_m is of order ε) we have

$$D^*_{m,m} = \frac{\Theta^j p_m}{2-\Theta^j}\left(1 - \frac{4\bar{\mathcal{S}}\hat{\varepsilon}_s - 3\Theta^j\left(1-\bar{\mathcal{S}}\right)\hat{\varepsilon}_j}{\left(1-\bar{\mathcal{S}}\right)\left(2-\Theta^j\right)}\right) \approx \frac{\Theta^j}{2-\Theta^j}p_m. \tag{A29}$$

Since we have $D^*_{m,m} \geqslant 0$ (Equation (A29)), the first term in Equation (A28) is always positive, and only V matters for the determination of the equilibrium.

Evolutionarily Stable Reproductive Effort Injecting Equation (A5) into Equation (A28), and solving $\Delta p_m = 0$ for e_0 yields the following ESS for the reproductive effort (e^*) :

$$e^* = \left[\frac{\mathcal{B} + \sqrt{\left(-\mathcal{B}\right)^2 + 4\left(1-S\right)\left(x-1\right)\left(1-\delta_a\right)\left[1-\delta_a\left(1-S\right)\right]\left(1-\alpha\delta_j\right)^2}}{2S\left(x-1\right)\left(1-\delta_a\right)\left(1-\alpha\delta_j\right)}\right]^{\frac{1}{x}} \tag{A30}$$

with $\mathcal{B} = \left(1-S\right)\left(2-x\right)\left(1-\alpha\delta_j\right) + \delta_a\left[\alpha x\left(1-\delta_j\right) - \left[1-S\left(2-x\right)\right]\left(1-\alpha\delta_j\right)\right]$.

1.6.3 Recursions for the evolution of self-fertilisation in a perennial population

In this section, we detail the mathematical derivations for the evolution of self-fertilisation in a perennial population. Individuals reproduce by self-fertilisation at a rate given by their genotype at a single biallelic modifier locus, and by random mating

otherwise. At the modifier, alleles M and m are assumed to be codominant, and code for selfing rates α_0 and $\alpha_0 + a$, respectively. Allele m is assumed to be a rare mutant with a weak effect (*i.e.* a is small). The theoretical framework we use is the same as in Appendix II.

Reproduction. Let us split the population into two groups : the *selfed* and the *outcrossed* individuals. The frequency of allele m in each of these groups are denoted $p_{m,s}$ and $p_{m,o}$, respectively, so that its frequency in the whole population is given by

$$p_m = \Theta \, p_{m,s} + (1 - \Theta) \, p_{m,o} \tag{A31}$$

where Θ is the proportion of selfed individuals in the population.

Frequency of the mutant among the selfed. The contribution of a given parent to the production of selfed offspring, α, is

$$\alpha = \alpha_0 + 2 \times \frac{a}{2} \left(X_m + \overset{*}{X}_m \right) = \alpha_0 + a \left(X_m + \overset{*}{X}_m \right), \tag{A32}$$

owing to the fact that increasing its selfing rate by $\frac{a}{2}$ induces a twofold transmission advantage in comparison with others (Fisher, 1941). Injecting ζ-variables into Equation (A32), we obtain the following expression for the relative contribution of a given parent to the production of offspring by self-fertilisation :

$$\tilde{\alpha} = \frac{\alpha}{\bar{\alpha}} = \frac{\alpha_0 + a p_m + a \left(\zeta_m + \overset{*}{\zeta}_m \right)}{\alpha_0 + a p_m} = 1 + \frac{a}{\alpha_0 + a p_m} \left(\zeta_m + \overset{*}{\zeta}_m \right). \tag{A33}$$

Thus, the frequency of allele m among the selfed offspring ($p^j_{m,s}$) is given by

$$p^j_{m,s} = \mathbb{E}\left[\tilde{\alpha}\ \frac{X_m + \overset{*}{X}_m}{2}\right] = p_m + \frac{a}{\alpha_0 + ap_m}\Big(p_m\left(1-p_m\right) + D_{m,m}\Big). \tag{A34}$$

Frequency of the mutant among the outcrossed. The contribution of a given parent to the production of outcrossed offspring, o, is

$$o = 1 - \alpha_0 - \frac{a}{2}\left(X_m + \overset{*}{X}_m\right), \tag{A35}$$

which yields the following expression for its relative contribution ($\widetilde{O}$) :

$$\tilde{o} = 1 - \frac{a}{2\left(1-\alpha_0 - ap_m\right)}\left(\zeta_m + \overset{*}{\zeta}_m\right). \tag{A36}$$

Hence, the frequency of allele m among the outcrossed offspring ($p^j_{m,o}$) is

$$p^j_{m,o} = \mathbb{E}\left[\tilde{o}\ \frac{X_m + \overset{*}{X}_m}{2}\right] = p_m - \frac{a}{2\left(1-\alpha_0 - ap_m\right)}\Big(p_m\left(1-p_m\right) + D_{m,m}\Big). \tag{A37}$$

Excess in homozygotes. Considering the whole offspring pool, we may neglect the effect of the mutant on the selfing rate. Thus, we may express the excess in homozygotes at the modifier among juveniles as

$$D^j_{m,m} \approx \frac{\Theta^j}{2}\Big(p_m\left(1-p_m\right) + D_{m,m}\Big), \tag{A38}$$

where Θ^j is the proportion of selfed individuals among the juveniles settling in the population (Equation (A3)).

Survival. Among the parents, allelic frequencies among the selfed and among the outcrossed do not vary, because there is no direct selection acting on the modifier

among them. Therefore,

$$p^s_{m,o} = p_{m,o} \quad \text{and} \quad p^s_{m,s} = p_{m,s}. \tag{A39}$$

Moreover, we have

$$D^s_{m,m} = D_{m,m}. \tag{A40}$$

Selfed individuals are however counterselected due to adult inbreeding depression. Thus, their proportion varies according to Equation (A2).

Next timestep. In the next timestep, assuming the proportion of selfed individuals remains close to its equilibrium value (Θ^*, Equation (A5)), the frequency of the mutant (p_m') and the excess in homozygotes ($D_{m,m}'$) are given by

$$\begin{cases} p_m' = \bar{\mathcal{S}}\ p^s_m + \left(1 - \bar{\mathcal{S}}\right)\ p^j_m \\ D_{m,m}' = \bar{\mathcal{S}}\ D^s_{m,m} + \left(1 - \bar{\mathcal{S}}\right) D^j_{m,m}, \end{cases} \tag{A41}$$

where $\bar{\mathcal{S}} = \mathcal{S}_o(e)\left[\Theta^*\ \ (1 - \delta_a) + 1 - \Theta^*\right]$ is the proportion of parents surviving until the next timestep. Using these recursions, we may express the change in allelic frequencies at the modifier (Δp_m) as

$$\Delta p_m = -\frac{a\left(1 - \alpha_0\delta_j\right) D_{m,m}\left(1 - \mathcal{S}_o(e)\right)}{2\left[\mathcal{S}_o(e)\left(1 - \alpha_0\right)\delta_a + \left(1 - \mathcal{S}_o(e)\right)\left(1 - \alpha_0\delta_j\right)\right]^2} \times T \tag{A42}$$

with $T = \left[1 - \mathcal{S}_o(e)\left(1 + \delta_a\right)\right]\left[\mathcal{S}_o(e)\delta_a - \left(1 - \mathcal{S}_o(e)\right)\left(1 - 2\delta_j\right)\right]$.

Separation of timescales approximation. Assuming homozygosity equilibrates quickly in comparison with allelic frequencies at the modifier, we obtain a quasi-equilibrium expression (Kirkpatrick et al., 2002) for the excess in homozygotes at the

modifier ($D^*_{m,m}$) by solving $D'_{m,m} - D_{m,m} = 0$,

$$D^*_{m,m} = p_m \left(1 - p_m\right) \frac{\alpha_0 \left(1 - \delta_j\right)}{2 - \alpha_0 \left(1 + \delta_j\right)} \geqslant 0. \tag{A43}$$

Using Equation (A43), one can see that the first term in Equation (A42) is always negative, and only T matters for the determination of the equilibria. Because T does not depend on α_0, there is only two situations : either $T > 0$ and full outcrossing is maintained ($\alpha^* = 0$), or $T < 0$ and full selfing is favoured ($\alpha^* = 1$). Hence, by solving $T < 0$, we obtain the following threshold for the evolution of self-fertilisation :

$$\delta_j < \frac{1 - \left(1 + \delta_a\right) \mathcal{S}_o(e)}{2 \left(1 - \mathcal{S}_o(e)\right)} \tag{A44}$$

1.6.4 Numerical analyses

In this section, numerical analyses results are presented along with our analytical predictions.

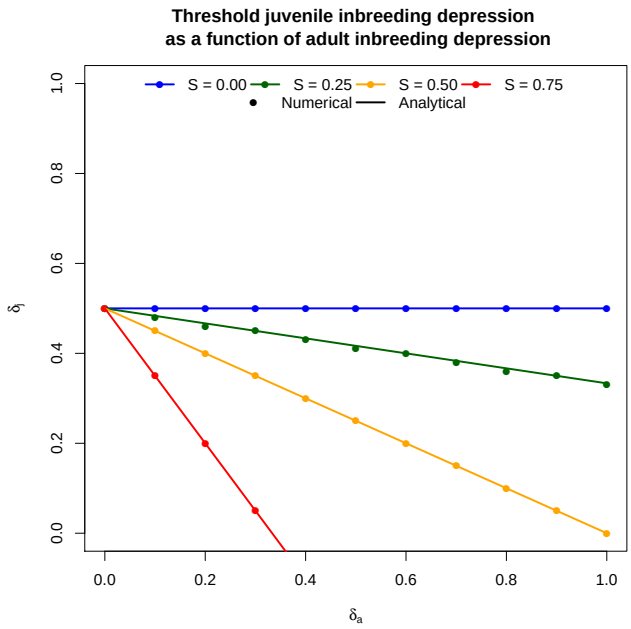

FIGURE A1 – Numerical analyses results for the evolution of self-fertilisation. Threshold juvenile ID is presented as a function of adult ID, for various survival probabilities $(\mathcal{S}_o(e) = 0\ ; 0.25\ ; 0.5\ ; 0.75)$.

For the evolution of self-fertilisation, the analysis was conducted starting from $\alpha_0 = 0$ and introducing mutants increasing the selfing rate. Good agreement was found between our analytical predictions and numerical analyses. Figure A1 shows some inbreeding depression threshold values obtained for various longevity situations. Either full selfing or strict outcrossing always evolved.

For the evolution of reproductive effort, two types of analyses were conducted : starting from $e_0 = 0$ and introducing mutants increasing reproductive effort (*upwards* analysis, presented in orange), and starting from $e_0 = 1$ and introducing mutants

decreasing reproductive effort (*downwards* analysis, presented in blue). Part of the results are presented in Figure A2. Discrepancies between upwards and downwards analyses were observed in situations involving high adult inbreeding depression and low extrinsic mortality for intermediate selfing rates (Fig. A2b). Such situations are never reached in the context of the joint evolution of lifespan and selfing, because inbreeding depression is too high for selfing to evolve and intermediate selfing rates are never evolutionarily stable. Besides, good agreement was found between analytical predictions and numerical analyses in most cases, especially for selfing rates close to 0 or 1 (Fig. A2), which are the states that matter for the joint evolution of lifespan and selfing. Therefore, we considered our analytical prediction to satisfactorily agree with numerical analyses.

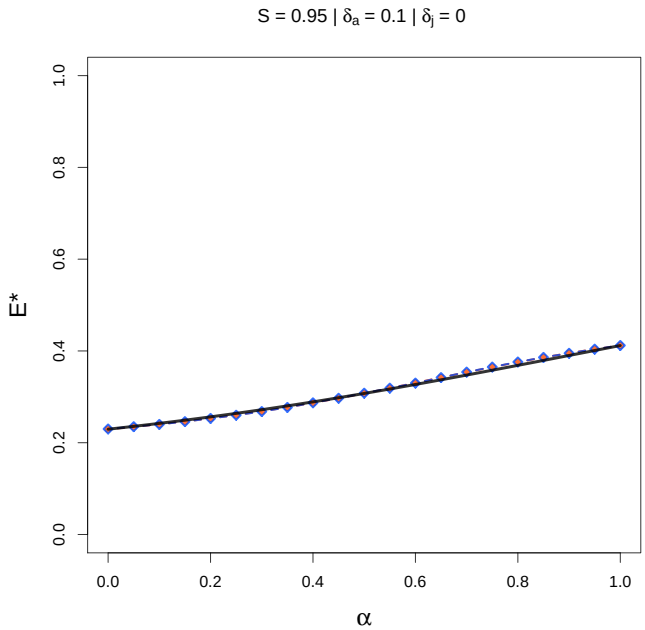

FIGURE A2 – Numerical analyses results for the evolution of reproductive effort. ESS reproductive effort is presented as a function of the selfing rate, in low ($S = 0.95$, Fig. (a) and (b)) and mild ($S = 0.75$, Figures (c) and (d)) extrinsic mortality conditions, with low ($\delta_a = 0.1$, Fig. (a) and (c)) and high ($\delta_a = 0.9$, Fig. (b) and (d)) adult inbreeding depression. Numerical analyses are presented in orange (*upwards*) and blue (*downwards*). Analytical predictions are presented in black.

1.6.5 Inbreeding depression affecting fecundity

In this appendix, we investigate the effect of inbreeding depression affecting fecundity, that is the contribution of selfed individuals to the gamete pool, on the coevolutionary dynamics. In Figure A3, we show the equilibria reached by simulations with different starting points, when setting $\delta_j = 0$ and $\delta_f = 0.1\ ;\ 0.4\ ;\ 0.7$ with $\delta_a = 0.4$ and $S = 0.8$. Solid lines represent the analytical expectations in the absence of inbreeding depression on fecundity ($\delta_f = 0$), and with $\delta_j = 0.1\ ;\ 0.4\ ;\ 0.7$.

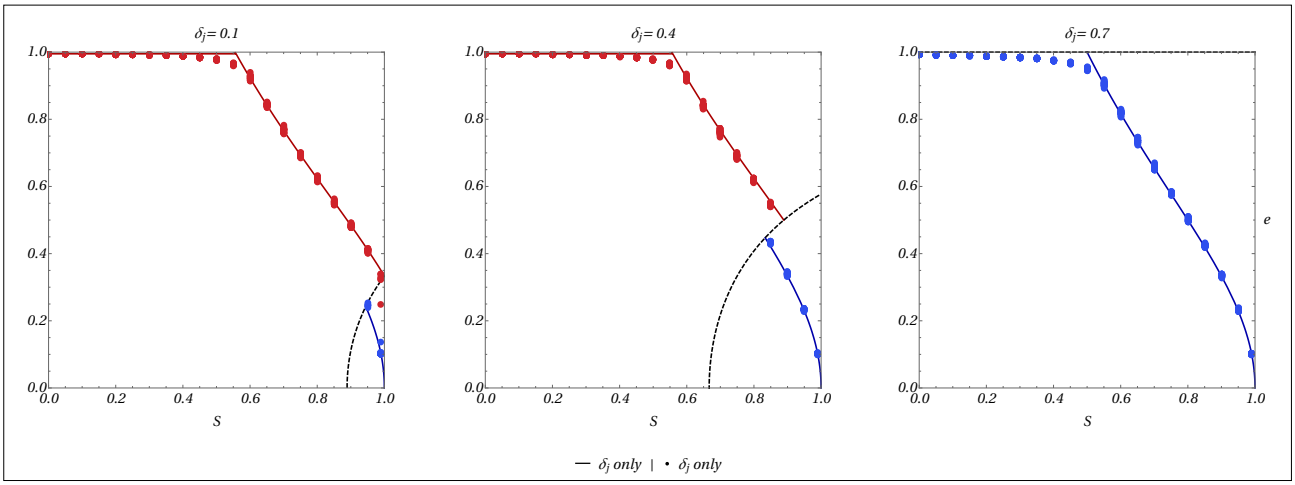

FIGURE A3 – All results presented here assume $\delta_a = 0.1$ and $x = 2$. Y-axis is the ESS reproductive effort (e^*), and X-axis is extrinsic mortality (S). The red line corresponds to the equilibrium reproductive effort when $\alpha = 1$ is reached. The blue line corresponds to the equilibrium reproductive effort when $\alpha = 0$ is reached. The dashed black line represents the threshold lifetime inbreeding depression under which outcrossing can be maintained. Dots depict the results of individual-centered simulations. When red dots mean $\alpha = 1$ is reached, while blue dots mean $\alpha = 0$ is reached.

Simulations results are well predicted by these analytical expectations, which shows that increasing δ_f or δ_j has the same effect : it only reduces the range of conditions under which the full selfing strategy can evolve. The mechanisms underlying these

similar behaviours are however slightly different. Indeed, while in both cases they are caused by a reduction of the participation of selfed individuals to reproduction, δ_j decreases the proportion of selfed individuals in the population. On the other hand, δ_f reduces the contribution of selfed adults to reproduction by reducing their contribution to the gamete pool.

Chapitre 2

On deleterious mutations in perennials : inbreeding depression, mutation load and life-history evolution

Auteurs : Thomas Lesaffre & Sylvain Billiard

Statut de publication : Le travail présenté dans ce chapitre est disponible sur *BioRXiv* (lien) et a récemment été accepté dans *The American Naturalist*.

2.1 Synthèse en français

2.1.1 Motivation

Dans le chapitre précédent, nous avons montré l'importance clef de la dépression de consanguinité pour l'évolution conjointe de la longévité et de l'autofécondation en supposant que la dépression de consanguinité était fixée par un paramètre. La magni-

tude de la dépression de consanguinité est influencée par le taux d'autofécondation, comme nous l'avons vu dans le deuxième chapitre d'introduction, mais également par l'histoire de vie des espèces, car elle est typiquement bien plus élevée chez les espèces pérennes que chez les annuelles parmi les Angiospermes (Duminil et al., 2009; Angeloni et al., 2011). Ainsi, les mécanismes génétiques qui la sous-tendent pourraient interagir avec la coévolution du régime de reproduction et de l'histoire de vie d'une manière qui reste pour l'heure inexplorée. En effet, si l'interaction entre régime de reproduction et dépression de consanguinité a fait l'objet d'une quantité importante de travaux théoriques et empiriques, les mécanismes responsables de l'augmentation de la dépression de consanguinité avec la pérennialité sont quant à eux mal compris. La dépression de consanguinité étant principalement générée par le maintien de mutations délétères récessives à l'équilibre mutation-sélection (Charlesworth and Willis, 2009), cette augmentation pourrait être causée par des différences soit dans la manière dont la sélection agit sur ces mutations, soit dans la façon dont ces mutations apparaissent dans la population. La lignée germinale n'étant pas séparée de la lignée somatique chez les plantes, Scofield and Schultz (2006) ont proposé que les espèces pérennes pourraient accumuler davantage de mutations somatiques au cours de leur croissance du fait de leur plus grande taille et longévité et les transmettre ensuite à leur descendance, générant ainsi une plus grande dépression de consanguinité chez ces espèces. Cette hypothèse fait l'objet du troisième chapitre. Dans le présent chapitre, nous explorons l'hypothèse que l'augmentation de la dépression de consanguinité observée chez les espèces plus longévives pourrait être expliquées par des différences de pression de sélection agissant sur les mutations.

2.1.2 Modèle

À cet effet, un modèle multilocus de génétique des populations est considéré. Dans ce modèle, les individus, supposés hermaphrodites et diploïdes, se reproduisent par-

tiellement par autofécondation et contribuent à la reproduction proportionnellement à leur taille. Ils survivent d'une saison de floraison à une autre avec une probabilité constante au cours de leur existence, et grandissent si ils survivent suivant un modèle physiologique de croissance à deux paramètres proposé par West et al. (2001), qui peut être résumé comme suit. La quantité de ressources qu'un individu est capable de prélever dans son environnement varie comme la puissance $\frac{3}{4}$ de sa taille en réponse aux contraintes allométriques dont nous discutions dans le premier chapitre d'introduction. Ces ressources sont ensuite consommées pour maintenir les unités physiologiques existantes, et les ressources restantes sont utilisées pour en produire de nouvelles. Les ressources nécessaires au maintien d'une unité sont contrôlées par un paramètre appelé coût de maintien qui ne varie pas, de sorte que la quantité totale de ressources nécessaires au maintien de l'organisme augmente linéairement avec sa taille. Les ressources nécessaires à la production d'une nouvelle unité sont quant à elles contrôlées par un coût de production qui ne varie pas non plus. Ainsi, la croissance d'un individu sature naturellement lorsque la quantité de ressources nécessaire au maintien de l'organisme existant équivaut aux ressources qu'il est capable de prélever.

Dans cette population, que l'on suppose de grande taille constante de sorte que le recrutement de juvéniles ne s'opère qu'en remplacement des adultes décédés au pas de temps précédent, la dynamique d'une infinité de loci auxquels ségrégent deux allèles, un allèle sain et un allèle délétère récessif, est considérée. À chaque locus, l'allèle sain mute à un certain taux pendant la méiose vers l'allèle délétère (pas de mutation somatique ni de mutation reverse). Plutôt que de diminuer directement la fitness d'une proportion constante, comme dans la plupart des modèles de fardeau génétique (e.g. Charlesworth et al., 1990; Roze, 2015), cet allèle affecte un phénotype qui affecte à son tour la fitness de l'individu, ce qui permet de décrire une carte génotype-phénotype-fitness complète. Trois types de mutations sont étudiés dans trois modèles séparés. Des mutations qui diminuent les chances de survie de leur porteur, et des mutations qui altèrent sa

croissance en augmentant soit le coût de maintien soit, le coût de production qui sont les paramètres du modèle physiologique de West et al. (2001). Une approche analytique inspirée des travaux de Roze (2015) prenant en compte les effets d'interférence sélective entre paires de loci est développée, et les approximations obtenues pour les propriétés de la population en termes de fardeau génétique et de dépression de consanguinité à l'équilibre mutation-sélection sont comparées aux résultats de simulations individu-centrées.

2.1.3 Résultats clefs

2.1.3.1 Effets sélectifs des mutations, nombre de mutations maintenues à l'équilibre et conséquences pour l'évolution des histoires de vie

Effets sélectifs des mutations. Les trois types de mutations ont des effets sélectifs différents, qui varient avec la longévité des individus. Les mutations qui réduisent la survie des individus affectent la fitness d'autant plus fort que la longévité initiale des individus (c'est-à-dire en l'absence de mutations) est grande, et atteignent rapidement des effets sélectifs élevés. Les mutations affectant la croissance conservent quant à elles des effets sélectifs plus modérés mais qualitativement très différents suivant le coût physiologique qu'elles affectent. Les mutations qui augmentent le coût de maintien créent un retard de croissance qui s'amplifie avec l'âge des individus et prend sa valeur la plus grande lorsque les individus atteignent leur taille maximale. Leur effet sur la fitness est donc le plus fort dans les classes d'âge les plus élevées et tend à augmenter avec la longévité. Les mutations augmentant le coût de production n'ont, elles, pas d'effet sur la taille maximale atteinte par les individus mais altèrent la vitesse à laquelle ils atteignent cette dernière. Ainsi, ces mutations diminuent la fécondité des premières classes d'âge et deviennent graduellement neutres dans les classes d'âge plus élevées, de sorte que leur effet sur la fitness a tendance à diminuer avec la longévité.

Mutations maintenues à l'équilibre. Ces différences d'intensité de la sélection conduisent au maintien de quantités différentes de mutations à l'équilibre mutation-sélection suivant le type de mutation et la longévité considérée. Les mutations affectant le coût de maintien sont maintenues en nombres moins importants lorsque la longévité augmente car elles sont plus efficacement contre-sélectionnées, contrairement aux mutations affectant le coût de production qui ont un effet sur la fitness de moins en moins élevé et sont donc maintenues plus nombreuses chez les espèces plus longévives. Quant aux mutations affectant la survie, le nombre de mutations maintenues décroît drastiquement avec la longévité car l'intensité de la sélection augmente fortement avec cette dernière.

Conséquences pour l'évolution des histoires de vie. Les différences dans le nombre de mutations maintenues en fonction de la longévité lorsque ces dernières affectent les coûts physiologiques de croissance résulte en une réduction du coût de maintien et une augmentation du coût de production à l'équilibre mutation-sélection dans chez les espèces à plus longue durée de vie. Ainsi, les espèces plus longévives croissent plus lentement mais atteignent des tailles plus élevées dans notre modèle, non pas du fait d'une évolution conjointe de la croissance et de la longévité, mais simplement parce que la variation de la longévité induit variation des pressions de sélection pesant sur le fardeau génétique affectant la croissance. Ce résultat suggère que les traits d'histoire de vie pourraient parfois évoluer conjointement indépendamment d'éventuels trade-offs, car un changement dans un trait donné pourrait altérer l'efficacité de la purge de mutations délétères affectant d'autres traits.

2.1.3.2 Dépression de consanguinité

Mutations affectant la croissance. Dans le cas des mutations affectant la croissance, les différences de nombre de mutations maintenues en fonction de la longévité se

traduisent en des différences de dépression de consanguinité quantitativement faibles. La faiblesse de ces différences peut être expliquée par leurs effets modérés sur la fitness, car la magnitude de la dépression de consanguinité est indépendante de l'intensité de la sélection lorsque cette dernière est faible (Bataillon and Kirkpatrick, 2000). Des différences qualitatives sont cependant notables. En effet, la magnitude de la dépression de consanguinité à l'équilibre varie dans un sens différent avec la longévité suivant le taux d'autofécondation considéré. Lorsque celui-ci est faible, la dépression de consanguinité générée par les mutations sur le coût de maintien diminue avec la longévité et celle générée par les mutations affectant le coût de production augmente. Le patron inverse est observé pour des taux d'autofécondation élevés.

Mutations affectant la survie. En ce qui concerne les mutations affectant la survie, des différences de dépression de consanguinité plus marquées sont observées en fonction de l'espérance de vie, du fait du plus fort effet sélectif des mutations. De manière importante, nous observons que les mutations affectant la survie ont des effets sur la fitness si forts qu'elles sont plus efficacement purgées de la population et résultent en une diminution de la dépression de consanguinité maintenue à l'équilibre en fonction de la longévité. Ce résultat vient contredire l'idée que nous développions au premier chapitre, selon laquelle la dépression de consanguinité agissant sur la survie pourrait favoriser le maintien de l'allofécondation chez les espèces pérennes. En effet, si la sélection contre les mutations et donc particulièrement contre les individus issus d'autofécondation augmente fortement avec la longévité, cela résulte en une purge de la dépression de consanguinité et non en son augmentation exponentielle comme le suggéraient des modèles basés sur dépression de consanguinité constante.

Perspectives. À travers les trois types de mutations étudiés dans ce chapitre, nous avons vu que les différences de pression de sélection affectant les mutations ont des effets soit quantitativement faibles soit contraires aux observations empiriques, qui

prédisent une augmentation marquée de la dépression de consanguinité chez les espèces plus longévives. À première vue, nos résultats représentent donc un argument en faveur de l'hypothèse d'accumulation de mutations somatiques dont nous discutions plus haut. Cependant, ils suggèrent que des variations suffisamment fortes de l'intensité de la sélection peuvent générer des différences significatives de dépression de consanguinité à l'équilibre mutation-sélection, comme nous l'avons vu dans le cas de mutations affectant la survie. De plus, le seul cas dans lequel nous obtenons une augmentation de la dépression de consanguinité avec la longévité est celui des mutations affectant le coût de production, c'est-à-dire dont l'effet sélectif décroît dans les espèces plus longévives. Ainsi, des mutations suffisamment délétères dont l'effet diminue avec la longévité pourraient en principe générer l'augmentation de dépression de consanguinité empiriquement observée, indépendamment de la manière dont elles sont produites. Nous pouvons donc nous demander sur quel(s) trait(s) une sélection forte et décroissante avec la longévité devrait être observée.

2.2 Introduction

Perennials, which make up the majority of Angiosperms ($\sim 70\%$, Munoz et al., 2016), typically present much higher levels of inbreeding depression than annuals. Indeed, meta-analyses found inbreeding depression to span from $\delta \approx 0.2$ on average in short-lived herbaceous species to $\delta \approx 0.5$ in long-lived herbaceous species and shrubs, and $\delta \approx 0.6$ in woody species (Duminil et al., 2009; Angeloni et al., 2011). Inbreeding depression, defined as the reduction in fitness of inbred relative to outbred individuals, is thought to be mainly due to the increased homozygosity of inbred individuals for recessive deleterious mutations segregating at low frequencies in populations (Charlesworth and Charlesworth, 1987; Charlesworth and Willis, 2009). Why inbreeding depression is higher in more long-lived species is still poorly understood, despite the potential signi-

ficance of this pattern for important evolutionary questions, such as mating systems or dispersal rates evolution (Barrett and Harder, 1996; Roze and Rousset, 2005; Epinat and Lenormand, 2009; Duputié and Massol, 2013), and for more applied issues, as many cultivated species are perennial (e.g. fruit trees in general) and efforts are being made to develop perennial grain crops (DeHaan and Van Tassel, 2014).

Two hypotheses have been proposed to explain this pattern. The first hypothesis was formally put forward by Scofield and Schultz (2006). In plants, mutations occurring in somatic tissues may be passed onto the offspring, because they do not have a segregated germline. Thus, Scofield and Schultz (2006) proposed that more long-lived species may accumulate more somatic mutations as they grow and transmit them to their offspring, thereby generating the increase in inbreeding depression observed in such species. Phenotypic data in a long-lived clonal shrub (*Vaccinium angustifolium*, Bobiwash et al., 2013), and genomic results in *Quercus robur* (Plomion et al., 2018) demonstrated that some somatic mutations can indeed be passed onto the offspring. However, the number of detected heritable somatic mutations is low (Schmid-Siegert et al., 2017; Plomion et al., 2018), and recent studies have concluded that the number of cell divisions from embryonic cells to gametes production may be much lower than previously thought (Burian et al., 2016; Watson et al., 2016; Burian et al., 2016; Schmid-Siegert et al., 2017; Lanfear, 2018), due for instance to early specification and quiescence mechanisms of axillary meristems cells, resulting in little opportunity for heritable mutations to accumulate during plant growth. Furthermore, intraorganismal selection is expected to efficiently purge deleterious somatic mutations, resulting in little to no somatically generated mutation load at the population level (Otto and Orive, 1995). Hence, although somatic mutations can be inherited and contribute to the mutation load in plants, their relative significance compared with meiotic mutations remains unclear (Schoen and Schultz, 2019).

The second hypothesis stems from the observation that inbreeding depression is

typically expressed across multiple life stages in Angiosperms (Husband and Schemske, 1996; Winn et al., 2011; Angeloni et al., 2011). It posits that increased inbreeding depression in more long-lived species could be explained by the fact that mutations, regardless of whether they are produced during mitosis or meiosis, may differ in the way they affect fitness in annual and perennial populations, through the life stages at which they are expressed. Most theoretical studies of the mutation load focused on the case of mutations affecting fitness on a strictly linear fitness landscape (that is, fitness is the trait, e.g. Charlesworth et al., 1990; Roze, 2015), or through an abstract trait (or set of traits) under stabilising selection on a gaussian fitness landscape (e.g. Roze and Blanckaert, 2014; Abu Awad and Roze, 2018). In these cases, inbreeding depression in annual and perennial populations is not expected to differ (Charlesworth, 1980). On the other hand, the dynamics of mutations affecting other aspects of individuals' life cycle, such as survival or growth, were seldom investigated. Morgan (2001) investigated the dynamics of mutations affecting survival between mating events in a perennial populations. They concluded that inbreeding depression should sharply decrease as life expectancy increases, and even become negative for long-lived species (outbreeding depression). However, Morgan (2001) studied mutations with a strong effect on fitness, and assumed no age-structure, that is, individuals did not differ in fecundity or survival probability with age. Yet, such variations with respect to age in both survival and fecundity are observed in perennials. Indeed, while juveniles typically suffer from very high mortality rates, established individuals tend to experience rather low mortalities with particularly slow senescence (Petit and Hampe, 2006). Furthermore, fecundity usually increases dramatically with age in perennials (Franco and Silvertown, 1996), due to the positive scaling of reproductive output with size in plants (Klinkhamer et al., 1985; Weiner et al., 2009). Mutations slowing their bearer's growth could therefore play a role in generating higher inbreeding depression in more long-lived species, as growth delays impact individuals' fecundities negatively and with varying intensity with

respect to age or size. This latter aspect of age-structuration in perennials was, to our knowledge, never tackled theoretically.

The present study aims to study the second hypothesis, that is, we investigate the behaviour of meiotic mutations affecting fitness differently with respect to life-history, putting aside somatic mutations accumulation. Namely, we study meiotic mutations affecting growth or survival in a partially selfing population, in which individuals grow as they age and fecundity is proportional to size, but survival between mating events is assumed to not depend on age. We combine a physiological growth model (West et al., 2001) and multilocus population genetics approaches (Barton and Turelli, 1991; Kirkpatrick et al., 2002) in order to describe a full genotype-to-phenotype-to-fitness map, where the fitness landscape emerges from developmental processes instead of being assumed *a priori*. We study the behaviour of different types of mutations affecting growth or survival, and explore their consequences in terms of inbreeding depression and mutation load (Crow, 1958). Then, we show the role deleterious mutations maintained at mutation-selection balance may play in the joint evolution of growth and survival strategies.

2.3 Model outline and methods

Demography. We consider a large population of diploid hermaphrodites, which may survive from one mating season to another with a probability S, assumed to be constant with respect to age and size. If they survive, individuals grow between mating events, following a physiological growth model described briefly in the next paragraph. If they die, juveniles, which are assumed to be produced in large excess compared with the resources available for establishment, are recruited to replace the dead, so that population size is kept constant (Fig. 2.1). We assume that juveniles are produced each year and do not survive to the next mating season if they are not recruited (no dormancy).

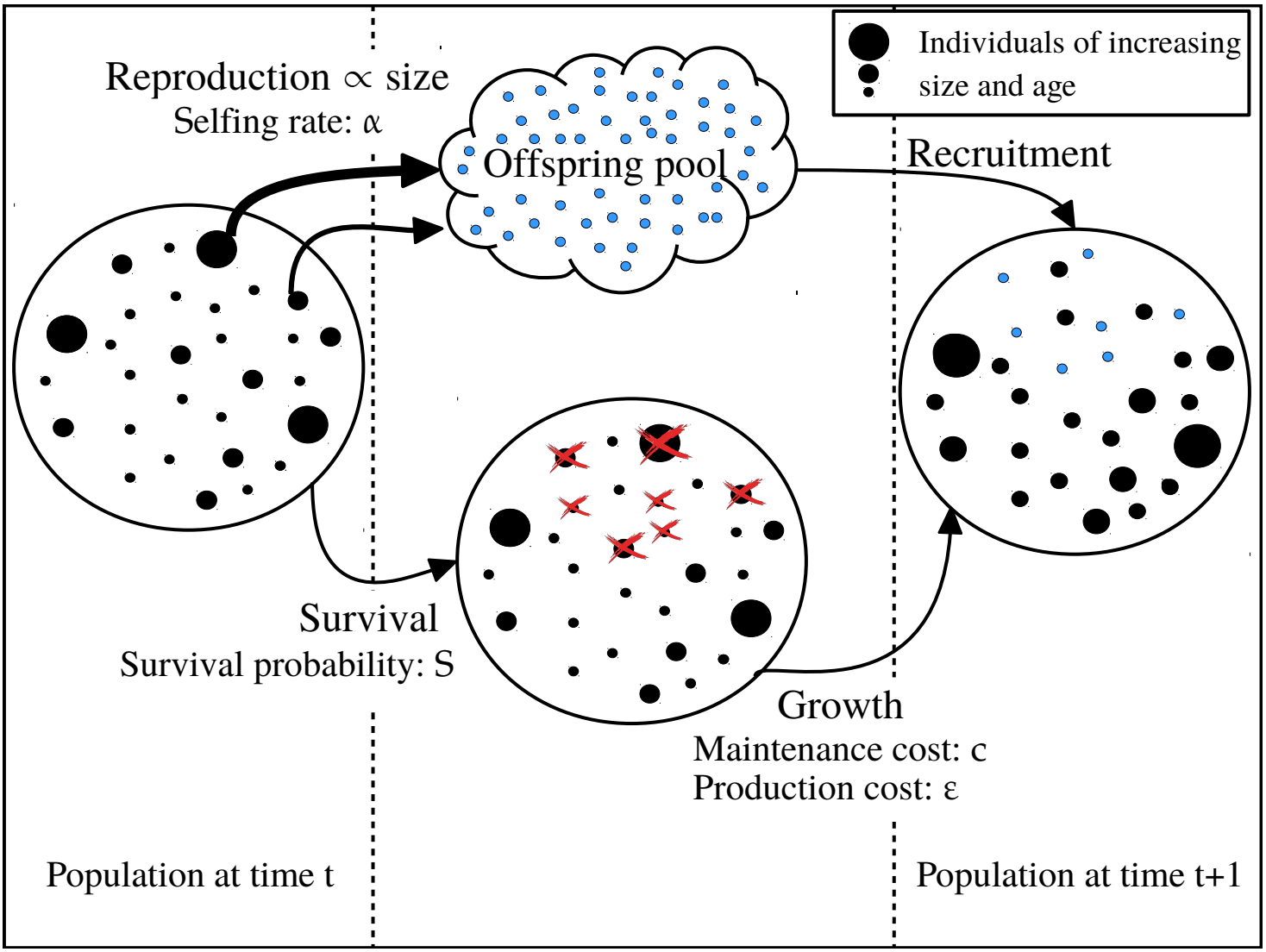

FIGURE 2.1 – Demographic events over the course of one timestep. Deceased individuals are marked by a red cross, juveniles are depicted in light blue. Larger dots depict larger and older individuals.

Each juvenile has a probability J of being recruited. During reproduction, individuals are assumed to contribute to the gamete pool in proportion to their size (the larger an individual, the larger its contribution to the gamete pool), and to self-fertilise at a fixed rate α.

Growth model. We consider the growth model developed by West et al. (2001). The energy available for growth and maintenance at age t, B_t, is assumed to scale as a $^3/_4$-power law of body size as a result of allometry (Peters, 1983), so that

$$B_t = B_0 \, G_t^{3/4}, \tag{2.1}$$

where B_0 is the basal metabolic rate and G_t is body size at age t. This energy can be subdivided into the energy required to maintain the existing body, controlled by a maintenance cost c, and the energy available to produce new body parts, controlled by a production cost ε, so that growth is fully described by the following differential equation

$$\frac{dG_t}{dt} = \frac{B_0}{\varepsilon} G_t^{3/4} - \frac{c}{\varepsilon} \, G_t. \tag{2.2}$$

Under this model, individual size naturally saturates when the energy required to maintain the existing body equals the available energy (Fig. 2.2a-2.2b). Further details regarding the growth model are given in Appendix 2.6.1.

Genetic assumptions. Mutations are assumed to occur at rate U (per haploid genome) at a large number of loci, which recombine at constant rate $0 \leqslant r \leqslant \frac{1}{2}$. In three separate models, we consider mutations affecting three different traits, that is we consider mutations affecting one trait at a time (no pleiotropy). Mutations may affect growth by increasing either their bearer's maintenance cost (c) or production cost (ε), or they may affect its survival. When mutations affect survival, they are assumed to decrease both their bearer's probability of being recruited as a juvenile

(J) and its adult survival probability (S). The effect of mutations is denoted s, with a dominance coefficient $0 \leqslant h \leqslant \frac{1}{2}$. Loci affect traits multiplicatively, so that for any trait z ($z \in \{c, \varepsilon, S\}$), we have

$$z = z_0 \left(1 \pm s\right)^{H_o} \left(1 \pm sh\right)^{H_e}, \tag{2.3}$$

where H_o (resp. H_e) is the number of homozygous (resp. heterozygous) mutations born by the considered individual. The $\pm$ sign is used as a general notation to indicate the fact that depending on the trait, deleterious mutations may increase it, as is the case for the maintenance and the production cost ; or decrease it, as is the case for survival. Note that s is the phenotypic effect of mutations here, not to be confused with $\bar{s}_z$ and $\hat{\bar{s}}_z$, the selection coefficients associated with mutations which will be defined in the following.

Approximation of the expected number of mutations, inbreeding depression and mutation load. Our analytical model is built using the framework introduced by Barton and Turelli (1991) and generalized in Kirkpatrick et al. (2002). In this framework, the genetic dynamic of the population is described using allelic frequencies and genetic associations instead of genotypic frequencies. To do so, we define X_i and $\overset{*}{X}_i$ as the indicator variables associated with the paternally and maternally inherited allele at the ith locus, respectively. These variables are worth 1 if the deleterious allele is present at the considered position, and 0 otherwise. Thus, Equation (2.3), which depicts the phenotypic effect of mutations, can be rewritten in terms of these variables as

$$z = z_0 \times \prod_i \left(1 \pm sh\left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)X_i\overset{*}{X}_i\right). \tag{2.4}$$

Indicator variables can in turn be used to define centered variables ζ_i and $\overset{*}{\zeta}_i$, which are given by

$$\zeta_i = X_i - \mathbb{E}[X_i] = X_i - p_i, \text{ and } \overset{*}{\zeta}_i = \overset{*}{X}_i - \mathbb{E}[\overset{*}{X}_i] = \overset{*}{X}_i - p_i. \tag{2.5}$$

By definition, we have $\mathbb{E}[\zeta_i] = \mathbb{E}[\overset{*}{\zeta}_i] = 0$, and expectations of products of these variables can be used to quantify how the population deviates from Hardy-Weinberg's equilibrium, that is genetic associations in the sense of Barton and Turelli (1991), such as excesses in homozygotes, or linkage and identity disequilibria between loci for instance (see Appendix 2.6.2.1 for details).

In order to obtain an analytical prediction of the average number of mutations per haploid genome at mutation-selection equilibrium, one needs to find the point at which the rate at which mutations enter the population through mutation during meiosis equals the rate at which mutations are removed from the population by selection. This can be done by following the change in allelic frequencies at selected loci between two timesteps, Δp_i, and looking for the point at which we have $\Delta p_i = 0$. We do so using two different approaches. In the first approach, we make the assumption that selective pressures acting on mutations at various life-stages can be summarised into a single lifetime fitness expression, so that the population can be studied as an adequately rescaled annual population (the Lifetime Fitness approach, LF - Appendix 2.6.3.1 and FIG. S1). For each type of mutation, this approach allows us to gain insights into the way selection acts on mutations, by summarising it into a single lifetime selection coefficient, $\bar{s}_z$, where $z = c$, ε or S depending on the considered trait (not to be confused with s, the phenotypic effect of mutation. See Appendix **??** for details). Besides, reasoning in terms of lifetime fitness is paramount to compute key quantities such as inbreeding depression or the mutation load. However, the LF approach fails to account for genetic associations correctly because it assumes that all selection occurs

before reproduction at each timestep (see FIG. S1), so that it overlooks the fact that selection may occur both in parents and juveniles, that is before and after reproduction, respectively, which induces different effects of selection on genetic associations in these two subpopulations. Thus, in order to obtain approximations of the expected number of mutations per haploid genome accounting for genetic associations, we also use a second approach where we study each step of the life cycle successively (the Life Cycle approach, LC - Appendix 2.6.3.2 and FIG. S2). We assume that the phenotypic effect of mutations is weak and that the number of segregating mutations is large, following the work of Roze (2015) that we adapted to the case of many mutations affecting a trait rather than fitness directly, in an age- and size-structured population.

Simulations. In order to assess the validity of our analytical approximations, individual-centered simulations were run assuming diploid individuals depicted by two linear chromosomes of length λ (in cM), along which mutations occur stochastically during meiosis (Roze and Michod, 2010). The number of mutations occurring is sampled from a Poisson distribution with mean U, and their position on chromosomes are sampled from a uniform distribution. Recombination is modeled by exchanging segments between chromosomes. Similar to mutations, the number of crossing-overs is sampled from a Poisson distribution with mean λ, while their positions along chromosomes are randomly drawn in a uniform distribution. The population has a constant size N. At each timestep, an individual survives with probability S, which can depend on its genotype in the case of mutations affecting survival. If the individual does survive, it grows deterministically depending on its age and individual physiological growth costs (FIG. 2.2, Equation A5). If it does not, it is replaced by an offspring generated from the parental population, which includes the dead individual, and in which parents are chosen with a probability proportional to their size. The offspring is produced by self-fertilisation with probability α, and by random mating otherwise. We measure the average of the

trait affected by mutations, the average number of mutations per haploid genome in the population, and inbreeding depression. Inbreeding depression is measured as the relative difference in lifetime reproductive success of selfed and outcrossed individuals. Individuals lifetime reproductive success is obtained by counting the number of times they are chosen as parents before they die. The mutation load in the sense of Crow (1958), that is the decrease in mean fitness of the population compared with a population with no mutations, is measured using the population average of the trait affected by mutations. This average is used to compute the expected mean lifetime fitness in the population (Equation A35). Then, this quantity is compared to the expected mean lifetime fitness when mutations are absent. Five replicates were run for each parameter set showing a very small standard deviation, indicating that our simulations are representative of the average behaviour. All programs are available from GitHub (links are available in the Data and Code Accessibility Statement section).

2.4 Results

In what follows, results will be presented as functions of the life expectancy of the population, E, that is the average number of reproductive cycles an individual is expected to live through before dying. Since survival is not assumed to vary with respect to age nor size, the probability of an average individual to survive up to exactly age k then die follows a geometric distribution of parameter $\bar{S}$, where $\bar{S}$ is the mean survival probability in the population, which is simply $\bar{S} = S$ when mutations affect growth, and depends on mutations segregating in the population when said mutations affect survival. Thus, we have

$$E = \frac{1}{1 - \bar{S}}. \tag{2.6}$$

2.4.1 Mutation-selection equilibrium

To study the change in allelic frequencies at selected loci, one needs to quantify the intensity of selection acting on mutations. Generally speaking, the intensity of selection depends on how mutations modify the fitness of their bearer, and how the fitness of their bearer then compares with the fitnesses of all the individuals they are competing with, that is the distribution of fitness in the population. Hence, the intensity of selection faced by a mutation segregating at a given locus may depend on the presence of mutations at other loci, since said mutations may alter both the distribution of fitness in the population and how a mutation at the considered locus affects its bearer's fitness. This second aspect is referred to as epistasis. There are numerous ways by which epistasis may be generated between loci (Phillips, 2008), which can roughly be split into two categories : genotype-to-phenotype epistasis on the one hand, and phenotype-to-fitness epistasis on the other hand. Genotype-to-phenotype epistasis refers to loci which influence each other's expression in the determination of a phenotype, say genes involved in a common metabolic pathway, while phenotype-to-fitness epistasis refers to situations where loci, even if they influence phenotype independently, influence each other because the way the phenotype affects fitness is non-linear.

In our model, we assume that mutations affect a phenotype multiplicatively, so that genotype-to-phenotype epistasis does not occur. On the other hand, we do not impose any constraint on the shape of the fitness landscape. Instead, we let the genotype-to-phenotype-to-fitness map arise from biological assumptions. The resulting fitness landscape is non-linear, convex, and has singularities at the optima (Equation A35 in Appendix 2.6.3.1, Fig. S3). Thus, phenotype-to-fitness epistatic interactions are generated, and have to be accounted for to obtain accurate mutation-selection balance approximations, so that quantifying the intensity of selection acting on a mutation at a given locus requires accounting for the average number of mutations it may be in

epistatic interaction with. This can be done by following the change in the trait average as mutations segregate (Equation 2.7, see Appendices **??** and **??** for the derivation of this result under the Lifetime Fitness and the Life Cycle approach, respectively). We find this trait average at mutation-selection balance, when the change in allele frequency is zero.

$$\begin{cases} \Delta p_i\left(\bar{z}\right) = 0 \\ \bar{z}\left(p_i\right) = z_0 \times \exp\left(\sum_i \overline{Z}_i\right) \times \left(1 + \sum_{j \neq i} \frac{\overline{\Delta Z_i \times \Delta Z_j}}{2}\right). \end{cases} \tag{2.7}$$

In the first line of Equation (2.7), Δp_i is the change in the frequency of the mutant allele at the ith locus, which depends on the trait average $\bar{z}$. On the second line, $\bar{z}$ is the average of the trait in the population, which depends on the number of mutations per haploid genome. This number of mutations is obtained by summing allelic frequencies over every relevant loci, that is $n = \sum_i p_i$ (Kimura and Maruyama, 1966). On this second line, the $\pm$ sign indicates the fact that deleterious mutations may increase the average trait of the population, as it is the case for the maintenance and the production cost, or decrease it, as it is the case for survival. The term $\sum_i \overline{Z}_i$ quantifies the effect on the trait of the n mutations per haploid genome borne on average by individuals, neglecting the effects of genetic associations between loci, while the term $\overline{\Delta Z_i \times \Delta Z_j}$ quantifies the effect of pairwise associations between loci on the trait average. Associations of higher order are neglected. Solving Equation (2.7) for p_i and $\bar{z}$ allows us to obtain predictions for the average of the trait and for the average number of mutations per haploid genome maintained at mutation-selection equilibrium, which we then use to compute the inbreeding depression and mutation load expected at equilibrium.

2.4.2 Fitness effect of mutations neglecting genetic associations

Under the Lifetime Fitness approach, to leading order in s so that genetic associations between loci can be neglected, Equation (2.7) simplifies into

$$\begin{cases} U + \bar{s}_z\left(\bar{z}\right)\left(h\sum_i p_i + (1-h)F\sum_i p_i\right) = 0 \\ \bar{z}\left(p_i\right) = z_0 \times \exp\left[\pm s\left(2h\sum_i p_i + (1-2h)F\sum_i p_i\right)\right], \end{cases} \tag{2.8}$$

where $F = \frac{\alpha}{2-\alpha}$ is the equilibrium inbreeding coefficient. The derivation of this result is explained in Appendix **??** (Equation A44). The first line of Equation (2.8) shows that when we neglect genetic associations between loci, the selective pressures acting on mutations are encapsulated in a single lifetime selection coefficient $\bar{s}_z$, which is specific to each model and emerges from the application of the Taylor expansion methods presented in Appendix **??** to lifetime fitness. Its expression for each type of mutation studied is given in Appendix **??**.

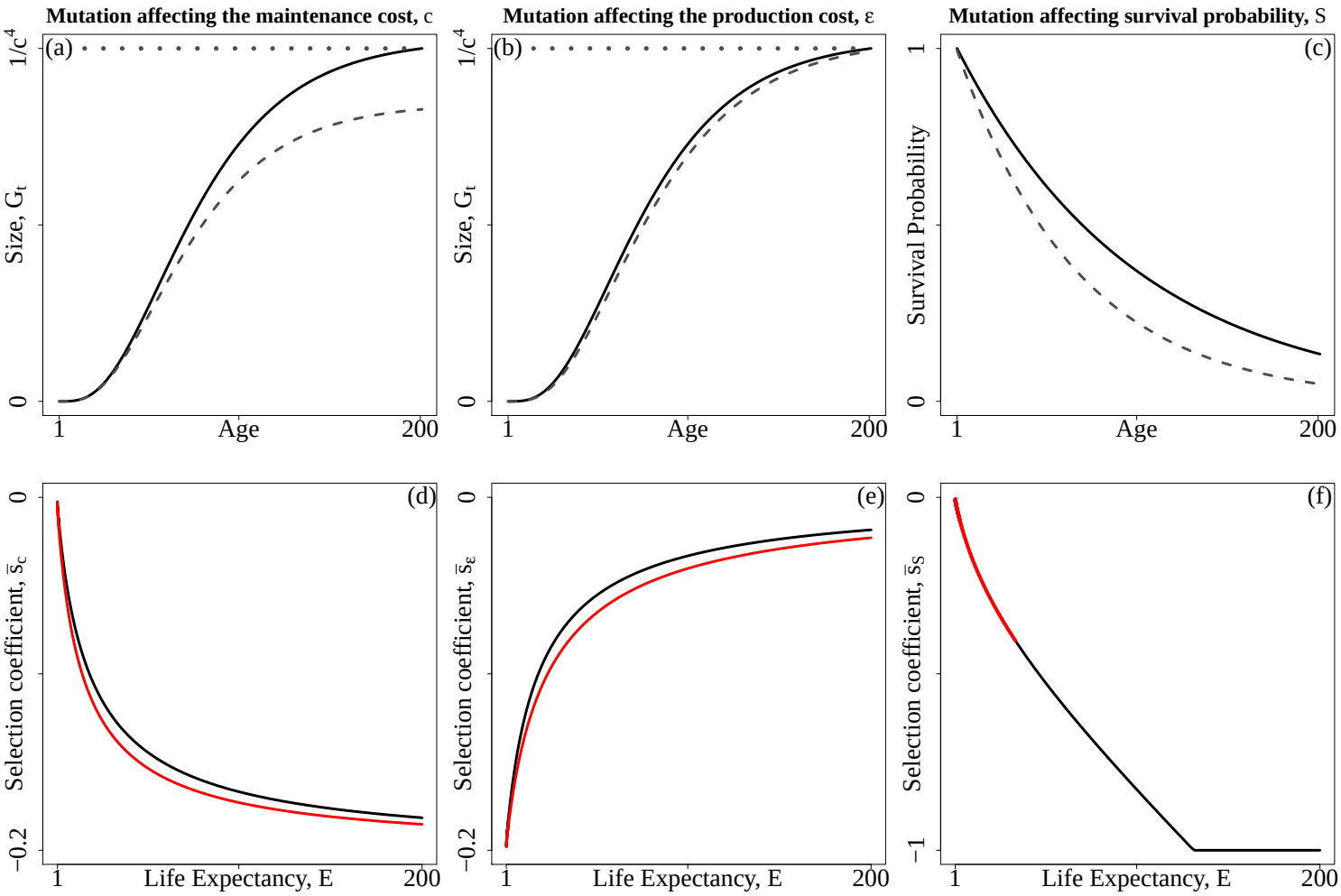

FIGURE 2.2 – Phenotypic and leading order fitness effect of a mutation. Top row : phenotypic effects on growth or survival are presented as a function of age, the dashed line depicts the mutant phenotype while the solid line depicts the wild type ($c = 0.001$, $\varepsilon = 0.01$ and $s = 0.05$ for mutations affecting growth, and $S = 0.99$, $s = 0.005$ for mutations affecting survival). The dotted gray lines depict maximal size. Bottom row : Resulting effects of mutations on lifetime fitness as a function of life expectancy in a single locus case (black lines, $n = 0$), and 10 mutations segregate at other loci (red lines, $n = 10$). Note that the phenotypic effect of mutations (s) differs between mutations affecting growth and survival.

The phenotypic effect of each type of mutation (FIG. 2.2, top row), and the resulting leading order lifetime selection coefficients are presented in FIG. 2.2 (bottom row). These coefficients are all negative because we consider deleterious mutations,

and depend on the population average of the trait, owing to the epistatic mechanisms described in the former section. Although it is not possible to fully disentangle the effects of epistasis from the effects of direct selection at the locus, one may gain insight into the sign of epistasis by studying the direction in which $\bar{s}_z$ coefficients are changed when the number of segregating mutations increases in the population. The $\bar{s}_z$ coefficients plotted in black were obtained assuming mutations are absent (*i.e.* $\bar{z} = z_0$), while those plotted in red were obtained assuming $n = \sum_i p_i = 10$ mutations were segregating at surrounding loci. In other words, the coefficients plotted in black in FIG. 2.2 (bottom row) represent the intensity of selection a mutation would face if a single locus was modeled, while the coefficients plotted in red depict the intensity of selection a mutation would face when in interaction with ten mutations at other loci.

Mutations affecting the maintenance cost c cause size differences to increase as individuals age (FIG. 2.2a). Hence, their fitness effect increases with life expectancy (FIG. 2.2d). Furthermore, epistatic interactions cause selection against mutations to increase that is, we observe negative epistasis. On the contrary, mutations affecting the production cost ε do not affect individuals' maximal size, as they asymptotically tend to the same size, but rather the speed at which they reach it (FIG. 2.2b). Therefore, the growth delay mutated individuals accumulate in early years fades away in older individuals. This causes selection against mutations affecting ε to decrease with respect to life expectancy, because they become gradually neutral in older age-classes. However, epistasis also causes selection against mutations to increase, so that it is negative in this case as well.

Mutations affecting survival cause individuals to perform less mating events in a lifetime. They also cause mutated individuals to perform less well during mating events, because they tend to be younger than unmutated individuals (FIG. 2.2c). Thus, age-structure increases selection against mutations affecting survival. Furthermore, selection against mutations affecting survival strongly increases as life expectancy increases

(FIG. 2.2f), contrary to mutations affecting growth costs, whose fitness effects remain moderate (FIG. 2.2d-e). Besides, life expectancy decreases as the number of segregating mutations increases (because the mean survival probability $\bar{S}$ decreases), which reduces the intensity of selection acting on mutations, thereby generating strong positive epistasis (FIG. 2.2f, in this panel the red and black lines overlie because the trait affected by mutations is plotted on the x-axis, contrary to mutations affecting growth).

Overall, the results described in FIG. 2.2 show that mutations affecting different traits on the same genotype-to-phenotype-to-fitness map face very different selective pressures, both in magnitude and in the way they vary with life expectancy.

2.4.3 Average number of mutations, inbreeding depression and mutation load

The intensity of selection acting on mutations does not depend on the values of c and ε, but on the ratio $\frac{c}{\varepsilon}$ (Appendix **??**). Thus, in what follows, results are described using this ratio. FIGURE 2.3 presents results for the average number of mutations per haploid genome maintained at equilibrium, and the resulting inbreeding depression and mutation load for $h = 0.25$ and $\frac{c}{\varepsilon} = 1$ (other parameter sets are shown in appendix as results are qualitatively similar, FIG. S5 to S12). Analytical predictions, which are depicted by solid lines, are obtained by solving Equation (2.7) numerically using the LC approach accounting for pairwise genetic associations (Appendix 2.6.3.2), and dots depict simulations results. Results obtained with the LF approach (lighter lines, Appendix 2.6.3.1) are also presented on the first row of FIG. 2.3.

The first row in FIG. 2.3 presents the log-scaled average number of mutations per haploid genome maintained at mutation-selection balance. For mutations affecting the maintenance cost and survival, the average number of mutations per haploid genome maintained at equilibrium ($n = \sum_i p_i$) decreases as life expectancy increases (FIG. 2.3a

and 2.3c), because selection against mutations increases. This effect is more marked for mutations affecting survival because selection is stronger in this case (FIG. 2.2d and 2.2f). Conversely, n increases as life expectancy increases for mutations affecting the production cost, because selection against these mutations weakens as life expectancy increases (FIG. 2.2e). In every case, n decreases as the selfing rate increases due to the purging effect of self-fertilisation. The LF approach performs just as well as the LC approach for mutations affecting growth costs, but tends to fail for mutations affecting survival due to the fact that it overlooks important aspects of the model, as discussed above.

The middle row in FIG. 2.3 shows the levels of inbreeding depression generated by mutations maintained at mutation-selection balance. Large differences in the number of deleterious mutations maintained per genome do not translate into strong variations in inbreeding depression with respect to life expectancy. However, variations are still observed and differ between mutation types. Indeed, inbreeding depression always decreases with life expectancy when mutations affect survival, irrespective of the selfing rate (FIG. 2.3f), contrary to mutations affecting growth. Mutations affecting c generate slightly higher inbreeding depression in more short-lived species at low selfing rate, but this pattern is reversed at higher selfing rates (FIG. 2.3d). Conversely, mutations affecting ε generate higher inbreeding depression in more long-lived species for low α and this pattern is reversed for high α (FIG. 2.3e). This is the result of the interaction between life expectancy, the shape of the fitness landscape and selfing, which is difficult to disentangle as these elements interact in non-trivial ways. Indeed, the shape of the fitness landscape changes with respect to life expectancy, which causes variations in the intensity of selection and epistatic interactions, and selfing changes the distribution of fitness by altering homozygosity and by promoting the purging of deleterious mutations in doing so, so that it too alters epistatic interactions.

The bottom row in FIG. 2.3 presents the mutation load measured in the sense

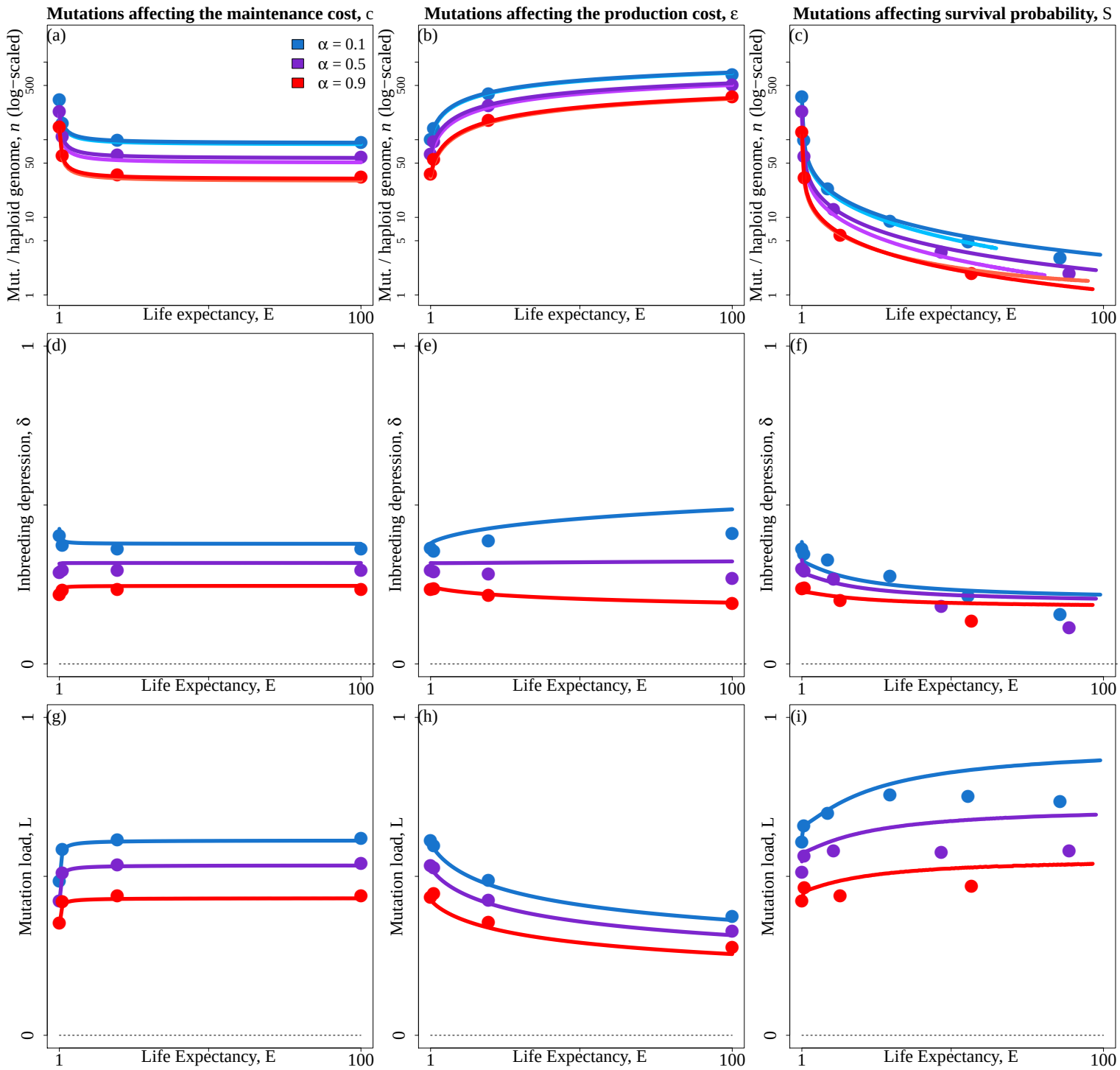

FIGURE 2.3 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : averaged simulation results. Lines : analytical predictions accounting for genetic associations between loci using the LC approach (darker lines) and the LF approach (lighter lines). Parameters shown here are $\frac{c}{\varepsilon} = 1$, $U = 0.5$, $s = 0.005$, $h = 0.25$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

of Crow (1958) which results from the mutations segregating at mutation-selection balance. Intuitively, one would assume that having more deleterious mutations segregating in the population should lead to a higher mutation load. Yet, the mutation load is lower when mutations are more numerous for all three mutation types in our model. Importantly, this result is not a mere reflection of differences in the absolute strength of selection acting on mutations in populations with different life expectancies. Otherwise, changing the phenotypic effect of mutations in a given model and for a given parameters set would also change the mean phenotypic deviation from the optimum, and therefore the mutation load. This is not what we observe. Indeed, FIGURE S4 in Appendix 2.6.5 shows that when the effect of mutations is made ten times larger, this does not cause differences in the equilibrium phenotypic deviation, although the number of segregating mutations at equilibrium was considerably lower (because selection was stronger). This observation is consistent with results obtained by previous authors (Bataillon and Kirkpatrick, 2000), who showed that when the population exceeds a particular size, the mutation load becomes independent from the strength of selection acting on mutations. We argue that the differences we observe in terms of mutation load are imputable to differences in the shape of the fitness landscapes (FIG. S3), which are not fully captured by differences in the efficacy of selection.

2.4.4 Consequences for life-history evolution

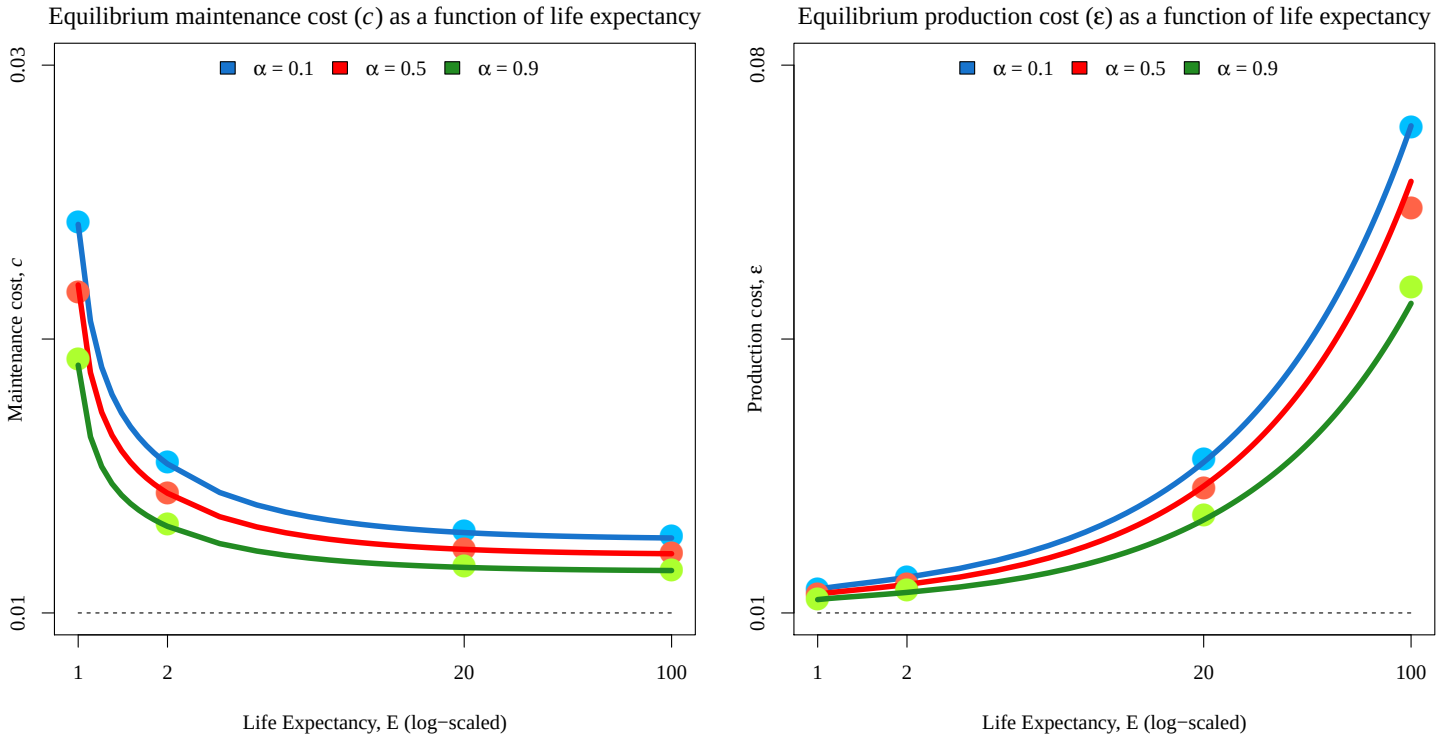

FIGURE 2.4 – Equilibrium maintenance (left) and production (right) costs as function of life expectancy (log-scaled), for various selfing rates. Dots depict simulation results while lines depict analytical predictions. Parameters shown here are $c = \varepsilon = 0.01$, $h = 0.25$, $s = 0.005$.

The mutations segregating in the population cause the mean phenotype to deviate from its initial value. As the intensity of selection acting on mutations depends on this mean phenotype, this leads to a joint equilibrium for both the mean phenotype and the number of mutations segregating in the population (Appendices 2.6.2 and 2.6.3). This equilibrium varies with respect to life expectancy and mating system. Increasing the selfing rate slightly decreases the deviation of the mean phenotype from its initial value, because selfing induces a better purging of mutations. More importantly, the equilibrium phenotypes vary significantly between populations with different life expectancies. Indeed, as species become more long-lived, the equilibrium mainte-

nance cost decreases, so that maximal size increases, and the equilibrium production cost increases, so that growth is slowed down. This means that if, for any reason, life expectancy changes in a species, the growth strategy should also be changed as a consequence of the selective pressures acting on deleterious mutations maintained at mutation-selection balance being altered.

2.5 Discussion

Inbreeding depression. In this paper, we set off to study the magnitude and variation of inbreeding depression generated by mutations affecting survival or growth in relation to life expectancy. We showed that inbreeding depression at mutation-selection balance varies weakly with respect to life expectancy when mutations affect growth, while inbreeding depression decreases more sharply as life expectancy increases in the case of mutations affecting survival. We showed that these differences between mutation types can be attributed to differences in the intensity of selection and how it varies with respect to life expectancy. In any case, our results only agree with empirical data in a limited region of the parameters space, that is when selfing is rare and selective pressures acting on mutations decrease as life expectancy increases, as it is the case for mutations affecting the production cost most noticeably. Although these results may at first glance seem to point to a more preponderant role for somatic mutations accumulation in generating the sharp increase in inbreeding depression observed in more long-lived species, we argue in what follows that they may rather point us towards the type of traits mutations should affect in order to produce this empirically observed pattern, irrespective of the way mutations are produced.

As stated above, differences in the magnitude of inbreeding depression with respect to life expectancy were quantitatively small in every investigated case when mutations affected growth, that is either the maintenance or the production cost, even for low

dominance coefficients and high mutation rates. This result can be explained by the moderate variations in the intensity of selection on growth-related traits with respect to life expectancy, and by the fact that we only considered mutations with weak phenotypic effects, which in this case translated into weak fitness effects. Therefore, this result is consistent with previous population genetics work that showed that inbreeding depression becomes independent of the intensity of selection when selection is weak and the population is large (e.g. Charlesworth et al., 1990; Bataillon and Kirkpatrick, 2000; Roze, 2015). As for mutations affecting survival, their dynamics were studied by Morgan (2001) prior to this study in a perennial, but not age-structured population. They concluded that inbreeding depression should quickly decrease as life expectancy increases, and that significant outbreeding depression should be observed in long-lived species. They argued this result could be attributed to the greater variance in fitness observed among the offspring produced by self-fertilisation, which led to a higher mean lifetime fitness among them. Although they did not state it explicitly, this result stems from the fact that a very small number of mutations were maintained at mutation-selection balance in long-lived species under the parameter sets they investigated. Indeed, Morgan (2001) considered large phenotypic effects of mutations ($s = 0.1$ and $s = 1.0$), which resulted in tremendous lifetime fitness effects and therefore in the maintenance of almost no mutations at equilibrium. This led variance in fitness to dominate over other aspects. Here, we investigated mutations with much lower phenotypic effects ($s = 0.005$), and did not observe the same patterns. In fact, we show that although inbreeding depression still decreases with life expectancy, this decrease is considerably reduced, and outbreeding depression is no longer observed. Thus, if substantial inbreeding depression is observed on traits related to survival in long-lived species, we conclude that this should be caused by mutations with very weak phenotypic effects, or whose effects are limited to early stages of life, as mutations would otherwise not be maintained in the population.

Taken together, our results and those of Morgan (2001) show that variations in the fitness effect of mutations with respect to life expectancy can generate differences in inbreeding depression, provided that these variations are strong enough. Furthermore, we showed that inbreeding depression maintained at mutation-selection balance decreases when the fitness effect of mutations increases, as said mutations are more efficiently purged from the population. This implies that mutations facing sufficiently strong selective pressures, which decrease sharply as life expectancy increases, could in principle generate the increase in inbreeding depression observed in more long-lived species. In the light of this result, we propose that asking on what traits selection is expected to be strong and to decrease significantly as life expectancy increases in plants is a track worth following in order to further our understanding of the mechanisms underlying inbreeding depression. Importantly, this result is not in contradiction with the somatic mutations accumulation hypothesis formulated by Scofield and Schultz (2006). Indeed, the intensity of selection acting on mutations matters for their persistence in the population and for the magnitude of inbreeding depression generated by them, regardless of the way mutations are produced.

Mutation load. Contrary to what one might intuitively expect, we showed that more mutations maintained at mutation-selection balance led to a lower mutation load in all three of our models, and that this result is not a straightforward consequence of differences in the intensity of selection acting on mutations, in agreement with results obtained by previous authors (Bataillon and Kirkpatrick, 2000). Instead, it is a consequence of differences in the shape of fitness landscapes between species with different life expectancies on the one hand, and of the interaction between the mean phenotype and the intensity of selection on the other hand. This result highlights the fact that incorporating a phenotypic dimension to population genetics studies may lead to counter-intuitive results. In particular, when fitness landscapes are obtained as

the result of biological assumptions, and not arbitrarily assumed to be of a particular shape, unusual interactions with various aspects of species life-histories may result in novel predictions. For instance, in the present study, we showed that the mutation load may behave very differently depending on the trait affected by mutations and life expectancy. Therefore, we conclude that comparing mutation loads between species with contrasting life-histories may sometimes be misleading, and would likely require trait-specific approaches.

Epistasis emerging from the fitness landscape. The phenotypic dimension of the genotype-to-phenotype-to-fitness map is usually overlooked in mutation load dynamics studies. Indeed, most theoretical investigations interested in such matters consider mutations affecting fitness multiplicatively (e.g. Kondrashov, 1985; Charlesworth et al., 1990; Roze, 2015), so that fitness is the phenotype affected by mutations and the phenotype-to-fitness map (the fitness landscape) is strictly linear. Hence, neither phenotype-to-fitness epistasis nor genotype-to-phenotype epistasis occur in these models. Some authors however considered the effects of epistasis on mutation load dynamics.

Charlesworth et al. (1991) studied uniformly deleterious mutations under synergistic selection. Because they assumed mutations to affect fitness directly too, no phenotype-to-fitness epistasis occured in their model. On the other hand, genotype-to-phenotype epistasis occured because they assumed mutations to affect fitness synergistically. Abu Awad and Roze (2019) generalised Charlesworth et al. (1991)'s results, showing that their model is equivalent to assuming fixed negative pairwise epistasis between loci. Furthermore, they showed that different forms of pairwise epistasis have different effects : while additive-by-additive epistasis lowers the frequency of mutations by increasing selection, additive-by-dominance and dominance-by-dominance epistasis tend to increase inbreeding depression by making homozygotes less fit.

Phenotype-to-fitness epistasis was also considered in the case of mutations affecting an abstract trait (or set of traits) under stabilising selection (e.g. Abu Awad and Roze, 2018, 2019). These studies assumed mutations to affect the trait(s) additively, so that no genotype-to-phenotype epistasis occured, but assumed the fitness landscape to be gaussian or quasi-gaussian (Fisher, 1930), so that phenotype-to-fitness epistasis occured as the landscape was non-linear. However, because the fitness landscape was symmetrical with respect to the optimal phenotype and mutations occured in both directions at the same rate, epistasis was null on average (Gros et al., 2009). Yet, it influenced mutation load dynamics through variance effects, by generating associations between alleles with compensatory effects, and decreasing the efficiency of selection on deleterious alleles when these associations stay moderate.

In this paper, we studied the dynamics of mutations affecting growth and survival using a physiological growth model (West et al., 2001). Hence, the fitness effect of mutations emerged from biological assumptions instead of being assumed a priori. The resulting fitness landscape had singularities at the optimal phenotypes, as in the absence of trade-offs, immortality ($S = 1$) is the optimal survival strategy and leads to an infinite life expectancy, and null growth costs ($c = \varepsilon = 0$) are the optimal growth strategy, which leads to infinitely large individuals. Furthermore, these optimal phenotypes were situated on the boundaries of the landscape for all three traits, as negative growth costs or survival probabilities greater than one are meaningless. As a consequence, mutations were always under directional selection in our model. We hence focused on the case of mutations increasing growth costs or decreasing survival, that is deleterious mutations, as opposite mutations would always be favoured and eventually reach fixation. Here, the fitness landscape was non-linear (Fig. S3). Hence, although we assumed mutations to affect phenotypes multiplicatively, so that no genotype-to-phenotype epistasis occured, phenotype-to-fitness epistasis was generated. Contrary to models assuming stabilising selection (e.g. Abu Awad and Roze, 2018, 2019), epistasis

was non-zero on average in our model, and its direction (positive or negative) differed depending on the phenotype affected by mutations. It was however difficult to isolate the effects of epistasis in our model because pairwise epistatic interactions are not sufficient to capture these effects, as considering only pairwise interactions is equivalent to linearising the effect of mutations on the trait average, which only works when very few mutations segregate in the population.

Life-history evolution. Plants vary widely in life expectancy and stature, with life expectancies ranging from a few weeks to hundreds, possibly thousands of years, and stature spanning several orders of magnitude across Tracheophytes (Ehrlén and Lehtilä, 2002). These variations are correlated. Indeed, long-lived species tend to grow slower than short-lived species (Salguero-Gómez et al., 2016). In life-history traits evolution theory, this type of correlation is usually interpreted in terms of trade-offs, with populations evolving towards the evolutionarily stable allocation of resources between growth, survival and reproduction, given a number of constraints (Stearns, 1992). In this paper, we have shown that the equilibrium maintenance and production costs differ between life expectancies, causing more long-lived species to grow slower but ultimately larger than more short-lived species, as commonly expected. However, the mechanism underlying this result is completely different. Indeed, in our model, life-history traits do not evolve jointly in response to trade-offs. Instead, the equilibrium growth costs are modified when life expectancy varies because the selective pressures acting on the many deleterious mutations affecting these traits are altered, leading to a more or less efficient purging of said mutations and thereby phenotypic differences at mutation-selection equilibrium. Thus, our results suggest that life-history traits may sometimes evolve jointly regardless of trade-offs, because a change in a given trait may alter the efficiency of purging of deleterious mutations affecting other traits.

Acknowledgements

The authors thank Sylvain Glémin, Florence Débarre, Vincent Castric and Denis Roze for taking the time to discuss this work. This work was funded by the European Research Council (NOVEL project, grant #648321). The authors also thank the Région Hauts-de-France, and the Ministère de l'Enseignement Supérieur et de la Recherche (CPER Climibio), and the European Fund for Regional Economic Development for their financial support. Finally, the authors would like to thank the Editor and Associate Editor along with the both anonymous reviewers for the time and effort they devoted to the improvement of this article.

Author Contribution Statement

Both authors contributed to the conceptualization of the work presented in this article. Mathematical developments and individual-centered simulations were carried out by TL. Both authors contributed to the analysis of the results and to writing and editing the manuscript. SB supervised the work and acquired the funding.

Data and Code Accessibility Statement

All individual-centered simulation and numerical computation programs are available on GitHub (link).

2.6 Appendices

2.6.1 Growth model

At any age t, we assume that an individual is composed of G_t identical units (say, branches or buds), and has a resting metabolic rate B_t. This metabolic rate can be subdivided into energies spent on maintenance and growth as follows (West et al., 2001),

$$B_t = c\ G_t + \varepsilon\ \frac{dG_t}{dt}, \tag{A1}$$

where c is the energy required to maintain a single unit, and ε is the energy required to create a new unit. Moreover, the resting metabolic rate can be expressed as

$$B_t = B_0\ {G_t}^x, \tag{A2}$$

where B_0 is the species' basal metabolic rate and x is the allometric coefficient (Peters, 1983). In vascular plants and animals, the value $x = {}^3\!/_4$ is generally used in growth models on the basis of both theoretical (e.g. West et al., 1997; **?**), and empirical arguments (e.g. Peters, 1983; Enquist et al., 1998; West et al., 2001), and will thus be used in what follows. Different exponent values would however likely yield qualitatively similar results, provided that they remain positive and smaller than one, because the rate at which the metabolic rate $B_0\ {G_t}^x$ increases with respect to size would still slow down as size increases so that the energy available to the individual would at some point equal the energy required for its maintenance $c\ G_t$, thereby reaching a plateau which is given by

$$B_0 G_t^x = c\ G_t \Leftrightarrow G_t = \left(\frac{B_0}{c}\right)^{\frac{1}{1-x}}. \tag{A3}$$

Using Equation (A2), Equation (A1) can be rearranged into the following differen-

tial equation

$$\frac{dG}{dt} = \frac{1}{\varepsilon}\left(B_0 {G_t}^{\frac{3}{4}} - c\, G_t\right). \tag{A4}$$

Solving Equation (A4) for G_t, and setting $B_0 = 1$ for convenience, we obtain

$$G_t = \frac{e^{-\frac{c}{\varepsilon}(t+1)}}{c^4}\left(e^{\frac{c}{4\varepsilon}(t+1)} - 1\right)^4. \tag{A5}$$

This growth function naturally saturates, when the energy required to maintain existing units becomes too large for new units to be produced. The size at which it saturates is given by

$$\lim_{t\to\infty} G_t = \frac{1}{c^4}, \tag{A6}$$

which is equal to the plateau given in Equation (A3), with $B_0 = 1$ and $x = \frac{3}{4}$.

2.6.2 General recursions for the effects of selection and reproduction under partial selfing

In this paper, we use two different approaches to study our model. The first approach makes the assumption that an age-structured population can be viewed as a rescaled annual population (the Lifetime Fitness approach, LF), so that individuals first undergo lifetime selection, then reproduce. The second method proceeds in a more detailed manner, by studying each step of the life cycle successively (the Life Cycle approach, LC). In both approaches, the objective is to derive an approximation for the change in frequency of deleterious alleles at the selected loci, Δp_i, in order to find the expected number of mutations maintained at mutation-selection balance in the population by solving $\Delta p_i = 0$.

Irrespective of the method, one may derive a general approximation of the relative fitness for an individual as a function of its genotype during any given selection stage. In this section, after introducing our theoretical framework (Barton and Turelli, 1991; Kirkpatrick et al., 2002), we derive such an approximation assuming fitness is a function of a trait z, and compute general recursions for the effects of selection on allelic frequencies and genetic associations. Then, we compute general recursions for the effects of reproduction.

2.6.2.1 Defining variables

Let X_i and $\overset{*}{X}_i$ be the indicator variables associated with the paternally and maternally inherited alleles at the ith locus. These variables are worth 1 when allele a is present at the position they are associated with, and 0 otherwise. Let us also define the centered variables associated with the paternally and maternally inherited alleles

at the ith locus, ζ_i and $\overset{*}{\zeta}_i$, as

$$\zeta_i = X_i - \mathbb{E}\left[X_i\right] = X_i - p_i \text{ and, } \overset{*}{\zeta}_i = \overset{*}{X}_i - \mathbb{E}\left[\overset{*}{X}_i\right] = \overset{*}{X}_i - p_i, \tag{A7}$$

where $\mathbb{E}\left[X_i\right] = \mathbb{E}[\overset{*}{X}_i] = p_i$ is allele a's frequency at the ith locus. The expectation of products of these centered variables allows one to quantify genetic associations between positions in the genome, that is, these variables allow one to quantify to what extent the population deviates from Hardy-Weinberg's equilibrium. For example, the expectation $\mathbb{E}\left[\zeta_i \times \overset{*}{\zeta}_i\right]$ measures to what extent the homozygosity deviates from the panmictic expectation at locus i.

Condensed notations and simplification rule. To make recursions clearer, let us introduce the notation

$$D_{\mathbb{U},\mathbb{V}} = \mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}}\right] = \mathbb{E}\left[\prod_{i\in\mathbb{U}}\left(\zeta_{\mathbb{U}}\right) \times \prod_{i\in\mathbb{V}}\left(\overset{*}{\zeta}_{\mathbb{V}}\right)\right], \tag{A8}$$

where $\mathbb{U}$ and $\mathbb{V}$ are sets of positions paternally and maternally inherited, respectively (in the previous example, $\mathbb{U} = \{i\}$ and $\mathbb{V} = \{i\}$), so that the excess in homozygotes at locus i is denoted $D_{i,i}$. Sets $\mathbb{U}$ and $\mathbb{V}$ can contain positions situated at the same locus or at different loci, and can be empty. For instance, Association $D_{ij,\varnothing}$ which in terms of sets corresponds to $\mathbb{U} = \{i, j\}$, and $\mathbb{V} = \{\varnothing\}$, quantifies the linkage disequilibrium between alleles at loci i and j on the paternally inherited chromosome.

Besides, individuals are hermaphroditic and no sex-specific effect is assumed in our model, so that we always have

$$D_{\mathbb{U},\mathbb{V}} = D_{\mathbb{V},\mathbb{U}} \text{ for all sets } \mathbb{U} \text{ and } \mathbb{V}, \tag{A9}$$

hence in our former example, we have $D_{ij,\varnothing} = D_{\varnothing,ij}$. Again, to make recursions shorter,

we will use the notation

$$\widetilde{D}_{\mathbb{U},\mathbb{V}} = \frac{D_{\mathbb{U},\mathbb{V}} + D_{\mathbb{V},\mathbb{U}}}{2}, \tag{A10}$$

and when possible, we will omit empty sets. For example, we have

$$\widetilde{D}_{ij} = \frac{D_{ij,\varnothing} + D_{\varnothing,ij}}{2}. \tag{A11}$$

Finally, as shown in Kirkpatrick et al. (2002), repeated indexes sometimes appear in recursions, and can be dealt with using the relationship

$$D_{\mathbb{U}ii,\mathbb{V}} = (1-2p_i)D_{\mathbb{U}i,\mathbb{V}} + p_i q_i D_{\mathbb{U},\mathbb{V}}, \text{ where } q_i = 1 - p_i. \tag{A12}$$

Genetic associations relevant to our model. When mating is not random, as it is the case in our model because of self-fertilisation, non-zero genetic associations are maintained between and within loci at equilibrium. Since we model an infinite number of loci, accounting for all associations arising between them is not possible as infinitely many combinations of genetic positions would have to be considered. Luckily, most genetic associations remain negligible at equilibrium when selection is weak, so that we may confine ourselves to taking into account within-locus and pairwise between-loci associations.

Furthermore, different kinds of genetic associations are affected by selection differently, so that further simplifications can be applied (Roze, 2015). Indeed, association $\widetilde{D}_{i,i}$, which depicts the excess in homozygotes at locus i, is generated by inbreeding even in the absence of selection, and direct selection acting on this association can be neglected when mutations segregating at other loci are numerous, because indirect selection becomes much stronger. Besides, $\widetilde{D}_{ij,ij}$, which reflects correlations in homozygosity between loci i and j, is also generated by inbreeding even under neutrality, and is little affected by selection, so that we can neglect the effects of selection on this

association.

Other between-loci associations are generated both by inbreeding and selection. Indeed, $\widetilde{D}_{ij,j}$ is generated by selection at locus i and is of order s, while $\widetilde{D}_{ij}$ and $\widetilde{D}_{i,j}$ result from selection affecting both loci, and are of order s^2. Following Roze (2015), we obtain leading order approximations in s for these associations, so that $\widetilde{D}^*_{ij} = \widetilde{D}^*_{i,j} \approx 0$ at quasi-linkage equilibrium. Thus, in what follows, we will only be concerned with associations $\widetilde{D}_{i,i}$, $\widetilde{D}_{ij,j}$ and $\widetilde{D}_{ij,ij}$, and neglect the rest of them.

2.6.2.2 Selection

Deriving relative fitness. The results presented here are used to derive the results in Appendix 2.6.3. For any trait z, assuming mutations affect this trait multiplicatively, an individual's trait value can be written in terms of indicator variables as

$$z = z_0 \times \prod_i \left(1 \pm sh\left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)X_i\overset{*}{X}_i\right), \tag{A13}$$

where $\pm$ is replaced with $+$ or $-$ depending on whether mutations are assumed to increase (e.g. physiological growth costs) or decrease (e.g. survival probability) the trait, s is the phenotypic effect of the mutant allele, and h is its dominance coefficient. The log-value of the trait is then

$$\ln z = \ln z_0 + \sum_i \ln\left(1 \pm sh\left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)X_i\overset{*}{X}_i\right), \tag{A14}$$

which, under the assumption that selection is weak, *i.e.* that s is small, can be approximated by

$$\ln z \approx \ln z_0 \pm sh\sum_i \left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)\sum_i X_i\overset{*}{X}_i \tag{A15}$$

to leading order in s. Injecting ζ-variables into Equation (A15), it can then be rearran-

ged into

$$\ln z \approx \ln z_0 + \sum_i Z_i \tag{A16}$$

with $Z_i = \pm sh \sum_i \left(\zeta_i + \overset{*}{\zeta}_i + 2p_i\right) \pm s(1-2h) \sum_i \left(\widetilde{\zeta}_{i,i} + p_i(\zeta_i + \overset{*}{\zeta}_i) + p_i^2\right)$. Hence, the mean log-value is given by

$$\overline{\ln z} = \ln z_0 + \sum_i \overline{Z}_i \text{ with, } \overline{Z}_i = \pm 2sh \sum_i p_i \pm s(1-2h) \sum_i \left(\widetilde{D}_{i,i} + p_i^2\right), \tag{A17}$$

and one may express the trait value z as

$$z = \exp(\ln z) = \exp\left(\overline{\ln z} + \ln z - \overline{\ln z}\right) = z_0 \ \exp\left(\sum_i \left(\overline{Z}_i + \Delta Z_i\right)\right), \tag{A18}$$

with $\Delta Z_i = Z_i - \overline{Z}_i$. Therefore, the trait average, $\bar{z}$, is simply

$$\bar{z} \approx z_0 \exp\left(\sum_i \overline{Z}_i\right). \tag{A19}$$

All else being fixed, an individual's fitness during a given selection stage, $w(z)$, is a function of its trait value z. The mean fitness is a constant, given by $\bar{w} = w(\bar{z})$, and the relative fitness $\frac{w(z)}{\bar{w}}$, to second order in s, is given by

$$\frac{w}{\bar{w}}(z) = 1 + \frac{w'(\bar{z})}{\bar{w}} \sum_i \Delta Z_i + \frac{w''(\bar{z})}{2\ \bar{w}} \sum_i \sum_j \Delta Z_i \times \Delta Z_j, \tag{A20}$$

where w' and w'' are the first and second derivatives of w with respect to ΔZ_i. Neglecting terms in $i = j$, and denoting $\bar{s}_z(\bar{z}) = \pm s \frac{w'(\bar{z})}{\bar{w}}$ and $\hat{\bar{s}}_z(\bar{z}) = \frac{w''(\bar{z})}{2\bar{w}}(\pm s)^2$, this

is

$$\frac{w}{\overline{w}}(z) \approx 1+\bar{s}(\bar{z})\sum_i\left(h\left(\zeta_i+\overset{*}{\zeta}_i\right)+(1-2h)\left(\widetilde{\zeta}_{i,i}-\widetilde{D}_{i,i}\right)\right)$$
$$+\hat{\bar{s}}(\bar{z})\Bigg[h^2\sum_{i<j}\left(\zeta_i+\overset{*}{\zeta}_i\right)\left(\zeta_j+\overset{*}{\zeta}_j\right)+h(1-2h)\sum_{i\neq j}\left(\zeta_i+\overset{*}{\zeta}_i\right)\left(\widetilde{\zeta}_{j,j}-\widetilde{D}_{j,j}\right)$$
$$+(1-2h)^2\sum_{i<j}\left(\left(\widetilde{\zeta}_{i,i}-\widetilde{D}_{i,i}\right)\left(\widetilde{\zeta}_{j,j}-\widetilde{D}_{j,j}\right)-\left(\widetilde{D}_{ij,ij}-\widetilde{D}_{i,i}\widetilde{D}_{j,j}\right)\right)\Bigg]. \tag{A21}$$

The leading order selection coefficient ($\bar{s}_z$) encapsulates the effects of selection acting directly on loci, while the second order selection coefficient ($\hat{\bar{s}}_z$) quantifies the effects of indirect selection between pairs of loci. We neglect higher order selection coefficients. Note that $\bar{s}_z < 0$ in all our models since we consider deleterious mutations.

Allelic frequencies change due to selection. The change in allelic frequencies at the ith selected locus owing to selection at a given stage is

$$\Delta^s p_i = \mathbb{E}\left[\frac{w}{\overline{w}}\ \frac{X_i+\overset{*}{X}_i}{2}\right]-p_i, \tag{A22}$$

which, using Equation (A21), to second order in s and neglecting terms in p_i^2, that is assuming mutations are rare, yields

$$\Delta^s p_i = \bar{s}_z\left(hp_i+(1-h)\widetilde{D}_{i,i}\right)+\bar{s}_z(1-2h)\sum_{j\neq i}\widetilde{D}_{ij,j}$$
$$+\hat{\bar{s}}_z(1-h)(1-2h)\sum_{j\neq i}\left(\widetilde{D}_{ij,ij}-\widetilde{D}_{i,i}\widetilde{D}_{j,j}\right)+o\left(s^3\right), \tag{A23}$$

Effect of selection on associations between selected loci. We compute the effect of selection on genetic associations to leading order in s (first line in Equation

A21). As shown by Roze (2015), the effects of selection on $\widetilde{D}_{ij,ij}$ can be neglected and $\widetilde{D}_{ij}$ and $\widetilde{D}_{i,j}$ are of order s^2. Hence, we only need to consider the effects of selection on $\widetilde{D}_{ij,j}$ and $\widetilde{D}_{i,i}$, which can be computed using Equation (A24).

$$\widetilde{D}^s_{ij,j} = \mathbb{E}\left[\frac{w}{\overline{w}}\,\widetilde{\zeta}_{ij,j}\right] - \Delta^s p_i \widetilde{D}^s_{j,j} \text{ and, } \widetilde{D}^s_{i,i} = \mathbb{E}\left[\frac{w}{\overline{w}}\,\widetilde{\zeta}_{i,i}\right] + \left(\Delta^s p_i\right)^2. \tag{A24}$$

To leading order, $\Delta^s p_i$ simplifies to

$$\Delta^s p_i = \bar{s}_z\left(h\ p_i - (1-h)\widetilde{D}_{i,i}\right) + o\left(s^2\right). \tag{A25}$$

Hence, $\widetilde{D}^s_{i,i}$ and $\widetilde{D}^s_{ij,j}$ are given by

$$\widetilde{D}^s_{i,i} \approx \widetilde{D}_{i,i} + \bar{s}_z(1-2h)\sum_{j\neq i}\left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}\right), \tag{A26}$$

and,

$$\widetilde{D}^s_{ij,j} \approx \widetilde{D}_{ij,j} + \bar{s}_z(1-h)\left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}\right). \tag{A27}$$

2.6.2.3 Reproduction and mutation under partial selfing

As we model infinitely many loci, we may assume that mutations never occur two times at the same locus. Since they are not affected by recombination and self-fertilisation, allelic frequencies following reproduction are simply given by

$$p^r_i = p_i + U\left(1-p_i\right) \approx U + p_i, \tag{A28}$$

where p_i here depicts the frequency of the deleterious allele at locus i just before reproduction. On the other hand, while we neglect the effects of mutation on genetic associations, they are affected by recombination and selfing. Following reproduction,

denoting $F = \frac{\alpha}{2-\alpha}$, we have

$$\widetilde{D}^r_{i,i} = \alpha\ \mathbb{E}\left[\frac{1}{4}\zeta_{ii,\varnothing} + \frac{1}{2}\zeta_{i,i} + \frac{1}{4}\zeta_{\varnothing,ii}\right] = \frac{F}{1+F}\left(\widetilde{D}_{i,i} + p_i q_i\right), \tag{A29}$$

$$\widetilde{D}^r_{ij,j} = \alpha\,\mathbb{E}\left[\left((1-r)\widetilde{\zeta}_{ij} + r\ \widetilde{\zeta}_{i,j}\right)\ \frac{\zeta_j + \overset{*}{\zeta}_j}{2}\right] = \frac{F}{1+F}\left(\widetilde{D}_{ij,j} + (1-r)\widetilde{D}_{ij} + r\widetilde{D}_{i,j}\right) \approx \frac{F}{1+F}\widetilde{D}_{ij,j}, \tag{A30}$$

and,

$$\widetilde{D}^r_{ij,ij} = \alpha\,\mathbb{E}\left[\left((1-r)\widetilde{\zeta}_{ij} + r\ \widetilde{\zeta}_{i,j}\right)^2\right] = \frac{F}{1+F}\left[(1-R)\left(\widetilde{D}_{ij,ij} + p_i q_i p_j q_j\right) + R\left(p_i q_i \widetilde{D}_{j,j} + p_j q_j \widetilde{D}_{i,i}\right)\right], \tag{A31}$$

with $R = 2r(1-r)$ and where $D_{\mathbb{U},\mathbb{V}}$ associations and allelic frequencies p_i are measured just before reproduction.

2.6.3 Analytical results

In Appendix 2.6.2, we derived general recursions for the effects of reproduction (involving mutation, non-random mating and recombination) and selection at any given selection stage. In this section, our aim is to apply these general recursions to our different models and analytical approaches. For each analytical approach, we begin with deriving the leading and second order selection coefficients arising from the models, then present analytical results at quasi-linkage equilibrium and explain how they are used to obtain an approximation for the average number of mutations per haploid genome.

2.6.3.1 Lifetime fitness approach

In the LF approach, we make the assumption that all the selective pressures acting on mutations can be summarised into a single lifetime fitness expression, so that the population can be studied as a rescaled annual population. Thus, in this analytical model, timesteps begin with lifetime selection, and are followed by reproduction (Fig. S1).

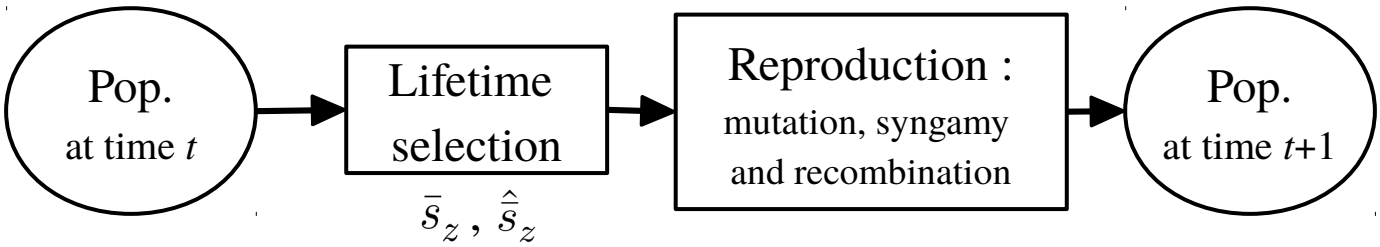

FIGURE S1 – Steps followed over the course of one timestep with the LF approach

Lifetime fitness and lifetime selection coefficients In this paragraph, we compute the lifetime reproductive output of an individual given its phenotype, which is

used as a measure of lifetime fitness, and derive the leading and second order lifetime selection coefficients resulting from this expression.

Lifetime fitness. The size of an individual at age t (G_t) is the solution of Equation (2.2). Given the individual's maintenance and production costs, c and ε, it is

$$G_t = \frac{e^{-\frac{c}{\varepsilon}(t+1)}}{c^4}\left(e^{\frac{c}{4\varepsilon}(t+1)} - 1\right)^4. \tag{A32}$$

Hence, the contribution to reproduction of an individual up to age τ, $\bar{G}_\tau$, is given by

$$\bar{G}_\tau = \sum_{t=0}^{\tau} G_t. \tag{A33}$$

Furthermore, the probability that, given its survival probability S, an individual lives up to exactly age τ then dies ($\bar{S}_\tau$) is

$$\bar{S}_\tau = S^\tau(1-S). \tag{A34}$$

Hence, the lifetime fitness of an individual (W) is given by

$$W = J \times \sum_{\tau=0}^{\infty} \bar{S}_\tau \bar{G}_\tau = \frac{J}{c^4}\left(\frac{1}{1-S} - \frac{4}{e^{\frac{c}{4\varepsilon}} - S} + \frac{6}{e^{\frac{c}{2\varepsilon}} - S} - \frac{4}{e^{\frac{3c}{4\varepsilon}} - S} + \frac{1}{e^{\frac{c}{\varepsilon}} - S}\right), \tag{A35}$$

where J is the recruitment probability of the individual as a juvenile.

Selection coefficients. Let us now write down the lifetime selection coefficients obtained with the method described in Appendix 2.6.2.2. For any trait z, these coefficients are obtained by applying Equation (A20) to Equation (A35). We show that they depend on the ratio $\frac{c}{\varepsilon}$, rather than c and ε absolute values. In what follows, we set $J = 1$ for convenience.

Mutations altering c. The leading order lifetime selection coefficient for mutations altering c is given by

$$\bar{s}_c = s\frac{-4V_c + c\dfrac{\partial V_c}{\partial c}}{V_c} \quad \text{with,} \quad V_c = \frac{1}{1-S} - \frac{4}{e^{\frac{\bar{c}}{4\varepsilon}} - S} + \frac{6}{e^{\frac{\bar{c}}{2\varepsilon}} - S} - \frac{4}{e^{\frac{3\bar{c}}{4\varepsilon}} - S} + \frac{1}{e^{\frac{\bar{c}}{\varepsilon}} - S}. \tag{A36}$$

As for the second order lifetime selection coefficient, $\hat{\bar{s}}_c$, it is given by

$$\hat{\bar{s}}_c = \frac{s^2}{V_c}\left[8V_c - c\frac{7}{2}\frac{\partial V_c}{\partial c} + \frac{\bar{c}^2}{2\varepsilon^2}\left(\frac{e^{\frac{\bar{c}}{\varepsilon}}\left(e^{\frac{\bar{c}}{\varepsilon}} + S\right)}{\left(e^{\frac{\bar{c}}{\varepsilon}} - S\right)^3} - \frac{9e^{\frac{3\bar{c}}{4\varepsilon}}\left(e^{\frac{3\bar{c}}{4\varepsilon}} + S\right)}{4\left(e^{\frac{3\bar{c}}{4\varepsilon}} - S\right)^3} + \frac{3e^{\frac{\bar{c}}{2\varepsilon}}\left(e^{\frac{\bar{c}}{2\varepsilon}} + S\right)}{2\left(e^{\frac{\bar{c}}{2\varepsilon}} - S\right)^3} - \frac{e^{\frac{\bar{c}}{4\varepsilon}}\left(e^{\frac{\bar{c}}{4\varepsilon}} + S\right)}{4\left(e^{\frac{\bar{c}}{4\varepsilon}} - S\right)^3}\right)\right] \tag{A37}$$

Making the variable change $c = \varepsilon\,\frac{c}{\varepsilon}$ in Equations (A36) and (A37), so that it now depends of $\frac{c}{\varepsilon}$ instead of c, we obtain that

$$\frac{\partial \bar{s}_c}{\partial \varepsilon} = \frac{\partial \hat{\bar{s}}_c}{\partial \varepsilon} = 0, \tag{A38}$$

which demonstrates that $\bar{s}_c$ and $\hat{\bar{s}}_c$ depend only on the ratio $\frac{c}{\varepsilon}$. The same reasoning is followed for all three mutation types and will thus not be repeated.

Mutations altering ε. In this case, the leading order lifetime selection coefficient is

$$\bar{s}_\varepsilon = s\frac{\varepsilon}{V_\varepsilon}\frac{\partial V_\varepsilon}{\partial \varepsilon} \quad \text{with,} \quad V_\varepsilon = \frac{1}{1-S} - \frac{4}{e^{\frac{c}{4\bar{\varepsilon}}} - S} + \frac{6}{e^{\frac{c}{2\bar{\varepsilon}}} - S} - \frac{4}{e^{\frac{3c}{4\bar{\varepsilon}}} - S} + \frac{1}{e^{\frac{c}{\bar{\varepsilon}}} - S}, \tag{A39}$$

and the second order lifetime selection coefficient is given by

$$\hat{\bar{s}}_\varepsilon = -\frac{s^2}{2V_\varepsilon}\left[\varepsilon\frac{\partial V_\varepsilon}{\partial \varepsilon} + \frac{c^2 e^{\frac{c}{4\bar{\varepsilon}}}}{4\bar{\varepsilon}^2}\left(\frac{e^{\frac{c}{4\bar{\varepsilon}}} + S}{\left(e^{\frac{c}{4\bar{\varepsilon}}} - S\right)^3} - \frac{6e^{\frac{c}{4\bar{\varepsilon}}}\left(e^{\frac{c}{2\bar{\varepsilon}}} + S\right)}{\left(e^{\frac{c}{2\bar{\varepsilon}}} - S\right)^3} + \frac{9e^{\frac{c}{2\bar{\varepsilon}}}\left(e^{\frac{3c}{4\bar{\varepsilon}}} + S\right)}{\left(e^{\frac{3c}{4\bar{\varepsilon}}} - S\right)^3} - \frac{4e^{\frac{3c}{4\bar{\varepsilon}}}\left(e^{\frac{c}{\bar{\varepsilon}}} + S\right)}{\left(e^{\frac{c}{\bar{\varepsilon}}} - S\right)^3}\right)\right]. \tag{A40}$$

Mutations altering S. Finally, the leading order lifetime selection coefficient for mutations affecting survival is given by

$$\bar{s}_S = s\frac{V_S + S\frac{\partial V_S}{\partial S}}{V_S} \quad \text{with,} \quad V_S = \frac{1}{1-\bar{S}} - \frac{4}{e^{\frac{c}{4\varepsilon}} - \bar{S}} + \frac{6}{e^{\frac{c}{2\varepsilon}} - \bar{S}} - \frac{4}{e^{\frac{3c}{4\varepsilon}} - \bar{S}} + \frac{1}{e^{\frac{c}{\varepsilon}} - \bar{S}}, \tag{A41}$$

while the second order lifetime selection coefficient is

$$\begin{aligned}\hat{\bar{s}}_S = \frac{s^2}{2V_S}\Bigg[V_S + 2\frac{\partial V_S}{\partial S} + S^2\frac{\partial^2 V_S}{\partial S^2} \\ + \bar{S}\left(\frac{1}{\left(1-\bar{S}\right)^3} + \frac{e^{\frac{c}{\varepsilon}}}{\left(e^{\frac{c}{\varepsilon}} - \bar{S}\right)^3} - \frac{4e^{\frac{c}{4\varepsilon}}}{\left(e^{\frac{c}{4\varepsilon}} - \bar{S}\right)^3} + \frac{6e^{\frac{c}{2\varepsilon}}}{\left(e^{\frac{c}{2\varepsilon}} - \bar{S}\right)^3} - \frac{4e^{\frac{3c}{4\varepsilon}}}{\left(e^{\frac{3c}{4\varepsilon}} - \bar{S}\right)^3}\right)\Bigg]\end{aligned} \tag{A42}$$

Quasi-linkage equilibrium results To obtain an approximation for the number of mutations maintained at equilibrium, we need to find the point at which the amount of mutations entering the population due to mutation is compensated by the amount of mutations removed from the population by selection, that is the mutation-selection balance. To do so, one has to find the point at which the variation in allelic frequencies at the selected loci is worth zero. In the LF approach, selection is followed by reproduction, so that the change in allelic frequencies at the selected loci is given by Equations (A23) and (A28), that is

$$\begin{aligned}\Delta p_i = \bar{s}_z\left(hp_i + (1-h)\widetilde{D}_{i,i}\right) + \bar{s}_z(1-2h)\sum_{j\neq i}\widetilde{D}_{ij,j} \\ + \hat{\bar{s}}_z(1-h)(1-2h)\sum_{j\neq i}\left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}\right) + U.\end{aligned} \tag{A43}$$

The leading order approximation of Equation (A43) is presented in Equation (2.8) in the main text. As mentioned in Appendix 2.6.2.1, association $\widetilde{D}_{ij,j}$ is of order s, so that the second term in Equation (A43) is of order s^2, while the third term is of order

$\hat{\bar{s}}_z$, which is of order s^2. Hence, both these terms can be neglected to leading order in s and Equation (A43) simplifies into

$$\Delta p_i = \bar{s}_z \left(h p_i + (1-h)\widetilde{D}_{i,i} \right) + U, \tag{A44}$$

which corresponds to the first line of Equation (2.8) in the main text.

In order to solve solve $\Delta p_i = 0$, we need to obtain equilibrium expressions for genetic associations appearing in Equation (A43).

Equilibrium genetic associations under neutrality. As mentioned above, the effects of selection on $\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}$ can be neglected, as it is generated by inbreeding even in the absence of selection. Assuming neutrality, neglecting terms in p_i^2, and using Equation (A29), we have

$$\Delta^{neut}\widetilde{D}_{i,i} = \frac{F}{1+F}\left(p_i + \widetilde{D}_{i,i}\right) - \widetilde{D}_{i,i}, \tag{A45}$$

which at equilibrium yields

$$\Delta^{neut}\widetilde{D}_{i,i} = 0 \Leftrightarrow \widetilde{D}_{i,i}^{*|neut} = F p_i. \tag{A46}$$

Similarly, using Equation (A31), we obtain

$$\Delta^{neut}\widetilde{D}_{ij,ij} = \frac{F}{1+F}\left[(1-R)\left(\widetilde{D}_{ij,ij} + p_i p_j\right) + R\left(p_i\widetilde{D}_{j,j} + p_j\widetilde{D}_{i,i}\right)\right] - \widetilde{D}_{ij,ij}, \tag{A47}$$

which assuming free recombination ($r = \frac{1}{2}$) and using Equation (A46), yields

$$\Delta^{neut}\widetilde{D}_{ij,ij} = 0 \Leftrightarrow D_{ij,ij}^{*|neut} = \frac{F(1+2F)}{2+F} p_i p_j, \tag{A48}$$

so that at equilibrium, the term $\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}$ can be approximated by

$$\widetilde{D}^*_{ij,ij} - \widetilde{D}^*_{i,i}\widetilde{D}^*_{j,j} \approx \frac{F\,(1-F^2)}{2+F} p_i p_j. \tag{A49}$$

Equilibrium genetic associations to leading order in s. Using Equations (A26) and (A29), we have

$$\Delta\widetilde{D}_{i,i} = \frac{F}{1+F}\left[\widetilde{D}_{i,i} + \bar{s}_z(1-2h)\sum_{j\neq i}\left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}\right) + p_i + \bar{s}_z\left(h\ p_i - (1-h)\widetilde{D}_{i,i}\right)\right] - \widetilde{D}_{i,i}, \tag{A50}$$

which, when replacing $\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}$ with Equation (A49), and neglecting direct selection effects (which is reasonable when the number of mutations segregating is large ; Roze, 2015), yields

$$\Delta\widetilde{D}_{i,i} = 0 \Leftrightarrow \widetilde{D}^*_{i,i} \approx F\left(1 + \bar{s}_z(1-2h)\frac{F\,(1-F^2)}{2+F}\sum_{j\neq i} p_j\right) p_i. \tag{A51}$$

Finally, using Equations (A27) and (A30), and injecting Equation (A49), we have

$$\Delta\widetilde{D}_{ij,j} = \frac{F}{1+F}\left(\widetilde{D}_{ij,j} + \bar{s}_z(1-h)\frac{F\,(1-F^2)}{2+F} p_i\right) - \widetilde{D}_{ij,j}, \tag{A52}$$

which yields, at equilibrium,

$$\Delta\widetilde{D}_{ij,j} = 0 \Leftrightarrow \widetilde{D}^*_{ij,j} = \bar{s}_z(1-h)\frac{F\,(1-F^2)}{2+F} F p_i p_j \tag{A53}$$

Mutation-selection balance. The number of mutations per haploid genome maintained at mutation selection balance is the solution of $\Delta p_i = 0$. However, even after injecting Equations (A49), (A51) and (A53) into Equation (A43), it still cannot be solved explicitly, as it also depends on p_i through lifetime selection coefficients, which are functions of the mean trait of the population. Thus, we have to solve the

system given by Equation (2.7) in the main text numerically, which we do in a *Wolfram Language Script*, that is the script version of a *Mathematica* notebook, available from GitHub (links are available from the journal office).

2.6.3.2 Life cycle approach

In the LC approach, selection at different stages is considered separately, thus we denote the selection coefficients associated with selection phase k, when mutations affect trait z, as $\bar{s}_z^k$ and $\hat{\bar{s}}_z^k$ (FIG. S2). The stages at which selection occurs in our model are gametes production (denoted by $k = g$), juvenile recruitment (denoted $k = j$) and adult survival ($k = s$).

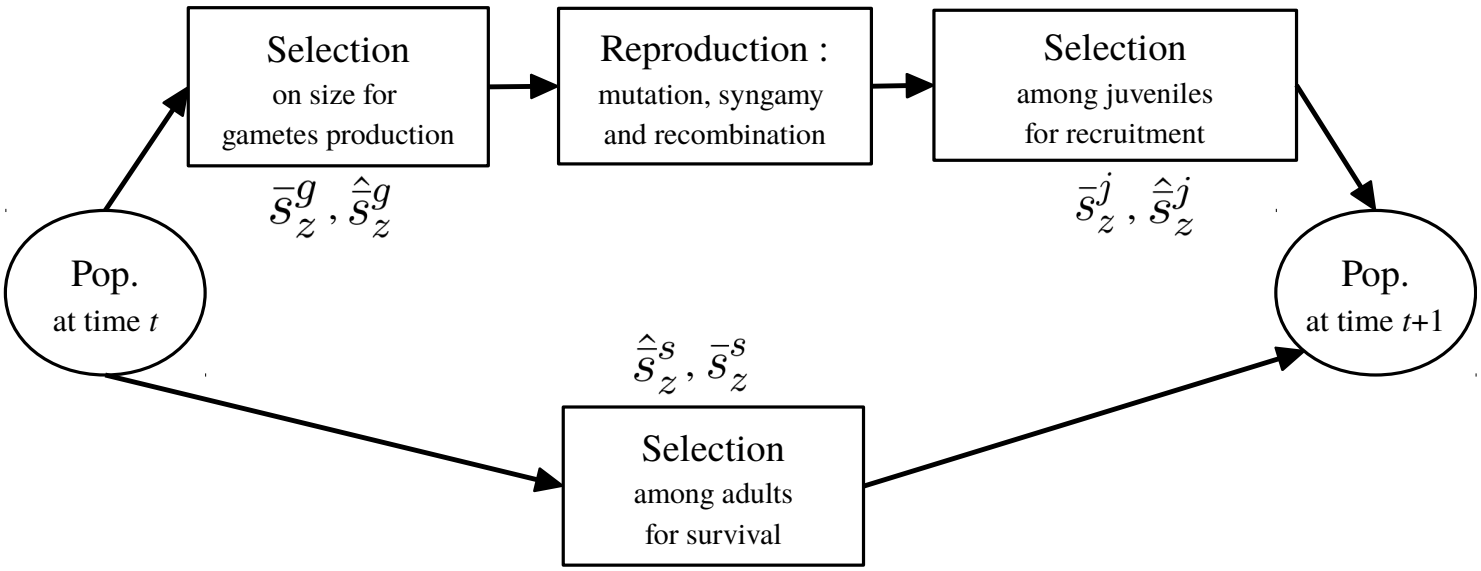

FIGURE S2 – Steps followed over the course of one timestep with the LF approach

Stage-specific selection coefficients

Gametes production. Selection on gametes production occurs in all three models ($k = g$) since fecundity is assumed to be proportional to size. Size at age t (G_t) is the solution of Equation (2.2). It is given by

$$G_t = \frac{e^{-\frac{c}{\varepsilon}(t+1)}}{c^4}\left(e^{\frac{c}{4\varepsilon}(t+1)} - 1\right)^4. \tag{A54}$$

Thus, the average size of an individual given its genotype, which gives its fitness during this selection phase, is

$$w_g = \sum_{t=0}^{\infty} S^t(1-S)\, G_t = \frac{1-S}{c^4}\left(\frac{1}{1-S} - \frac{4}{e^{\frac{c}{4\varepsilon}} - S} + \frac{6}{e^{\frac{c}{2\varepsilon}} - S} - \frac{4}{e^{\frac{3c}{4\varepsilon}} - S} + \frac{1}{e^{\frac{c}{\varepsilon}} - S}\right). \tag{A55}$$

For any trait z, the selection coefficients at this stage, $\bar{s}_z^g$ and $\hat{\bar{s}}_z^g$ are obtained by applying Equation (A20) to Equation (A55).

Mutations affecting growth. In the case of mutations affecting growth, that is affecting the maintenance cost c or the production cost ε, these selection coefficients are equal to the lifetime selection coefficients presented in Appendix 2.6.3.1, that is, they are given by Equations (A36) and (A37) for mutations affecting c, and Equations (A39) and (A40) for mutations affecting ε.

Mutations affecting survival. In the case of mutations affecting survival, the selection coefficients are given by

$$\bar{s}_S^g = -s\frac{S\left(1-\bar{S}\right)\frac{\partial V_S}{\partial S} - \bar{S}V_S}{\left(1-\bar{S}\right)V_S}, \text{ with } V_S = \frac{1}{1-\bar{S}} - \frac{4}{e^{\frac{c}{4\varepsilon}} - \bar{S}} + \frac{6}{e^{\frac{c}{2\varepsilon}} - \bar{S}} - \frac{4}{e^{\frac{3c}{4\varepsilon}} - \bar{S}} + \frac{1}{e^{\frac{c}{\varepsilon}} - \bar{S}}, \tag{A56}$$

and,

$$\begin{aligned}\hat{\bar{s}}_S^g = -\frac{s^2}{\left(1-\bar{S}\right)V_S}\Bigg[&\bar{S}\left(\frac{V_S}{2} + S\frac{\partial V_S}{\partial S}\right)\\ &- \bar{S}\left(1-\bar{S}\right)\left(\frac{1+\bar{S}}{2\left(1-\bar{S}\right)^3} - 2\frac{e^{\frac{c}{4\varepsilon}} + \bar{S}}{\left(e^{\frac{c}{4\varepsilon}} - \bar{S}\right)^3} + 3\frac{e^{\frac{c}{2\varepsilon}} + \bar{S}}{\left(e^{\frac{c}{2\varepsilon}} - \bar{S}\right)^3} - 2\frac{e^{\frac{3c}{4\varepsilon}} + \bar{S}}{\left(e^{\frac{3c}{4\varepsilon}} - \bar{S}\right)^3} + \frac{e^{\frac{c}{\varepsilon}} + \bar{S}}{\left(e^{\frac{c}{\varepsilon}} - \bar{S}\right)^3}\right)\Bigg].\end{aligned} \tag{A57}$$

Survival and recruitment. Selection on survival occurs in juveniles ($k = j$) and in adults ($k = s$). Individual fitnesses during survival selection stages are given by their survival probability. Hence, when mutations affect the maintenance or the production cost (*i.e.* not survival), we have $\bar{s}_c^j = \bar{s}_c^s = \hat{\bar{s}}_c^j = \hat{\bar{s}}_c^s = 0$ and $\bar{s}_\varepsilon^j = \bar{s}_\varepsilon^s = \hat{\bar{s}}_\varepsilon^j = \hat{\bar{s}}_\varepsilon^s = 0$. On the other hand, when mutations affect survival given by $\bar{s}_S^j = s$ and $\hat{\bar{s}}_S^j = s^2$ in juveniles, and $\bar{s}_S^s = s$ and $\hat{\bar{s}}_S^s = s^2$ in adults.

Quasi-linkage equilibrium results

Allelic frequency change. In the LC approach, we distinguish between selection occuring in juveniles and in adults, so that the change in allelic frequency at the ith locus, Δp_i, is given by

$$\Delta p_i = (1 - \bar{S})\Delta^{juv} p_i + \bar{S}\Delta^{surv} p_i, \tag{A58}$$

where $\bar{S} = S$ when mutations affect growth, and depends on the genetic composition of the population when they affect survival.

Among juveniles, we first have selection for gametes production, which is followed by reproduction, and selection for recruitment. Applying Equations (A23) and (A28), to these successive stages, one obtains, to second order in s,

$$\begin{aligned}\Delta^{juv} p_i = U(1 + \bar{s}_z^j h) - \left[\bar{s}_z^g h + \bar{s}_z^j \left(h + (1-h)\frac{F}{1+F}\right) + \bar{s}_z^g \bar{s}_z^j h \left(h + (1-h)\frac{F}{1+F}\right)\right] p_i \\ + \left[\bar{s}_z^g + \bar{s}_z^j \frac{F}{1+F} + \bar{s}_z^g \bar{s}_z^j \frac{h+F}{1+F}\right](1-h)\widetilde{D}_{i,i} \\ + \left[\bar{s}_z^g + \bar{s}_z^j \frac{F}{1+F}\right](1-2h)\sum_{j\neq i}\widetilde{D}_{ij,j} \\ + \left[\hat{\bar{s}}_z^j + \hat{\bar{s}}_z^g + 2\bar{s}_z^g \bar{s}_z^j \frac{F}{1+F}\right](1-h)(1-2h)\sum_{j\neq i}\left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i}\widetilde{D}_{j,j}\right).\end{aligned} \tag{A59}$$

Besides, among surviving adults, only selection for survival occurs, so that to second

order in s we have

$$\Delta^{surv} p_i = \bar{s}_z^s h p_i + \bar{s}_z^s (1-h) \widetilde{D}_{i,i} + \bar{s}_z^s (1-2h) \sum_{j \neq i} \widetilde{D}_{ij,j} + \hat{\bar{s}}_z^s (1-h)(1-2h) \sum_{j \neq i} \left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i} \widetilde{D}_{j,j} \right). \tag{A60}$$

Putting Equations (A59) and (A60) together, this yields

$$\begin{aligned}
\Delta p_i = & \left(1 - \bar{S}\right) U (1 + \bar{s}_z^j h) \\
& - \left\{ \bar{S}\ \bar{s}_z^s h + \left(1 - \bar{S}\right) \left[\bar{s}_z^g h + \bar{s}_z^j \left(h + (1-h) \frac{F}{1+F} \right) + \bar{s}_z^g \bar{s}_z^j h \left(h + (1-h) \frac{F}{1+F} \right) \right] \right\} p_i \\
& + \left[\bar{S}\ \bar{s}_z^s + \left(1 - \bar{S}\right) \left(\bar{s}_z^g + \bar{s}_z^j \frac{F}{1+F} + \bar{s}_z^g \bar{s}_z^j \frac{h+F}{1+F} \right) \right] (1-h) \widetilde{D}_{i,i} \\
& + \left[\bar{S}\ \bar{s}_z^s + \left(1 - \bar{S}\right) \left(\bar{s}_z^g + \bar{s}_z^j \frac{F}{1+F} \right) \right] (1-2h) \sum_{j \neq i} \widetilde{D}_{ij,j} \\
& + \left[\bar{S}\ \hat{\bar{s}}_z^s + \left(1 - \bar{S}\right) \left(\hat{\bar{s}}_z^j + \hat{\bar{s}}_z^g + 2 \bar{s}_z^g \bar{s}_z^j \frac{F}{1+F} \right) \right] (1-h)(1-2h) \sum_{j \neq i} \left(\widetilde{D}_{ij,ij} - \widetilde{D}_{i,i} \widetilde{D}_{j,j} \right).
\end{aligned} \tag{A61}$$

Genetic associations at quasi-linkage equilibrium. Similar to the LF approach, we know need to derive QLE approximations for genetic associations. In the case of associations derived under neutrality, the associations remain unchanged so that we still have

$$\widetilde{D}^*_{ij,ij} - \widetilde{D}^*_{i,i} \widetilde{D}^*_{j,j} \approx \frac{F\left(1 - F^2\right)}{2 + F} p_i p_j. \tag{A62}$$

In the case of genetic associations affected by selection on the other hand, results differ from the LF approach. By the same logic as for allelic frequency change, the changes in these associations over the course of one timestep are given by

$$\Delta \widetilde{D}_{i,i} = (1-\bar{S}) \Delta^{juv} \widetilde{D}_{i,i} + \bar{S} \Delta^{surv} \widetilde{D}_{i,i}, \quad \text{and,} \quad \Delta \widetilde{D}_{ij,j} = (1-\bar{S}) \Delta^{juv} \widetilde{D}_{ij,j} + \bar{S} \Delta^{surv} \widetilde{D}_{ij,j}. \tag{A63}$$

Applying the methods described in Appendix 2.6.2 to the case of association $\widetilde{D}_{i,i}$, we have

$$\Delta^{juv}\widetilde{D}_{i,i} = \frac{Fp_i - \widetilde{D}_{i,i}}{1+F} + \left[\bar{s}_z^j + \hat{s}_z^g \frac{F}{1+F}\right](1-2h)\frac{F\left(1-F^2\right)}{2+F} p_i \sum_{j\neq i} p_j, \tag{A64}$$

and,

$$\Delta^{surv}\widetilde{D}_{i,i} = \bar{s}_z^s (1-2h)\frac{F\left(1-F^2\right)}{2+F} p_i \sum_{j\neq i} p_j, \tag{A65}$$

which yields

$$\begin{aligned}\Delta\widetilde{D}_{i,i} &= \left(1-\bar{S}\right)\frac{Fp_i - \widetilde{D}_{i,i}}{1+F} + \left[\bar{S}\ \bar{s}_z^s + \left(1-\bar{S}\right)\left(\bar{s}_z^j + \hat{s}_z^g \frac{F}{1+F}\right)\right](1-2h)\frac{F\left(1-F^2\right)}{2+F} p_i \sum_{j\neq i} p_j, \\ \Delta\widetilde{D}_{i,i} = 0 \Leftrightarrow \widetilde{D}_{i,i}^* &= \left\{F + \left[F\ \bar{s}_z^g + (1+F)\left(\bar{s}_z^j + \frac{\bar{S}}{1-\bar{S}}\bar{s}_z^s\right)\right](1-2h)\frac{F\left(1-F^2\right)}{2+F}\sum_{j\neq i} p_j\right\} p_i.\end{aligned} \tag{A66}$$

As for the $\widetilde{D}_{ij,j}$ association, we have

$$\Delta^{juv}\widetilde{D}_{ij,j} = \frac{\widetilde{D}_{ij,j}}{1+F} + \left(\bar{s}_z^j + \frac{F}{1+F}\bar{s}_z^g\right)(1-h)\frac{F\left(1-F^2\right)}{2+F} p_i p_j, \tag{A67}$$

and,

$$\Delta^{surv}\widetilde{D}_{ij,j} = \hat{s}_z^s (1-h)\frac{F\left(1-F^2\right)}{2+F} p_i p_j, \tag{A68}$$

which yields

$$\begin{aligned}\Delta\widetilde{D}_{ij,j} &= \left(1-\bar{S}\right)\frac{\widetilde{D}_{ij,j}}{1+F} + \left[\bar{S}\ \bar{s}_z^s + \left(1-\bar{S}\right)\left(\bar{s}_z^j + \frac{F}{1+F}\bar{s}_z^g\right)\right](1-h)\frac{F\left(1-F^2\right)}{2+F} p_i p_j, \\ \Delta\widetilde{D}_{ij,j} = 0 \Leftrightarrow \widetilde{D}_{ij,j}^* &= \left[F\ \bar{s}_z^g + (1+F)\left(\bar{s}_z^j + \frac{\bar{S}}{1-\bar{S}}\bar{s}_z^s\right)\right](1-h)\frac{F\left(1-F^2\right)}{2+F} p_i p_j\end{aligned} \tag{A69}$$

Mutation-selection balance. Solving $\Delta p_i = 0$ is done numerically in the same way as for the LF approach, using a *Wolfram Language Script* available from GitHub

(links are available from the journal office).

2.6.3.3 Inbreeding depression and mutation load

In this section, we describe how an approximation for inbreeding depression and for the mutation load (Crow, 1958) can be obtained given the average number of mutations per haploid genome.

Inbreeding depression, δ. Inbreeding depression is defined as the relative decrease in lifetime fitness of selfed individuals relative to outcrossed ones (Charlesworth and Charlesworth, 1987). In our case, lifetime fitness is directly dependent individuals' value for the trait z, through Equation (A35).

$$\delta = 1 - \frac{W(\bar{z}_{self})}{W(\bar{z}_{out})}. \tag{A70}$$

Thus, to compute an approximation for inbreeding depression, we need to obtain the average trait of selfed and outcrossed individuals, $\bar{z}_{self}$ and $\bar{z}_{out}$. In the general population, the average of the trait is given by Equation (A19), that is

$$\bar{z} \approx z_0 \ \exp\left[\pm s\left(2h\sum_i p_i + (1-2h)\sum_i \widetilde{D}_{i,i}\right)\right]. \tag{A71}$$

Among the outcrossed, we have $\widetilde{D}_{i,i}^{out} = 0$ since they are produced by random mating. Hence,

$$\bar{z}_{out} \approx z_0 \exp\left(\pm 2sh\sum_i p_i\right) \tag{A72}$$

On the other hand, $\widetilde{D}_{i,i}^{self}$ is given by

$$\widetilde{D}_{i,i}^{self} = \mathbb{E}\left[\frac{1}{4}\zeta_{ii,\varnothing} + \frac{1}{2}\zeta_{i,i} + \frac{1}{4}\zeta_{\varnothing,ii}\right] = \frac{1}{2}\left(p_i + \widetilde{D}_{i,i}\right) \approx \frac{1}{2}\left(1+F\right)p_i, \tag{A73}$$

so that,

$$\bar{z}_{self} \approx z_0 \exp\left[\pm s\left(2h + \frac{1}{2}(1-2h)\,(1+F)\right)\sum_i p_i\right] = z_0 \exp\left[\pm s(1-2h)\frac{1+F}{2}\sum_i p_i\right]. \tag{A74}$$

Mutation Load, L. The mutation load was defined by Crow (1958) as the decrease in mean fitness of the population compared to an optimal population bearing no mutation. Hence, in our model it is given by

$$L = 1 - \frac{W(\bar{z})}{W(z_0)}, \tag{A75}$$

where $W(z)$ is the lifetime fitness given by Equation (A35).

2.6.4 Fitness landscapes

In FIGURE S3, lifetime fitness is presented as a function of phenotypic deviation ΔP. For each trait z, we have $z = z_0\,(1 + \Delta P)$. The resulting fitness landscape for each set of parameters is rescaled by its maximal value, which is always obtained for $\Delta P = 0$.

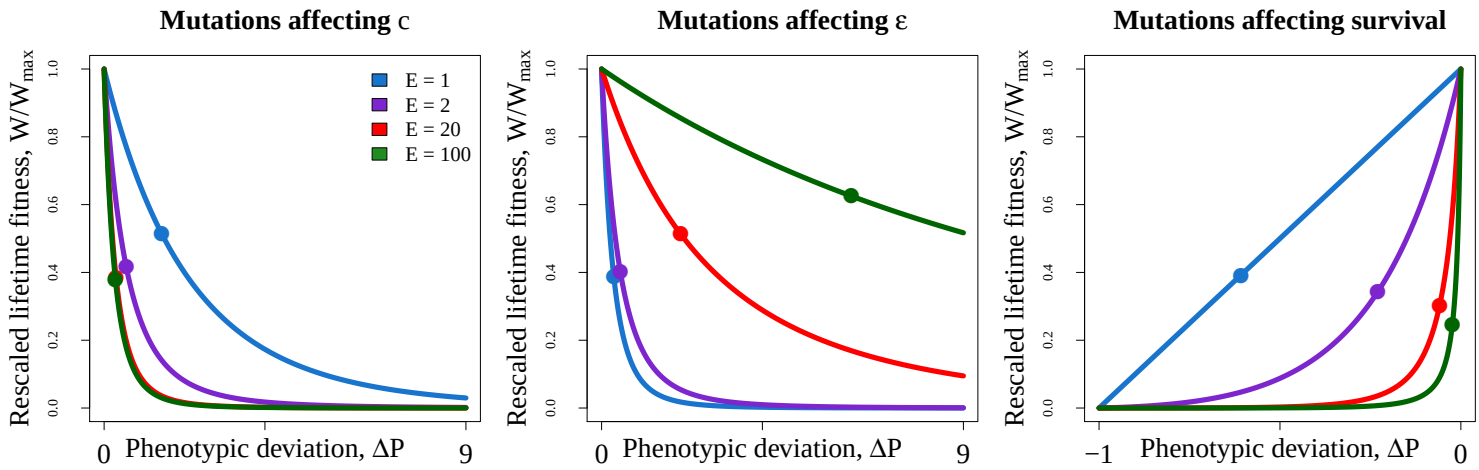

FIGURE S3 – Rescaled fitness landscapes for each mutation type, for initial life expectancies ($E = 1;\ 2;\ 20;\ 100$). Lines depict fitness, while dots depict the mutation-selection equilibrium fitness reached for $\frac{c}{\varepsilon} = 1$, $h = 0.25$, $s = 0.005$, and $U = 0.5$.

2.6.5 Comparison of mean phenotypic deviation for different magnitude of effect of mutations

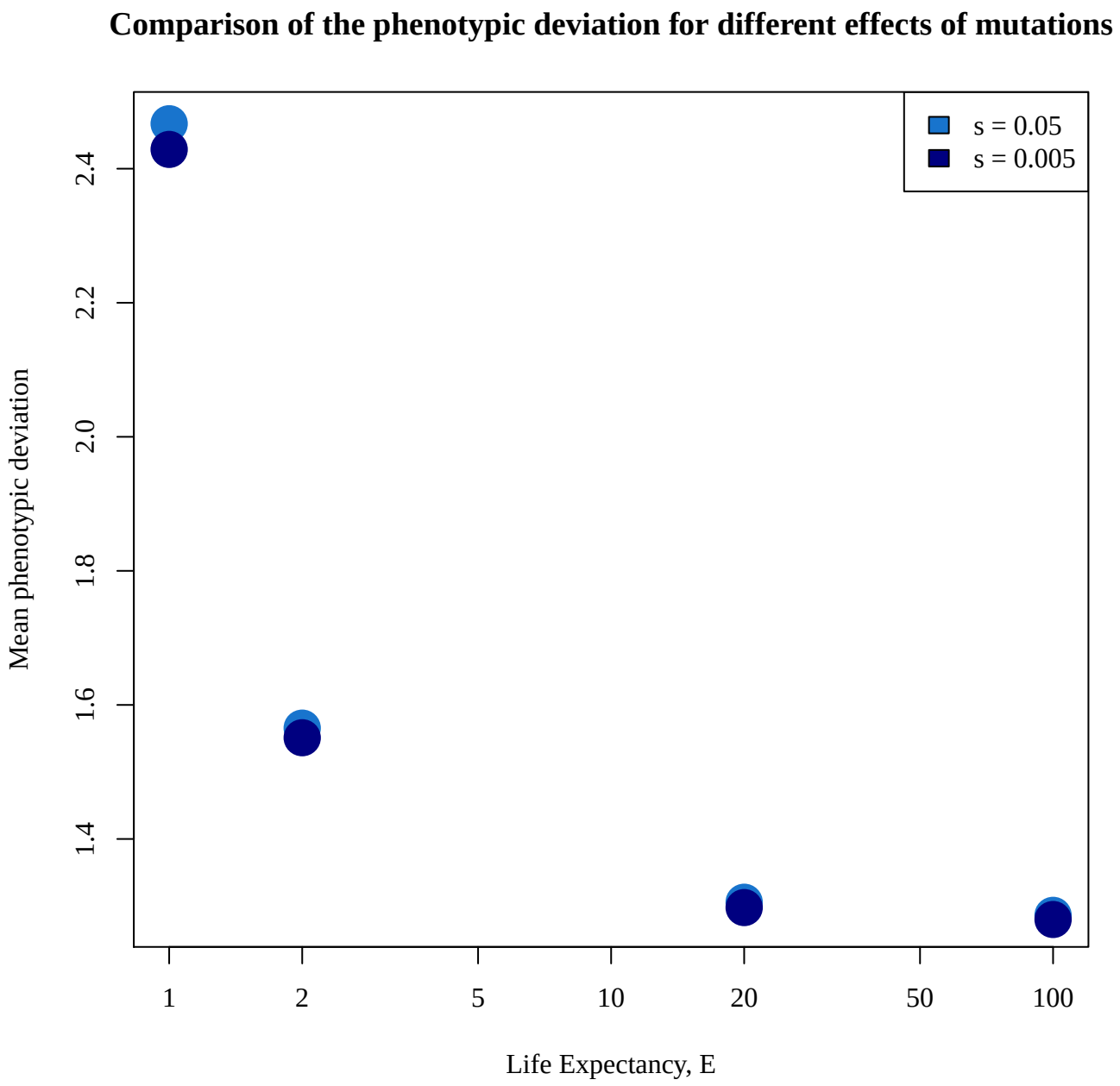

FIGURE S4 – Comparison of the mean phenotypic deviation, for mutations affecting the maintenance cost (c), for two different magnitude of effect of mutations ($s = 0.05$ and $s = 0.005$). Parameters used are $\frac{c}{\varepsilon} = 1$, $h = 0.25$, $U = 0.5$.

2.6.6 Mutations per haploid genome, inbreeding depression and mutation load for other parameter values

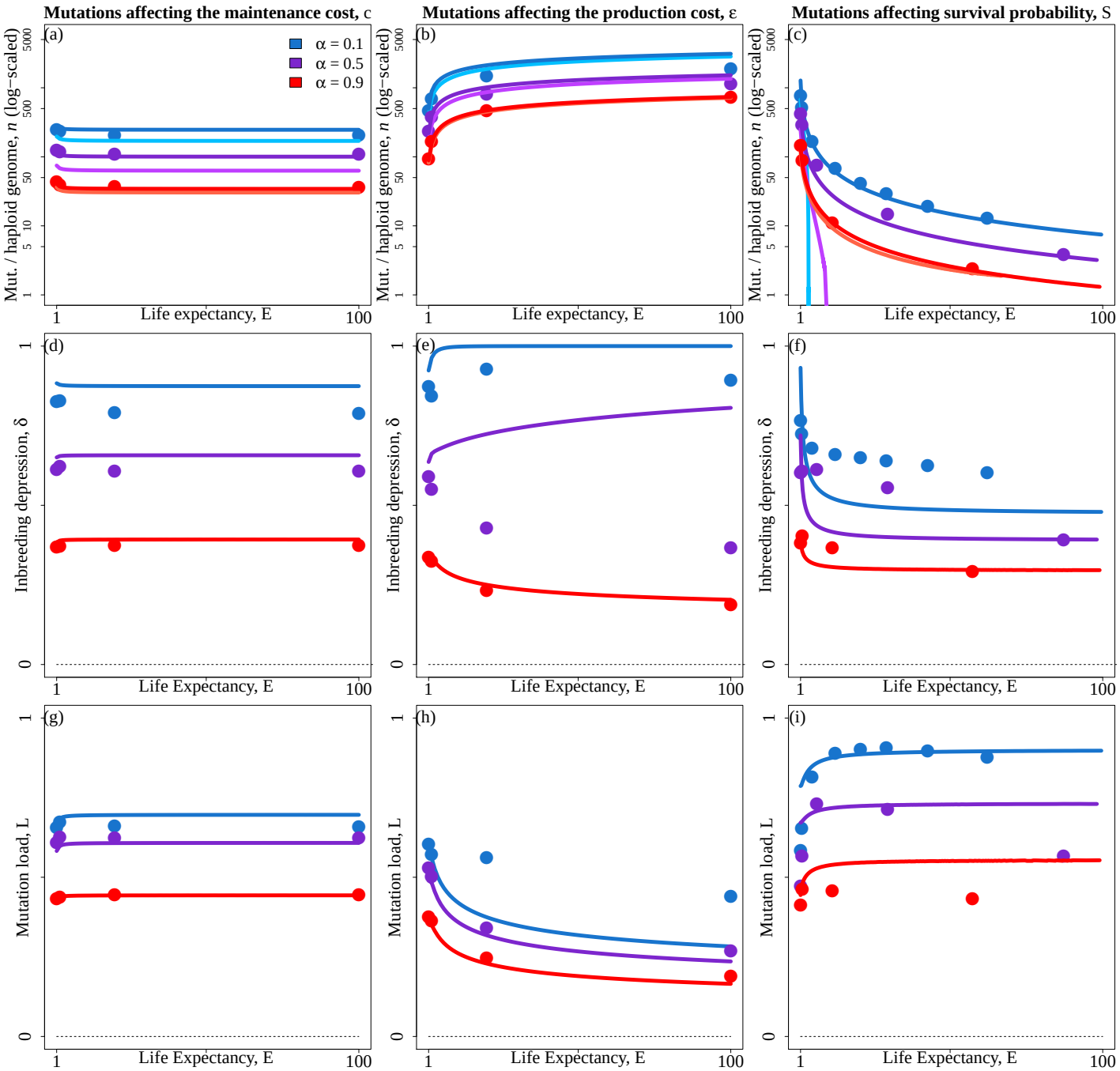

FIGURE S5 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 10$, $U = 0.5$, $s = 0.005$, $h = 0.10$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

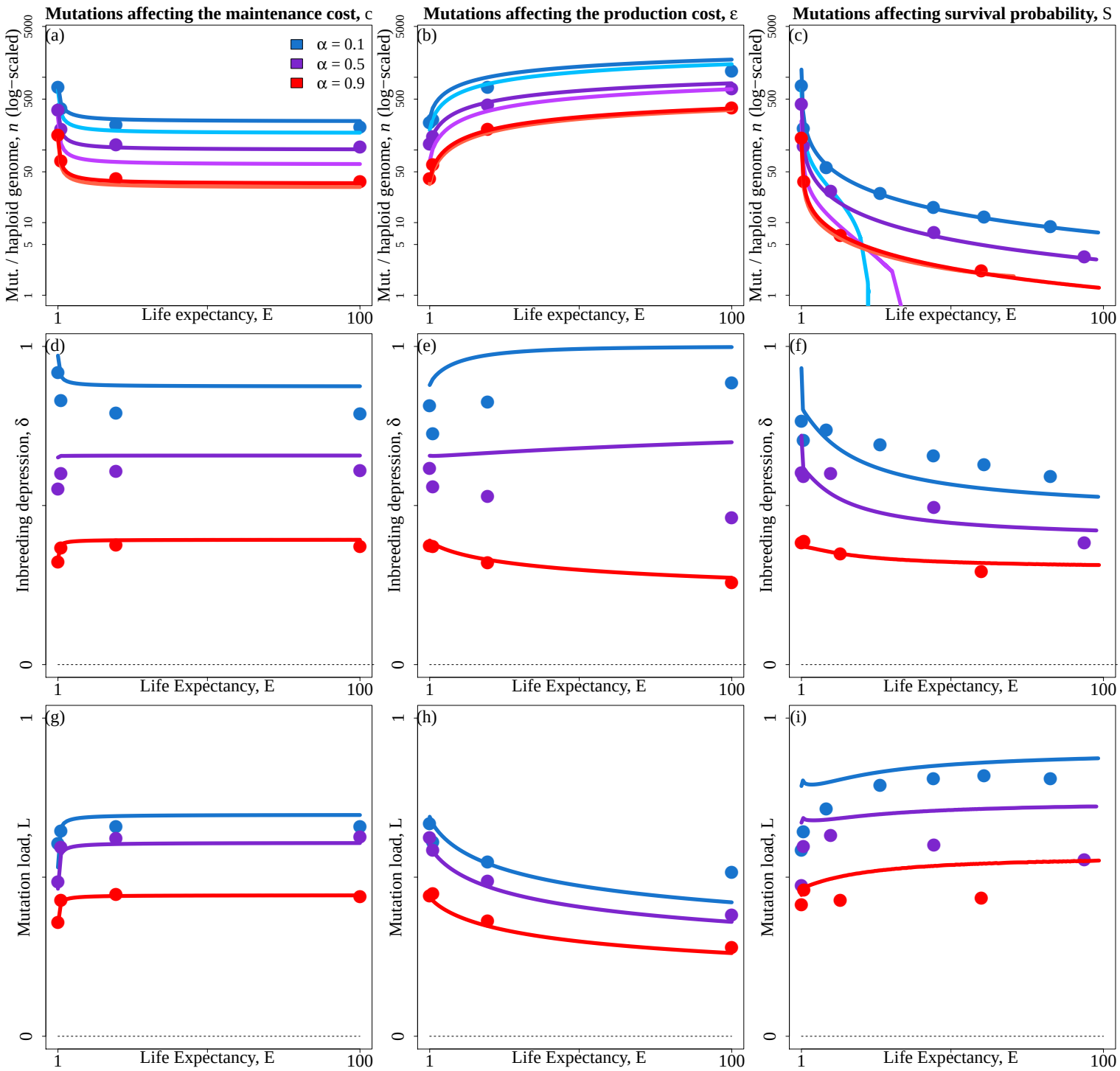

FIGURE S6 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 1$, $U = 0.5$, $s = 0.005$, $h = 0.10$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

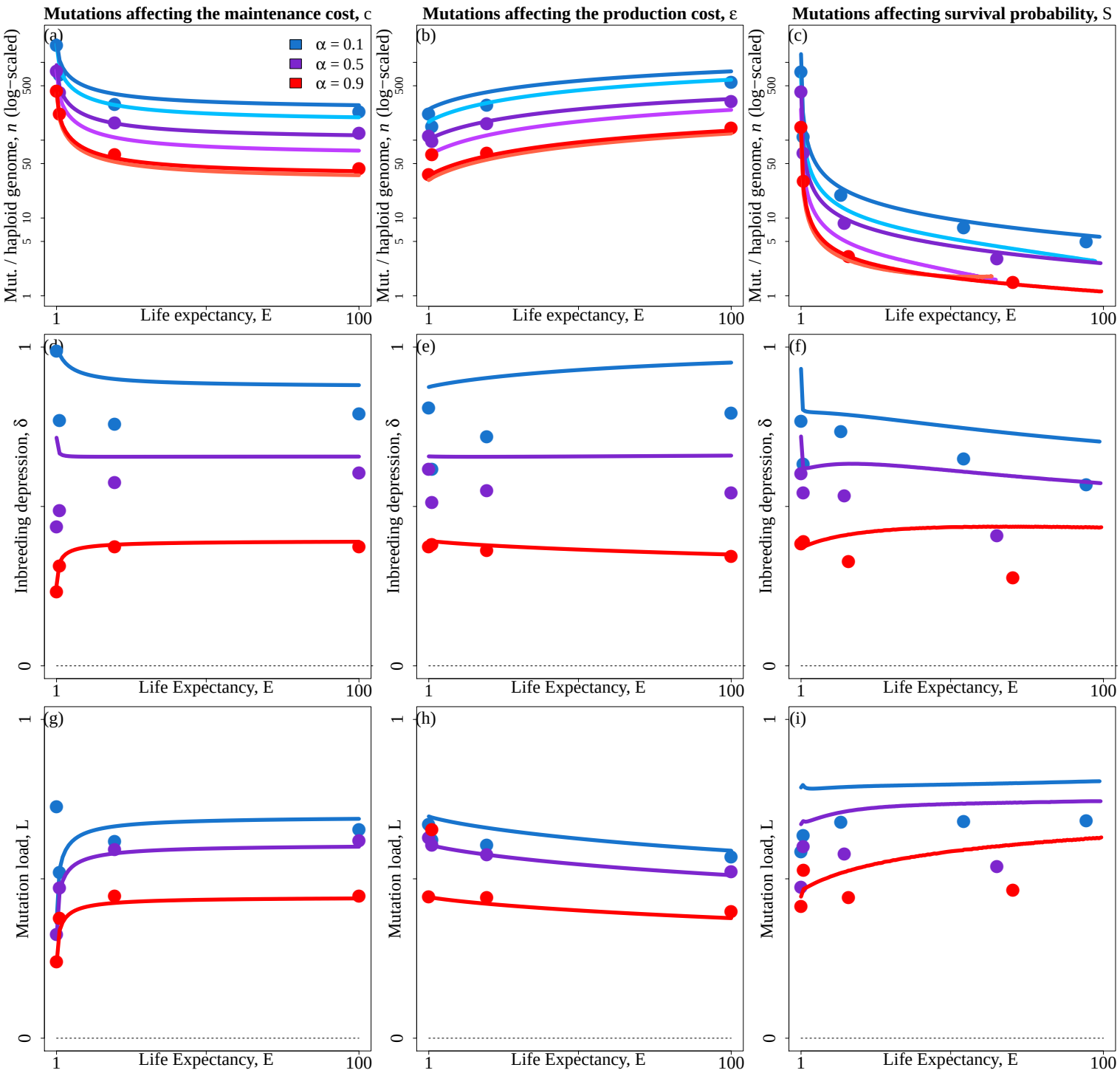

FIGURE S7 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 0.1$, $U = 0.5$, $s = 0.005$, $h = 0.10$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

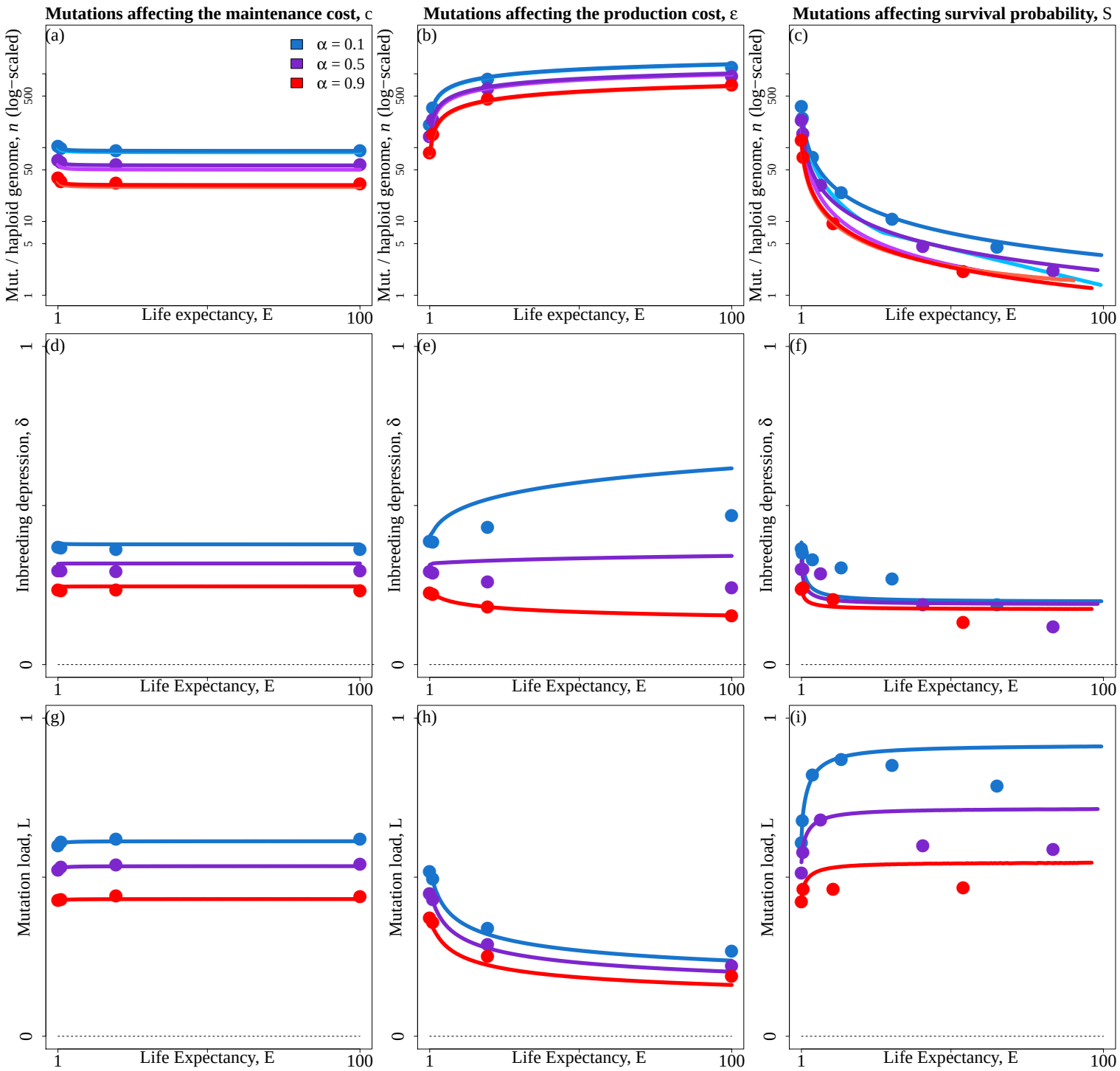

FIGURE S8 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 10$, $U = 0.5$, $s = 0.005$, $h = 0.25$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

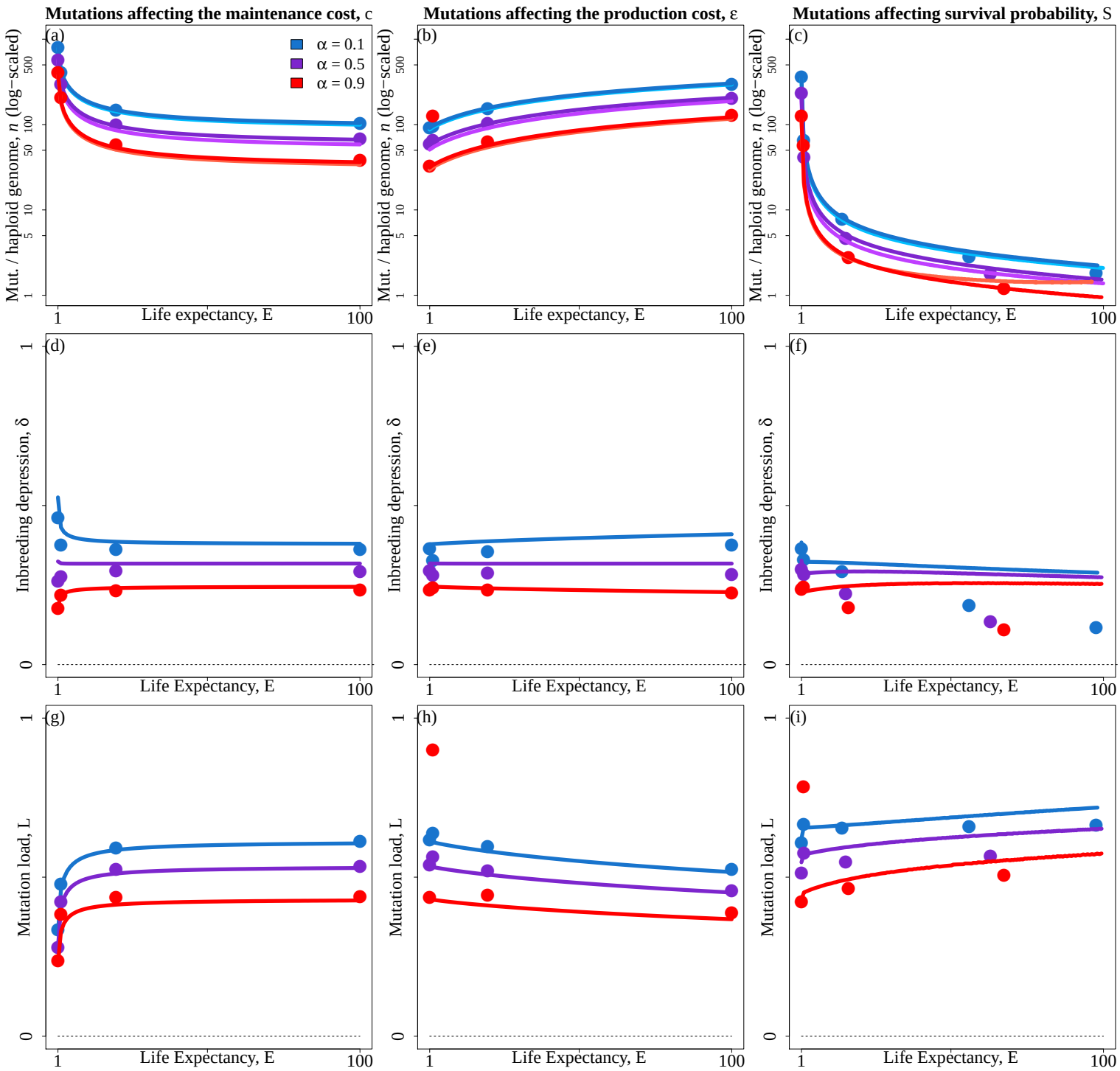

FIGURE S9 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 0.1$, $U = 0.5$, $s = 0.005$, $h = 0.250$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

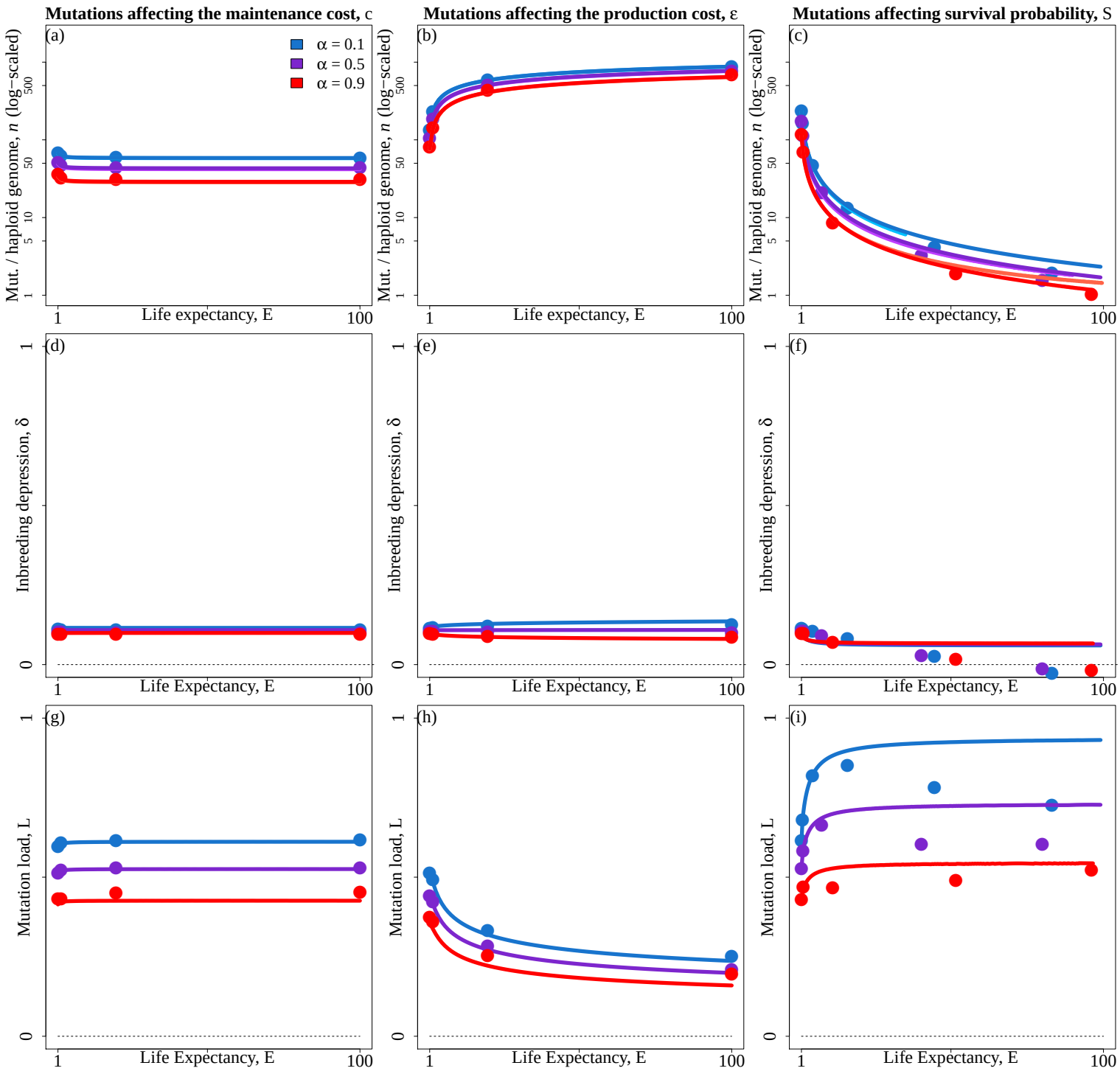

FIGURE S10 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 10$, $U = 0.5$, $s = 0.005$, $h = 0.40$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

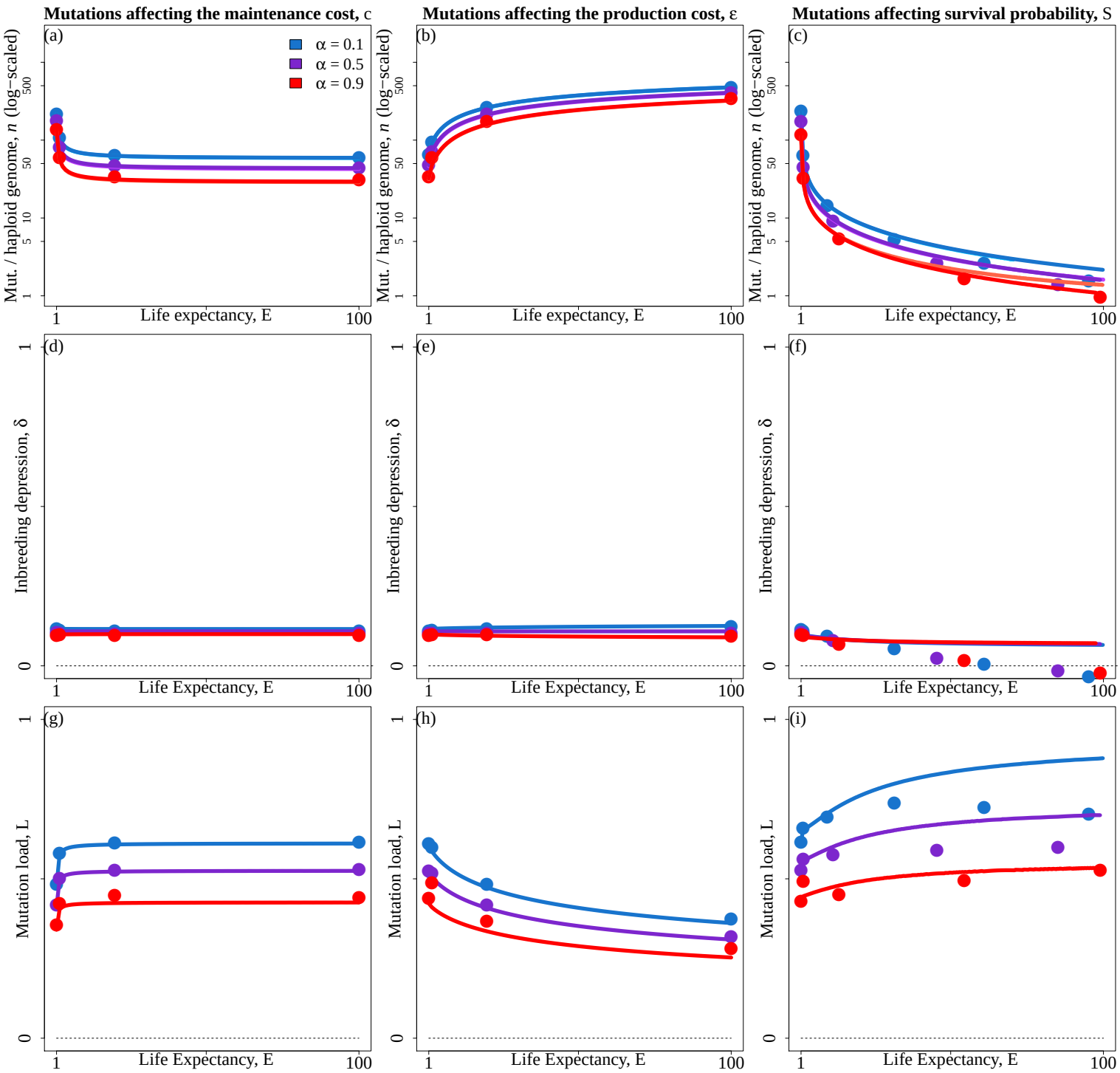

FIGURE S11 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 1$, $U = 0.5$, $s = 0.005$, $h = 0.40$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

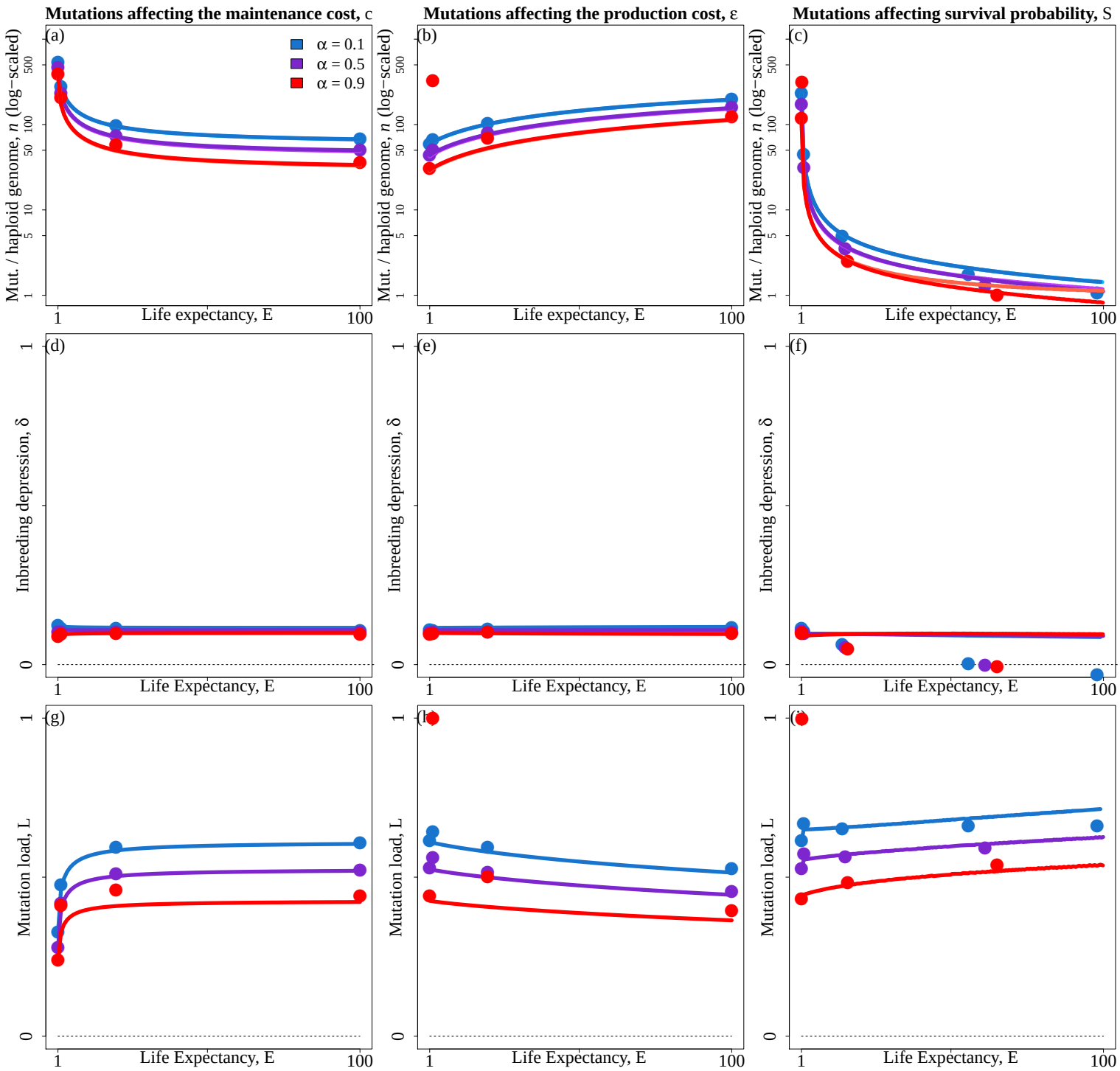

FIGURE S12 – Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of life expectancy (E), for three selfing rates : $\alpha = 0.1$ (blue), $\alpha = 0.5$ (purple) and $\alpha = 0.9$ (red). Each column corresponds to one type of mutation. Dots : simulation results. Lighter lines : LF approach predictions. Darker lines : LC approach predictions. Parameters shown here are $\frac{c}{\varepsilon} = 0.1$, $U = 0.5$, $s = 0.005$, $h = 0.40$. Five replicates were run for each parameter set and results were averaged over all of them. Standard deviation bars are plotted but are hidden behind the dots depicting the averages.

Chapitre 3

Population-level consequences of inheritable somatic mutations and the evolution of mutation rates in plants

Auteur : Thomas Lesaffre

Statut de publication : Le travail présenté dans ce chapitre est disponible sur *BioRXiv* (lien).

3.1 Synthèse en français

3.1.1 Motivation

Parce que la lignée germinale n'est pas séparée de la lignée somatique chez les végétaux, Scofield and Schultz (2006) ont proposé que les niveaux de dépression de consanguinité plus élevés observés chez les espèces plus longévives et de plus grandes

tailles pouvaient être expliqués par une accumulation plus importante de mutations somatiques héritables chez ces espèces. Cette idée et les conséquences populationnelles d'une telle accumulation d'une manière générale n'ont toutefois jamais été considérées du point de vue théorique. Un déterminant majeur de l'importance de ce mécanisme est le taux de mutation, dont l'évolution chez les plantes, prenant en compte la non-séparation de la lignée germinale, n'a pas non plus reçu de traitement théorique. L'objectif de ce chapitre est de produire le modèle le plus simple possible d'accumulation de mutations somatiques afin d'abord de considérer l'évolution du taux de mutation suivant le degré de pérennialité, puis d'étudier les propriétés génétiques de la population à l'équilibre mutation-sélection en prenant en compte cette évolution.

3.1.2 Modèle

Démographie & modèle de développement. Le modèle de développement le plus simple possible est considéré ici. Les individus sont consitués d'une unique tige, qui gagne une section entre chaque saison de floraison. De la même manière que dans les modèles précédents, ils sont diploïdes et hermaphrodites, et survivent d'une saison de floraison à l'autre avec une probabilité constante. Les juvéniles ne sont recrutés qu'en remplacement des adultes décédés, de sorte que la population conserve une taille constante supposée grande. Chacune des sections qui constituent un individu produit des gamètes et contribue donc à la production de graines. Un individu se reproduit partiellement par autofécondation, pour partie obligatoirement au sein de la même section et aléatoirement entre les sections qui le constituent pour le reste.

Modèle génétique. Chaque nouvelle section produite acquiert des mutations délétères récessives à une infinité de loci bialléliques affectant sa fécondité, c'est-à-dire sa contribution au pool de gamètes, à un taux qui dépend du génotype de l'individu à un locus modifieur qui ne mute, lui, que pendant la méiose. Des mutations délétères

se produisent également durant la méiose. Cependant, cette dernière faisant intervenir des mécanismes génétiques différents, susceptibles d'augmenter le taux de mutation par rapport aux mitoses exécutées pendant la croissance, comme par exemple la production de cassures double brins qui sont des événements particulièrement mutagènes, une constante de proportionnalité est introduite entre le taux de mutation méiotique et mitotique pour décrire la mutagénicité relative de la méiose par rapport à la croissance. À la suite de différents auteurs (Kimura, 1967; Gervais and Roze, 2017), nous introduisons également une fonction de coût à la fidélité de réplication qui réduit la fécondité d'un individu d'autant plus fortement que son taux de mutation méiotique est faible. Ainsi, l'évolution du taux de mutation est le résultat de l'interaction entre le coût à la fidélité de réplication d'une part et la sélection indirecte exercée par les mutations délétères ségrégeant aux autres loci sur le modifieur d'autre part.

3.1.3 Résultats clefs

Évolution du taux de mutation. Les espèces plus longévives atteignent des tailles plus élevées, et présentent de ce fait davantage d'opportunités pour accumuler des mutations somatiques et les transmettre à leur descendance. Ainsi, un allèle augmentant le taux de mutation a tendance à être associé avec davantage de mutations délétères chez ces espèces, et la sélection indirecte exercée par ces dernières induit une diminution du taux de mutation à l'ESS. Le taux de mutation *par événement mutagène* est donc plus faible chez les espèces plus longévives. Cependant, du fait de leurs plus grandes opportunités pour accumuler des mutations somatiques, le taux de mutation *total*, c'est-à-dire le taux auquel des mutations apparaissent dans la population à la fois par mitose et par méiose, demeure plus élevé chez ces espèces.

Propriétés de la population à l'équilibre mutation-sélection. En cohérence avec l'augmentation du taux de mutation total avec la pérennialité, le nombre de mutations

maintenues à l'équilibre mutation-sélection est plus élevé chez les espèces plus longévives, menant de ce fait à une dépression de consanguinité plus élevée chez ces espèces. Néanmoins, la magnitude de l'augmentation du taux de mutation total et donc de la dépression de consanguinité avec la pérennialité dépend de la mutagénicité relative de la méiose et des mitoses se produisant pendant la croissance. En effet, plus cette dernière est élevée et plus l'augmentation de la dépression de consanguinité avec la pérennialité est ténue. Ce résultat est expliqué par le plus faible taux de mutation par événement mutagène obtenu à l'ESS dans ce cas. Ainsi, les résultats de ce modèle suggèrent que l'accumulation de mutations somatiques ne pourra jouer un rôle important dans l'augmentation observée de la dépression de consanguinité avec la pérennialité que si les mitoses se produisant au cours de la croissance sont suffisamment mutagène par rapport à la méiose. Ces conclusions reposent toutefois sur un modèle très simpliste où la différence de production de mutations pendant la méiose et la croissance n'est contrôlée que par une simple constante de proportionnalité. Afin de mieux comprendre l'importance des mutations somatiques dans la dynamique génétique des espèces pérennes, il faudra s'attacher à étudier et modéliser les mécanismes qui pourraient générer une différence de mutagénicité de ces processus, en repartant des caractéristiques de la croissance et de la division cellulaire chez les plantes.

3.2 Introduction

Plant growth is fueled by cell divisions occuring in meristems. Each shoot is produced by an apical meristem and may bear axillary meristems, which are typically situated in the axils of leaves and grow out to become the apical meristem of a new shoot upon activation (Burian et al., 2016). As meristematic cells generate all the tissues constituting the shoot, any mutation occuring in a meristematic cell will be borne by all the cells it gave rise to, leading to genetic mosaicism within individual plants. Fur-

thermore, because meristems also give rise to reproductive tissues, mutations occuring during growth before the differentiation of the germline, that is somatic mutations, may be present in the gametes and hence be inherited (Lanfear, 2018). All else being equal, it follows that the larger and the older a given plant grows, the more somatic mutations it should accumulate and transmit to its offspring, potentially leading to a higher mutation load in more long-lived and larger species since it is thought that most mutations are deleterious (Eyre-Walker and Keightley, 2007).

Inbreeding depression, that is the decrease in fitness of inbred relative to outbred individuals (Charlesworth and Charlesworth, 1987), is thought to be mostly generated by recessive deleterious mutations maintained at mutation-selection balance in populations (Charlesworth and Willis, 2009). Hence, Scofield and Schultz (2006) proposed that somatic mutations accumulation could lead to higher inbreeding depression in larger and more long-lived species. Consistent with this view, inbreeding depression was indeed shown to increase strongly as life expectancy increases across plant species (Duminil et al., 2009; Angeloni et al., 2011). Furthermore, Bobiwash et al. (2013) showed that substantial inbreeding depression was generated by somatic mutations in a study performed at the phenotypic level in old *Vaccinium angustifolium* clones. To my knowledge, this is, however, the only empirical test of Scofield and Schultz (2006)'s idea. Besides, recent theoretical investigations have shown that variations in inbreeding depression can in principle be generated by differences in the fitness effect of mutations between species with contrastring life-histories (Lesaffre and Billiard, 2020), so that somatic mutations accumulation may not always be needed to explain variations in the magnitude of inbreeding depression across plant species. Moreover, theoretical investigations of the population-level consequences of somatic mutations accumulation are lacking, so that their role in the maintenance of high inbreeding depression in long-lived species remains poorly understood. Indeed, theoretical studies regarding somatic mutations in plants either focused on the case of favorable mutations, conferring resistance

against herbivores (e.g. Antolin and Strobeck, 1985), or studied the fate of deleterious mutations subject to intra-organismal selection (Otto and Orive, 1995; Pineda-Krch and Lehtilä, 2002), but never considered the population-level consequences of recessive deleterious mutations (Schoen and Schultz, 2019). In summary, deleterious somatic mutations accumulation has been proposed as a mechanism to explain the rarity of selfing species among long-lived plants (Scofield and Schultz, 2006), consistent with empirical measures of inbreeding depression, but theoretical support for this idea remains scarce.

An important determinant of the consequences of somatic mutations accumulation is the rate at which said mutations accumulate during growth, that is the somatic mutation rate, which is defined here as the number of mutations occuring per unit of vegetative growth. This rate is likely influenced by evolutionary mechanisms similar to those affecting mutation rates in general. For example, Kimura (1967) showed that mutation rates should be shaped by the opposition between the increase in the number of deleterious mutations borne by individuals with higher mutation rates on the one hand, which causes indirect selection against genetic variants increasing mutation rates to increase, and the direct fitness cost there is to increasing the fidelity of DNA replication on the other hand. Besides, Lynch (2011) proposed that selection to decrease the mutation rate should become weaker than genetic drift at some point in finite populations, thereby favoring the persistence of non-zero mutation rates. Nevertheless, the inheritability of somatic mutations in plants and their intrinsic link with growth and life expectancy likely contribute to shape the evolution of mutation rates in a specific manner which was never tackled theoretically. Great interest was however taken in empirically detecting somatic mutations and comparing mutations rates in a variety of plants species ranging from the very short-lived *Arabidopsis thaliana* to ancient, centuries old trees. In an analysis performed across many plant families, Lanfear et al. (2013) showed that taller species among pairs of sister species have signficantly

lower rates of molecular evolution, measured as the number of substitutions per site per 10^6 years. They argued that contrary to animals, this pattern is not a mere reflection of differences in generation time, which would reflect different rates of genome copying per unit of time, because somatic genome copying events contribute to the inheritable genetic variation in plants. Instead, they proposed that this pattern may be due to slower growth in taller species, which results in a lower number of mitosis (and therefore mutations) per unit of time. Consistent with this view, it was shown at the cellular level that axillary meristems cells are set aside early during the growth of a shoot (Burian et al., 2016), resulting in a number of cell divisions increasing linearly with the number of branching events in trees although the number of terminal branches increases exponentially. Furthermore, multiple studies showed that somatic mutation rates tend to be considerably lower in taller, more long-lived species (Schmid-Siegert et al., 2017; Plomion et al., 2018; Hofmeister et al., 2019; Orr et al., 2020; Wang et al., 2019; Hanlon et al., 2019). For instance, Orr et al. (2020) found the somatic mutation rate per generation to be only ten times higher in *Eucalyptus melliodora* than in *Arabidopsis*, despite being > 100 times larger in size.

Thus, empirical evidence indicates that more long-lived species have acquired mechanisms to reduce the amount of mutations accumulated during growth on the one hand, but still present high levels of inbreeding depression on the other hand, which suggests that more long-lived species still accumulate more mutations despite above mentioned limiting mechanisms. The aim of the present study is to disentangle the relationship between these two observations. I first study the evolution of the mutation rate in plants, and then consider the number of mutations and the magnitude of inbreeding depression maintained at mutation-selection balance, given the evolutionarily stable mutation rate reached by the population. To do so, I extend the work of previous authors (Kimura, 1967; Gervais and Roze, 2017) to the case of a perennial population in which individuals grow as they age and accumulate mutations in doing

so. I obtain analytical predictions which I test against the output of individual-centered simulations. I show that the evolutionarily stable mutation rate should decrease in plants as life expectancy increases, because deleterious mutations have more time to accumulate in more long-lived species. Furthermore, I show that despite substantially lower per year mutation rates, more long-lived species still tend to accumulate larger amounts of deleterious mutations because of higher per generation, leading to higher levels of inbreeding depression in these species. However, the magnitude of this increase depends strongly on how mutagenic meiosis is relative to growth.

3.3 Methods

Model outline. I consider a large population of hermaphroditic diploids. Individuals survive between mating events with a constant probability S. Juveniles may only settle in replacement of deceased individuals, so that population size is kept constant. I assume that individuals are made of a trunk, which grows by one section between each flowering event.

Mutations at the selected loci occur both during meosis and somatic growth. The somatic mutation rate per unit of growth, that is per new section produced (u), of a given individual is determined by its genotype at a single modifier locus. At this locus, I consider the fate of a rare mutant (m) with a weak effect (ε) competing with a resident allele (M). I assume this mutant allele to be codominant with the resident, so that an individual's somatic mutation rate is given by $u_{MM} = u_0$, $u_{Mm} = u_0 + \varepsilon$, or $u_{mm} = u_0 + 2\varepsilon$, depending on its genotype at the modifier.

Although meiotic and somatic cell divisions likely share common features, so that meiotic and somatic mutation rates should evolve together to some extent, they also differ in various ways. For instance, recombination during meiosis causes additional breaks in DNA strands, which gives the opportunity for more mutations. Furthermore,

somatic and meiotic mutation rates are not defined on the same scale. Indeed, while the somatic mutation rate is usually defined as a number of mutations per unit of growth, meiotic mutation rates are defined at the scale of a whole reproductive event. Thus, the relationship between these two mutation rates is not straightforward, because different genetic events happen and divisions occur at different paces. For the sake of simplicity, I will assume that meiotic mutations are produced at rate γu, where γ is a positive real number which allows one to tune the intensity of meiotic mutation relative to somatic mutation. In other words, I assume there is a linear relationship between the two rates.

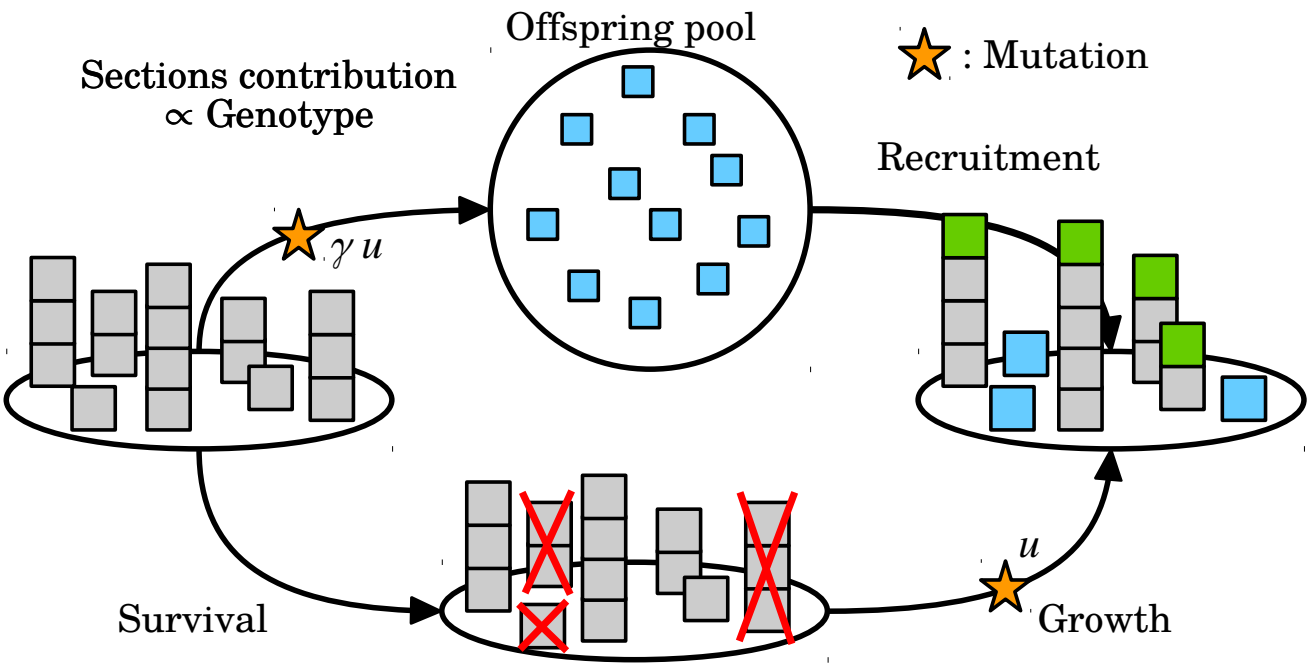

FIGURE 3.1 – Life cycle of the modeled population. Blue squares are offspring, which are made of a single section. Green squares depict the sections gained during growth by survivors. Yellow stars indicate the steps at which mutation occurs.

I assume that any section can contribute to reproduction (FIG. 3.1). Self-fertilisation occurs at rate α, a fraction σ of which imperatively occurs within the same section. The remaining fraction $1 - \sigma$ can occur between sections within the individual. A section's

fecundity is determined by its genotype at a very large number of biallelic loci acting multiplicatively. At these loci, allele 0 is an healthy allele, while allele 1 is a mutated allele which diminishes the section's fecundity by a proportion s. In heterozygotes, allele 1 expresses proportionally to its dominance coefficient h. Following previous authors (Gervais and Roze, 2017), I also introduce a DNA replication fidelity cost function, f, which is an increasing function of the meiotic mutation rate γu. Gervais and Roze (2017) considered a variety of cost functions and came to qualitatively similar conclusions in every case. Yet, most of their results were obtained using the cost function given in Equation (3.1),

$$f(\gamma u) = e^{-\frac{c}{\gamma u}}, \tag{3.1}$$

where c is the cost of replication fidelity, which I also use in this study. Thus, the fecundity of a section is given by

$$W = f(\gamma u) \times (1-s)^{n_{hom}}(1-sh)^{n_{het}}, \tag{3.2}$$

where n_{hom} and n_{het} are the number of mutations borne in the homozygous and heterozygous states, respectively.

Analytical methods. I use the theoretical framework described in Kirkpatrick et al. (2002) to study our model, which relies on indicator variables to describe individuals' multilocus genotypes. In the analytical model, I neglect the effect of the proportion of obligate within-section selfing (σ) since it will prove to have very little impact on the results. For the sake of brevity, derivations of the results presented in the following sections are detailed in Appendices 3.6.1.1 and 3.6.1.2 for results regarding the evolution of mutation rate and the mutation-selection equilibrium properties of the population given the evolutionarily stable mutation rate, respectively.

Individual-centered simulations. I ran individual-centered simulations to test the validity of analytical approximations. The simulation program was coded in C++11 and is available from GitHub (https://github.com/Thomas-Lesaffre/Somatic_mutations). In this program, individuals are represented by two chromosomes of length λ (expressed in cM) with the modifier situated at the center and along which mutations can occur at any position, so that infinitely many selected loci are effectively modeled (Roze and Michod, 2010).

Modeled loci. At the modifier, I assume that infinitely many alleles exist, coding for any value of $u \in [0, +\infty[$. Mutation occurs at rate $u_m = 10^{-3}$, and the value coded by the new allele is sampled from a Gaussian distribution centered on the former allele value with standard deviation $\sigma_m = 10^{-2}$, which is truncated at zero to prevent the modifier from going out of range. At selected loci, the number of mutations occuring on a chromosome during a given mutation event is sampled from a Poisson distribution with mean u (γu for meiosis), and their position is sampled from a uniform distribution. Recombination is modeled by exchanging segments between homologous chromosomes. The number of crossing-overs is sampled in a Poisson distribution with mean λ and their positions are sampled from a uniform distribution along chromosomes. Every time a mutation occurs, the age of the section at which it occured along the individual is stored, so that the genotype of any section within an individual can be reconstructed at any time from the individual genome. This method allows one to gain substantial computation time because mutations are stored only once per individual instead of being copied once for each new section.

Sequence of events. The population is kept of constant size, N. Between each mating event, individuals have a constant survival probability S. If they survive, they grow by one section, and mutations occur at rate u in this section. If they die, they are replaced by an offspring produced by the population. Any section within any individual

can be chosen as a parent, with a probability proportional to its fecundity (Equation 3.2). The offspring is produced by self-fertilisation with probability α, in which case the chosen section mates with itself with probability σ, and with any section within the same individual with probability $1-\sigma$. When selfing occurs between sections, a second parental section is selected within the individual. When the offspring is not produced by self-fertilisation, which occurs at rate $1-\alpha$, it is produced by random mating and a second parent is selected from the whole population. Mutation occurs at rate γu during meiosis.

Measurements. Once the equilibrium was reached, that is when both the mutation rate and the average number of mutations per chromosome were at equilibrium, the average number of mutations per chromosome in seeds, the average mutation rate and inbreeding depression are measured. Although individuals are chimeric in the model, I stuck to measuring inbreeding depression at the individual level to be in line with its definition. To do so, I counted how many times each individual was chosen as a parent before it died (*i.e.* I measured its lifetime reproductive success) and used this quantity as a measure of lifetime fitness. Individuals were marked as being produced by outcrossing (0), selfing within the same section (1), and selfing between sections within the same individual (2), so that I was able to measure fitness differences between these various categories of individuals. Namely, I measured inbreeding depression, that is the decrease in fitness of selfed individuals relative to the outcrossed (δ_{01}), and autogamy depression (Schultz and Scofield, 2009; Bobiwash et al., 2013), that is the decrease in fitness of within-section selfed individuals relative to between-sections ones (δ_{12}). Ten replicates were run for each parameter set. Simulations were kept running for 10^6 and 2×10^5 reproductive seasons for life expectancies lower and higher than 200 reproductive seasons, respectively. Results were averaged over the last 10^5 reproductive cycles (resp. 2×10^4) and the 95% confidence interval around the mean was also recorded.

3.4 Results

In what follows, life expectancy (E) will be used to discuss results instead of survival probability (S) for the sake of clarity and biological relevance. Given survival probability S, life expectancy can be computed as

$$E = \frac{1}{1-S}. \tag{3.3}$$

3.4.1 Evolutionarily stable mutation rate

Let us first study the evolution of the mutation rate. I show in Appendix 3.6.1.1 that the evolution of the mutation rate is the result of the opposition between the direct cost of DNA replication fidelity, which is higher when the mutation rate is lower, and the indirect selection caused by deleterious alleles which tend to be more frequently linked with modifier alleles increasing the mutation rate (Equation A23). The resulting evolutionarily stable mutation rate is given by

$$u^* = \sqrt{-\frac{c}{\gamma\ \hat{s}_{ind}}}, \tag{3.4}$$

where $\hat{s}_{ind}$ encapsulates the intensity of indirect selection acting on the modifier. Its expression is derived in Appendix 3.6.1.1. FIGURE 3.2 shows the evolutionarily stable mutation rate as a function of life expectancy (top row), along with the intensity of indirect selection (bottom row), for cases where $\gamma = 1$, $\gamma = 10$ and $\gamma = 100$. I chose to focus on cases where $\gamma \geqslant 1$, that is on cases where more mutations are produced during meiosis than during the development of a new section, on the basis of three lines of evidence. First, direct observations of plant development at the cellular level indicate that cells destined to form axillary meristems undergo much fewer divisions than other cells from the moment they are produced in the apical meristem, which suggests that the number of cell divisions per branching event, and therefore the number of oppor-

tunities for mutations to accumulate, may be lower than previously thought (Burian et al., 2016). Second, estimates of somatic mutation rates per unit of growth tend to be low (Orr et al., 2020). Third, to my knowledge, the only experiment comparing the mutagenicity of meiosis and mitosis was performed by Magni and Von Borstel (1962) in yeast. They found meiosis to be 6 to 20 times more mutagenic than mitosis, which further suggests that γ may tend to be greater than 1. Besides, performing simulations $\gamma < 1$ proved to be very challenging since the number of mutations accumulated in the population quickly became very high, causing simulations to run very slowly.

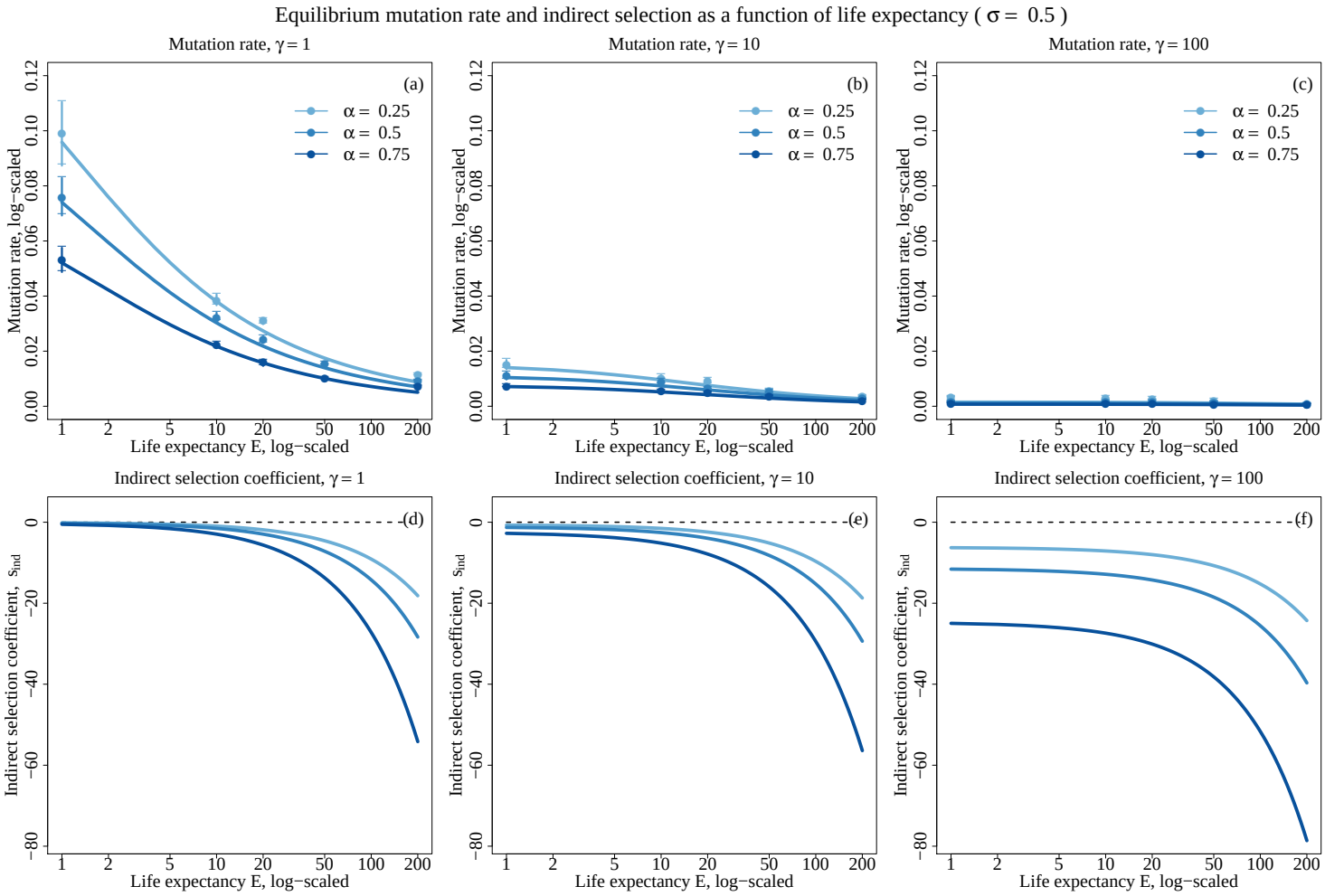

FIGURE 3.2 – Evolutionarily stable mutation rate (top) and intensity of indirect selection (bottom) as a function of life expectancy (log-scaled) for various selfing rates (colors) and for $\gamma = 1$ (left), $\gamma = 10$ (middle) and $\gamma = 100$ (right). Other parameters values are $s = 0.05$, $h = 0.3$, $c = 0.0014$, $\lambda = 20$, and $\sigma = 0.5$. Dots depict simulation results and error bars depict the 95% confidence intervals. Lines depict analytical predictions.

The evolutionarily stable mutation rate decreases with life expectancy for all γ values (FIG. 3.2a-c). In both cases, this is due to the greater number of opportunities to accumulate deleterious mutations in more long-lived species because they go trough more growth events, which in turn causes indirect selection to increase against alleles increasing the mutation rate because deleterious mutations become more numerous

(FIG. 3.2d-f). Furthermore, the evolutionarily stable mutation rate is much lower when γ is larger, because increasing γ decreases the cost of replication fidelity (Equation 3.1), and increases the intensity of indirect selection on alleles increasing the mutation rate.

The mutation rate also decreases as the selfing rate (α) increases, which may seem counter-intuitive since selfing tends to reduce the number of deleterious mutations segregating in the population through purging Roze (2015). However, self-fertilisation also causes genetic associations between selected loci and the modifier to increase, thereby increasing indirect selection and resulting in a decrease of the evolutionarily stable mutation rate when the selfing rate increases as shown by Gervais and Roze (2017). The results presented in FIG. 3.2 were obtained assuming half of selfing events occured imperatively within the same section ($\sigma = 0.5$). Cases with $\sigma = 0$ and $\sigma = 1$ were also investigated and yielded very similar results, which are presented in FIG. S1 and S2, respectively, in Appendix 3.6.2. I argue that the very small effect of σ on the results is due to the low evolutionarily stable mutation rate, which causes few somatic mutations to occur during growth, and to the fact that I assumed weak selection so that mutations have little effect on their bearer's fitness.

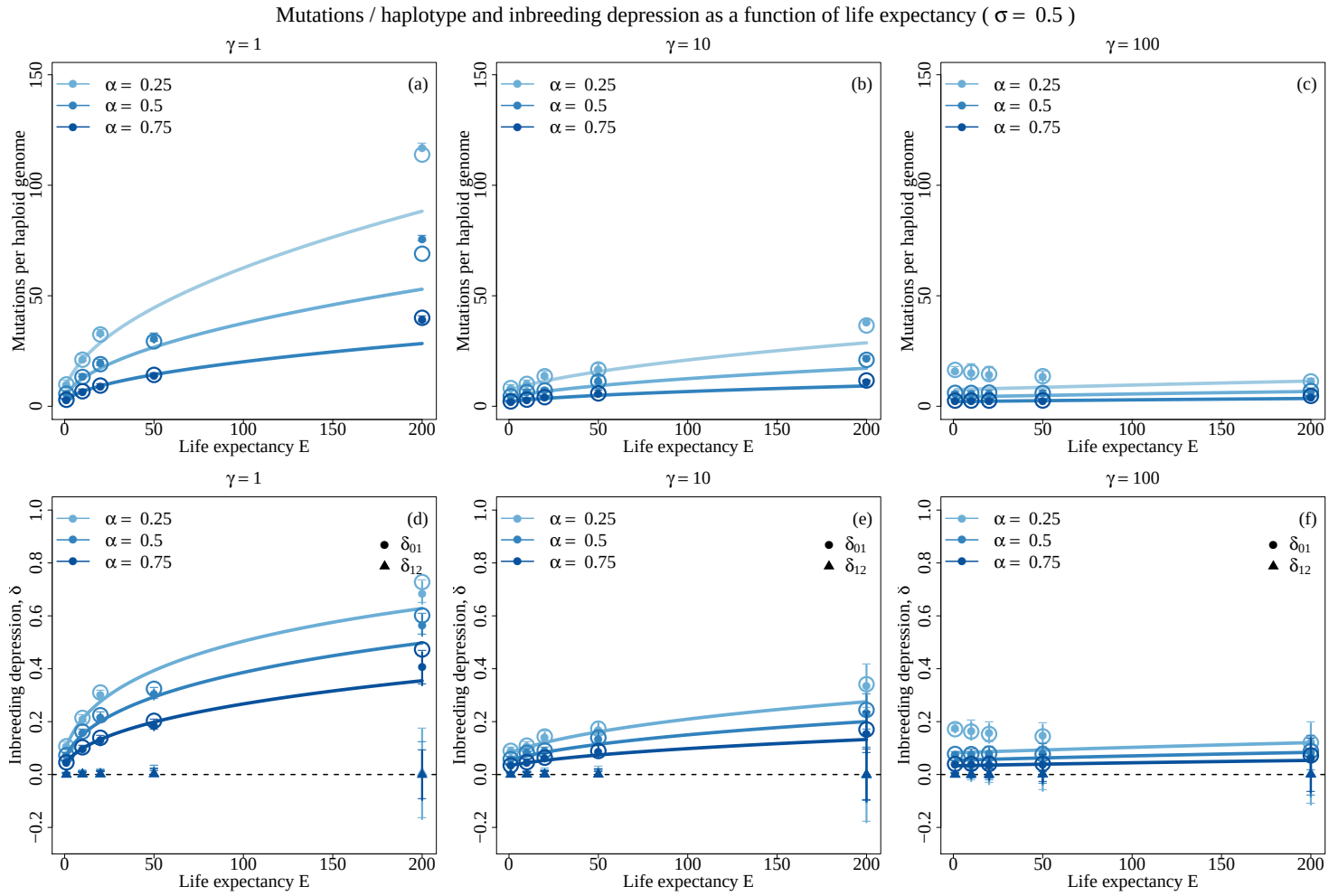

FIGURE 3.3 – Average number of mutations per haploid genome (top) and inbreeding depression (bottom) as a function of life expectancy (log-scaled) for various selfing rates (colors) and $\gamma = 1$ (left), $\gamma = 10$ (middle) and $\gamma = 100$ (right). Other parameters values are $s = 0.05$, $h = 0.3$, $c = 0.0014$, $\lambda = 20$, and $\sigma = 0.5$. Filled dots depict simulation results and error bars depict the 95% confidence intervals. Lines depict analytical predictions. Open circles depict the value predicted by our analytical model when the equilibrium mutation rate from simulations is used instead of Equation 3.4. On the bottom row, dots indicate inbreeding depression (δ_{01}), while triangles indicate autogamy depression (δ_{12}).

3.4.2 Mutation-selection balance

Once the mutation rate has reached an equilibrium and the population has reached mutation-selection balance, we show in Appendix 3.6.1.2 that a leading order approximation of the average number of mutations per haploid genome in juveniles (n) is given by

$$n \approx \frac{\hat{u}^*}{s\left[h + F(1-h)\right]} - u^* \frac{S}{1-S}, \tag{3.5}$$

where $u^* = \sqrt{-\frac{c}{\gamma\ \hat{s}_{ind}}}$, and $\hat{u}^* = \left(\frac{1}{1-S} + \gamma\right) u^*$ depicts the total mutation rate of the population over the course of one timestep, including both meiotic and somatic mutations. As for inbreeding depression, calculated between outcrossed and selfed individuals (δ_{01}), it is given by

$$\delta_{01} = 1 - \exp\left[-s(1-2h)\frac{1+F}{2}\left(\frac{\hat{u}^*}{s\left[h + F(1-h)\right]} - u^* \frac{S}{1-S}\right)\right], \tag{3.6}$$

where $F = \frac{\alpha}{2-\alpha}$, to leading order in s. Again, I neglect the impact of the proportion of selfing occuring within or between sections (σ) in the analytical model since it is negligible. FIGURE 3.3 shows the number of mutations per haploid genome among juveniles, (n, top row) and inbreeding and autogamy depression (δ_{01} and δ_{12}, bottom row) at mutation-selection balance. Deviations between analytical predictions (lines) and simulations results (dots) are observed. They can be explained by the slight differences between the predicted evolutionarily stable mutation rate and the equilibrium mutation rate reached by simulations, which build up large differences in n when life expectancy becomes high. Indeed, when the equilibrium mutation rate from the simulations is used to predict n instead of Equation (3.4), the agreement between predictions (open circles) and simulation results (dots) is restored.

The number of mutations maintained n increases as life expectancy increases in every case, due to the greater amount of opportunities for mutations to accumulate in more long-lived species. Indeed, in Equation (3.5), the denominator of the first term

shows that the intensity of selection is independent of life expectancy, while the total mutation rate $\hat{u}^*$ is a function of life expectancy. The increase of n with life expectancy becomes much lower when γ increases, to the point where it gets barely noticeable with $\gamma = 100$. Furthermore, n is lower when γ is higher despite the fact that many more mutations are produced during meiosis, because the evolutionarily stable mutation rate is much lower, so that the total mutation rate $\hat{u}^*$ is lower (FIG. 3.2, Equation 3.5). As a result, inbreeding depression is lower when γ is higher, and increases when life expectancy increases, but this increase becomes less and less sharp as γ increases. Besides, n is lower when the selfing rate increases, as expected under weak selection (Roze, 2015). Furthermore, consistent with the negligible effect σ had on the evolution of the mutation rate, almost no autogamy depression is generated (triangles in FIG. 3.2, bottom row).

3.5 Discussion

Evolution of the mutation rate. In this paper, I studied the evolution of the mutation rate when somatic mutations are assumed to be inheritable, as it is thought to be the case in plants (Scofield and Schultz, 2006; Lanfear, 2018). I showed that the evolutionarily stable mutation rate decreases as life expectancy increases because of the greater number of opportunities to accumulate mutations during growth in more long-lived species, which makes indirect selection against alleles increasing the mutation rate stronger. However, although the mutation rate per mutagenic event (u), that is per growth season or per meiosis in our model, decreased in more long-lived species, the total mutation rate ($\hat{u}$), that is the rate at which mutations entered the population through both somatic growth and meiosis, increased. Hence, my results indicate that while we should expect more efficient mechanisms reducing the accumulation of deleterious mutations during growth to evolve in more long-lived species, so that their per

unit of growth and per year mutation rate should be lower, their per generation mutation rates should still be higher. These predictions are in line with empirical evidence, which suggest that mutation rates per generation tend to be higher in more long-lived species although the mutation rates per unit of growth tend to be lower (Hofmeister et al., 2019; Hanlon et al., 2019; Orr et al., 2020).

I modeled the evolution of the mutation rate following the work of Kimura (1967), by assuming there is a direct fitness cost to DNA replication fidelity opposing the indirect selection generated by deleterious mutations linked to the modifier, so that the mutation rate was maintained greater than zero in response to a trade-off. An alternative mechanism, which is not mutually exclusive with the trade-off described above, was put forward by Lynch (2011). They proposed that selection should always act to reduce the mutation rate, down until it becomes so low that the selective advantage brought by any further reduction should be overwhelmed by genetic drift, thus maintaining non-zero mutation rates because alleles further decreasing the mutation rate should at some point become effectively neutral, thereby creating a lower bound for the evolution of the mutation rate (Lynch, 2011). This lower bound is inevitably influenced by effective population size, as it plays on the relative strength of selection and genetic drift. In the present model, I overlooked Lynch (2011)'s lower bound by assuming a large and fixed population size. Yet, effective population sizes are expected to be higher in more long-lived species in which generations overlap (Felsenstein, 1971; Charlesworth, 1980; Petit and Hampe, 2006), which implies the lower bound described by Lynch (2011) should be met for lower mutation rates in said species. Hence, I expect the decrease in the evolutionarily stable mutation rate described in this study to become sharper in conditions where Lynch (2011)'s lower bound is expected to matter for the evolution of the mutation rate.

Inbreeding depression. The larger total mutation rate in more long-lived species led to the maintenance of more mutations in the population at mutation-selection balance, and therefore to higher inbreeding depression in these species, consistent with results from meta-analyses which found inbreeding depression to increase in larger-statured, more long-lived species (Duminil et al., 2009; Angeloni et al., 2011). Importantly however, the magnitude of the increase in the total mutation rate, and therefore in inbreeding depression with life expectancy depended strongly on the relative mutagenicity of meiosis and growth, which was controlled by the γ parameter in this model. Indeed, while the increase in inbreeding depression was strong when γ was close to 1, that is when the same amount of mutation was produced during meiosis and during growth between two flowering seasons, it became smaller as γ increased, to the point of being barely noticeable for $\gamma = 100$. This was due to the decrease of the evolutionarily stable mutation rate as γ increased, which made the contribution of somatic mutations to the mutation load more and more negligible compared with meiotic mutations. Hence, according to my results, for somatic mutations to be the main driver of the empirically observed increase in inbreeding depression in more long-lived species, roughly the same amount of mutations should be produced during growth between two flowering seasons and during reproduction.

Mating system evolution. Inbreeding depression is thought to be one of the main factors preventing the evolution of self-fertilisation (Lande and Schemske, 1985; Barrett and Harder, 2017). In Angiosperms, consistent with the observed increase in inbreeding depression in more long-lived species, there exists a strong correlation between mating systems and life-histories. Indeed, many self-fertilising species are annuals whereas most long-lived species are strictly outcrossing (Barrett and Harder, 1996; Munoz et al., 2016). Thus, somatic mutations accumulation was proposed as an explanation for this correlation (Scofield and Schultz, 2006). While results presented in this study

indicate that inbreeding depression increases with respect to life expectancy due to somatic mutations accumulation, particularly when γ is small, this increase is tempered by the decrease of the evolutionarily stable mutation rate with life expectancy. Furthermore, in agreement with results obtained by Gervais and Roze (2017), I showed that the evolutionarily stable mutation rate decreases as the selfing rate increases because the modifier becomes more strongly associated with selected loci. These decreases of the mutation rate with respect to mating system and life expectancy, together with the purging effect of self-fertilisation (Roze, 2015), result in a substantial drop in the magnitude of inbreeding depression as the selfing rate increases in more long-lived species, potentially opening the way for the evolution of self-fertilisation. Hence, whether somatic mutations accumulation is sufficient to explain the correlation between life-history and mating system in Angiosperms when the mutation rate is allowed to evolve jointly with the mating system is an open question.

Autogamy depression. In order to empirically estimate the contribution of somatic mutations accumulation to inbreeding depression using phenotypic data, a method was developed by Schultz and Scofield (2009). This method, called the autogamy depression test, relies on the comparison of the fitnesses of individuals produced by selfing within an inflorescence with those of individuals produced by selfing between distant inflorescences on the plant's crown (Schultz and Scofield, 2009; Bobiwash et al., 2013). In this paper, I performed such test by measuring autogamy depression (δ_{12}). Contrary to inbreeding depression, I found autogamy depression to be almost null in every case, even in situations where the contribution of somatic mutations accumulation to inbreeding depression was high. This result can be explained by the low evolutionarily stable mutation rates, and by the fact that we only considered mutations with a weak fitness effect. It suggests that the autogamy depression test should only be able to detect mutations with a large fitness effect in large enough individuals,

where mutations have had time to accumulate. Thus, it implies that detecting no autogamy depression in a given population cannot be taken as evidence of a negligible contribution of somatic mutations accumulation to the population's mutation load.

Mutagenicity of growth and meiosis. The results presented above suggest that valuable insights into the evolutionary relevance of somatic mutations and the evolution of the mutation rate in plants could be gained by further investigating the γ parameter in this model, which depicts the relative mutagenicity of meiosis and growth between two flowering seasons, and is likely influenced by at least three important factors that were unaccounted for in this study. First, it is necessarily influenced by how mutagenic meiotic divisions are in comparison with mitotic divisions, about which little is known although one may expect meiotic divisions to generate more mutations, as they generate many more double strand DNA breaks which are required for recombination and are known to be particularly mutagenic events (Magni and Von Borstel, 1962; Arbel-Eden and Simchen, 2019). Second, it is influenced by the number of mitoses occurring between flowering seasons. This number depends on the growth habit of the considered species, because fast growing species undergo more mitoses per unit of time than slow growing species, and because the rate at which mitoses occur, and thus the growth rate, may interact with the evolution of the mutation rate. Indeed, investing in a higher fidelity of DNA replication may tend to slow down individual growth. Third, apart from mechanisms reducing the amount of mutations produced during growth, deleterious mutations may also be affected by intra-organismal selection, which may not only reduce the growth rate by eliminating mutated cells, but also efficiently purge deleterious mutations from the organism, so that little to no somatic mutation may be present in the gamete. This could in turn affect the evolution of the mutation rate. Little is known, however, about the actual efficacy of intra-organismal selection in removing deleterious mutations since it was seldom investigated theoretical (Otto and

Orive, 1995), and mostly empirically demonstrated to occur in the case of strongly beneficial mutations (e.g. Edwards et al., 1990; Simberloff and Leppanen, 2019). The various elements discussed above show that γ is an emerging property of the interaction between a variety of mechanisms, which advocates for the development of theoretical models treating it as such rather than as a fixed parameter, by incorporating growth, mutation and selection at the cellular level.

Acknowledgements

I would like thank Sylvain Billiard, Vincent Castric, Ludovic Maisonneuve, Denis Roze and Roman Stetsenko for taking the time to discuss this work and comment on the manuscript. This work was funded by the European Research Council (NOVEL project, grant #648321). The author also thanks the Région Hauts-de-France, and the Ministère de l'Enseignement Supérieur et de la Recherche (CPER Climibio), and the European Fund for Regional Economic Development for their financial support.

3.6 Appendices

3.6.1 Derivation of analytical results

In this appendix, we detail the derivation of analytical results presented in the main text. The first section is dedicated to obtaining an approximation for the evolutionarily stable mutation rate, while the second is dedicated to the derivation of various population genetics quantities at mutation-selection balance.

3.6.1.1 Mutation rate evolution

In this section, we present the details of the method used to obtain the evolutionarily stable mutation rate approximation presented in the main text.

Defining age-dependent variables

Modifier locus. The mutation rate of an individual, given its genotype at the modifier, is given by

$$u = u_0 + \varepsilon \left(\overline{X}_m + \overset{*}{\overline{X}}_m \right) = \bar{u} + \varepsilon \left(\overline{\zeta}_m + \overset{*}{\overline{\zeta}}_m \right), \tag{A1}$$

where u_0 is the mutation rate coded by the resident allele, $\bar{u}$ is the average mutation rate in the population, $\overline{X}_m$ and $\overset{*}{\overline{X}}_m$ are indicator variables of the presence of the mutant allele on the paternally and maternally inherited chromosomes, respectively, and $\overline{\zeta}_m$ and $\overset{*}{\overline{\zeta}}_m$ are their associated centered variables, with

$$\overline{\zeta}_m = \overline{X}_m - \mathbb{E}\left[\overline{X}_m\right] = \overline{X}_m - p_m, \tag{A2}$$

and respectively for $\overset{*}{\overline{\zeta}}_m$. The bar in the notations indicates that these variables are considered over the whole population and not a particular age-class (see below). Im-

portantly here, the genotype at the modifier of an individual does not vary between sections so that this notation is unnecessary in the above but is kept so that notations remain consistent all along derivations.

Selected loci. Since sections of various ages, that is sections bearing a different load of somatic mutations, contribute to reproduction, we define similar variables X_i^n and $\overset{*}{X}{}_i^n$ for the presence of deleterious mutations at the ith selected locus, in the nth section of individuals. Given the mutation rate u of an individual, its genotype in the $(n+1)^{th}$ section can be deduced from its genotype in the nth section using Equation (A3) :

$$X_i^{n+1} = X_i^n + u\left(1 - X_i^n\right). \tag{A3}$$

We can obtain the general term of this recursive sequence, which is given by

$$X_i^n = 1 - (1-u)^n(1 - X_i^0). \tag{A4}$$

Injecting $\zeta-$variables, developing u and rearranging Equation (A4), this yields,

$$\zeta_i^n = X_i^n - \mathbb{E}\left[X_i^n\right] = (1-\bar{u})^n\,\zeta_i^0 + n\varepsilon(1-\bar{u})^{n-1}(\overline{\zeta}_m + \overset{*}{\overline{\zeta}}_m) + o(\varepsilon^2). \tag{A5}$$

Summarising notations For the sake of clarity, we summarise notations here. The same notations will be applied to indicator variables (X) and centered variables (ζ), allelic frequencies (p) and genetic associations (D, defined below).

Genetic associations. Let us forget for a moment the existence of age-classes in order to define properly genetic associations. Genetic associations are expectations of products of ζ-variables, and are in fact covariances between sets of indicator variables. In other words, they quantify the extent to which the frequency of sets of alleles at a given set of genetic positions deviate from the panmictic expectation, which

assumes that alleles segregate independently both within and between loci. In general, we denote the genetic associations between the set of genetic positions $\mathbb{U}$ on the paternal chromosome and the set of genetic positions $\mathbb{V}$ on the maternal chromosome as

$$D_{\mathbb{U},\mathbb{V}} = \mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}}\right] = \mathbb{E}\left[\left(\prod_{l\in\mathbb{U}} \zeta_l\right) \times \left(\prod_{l\in\mathbb{V}} \overset{*}{\zeta}_l\right)\right], \tag{A6}$$

where the comma separates genetic positions situated on the paternal and maternal chromosomes. For instance, the association

$$D_{i,i} = \mathbb{E}\left[\zeta_{i,i}\right] = \mathbb{E}\left[\zeta_i \times \overset{*}{\zeta}_i\right] = \mathbb{E}\left[X_i \overset{*}{X}_i\right] - p_i^2, \tag{A7}$$

designates the excess in homozygotes at the ith locus. Sets $\mathbb{U}$ and $\mathbb{V}$ can be empty, so that genetic associations between genetic positions all situated on the same chromosome (e.g. linkage disequilibrium) may be considered. For example, the association

$$D_{ij,\varnothing} = \mathbb{E}\left[\zeta_{ij,\varnothing}\right] = \mathbb{E}\left[\zeta_i \times \zeta_j\right] = \mathbb{E}\left[X_i X_j\right] - p_i p_j, \tag{A8}$$

measures the linkage disequilibrium between loci i and j on the paternal chromosome. Since there is no separate sexes nor sex-specific effect in our model, we have

$$D_{\mathbb{U},\mathbb{V}} = D_{\mathbb{V},\mathbb{U}}. \tag{A9}$$

Hence, we define the condensed notation

$$\widetilde{D}_{\mathbb{U},\mathbb{V}} = \frac{D_{\mathbb{U},\mathbb{V}} + D_{\mathbb{V},\mathbb{U}}}{2}, \tag{A10}$$

in order to shorten recursions.

Keeping track of age-classes and steps in the sequence of events. We stated above that X_i^n (resp. $\overset{*}{X}{}_i^n$) designates the indicator variable associated with the paternal (resp. maternal) position at the ith selected locus in sections aged n. The same notation will be used for ζ-variables, genetic associations ($\widetilde{D}^n_{\mathbb{U},\mathbb{V}}$) and allelic frequencies (e.g. p_i^n). Importantly, while the genotype of sections at the modifier does not vary with age, selection may change allelic frequencies at the modifier differently in sections with different ages. Hence, the notation p_m^n will also sometimes be used to indicate the fact that we consider allelic frequencies at the modifier among sections aged n. Hence, denoting V a generic variable (which may be X, D, ζ or p), and $\mathbb{U}$ a generic set of genetic positions, $V_{\mathbb{U}}^n$ indicates that the variable V is considered over the set $\mathbb{U}$ of genetic positions among sections aged n. To keep track of the step in the sequence of events occurring over the course of one timestep we are looking at, that is for instance variables measured among parents, gametes or juveniles, we add an additional superscript separated from the age-class indication by a pipe. Denoting k a generic stage in the sequence of events, we thus have the notation $V_{\mathbb{U}}^{n|k}$. Finally, to indicate that we consider the average over all age-classes in the step k of the variable $V_{\mathbb{U}}^{n|k}$, we will use the notation $\overline{V_{\mathbb{U}}}^k$.

Deriving an approximation for fecundity We now derive an approximation for the fecundity of a section of age n as a function of its genotype, and relative to the entire parental population, that is sections of all ages. Let W_n be the fecundity of a section aged n and $\overline{W}$, the fecundity of sections averaged over all ages. To leading order in $\ln W_n - \ln \overline{W}$, we have

$$W_n = e^{\ln \overline{W} + \ln W_n - \ln \overline{W}} \approx W \left(1 + \ln W_n - \ln \overline{W}\right). \tag{A11}$$

Furthermore, we have

$$W = e^{\mathbb{E}\left[\ln \overline{W}\right] + \ln \overline{W} - \mathbb{E}\left[\ln \overline{W}\right]} \approx e^{\mathbb{E}\left[\ln \overline{W}\right]} \left(1 + \ln W - \mathbb{E}\left[\ln \overline{W}\right]\right), \tag{A12}$$

where $\mathbb{E}\left[\ln \overline{W}\right]$ is the mean of the log-fitness. Thus, neglecting products of fitnesses, we have

$$W_n \approx e^{\mathbb{E}\left[\ln \overline{W}\right]} \left(1 + \ln W_n - \mathbb{E}\left[\ln \overline{W}\right]\right), \tag{A13}$$

and since W is the average of W_n over all ages, the relative fecundity of a section aged n is given by

$$\frac{W_n}{\mathbb{E}\left[\overline{W}\right]} \approx 1 + \ln W_n - \mathbb{E}\left[\ln \overline{W}\right]. \tag{A14}$$

Let us now derive expressions for $\ln W_n$ and $\mathbb{E}\left[\ln \overline{W}\right]$. Expressed in terms of indicator variables, we have

$$\begin{aligned} W_n = f\Big(\gamma \bar{u} + \gamma \varepsilon (\overline{\zeta}_m + \overset{*}{\overline{\zeta}}_m)\Big) \\ \times \prod_i \left[1 - 2shp_i^n - sh(\zeta_i^n + \overset{*}{\zeta}_i^n) - s(1-2h)\left(\widetilde{\zeta}_{i,i}^n + p_i^n(\zeta_i^n + \overset{*}{\zeta}_i^n) + (p_i^n)^2\right)\right], \end{aligned} \tag{A15}$$

where $f(\gamma u)$ is the replication fidelity cost function, and γ allows one to tune the number of cell divisions occuring in meiosis relative to growth. Assuming selection to be weak and deleterious mutations to remain rare, so that terms in $s \times p_i^n$ can be neglected, and using the fact that $\ln(1+x) \approx x$ when x is small, the log-fitness of a section of age n can be approximated as follows

$$\ln W_n \approx \ln f\left(\gamma \bar{u} + \gamma \varepsilon (\overline{\zeta}_m + \overset{*}{\overline{\zeta}}_m)\right) - s \sum_i \left[2hp_i^n + h(\zeta_i^n + \overset{*}{\zeta}_i^n) + (1-2h)\widetilde{\zeta}_{i,i}^n\right], \tag{A16}$$

which to leading order in ε and in s, yields

$$\ln W_n \approx 1+\ln f(\gamma\bar{u})+\varepsilon\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})}(\overline{\zeta}_m+\overset{*}{\overline{\zeta}}_m)-2sh\sum_i p_i^n-sh\sum_i(\zeta_i^n+\overset{*}{\zeta}_i^n)-s(1-2h)\sum_i \widetilde{\zeta}_{i,i}^n. \tag{A17}$$

Since we have $\overline{W}=\sum_n S^n(1-S)W_n$, where S is the survival probability of individuals between mating events, we have

$$\ln \overline{W} \approx 1+\ln f(\gamma\bar{u})+\varepsilon\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})}(\overline{\zeta}_m+\overset{*}{\overline{\zeta}}_m)-2sh\sum_i \overline{p}_i-sh\sum_i(\overline{\zeta}_i+\overset{*}{\overline{\zeta}}_i)-s(1-2h)\sum_i \widetilde{\overline{\zeta}}_{i,i}, \tag{A18}$$

where the bar denotes the average over all ages, so that

$$\mathbb{E}\left[\ln \overline{W}\right] \approx \ln f(\gamma\bar{u}) - 2sh\sum_i \overline{p}_i - s(1-2h)\sum_i \widetilde{\overline{D}}_{i,i}, \tag{A19}$$

and,

$$\frac{W_n}{\mathbb{E}\left[\overline{W}\right]} \approx 1+\varepsilon\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})}\,(\overline{\zeta}_m+\overset{*}{\overline{\zeta}}_m)-2sh\sum_i(p_i^n-\overline{p}_i)-sh\sum_i(\zeta_i^n+\overset{*}{\zeta}_i^n)-s(1-2h)\sum_i(\widetilde{\zeta}_{i,i}^n-\widetilde{\overline{D}}_{i,i}). \tag{A20}$$

Computing the change in frequency at the modifier Since we consider the fate of a rare mutant allele, the only relevant change in its frequency is due to selection. Assuming deleterious mutations remain at low frequencies, the frequency of the mutant in gametes produced by sections aged n following selection is given by

$$p_m^{n|s} = \mathbb{E}\left[\frac{W_n}{\mathbb{E}\left[\overline{W}\right]}\times\frac{\overline{X}_m+\overset{*}{\overline{X}}_m}{2}\right] \approx \bar{p}_m+\varepsilon\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})}\left(\bar{p}_m\bar{q}_m+\widetilde{\overline{D}}_{m,m}\right)-sh\sum_i\left(\widetilde{D}_{mi}^n+\widetilde{D}_{m,i}^n\right) - s(1-2h)\sum_i \widetilde{D}_{mi,i}^n. \tag{A21}$$

Thus, the frequency of the mutant among juveniles, which is the same as among gametes, is given by

$$\bar{p}_m^j = \sum_n S^n(1-S)p_m^{n|s} = \bar{p}_m + \varepsilon\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})}\left(\bar{p}_m\bar{q}_m + \widetilde{\overline{D}}_{m,m}\right) - sh\sum_i\left(\widetilde{\overline{D}}_{mi} + \widetilde{\overline{D}}_{m,i}\right) - s(1-2h)\sum_i \widetilde{\overline{D}}_{mi,i}. \tag{A22}$$

Since no selection occurs in parents, we have $\bar{p}_m^a = \bar{p}_m$, and

$$\Delta\bar{p}_m = (1-S)\bar{p}_m^j + S\bar{p}_m^a - \bar{p}_m = (1-S)\left[\varepsilon\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})}(1+F)\bar{p}_m - sh\sum_i\left(\widetilde{\overline{D}}_{mi} + \widetilde{\overline{D}}_{m,i}\right) - s(1-2h)\sum_i \widetilde{\overline{D}}_{mi,i}\right], \tag{A23}$$

assuming $D_{m,m} \approx Fp_m$ (Roze, 2015). Equation (A23) can be broken down into two terms, one involving the cost function f, which depicts the direct selection acting on the modifier due to the replication fidelity cost, and a second term proportional to s, which depict the intensity of indirect selection acting on the modifier due to linked deleterious alleles.

Computing the indirect selection term In order to obtain an approximation for the indirect selection term, which we denote s_{ind} following Gervais and Roze (2017), we need to obtain expressions for the genetic associations appearing in it at quasi-linkage equilibrium. On top of selection, these associations will also be affected by mutation, recombination and the mating system.

Two-way genetic associations. Let us begin with deriving the changes occuring in two-way genetic associations $\widetilde{\overline{D}}_{mi}$ and $\widetilde{\overline{D}}_{m,i}$ over one timestep. The effects of selection on such associations in gametes produced by sections aged n can be computed

using

$$\widetilde{D}^{n|s}_{mi} = \mathbb{E}\left[\frac{W_n}{\mathbb{E}\left[\overline{W}\right]}\widetilde{\zeta}^n_{mi}\right] + \Delta^s p^n_m \Delta^s p^n_i \text{ and, } \widetilde{D}^{n|s}_{m,i} = \mathbb{E}\left[\frac{W_n}{\mathbb{E}\left[\overline{W}\right]}\widetilde{\zeta}^n_{m,i}\right] + \Delta^s p^n_m \Delta^s p^n_i,$$

where $\Delta^s p_m$ and $\Delta^s p_i$ are changes in frequencies at the modifier and at the ith selected locus, respectively. However, the product of changes in allelic frequencies is at best of order $\varepsilon \times s$, so that we may neglect it in this context. Thus, we have

$$\widetilde{D}^{n|s}_{mi} \approx \widetilde{D}^n_{mi}(1-sh) - s(1-h)\widetilde{D}^n_{mi,i} \text{ and, } \widetilde{D}^{n|s}_{m,i} \approx \widetilde{D}^n_{m,i}(1-sh) - s(1-h)\widetilde{D}^n_{mi,i}. \quad \text{(A24)}$$

Selection is followed by meitoic mutation, which occurs at rate γu. Thus, in gametes we have

$$\begin{aligned}\widetilde{D}^{n|g}_{mi} &= \frac{1}{2}\mathbb{E}\left[\zeta^{n|s}_m \times \left(X^{n|s}_i + \gamma u(1 - X^{n|s}_i) - p^{n|g}_i\right) + \overset{*}{\zeta}{}^{n|s}_m \times \left(\overset{*}{X}{}^{n|s}_i + \gamma u(1 - \overset{*}{X}{}^{n|s}_i) - p^{n|g}_i\right)\right] \\ \widetilde{D}^{n|g}_{m,i} &= \frac{1}{2}\mathbb{E}\left[\overset{*}{\zeta}{}^{n|s}_m \times \left(X^{n|s}_i + \gamma u(1 - X^{n|s}_i) - p^{n|g}_i\right) + \zeta^{n|s}_m \times \left(\overset{*}{X}{}^{n|s}_i + \gamma u(1 - \overset{*}{X}{}^{n|s}_i) - p^{n|g}_i\right)\right],\end{aligned} \quad \text{(A25)}$$

which to leading order in ε yields

$$\widetilde{D}^{n|g}_{mi} \approx (1-\gamma\bar{u})\widetilde{D}^{n|s}_{mi} + \gamma\varepsilon(1+F)\bar{p}_m \text{ and, } \widetilde{D}^{n|g}_{m,i} \approx (1-\gamma\bar{u})\widetilde{D}^{n|s}_{m,i} + \gamma\varepsilon(1+F)\bar{p}_m. \quad \text{(A26)}$$

Reproduction occurs in three ways, outcrossing at rate $1 - \alpha$, selfing within section at rate $\alpha(1 - \sigma)$, and selfing between sections at rate $\alpha\sigma$. While outcrossing does not generate genetic associations, both kinds of selfing do when associations involve genomic positions situated on different chromosomes. For the sake of simplicity, and because sections only differ from one another due to somatic mutations accumulation, which should be negligible at a given locus when the mutation rate is small, we will consider that selfing between and within sections have the same effect on genetic

associations, so that in seeds following reproduction we have

$$\widetilde{D}^{r}_{mi} = (1-S)\sum_{n} S^{n}\left((1-r)\widetilde{D}^{n|g}_{mi} + r\widetilde{D}^{n|g}_{m,i}\right) = (1-r)\widetilde{\overline{D}}^{g}_{mi} + r\widetilde{\overline{D}}^{g}_{m,i}, \tag{A27}$$

and,

$$\widetilde{D}^{r}_{m,i} = (1-S)\frac{\alpha}{2}\sum S^{n}\left(\widetilde{D}^{n|g}_{mi} + \widetilde{D}^{n|g}_{m,i}\right) = \frac{\alpha}{2}\left(\widetilde{\overline{D}}^{g}_{mi} + \widetilde{\overline{D}}^{g}_{m,i}\right). \tag{A28}$$

All individuals then undergo somatic mutation at rate u, which is incorporated in the exact same way as meiotic mutation was. Thus, we have

$$\widetilde{\overline{D}}^{a}_{mi} = (1-\bar{u})\widetilde{\overline{D}}_{mi} + \varepsilon(1+F)\bar{p}_{m}, \quad \text{and,} \quad \widetilde{\overline{D}}^{a}_{m,i} = (1-\bar{u})\widetilde{\overline{D}}_{m,i} + \varepsilon(1+F)\bar{p}_{m}, \tag{A29}$$

in adults and,

$$\widetilde{\overline{D}}^{j}_{mi} = (1-\bar{u})\widetilde{D}^{r}_{mi} + \varepsilon(1+F)\bar{p}_{m}, \quad \text{and,} \quad \widetilde{\overline{D}}^{j}_{m,i} = (1-\bar{u})\widetilde{D}^{r}_{m,i} + \varepsilon(1+F)\bar{p}_{m}, \tag{A30}$$

in juveniles.

Three-way association. Let us now turn to the derivation of the changes occuring the three-way association $\widetilde{\overline{D}}_{mi,i}$ over the course of one timestep. Similar to other associations, we have,

$$\widetilde{D}^{n|s}_{mi,i} = \mathbb{E}\left[\frac{W_n}{\mathbb{E}\left[\overline{W}\right]}\widetilde{D}^{n}_{mi,i}\right] - \Delta^{s} p^{n}_{i} D^{n|s}_{m,m} \approx \widetilde{D}^{n}_{mi,i}(1-s), \tag{A31}$$

then, following the same method as in Equation (A25), we obtain

$$\widetilde{D}^{n|g}_{mi,i} = (1-2\gamma\bar{u})\widetilde{D}^{n|s}_{mi,i}. \tag{A32}$$

Among juveniles, we thus have

$$\widetilde{D}^{j}_{mi,i} = (1-S)\frac{\alpha}{2}\sum_{n} S^{n}\left(\widetilde{D}^{n|g}_{mi,i} + (1-r)\widetilde{D}^{n|g}_{mi} + r\widetilde{D}^{n|g}_{m,i}\right) = \frac{\alpha}{2}\left(\widetilde{\overline{D}}^{g}_{mi,i} + (1-r)\widetilde{\overline{D}}^{g}_{mi} + r\widetilde{\overline{D}}^{g}_{m,i}\right). \tag{A33}$$

Among adults in the next timestep, we have

$$\widetilde{D}^{a}_{mi,i} = (1-2\bar{u})\widetilde{\overline{D}}_{mi,i}. \tag{A34}$$

QLE expressions. To obtain expressions for genetic associations at quasi-linkage equilibrium, we need to solve Equation (A35) for $\widetilde{D}^{0}_{mi}$, $\widetilde{D}^{0}_{m,i}$ and $\widetilde{D}^{0}_{mi,i}$,

$$\begin{cases} S\widetilde{D}^{a}_{mi} + (1-S)\widetilde{D}^{j}_{mi} - \widetilde{\overline{D}}_{mi} = 0 \\ S\widetilde{D}^{a}_{m,i} + (1-S)\widetilde{D}^{j}_{m,i} - \widetilde{\overline{D}}_{m,i} = 0 \\ S\widetilde{D}^{a}_{mi,i} + (1-S)\widetilde{D}^{j}_{mi,i} - \widetilde{\overline{D}}_{mi,i} = 0, \end{cases} \tag{A35}$$

which, neglecting terms in $\bar{u}$ and to leading order in ε and s, yields

$$\begin{aligned} \widetilde{\overline{D}}^{*}_{mi} &= \frac{\varepsilon(1+F)\left[1+r(1+F)+\gamma(1+2Fr)(1-S)\right]\bar{p}_m}{(1-S)\left[sh_e + r(1-F-s\left(h_e(1-2F)-F(2-F)\right)\right]}, \\ \widetilde{\overline{D}}^{*}_{m,i} &= \frac{\varepsilon(1+F)\left[r(1+F)+F\left(1+\gamma(1-S)(1+2r)\right)\right]\bar{p}_m}{(1-S)\left[sh_e + r(1-F-s\left(h_e(1-2F)-F(2-F)\right)\right]}, \\ \widetilde{\overline{D}}^{*}_{mi,i} &= \frac{\varepsilon(1+F)F(1+2Fr)[1+\gamma(1-S)]\bar{p}_m}{(1-S)\left[sh_e + r(1-F-s\left(h_e(1-2F)-F(2-F)\right)\right]}, \end{aligned} \tag{A36}$$

with $h_e = h + F(1-h)$.

Indirect selection term. Equation (A23) can be rearranged into

$$\Delta\bar{p}_m = \varepsilon(1+F)(1-S)\left[\gamma\frac{f'(\gamma\bar{u})}{f(\gamma\bar{u})} + s_{ind}\right]\bar{p}_m, \tag{A37}$$

where the indirect selection term s_{ind} is given by

$$s_{ind} = \frac{-sh\left(\widetilde{\overline{D}}_{mi} + \widetilde{\overline{D}}_{m,i}\right) - s(1-2h)\widetilde{\overline{D}}_{mi,i}}{\varepsilon(1+F)\bar{p}_m}. \tag{A38}$$

Using expressions presented in Equation (A36), this is

$$s_{ind} \approx -s\frac{h_e\Big(1+\gamma(1-S)\Big) + 2r\left[(1-2h)\Big(1+\gamma(1-S)\Big)F^2 + 2\gamma hF(1-S) + h(1+F)\right]}{(1-S)\left(sh_e + r(1-F) - sr\left[F(F-2) + h_e(1-2F)\right]\right)}. \tag{A39}$$

Integration over the genetic map. To account for the effect of the infinitely many loci present in the genome, one may integrate s_{ind} over the genetic map. The total length of the genetic map is λ (in cM), but since chromosomes are symetrical with respect to the centromere, we may focus one half of the genetic map, that is on the fraction of the chromosome situated between the centromere ($x = 0$), where the modifier is, and the tail of the chromosome ($x = \frac{\lambda}{2}$). Thus, the probability that a crossing-over occurs between position $x \in \left[0, \frac{\lambda}{2}\right]$ and the centromere, where the modifier is situated, is simply given by

$$p(x) = \frac{2x}{\lambda}. \tag{A40}$$

The number of crossing-overs during one meiosis event is drawn from a Poisson law with mean λ. Thus, the number of crossing-overs occuring on the $\left[0, \frac{\lambda}{2}\right]$ segment, n, follows a Poisson law with mean $\frac{\lambda}{2}$ and we have

$$\mathbb{P}\left(n \mid \lambda/2\right) = \frac{(\lambda/2)^n}{n!}e^{-\lambda/2}. \tag{A41}$$

Then, the probability k of the n crossing-overs occur between position x and the centromere follows a Binomial law with probability $p(x) = \frac{2x}{\lambda}$. Importantly, recombination

only effectively occurs when an uneven number of crossing-overs occurs on $\left[0, \frac{\lambda}{2}\right]$, which given n occurs with probability

$$1 - \sum_{k=0}^{n/2} \binom{n}{k/2} p(x)^{k/2} (1 - p(x))^{n-k/2}, \tag{A42}$$

so that the probability of recombination effectively occurring between position x and the centromere along $\left[0, \frac{\lambda}{2}\right]$ is given by

$$r_x = \sum_{n=0}^{\infty} \frac{(\lambda/2)^n}{n!} e^{-\lambda/2} \left(1 - \sum_{k=0}^{n/2} \binom{n}{k/2} p(x)^{k/2} (1 - p(x))^{n-k/2}\right) = \frac{1 - e^{-2x}}{2}. \tag{A43}$$

Plugging Equation (A43) in Equation (A39), that is swapping r for r_x, the indirect selection term accounting for the genetic map is given by

$$\hat{s}_{ind} = \frac{2}{\lambda} \int_0^{\lambda/2} -s \frac{h_e\Big(1 + \gamma(1 - S)\Big) + 2r_x \left[(1 - 2h)\Big(1 + \gamma(1 - S)\Big)F^2 + 2\gamma h F(1 - S) + h(1 + F)\right]}{(1 - S)\left(sh_e + r_x(1 - F) - sr_x\left[F(F - 2) + h_e(1 - 2F)\right]\right)} \, \mathrm{d}x \tag{A44}$$

Thus, using $f(\gamma u)$ cost function described in the main text, the change in frequency at the modifier can be written as

$$\Delta p_m = \frac{c}{\gamma u^2} + \hat{s}_{ind}, \tag{A45}$$

so that the evolutionarily stable mutation rate is given by

$$\Delta p_m = 0 \Leftrightarrow u^* = \sqrt{-\frac{c}{\gamma \hat{s}_{ind}}}. \tag{A46}$$

3.6.1.2 Mutation-selection balance

Once the population has reached its equilibrium mutation rate, it reaches mutation-selection equilibrium. The average number of mutations per haploid genome and the

resulting inbreeding depression can then be obtained by assuming the mutation rate is constant.

Average number of mutations per haploid genome

Among juveniles. Using Equation A20 in Appendix 3.6.1.2, and assuming the modifier is fixed at the evolutionarily stable mutation rate (*i.e.* $u = u^*$ and $\varepsilon = 0$), the frequency of the deleterious allele at the ith selected locus among gametes produced sections aged n due to selection is given by

$$p_i^{n|s} = \mathbb{E}\left[\left.\frac{W_n}{\mathbb{E}\left[\overline{W}\right]}\right|_{\varepsilon=0} \times \frac{X_i^n + \overset{*}{X}{}_i^n}{2}\right] = p_i^n - sh\left(\widetilde{D}_{ii}^n + \widetilde{D}_{i,i}^n\right) - s(1-2h)\widetilde{D}_{ii,i}^n, \quad \text{(A47)}$$

which noting that

$$s(1-h)\widetilde{D}_{ii,i}^n = s(1-h)(1-u)^{3n}(1-2p_i^0)\widetilde{D}_{i,i}^0 \approx s(1-h)(1-u)^n\widetilde{D}_{i,i}^n \approx s(1-h)\widetilde{D}_{i,i}^n, \quad \text{(A48)}$$

and that,

$$\widetilde{D}_{ii}^n = (1-u)^{2n}\widetilde{D}_{ii}^0 = (1-u)^{2n}p_i^0 q_i^0 \approx p_i^n, \quad \text{(A49)}$$

can be rearranged into

$$p_i^{n|s} \approx p_i^n - shp_i^n - s(1-h)\widetilde{D}_{i,i}^n. \quad \text{(A50)}$$

Thus, in juveniles following meiotic and somatic mutation with have

$$\bar{p}_i^j = u\,(1+\gamma) + \sum_n S^n(1-S)p_i^{n|s}. \quad \text{(A51)}$$

As for the excess in homozygotes among juveniles, we neglect the effects of selection and mutation so that we have

$$\widetilde{D}^{j}_{i,i} \approx \frac{\alpha}{2}\left(\bar{p}_i + \widetilde{\overline{D}}_{i,i}\right) \tag{A52}$$

Among adults. Among adults, selection does not have any effect, so that we simply have

$$\bar{p}^{a}_{i} = \bar{p}_i + u, \tag{A53}$$

and,

$$\widetilde{\overline{D}}^{a}_{i,i} = \widetilde{\overline{D}}_{i,i}. \tag{A54}$$

Equilibrium excess in homozygotes. Neglecting the effects of mutation and selection on homozygosity at selected loci, we have

$$\Delta\widetilde{\overline{D}}_{i,i} = (1-S)\frac{\alpha}{2}\left(\bar{p}_i + \widetilde{\overline{D}}_{i,i}\right) + S\widetilde{\overline{D}}_{i,i} - \widetilde{\overline{D}}_{i,i} = (1-S)\left[\frac{\alpha}{2}\bar{p}_i - \left(1-\frac{\alpha}{2}\right)\widetilde{\overline{D}}_{i,i}\right]. \tag{A55}$$

Thus, the equilibrium excess in homozygotes at selected loci is given by

$$\Delta\widetilde{\overline{D}}_{i,i} = 0 \Leftrightarrow \widetilde{\overline{D}}_{i,i} = \frac{\alpha}{2-\alpha}\bar{p}_i = F\bar{p}_i. \tag{A56}$$

Equilibrium number of mutations. The change in frequency of the deleterious allele at the ith locus is given by

$$\Delta\bar{p}_i = (1-S)\bar{p}^{j}_{i} + S\bar{p}^{a}_{i} - \bar{p}_i, \tag{A57}$$

so that at equilibrium, with $\widetilde{\overline{D}}_{i,i} = F\bar{p}_i$, we have

$$\Delta\bar{p}_i = 0 \Leftrightarrow \bar{p}^{*}_{i} = \frac{u\left(\gamma + \frac{1}{1-S}\right)}{s\left[h + F(1-h)\right]}. \tag{A58}$$

In simulations, we measure the average number of mutations per haploid genome in seeds, which can be obtained as follows :

$$p_i^{0|*} = u(1+\gamma) + \bar{p}_i^* - s\left[h + F(1-h)\right]\bar{p}_i^* = \bar{p}_i^* - u\frac{S}{1-S}. \tag{A59}$$

Inbreeding depression In this paper, we chose to measure inbreeding depression at the scale of whole individuals even though individuals are chimeric from the genetic point of view, in order to remain consistent with its classical definition. To obtain an analytical approximation for inbreeding depression, we thus need to compute the mean lifetime reproductive success of selfed and outcrossed individuals.

Lifetimes fitness expression. Since mutations have a null probability of occurring on both alleles at the same time at a locus, so that homozygotes are never created by mutation, the fecundity of a section aged i can be approximated as

$$W_i = W_0(1-sh)^{2ui}. \tag{A60}$$

Thus, the total fecundity of an individual aged k is

$$\hat{W}_k = W_0\frac{1-(1-sh)^{2u(k+1)}}{1-(1-sh)^{2u}}. \tag{A61}$$

Hence, lifetime fitness can be computed as

$$\hat{W} = \sum_n S^n(1-S)\sum_{k=0}^{n}\hat{W}_k = \frac{W_0}{(1-S)\left[1-(1-sh)^{2u}S\right]}, \tag{A62}$$

Inbreeding depression. Using Equation (A62), inbreeding depression can be expressed as

$$\delta = \mathbb{E}\left[1-\frac{\hat{W}^{self}}{\hat{W}^{out}}\right] = 1-\frac{\mathbb{E}\left[W_0^{self}\right]}{\mathbb{E}\left[W_0^{out}\right]}, \tag{A63}$$

where $\hat{W}^{self}$ and $\hat{W}^{out}$ are lifetime fitness measured among the selfed and the outcrossed, respectively. We now need to obtain leading order approximations for $\mathbb{E}\left[W_0^{self}\right]$ and $\mathbb{E}\left[W_0^{out}\right]$. For any subpopulation sub, we have

$$W_0^{sub} = e^{\overline{\ln W_0^{sub}}}\left(1 + \ln W_0^{sub} - \overline{\ln W_0^{sub}}\right), \tag{A64}$$

to leading order in $\ln W_0^{sub} - \overline{\ln W_0^{sub}}$, so that

$$\mathbb{E}\left[W_0^{sub}\right] \approx e^{\overline{\ln W_0^{sub}}}. \tag{A65}$$

Furthermore, from Equation (A16), we have

$$\overline{\ln W_0^{sub}} \approx -2sh\sum_i p_i^{0|sub} - s(1-2h)\sum_i \widetilde{D}_{i,i}^{0|sub}. \tag{A66}$$

In the case of selfed and outcrossed individuals, we have $p_i^{0|self} = p_i^{0|out} = p_i^0$. Besides, the excess in homozygotes among outcrossed individuals is $\widetilde{\overline{D}}_{i,i}^{out} = 0$, while among the selfed, it is given by $\widetilde{\overline{D}}_{i,i}^{self} = \frac{1}{2}\left(\bar{p}_i + \widetilde{\overline{D}}_{i,i}\right)$. Thus, we have

$$\mathbb{E}\left[W_0^{out}\right] = \exp\left(-2sh\sum_i p_i^{0|*}\right), \tag{A67}$$

and,

$$\mathbb{E}\left[W_0^{self}\right] = \exp\left(-2sh\sum_i p_i^{0|*} - s(1-2h)\sum_i \frac{1+F}{2}p_i^{0|*}\right). \tag{A68}$$

Hence, inbreeding depression in our model is given by

$$\delta \approx 1 - \exp\left[-s(1-2h)\frac{1+F}{2}\left(\frac{u\left[1+\gamma(1-S)\right]}{s(1-S)\left[h+F(1-h)\right]} - u\frac{S}{1-S}\right)\right]. \tag{A69}$$

3.6.2 Additional graphs

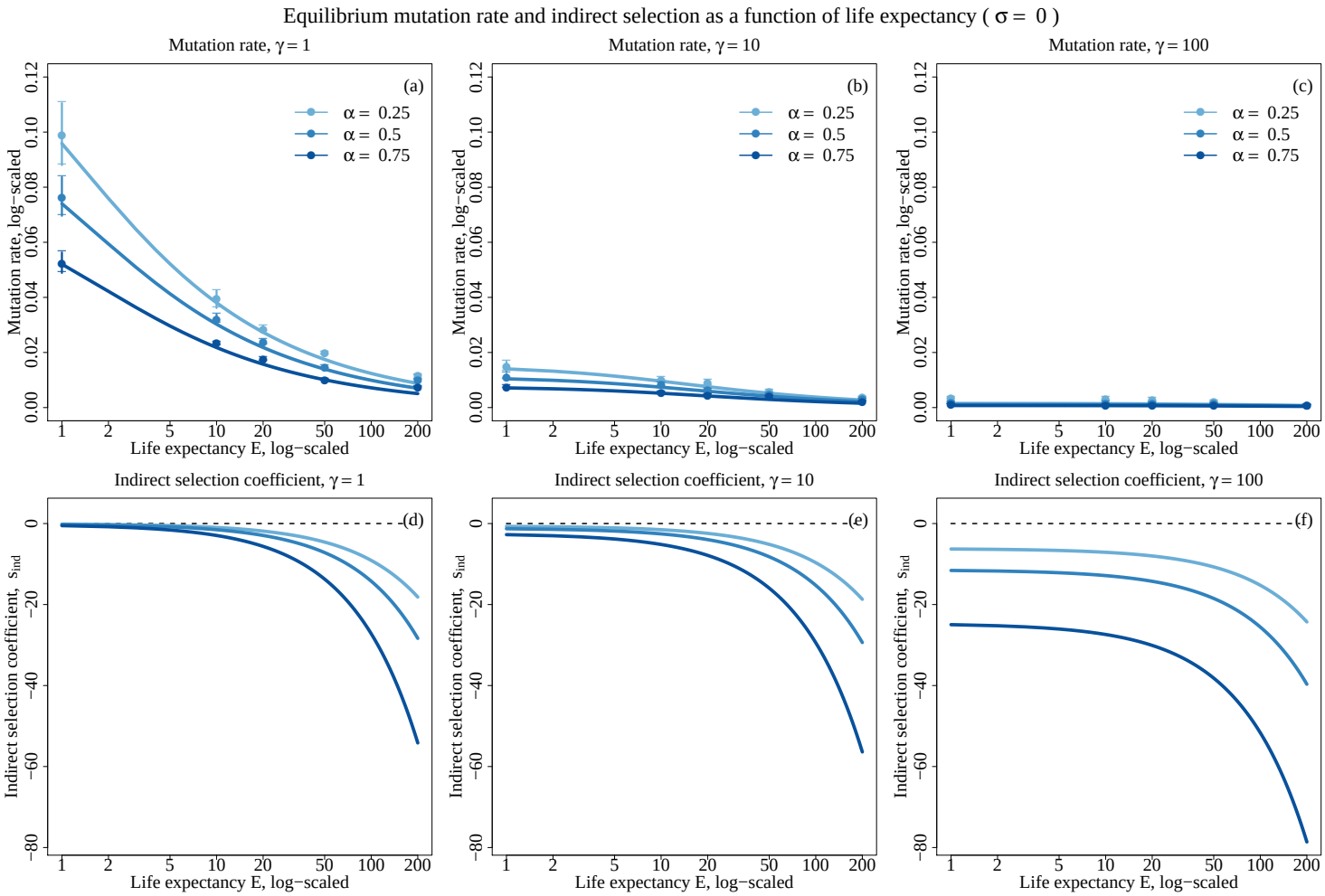

FIGURE S1 – Evolutionarily stable mutation rate (top) and intensity of indirect selection (bottom) as a function of life expectancy (log-scaled) for various selfing rates (colors) and for $\gamma = 1$ (left) and $\gamma = 10$ (right). Other parameters values are $s = 0.05$, $h = 0.3$, $c = 0.0014$, $\lambda = 20$, and $\sigma = 0$. Dots depict simulation results and error bars depict the 95% confidence intervals. Lines depict analytical predictions.

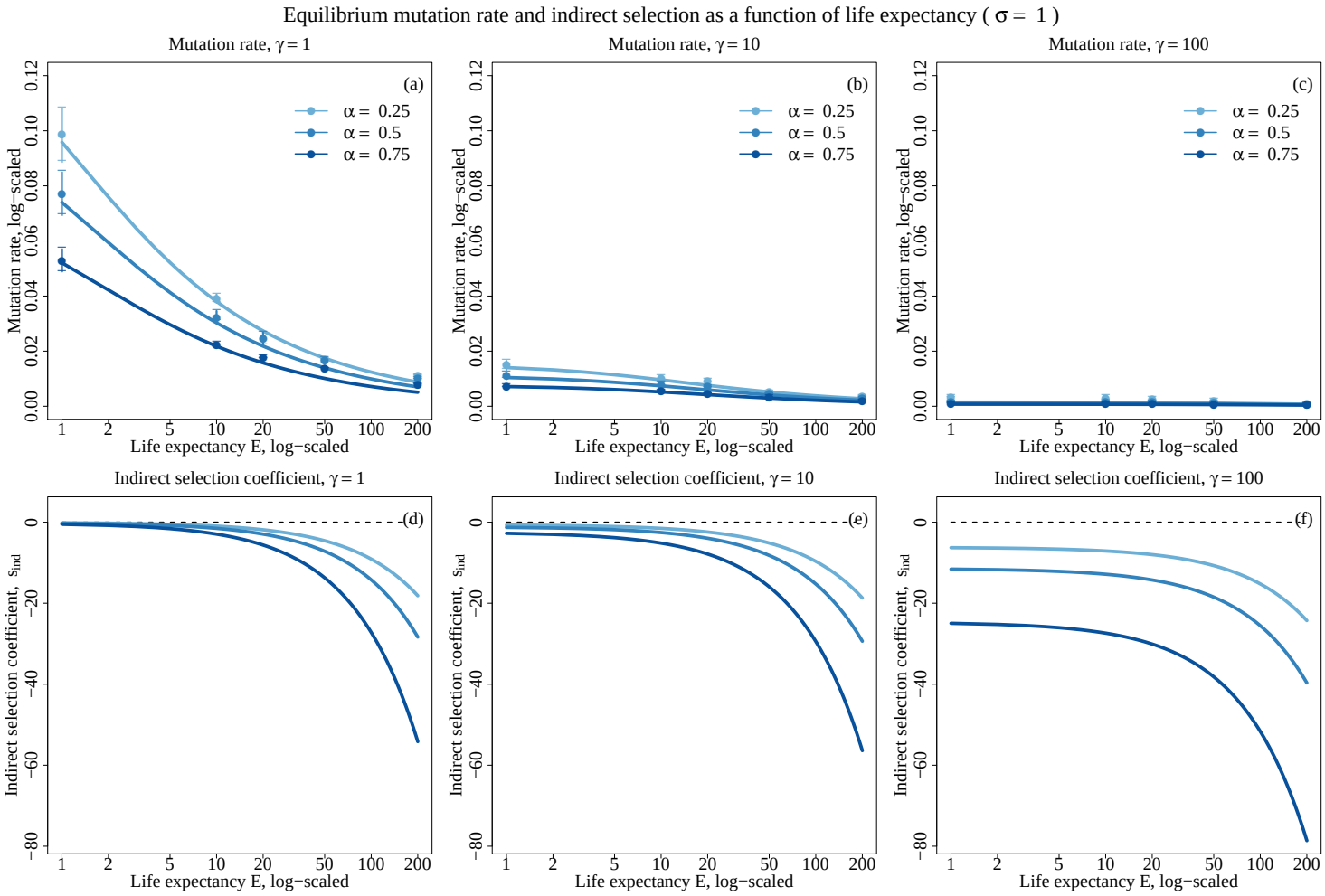

FIGURE S2 – Evolutionarily stable mutation rate (top) and intensity of indirect selection (bottom) as a function of life expectancy (log-scaled) for various selfing rates (colors) and for $\gamma = 1$ (left) and $\gamma = 10$ (right). Other parameters values are $s = 0.05$, $h = 0.3$, $c = 0.0014$, $\lambda = 20$, and $\sigma = 1$. Dots depict simulation results and error bars depict the 95% confidence intervals. Lines depict analytical predictions.

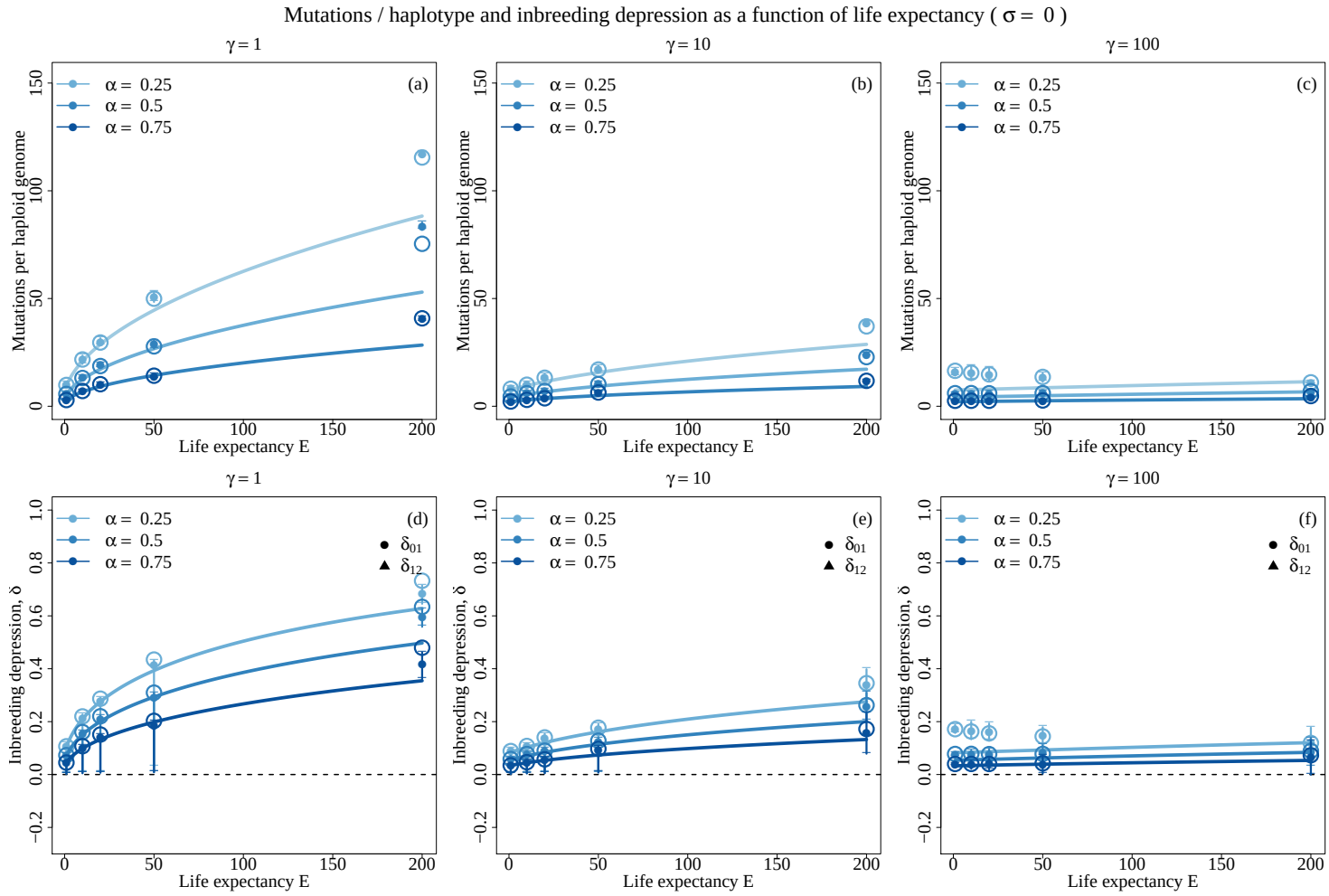

FIGURE S3 – Average number of mutations per haploid genome (top) and inbreeding depression (bottom) as a function of life expectancy (log-scaled) for various selfing rates (colors) and for $\gamma = 1$ (left) and $\gamma = 10$ (right). Other parameters values are $s = 0.05$, $h = 0.3$, $c = 0.0014$, $\lambda = 20$, and $\sigma = 0$. Filled dots depict simulation results and error bars depict the 95% confidence intervals. Lines depict analytical predictions. Open circles depict the value predicted by our analytical model when the equilibrium mutation rate from simulations is used instead of Equation 3.4. On the bottom row, dots indicate inbreeding depression (δ_{01}), while triangles indicate autogamy depression (δ_{12}).

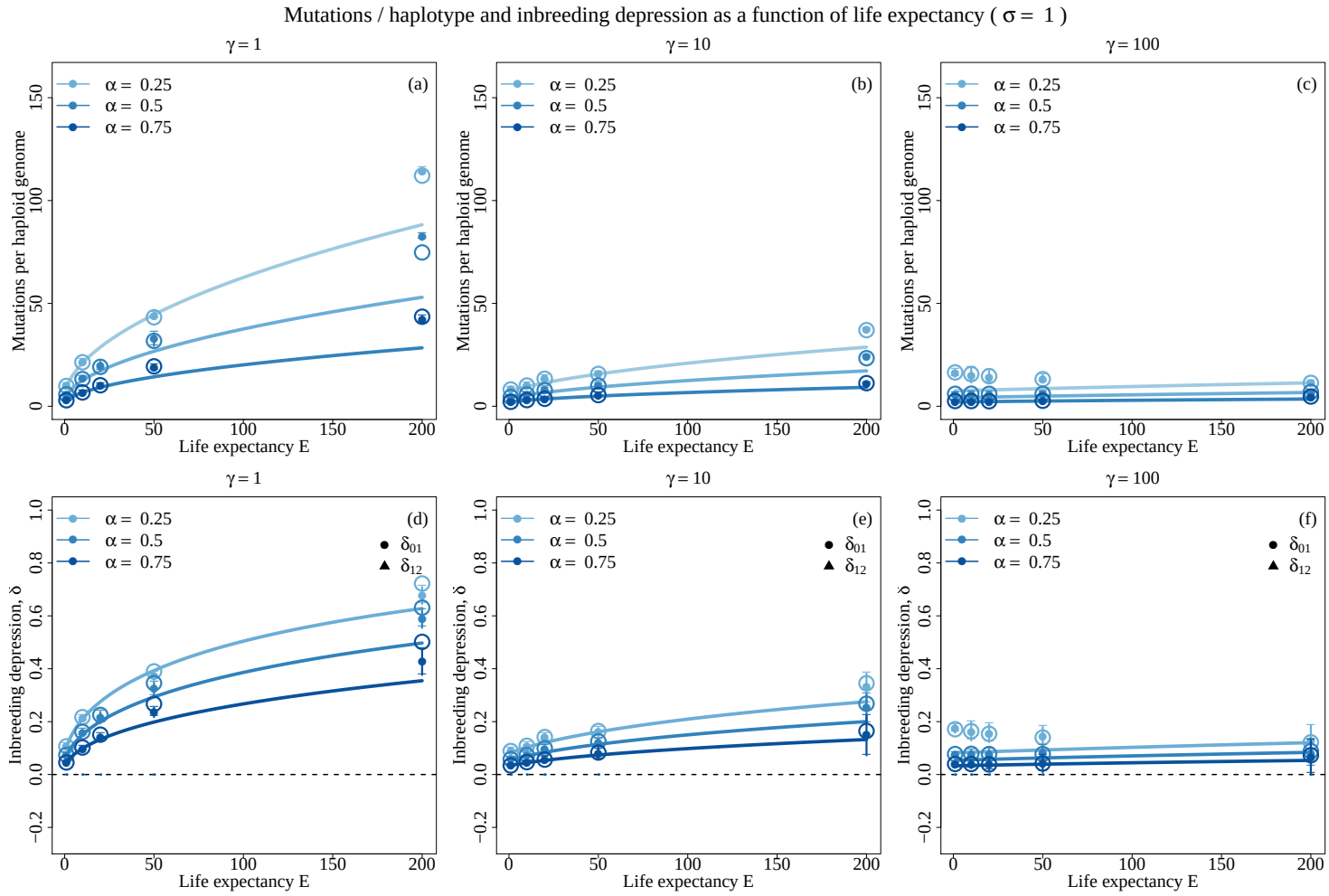

FIGURE S4 – Average number of mutations per haploid genome (top) and inbreeding depression (bottom) as a function of life expectancy (log-scaled) for various selfing rates (colors) and for $\gamma = 1$ (left) and $\gamma = 10$ (right). Other parameters values are $s = 0.05$, $h = 0.3$, $c = 0.0014$, $\lambda = 20$, and $\sigma = 1$. Filled dots depict simulation results and error bars depict the 95% confidence intervals. Lines depict analytical predictions. Open circles depict the value predicted by our analytical model when the equilibrium mutation rate from simulations is used instead of Equation 3.4. On the bottom row, dots indicate inbreeding depression (δ_{01}), while triangles indicate autogamy depression (δ_{12}).

Chapitre 4

(W.I.P) A theoretical evaluation of indirect inbreeding depression estimators

Auteurs : Thomas Lesaffre, Guillaume Lan-Fong, Sylvain Billiard

Statut de publication : Le travail présenté dans ce chapitre est à l'état d'ébauche.

4.1 Synthèse en français

4.1.1 Motivation

Dans les deux chapitres précédents, nous nous sommes intéressés à la dépression de consanguinité chez les espèces pérennes d'un point de vue théorique en supposant que le patron empiriquement observé d'augmentation de la dépression de consanguinité avec la pérennialité est un résultat fiable. Dans ce quatrième chapitre, nous explorons la possibilité que les estimations de la dépression de consanguinité souffrent de biais lorsqu'elles sont appliquées à des espèces pérennes. D'un point de vie empirique, me-

surer la dépression de consanguinité implique de mesurer la fitness d'individus issus d'auto- et d'allofécondation. Si une mesure directe de leur succès reproducteur total est possible dans le cas d'espèces à courtes durées de vie, une telle approche devient rapidement inenvisageable dès lors que l'espérance de vie des individus excède quelques mois. En outre, quantifier le succès reproducteur des individus par la voie mâle peut s'avérer difficile en populations naturelles. Des méthodes indirectes d'estimation de la dépression de consanguinité ont donc été développées afin de contourner ce problème. Deux d'entre elles sont couramment utilisées. La première consiste à choisir un trait ou un ensemble de traits comme proxy de fitness, et de s'en servir pour estimer la relation entre la fitness et le coefficient de consanguinité des individus. La pente de cette relation, appelée fardeau de consanguinité (*inbreeding load*), peut ensuite être utilisée pour obtenir une mesure de dépression de consanguinité. Lorsque cette méthode est utilisée, les résultats sont souvent discutés en termes de dépression de consanguinité affectant chaque trait. Étant donné que la dépression de consanguinité est une grandeur définie sur la fitness et non sur les phénotypes, cette manière de présenter les résultats fait implicitement l'hypothèse que la relation phénotype-fitness est linéaire. De plus, elle suppose que les traits influencent la fitness indépendamment les uns des autres, c'est-à-dire qu'aucun trade-off ne les lie de sorte que les différences entre individus pour un trait donné reflètent des différences de fitness et non différentes stratégies. La seconde méthode, proposée par Ritland (1990), s'affranchit des problèmes causés par la mesure de la fitness en se basant uniquement sur les variations du coefficient de consanguinité au sein de la population. Bien que l'estimateur de Ritland se fonde sur des récurrences de génétique des populations théorique faisant l'hypothèse que les générations sont non-chevauchantes et que la sélection ne se produit que parmi les juvéniles, il est souvent appliqué chez des espèces pérennes dont les générations sont chevauchantes. De plus, la sélection contre les individus consanguins ne se limite probablement pas à la phase juvénile, ce qui pourrait biaiser les prédictions de cet

estimateur. Dans ce chapitre, notre objectif est d'évaluer la validité des estimations obtenues par ces différentes méthodes indirectes lorsque leurs hypothèses ne sont pas vérifiées.

4.1.2 Modèle & résultats clefs

Le modèle utilisé dans ce but est le même que celui développé dans le deuxième chapitre, auquel des marqueurs neutres sont ajoutés afin de mesurer le niveau de consanguinité des individus. Ces marqueurs sont totalement indépendants des loci sous sélection et une forte diversité allélique y est maintenue par mutation. De nouveau, trois types de mutations affectant les coûts physiologiques de croissance ou la survie sont considérés dans trois modèles séparés. Les estimations obtenues par les méthodes de Ritland et de fardeau de consanguinité sont comparées à la véritable dépression de consanguinité endurée par les individus pour chacun des trois modèles.

4.1.2.1 Estimateur de Ritland

L'estimateur de Ritland est appliqué à chaque modèle de deux manières. D'abord, nous l'utilisons de la façon pour laquelle il a été pensé, c'est-à-dire en comparant le niveau de consanguinité des parents et des juvéniles produits au cours d'une saison de floraison. Ensuite, nous l'utilisons pour comparer la population complète entre deux pas de temps. Lorsque des mutations affectant la croissance sont considérées, nous devrions nous attendre à ne détecter aucune dépression de consanguinité car tous les individus survivent et sont recrutés avec la même probabilité. Pourtant, une dépression de consanguinité non-nulle est systématiquement détectée. Nous montrons que ce résultat peut être expliqué par la sélection qui s'exerce sur les individus consanguins au moment de contribuer au pool total de gamètes. Ainsi, les estimations basées sur la méthode de Ritland ne devraient pas être interprétées comme reflétant uniquement les effets de la dépression de consanguinité sur la survie car le moindre processus

provoquant une diminution de l'homozygotie parmi les juvéniles viendra les perturber. Le même biais, bien que moins visible car généré indirectement par la tendance des individus autofécondés à mourir plus jeunes que les allofécondés, est observé dans le cas de mutations affectant la survie.

4.1.2.2 Fardeau de consanguinité

Nous appliquons également une approche de fardeau de consanguinité aux trois modèles de deux manières différentes.

Pedigrees. Dans la première, nous construisons des pedigrees artificiels afin d'imiter une approche couramment utilisée dans le cas d'espèces strictement allofécondantes. Cela nous permet de créer trois classes de consanguinité, parmi lesquelles nous utilisons la taille à maturité comme mesure de fitness. Comme attendu, une dépression de consanguinité nulle est prédite dans le cas de mutations affectant la survie. Dans le cas de mutations affectant la croissance en revanche, une dépression de consanguinité non-nulle est prédite mais les résultats différent considérablement entre les deux types de mutations. En effet, la dépression de consanguinité prédite est plus faible que celle réellement endurée par les individus dans le cas des mutations affectant le coût de maintien, et considérablement plus élevée dans le cas de mutations affectant le coût de production. Ainsi, en fonction de la manière dont les mutations affectent la croissance, l'utilisation de la taille à maturité comme proxy de fitness donne des résultats radicalement différents et toujours éloignés de la vraie dépression de consanguinité. Ce résultat montre que s'intéresser aux classes d'âge les plus jeunes pour quantifier la dépression de consanguinité peut parfois mener à des résultats trompeurs.

Population complète. Nous appliquons également une approche de fardeau de consanguinité à la population complète. Tous les individus sont échantillonnés. Leur coefficient de consanguinité et leur taille sont mesurés. Le proxy de fitness utilisé ici

est donc la taille d'individus de tous les âges. La dépression de consanguinité prédite est de nouveau faiblement en accord avec la dépression de consanguinité réelle, bien qu'améliorée par rapport à celle prédite par les pedigrees.

4.2 Introduction

Inbreeding depression, that is the decrease in fitness of inbred individuals relative to the outbred, has been shown to occur in many animal and plant species (Charlesworth and Charlesworth, 1987). Because inbreeding depression can lead small populations to extinction (Hedrick and Kalinowski, 2000), and because it is central to our understanding of the evolution of many traits, such as mating systems or dispersal for instance (Epinat and Lenormand, 2009; Duputié and Massol, 2013; Barrett and Harder, 2017), accurately quantifying inbreeding depression is an important issue in evolutionary and in conservation biology (Hedrick and Garcia-Dorado, 2016).

In self-fertile species, the most straightforward approach to quantifying inbreeding depression consists in the direct comparison of fitnesses of selfed and outcrossed individuals, as Darwin did in his seminal experiments (Darwin, 1876). In outcrossing species on the other hand, this method cannot be applied, but an estimate of inbreeding depression is often obtained using the inbreeding load (Morton et al., 1956), that is the slope of the regression of log-fitness against the inbreeding coefficient (Wright, 1922), under the assumption that all factors influencing fitness act independently so that they have multiplicative effects (Charlesworth and Willis, 2009; Nietlisbach et al., 2019). Both of these approaches rely on one's ability to measure fitness, which one would ideally do by counting the total number of offspring sired by individuals before they die, that is their lifetime reproductive success. Although this can be achieved in short-lived species, directly measuring lifetime reproductive success quickly becomes tedious when the species' typical lifespan exceeds a few months, let alone several years. Therefore,

traits or combinations of traits are often used as fitness proxies (Ågren and Schemske, 1993; Younginger et al., 2017). While these traits are more easily measured, the use of such indirect quantifications of fitness involves making assumptions on how they relate to fitness. Indeed, when traits are used as fitness proxies, results are often presented as inbreeding depression affecting said traits, that is traits' relative variations between selfed and outcrossed individuals (e.g. Husband and Schemske, 1996; Charlesworth and Willis, 2009; Angeloni et al., 2011). Since inbreeding depression is defined as a relative difference in fitness and not in phenotype, presenting results this way implicitly makes the assumption that traits relate to fitness linearly. Furthermore, it assumes that traits relate to fitness independently, that is that there is no trade-offs linking said traits so that differences between individuals on a given trait reflect differences in fitness and not different strategies. Besides, individuals are seldom followed over several years, so that inbreeding depression is often estimated based on measurements acquired during a given year, which involves assuming that these measurements can be used to extrapolate on individuals' entire lifespan in the case of perennial species.

An alternative approach to those relying on fitness proxies was developed by Ritland (1990). Noting that selection against the inbred should cause a variations in homozygosity between old and young individuals, they constructed an inbreeding depression estimator which solely relies on variations in the inbreeding coefficient (Wright, 1922) between categories of individuals. This estimator was designed to capture inbreeding depression occurring among juveniles, and the population genetics theory underlying it therefore assumes that selection only occurs in juveniles and that generations do not overlap. Ritland (1990)'s estimator is often used, however, in cases where these assumptions are not met, that is in perennial species, where there is generations overlap and selection likely occurs at all stages of life (e.g. Michalski and Durka, 2007; Delmas et al., 2014; Van Etten et al., 2015). Yet, the reliability of Ritland (1990)'s estimators in these cases was, to our knowledge, never assessed theoretically.

In summary, the methods that have been developed to quantify inbreeding depression rely on assumptions, either on the link between phenotype and fitness or on the genetic dynamics followed by the population. In this paper, we ask whether the results produced by two methods, namely Ritland (1990)'s and inbreeding load approaches, can be trusted when their underlying assumptions are not met. To do so, we take advantage of recently published work to perform individual-centered simulations of populations with varying degrees of generations overlap (Lesaffre and Billiard, 2020), in which fecundity is proportional to size which is controlled by an allometric growth model (West et al., 2001). We model mutations affecting survival and growth at a very large number of loci along with independent neutral markers. We measure the true inbreeding depression at mutation-selection balance and apply the methods discussed above to the population in order to assess their performance. In particular, we focus on the consequences of our results for the empirical observed increase in inbreeding depression in more long-lived species (Duminil et al., 2009; Angeloni et al., 2011).

4.3 Methods

4.3.1 Model outline

4.3.1.1 Simulated population

We consider a large population of constant size N. Each timestep begins with reproduction. Individuals reproduce by self-fertilisation at rate α and by random mating otherwise, and are assumed to contribute to the total gamete pool in proportion to their size. Offspring are assumed to be produced in large excess compared to the spots available in the population. Individuals then survive to the next flowering season with a probability S which is independent of their age but may depend on their genotype (see section 4.3.1.3 below). If they survive, they grow according to an allometric growth

model described in section 4.3.1.2. Otherwise, they are replaced by an offspring recruited among those produced during the last reproduction episode (no seed-banking), so that the population size is kept constant.

4.3.1.2 Growth model

The growth model used in this study was developed by West et al. (2001) and is based on the allometric relationship between size and resource uptake (Enquist, 2002). The total energy available for growth and maintenance at age t, B_t, scales as a $\frac{3}{4}$ power law of body size, G_t, so that

$$B_t = B_0 G_t^{\frac{3}{4}}, \tag{4.1}$$

where B_0 is the basal metabolic rate. Part of this energy is then consumed to maintain the existing body parts, at a maintenance cost c per body part, and the remaining energy is devoted to producing new body parts at a production cost ε per body part. Hence, growth can be described by the differential equation

$$B_0 G_t^{\frac{3}{4}} = c\, G_t + \varepsilon \frac{\mathrm{d}G_t}{\mathrm{d}t}. \tag{4.2}$$

Under this model, individual size reaches a plateau when the energy consumed by maintenance equals the total available energy.

4.3.1.3 Genetic assumptions

Selected loci. Individual genomes are assumed to be composed of two linear chromosomes of length L (in cM) along which mutations occur at any position during meiosis. The position of mutations on chromosomes is sampled from a uniform distribution, and the number of mutations occurring per chromosome per meiosis follows a Poisson distribution with mean U, where U is the mutation rate per haploid genome.

Hence, an infinite number of biallelic loci is effectively modeled (Roze and Michod, 2010). Recombination between loci occurs in proportion to chromosome length L. The number of crossing-overs per gamete follows a Poisson distribution with mean L and their position is sampled from a uniform distribution. Once crossing-overs have been placed, segments of chromosomes are exchanged accordingly. In three separate models, mutations affecting three different traits are considered. In the first two models, mutations affect their bearer's growth by increasing its maintenance cost (c) in one model or its production cost (ε) in the other. In the third model, mutations reduce their bearer's probability of being recruted as a juvenile and its survival probability between flowering seasons. Loci affect traits multiplicatively, so that for any trait z, an individual's trait value is given by

$$z = z_0(1 \pm s)^{n_{\mathrm{Hom}}}(1 \pm sh)^{n_{\mathrm{Het}}}, \tag{4.3}$$

where z_0 is the trait value in the absence of mutations, and n_{Hom} and n_{Het} are the number of homozygous and heterozygous mutations borne by the individual, respectively. The parameter s denotes the phenotypic effect of mutations when homozygous, and $0 \leqslant h \leqslant \frac{1}{2}$ is their dominance coefficient. The $\pm$ sign indicates that depending on the trait affected by mutations, they may increase it, as with the growth costs (c and ε), or decrease it, as it is the case for survival.

Neutral markers. In order to apply the indirect measurement methods mentionned in the introduction, we also model neutral loci under an n-allele model. These loci are assumed to segregate independently from the selected loci in order to avoid any bias caused by genetic associations. n_{mark} neutral loci are modelled. At each locus d_{mark} distinct alleles segregate, and mutation occurs at rate u_{mark} at each marker to maintain allelic diversity. When an allele mutates, a new identity is sampled among the d_{mark} existing ones.

4.3.2 Measurements

Simulations are run for 5×10^5 timesteps and measurements are performed at each timestep. To ensure that mutation-selection equilibrium is reached at the selected loci, results are averaged over the last 10^5 timesteps and the 95% confidence interval around the average is computed.

4.3.2.1 True inbreeding depression

To quantify the true magnitude of inbreeding depression occuring in the population, we count the number of times each individual is chosen as a parent until it dies, that is, its lifetime reproductive output, and compute the relative difference in average lifetime reproductive success between selfed and outcrossed individuals.

4.3.2.2 Inbreeding coefficients

Wright (1922)'s inbreeding coefficient, that is the probability that the two alleles borne by an individual are identical by descent (Crow and Kimura, 1970), is measured at each timestep. For each individual, its inbreeding coefficient F_{ind} is measured using neutral markers, by comparing its observed proportion of heterozygous loci (H_{obs}) to the expected heterozygosity in the population H_{exp}, that is

$$F_{\text{ind}} = 1 - \frac{H_{\text{obs}}}{H_{\text{exp}}}, \tag{4.4}$$

where H_{exp} is given by

$$H_{\text{exp}} = \frac{1}{n_{\text{mark}}} \sum_{i=1}^{n_{\text{mark}}} \left(1 - \sum_{j=1}^{d_{\text{mark}}} {p_{ij}}^2 \right), \tag{4.5}$$

where p_{ij} is the frequency of the jth allele at the ith neutral locus in the population (Ritland, 1996). The inbreeding coefficient of any subpopulation is then obtained by

averaging the F_{ind} of the relevant individuals.

4.3.2.3 Ritland (1990)'s inbreeding depression estimator

Derivation. Ritland (1990)'s inbreeding depression estimate is derived as follows. Let $\bar{F}_{t-1}$ be the population's inbreeding coefficient at time $t-1$, and assume the population is annual, with adults self-fertilising at rate α. Then, the inbreeding coefficient among the offspring is given by

$$\bar{F}_t^R = \frac{\alpha}{2}(1 + \bar{F}_{t-1}). \tag{4.6}$$

Suppose now that reproduction is followed by selection occurring on juvenile survival so that a fraction w_s (resp w_o) of selfed (resp. outcrossed) offspring survives. Then, the inbreeding coefficient in the adult population at time t is given by

$$\bar{F}_t = \frac{\alpha w_s}{\alpha w_s + (1-\alpha) w_o} \bar{F}_t^R. \tag{4.7}$$

Denoting $\tilde{w} = \frac{w_s}{w_o}$ the relative fitness of selfed offspring compared to the outcrossed, equation (4.7) can be rearranged to obtain

$$\tilde{w} = \frac{2(1-\alpha)\bar{F}_t}{\alpha(1 + \bar{F}_{t-1} - 2\bar{F}_t)}. \tag{4.8}$$

Inbreeding depression can then be obtained as $\delta = 1 - \tilde{w}$.

Application to the simulated population. We use Ritland's estimator in two ways. First, it is applied to juveniles as it is intended, that is $\bar{F}_{t-1}$ is taken to be the inbreeding coefficient of the parental population and $\bar{F}_t$ that of the juveniles recruited during one timestep. Importantly however, to avoid any sampling bias due to differences in the number of offspring incorporated in the population over the course of one timestep, 500 juveniles are created using the parental population and $\bar{F}_t$ is measured among them rather than among those actually incorporated in the population. Second, Ritland's

estimate is applied to the whole population at time t, that is $\bar{F}_t$ is taken to be the inbreeding coefficient of the whole population, including surviving adults, instead of that of the juveniles.

4.3.2.4 Inbreeding load estimates

In addition to Ritland's estimator, we use inbreeding load approaches to estimate inbreeding depression. Under the assumption that the various factors influencing fitness act roughly independently from one another, the inbreeding load is measured as the slope of the log-fitness against the inbreeding coefficient, that is

$$\ln w_i = aF_i + b, \tag{4.9}$$

where w_i and F_i are the fitness and inbreeding coefficient of the ith individual (Morton et al., 1956). Inbreeding depression can be deduced from a and b by assuming that $F = \frac{1}{2}$ among the selfed and $F = 0$ among the outcrossed, so that their respective average fitnesses are given by

$$w_s = e^{\frac{a}{2}+b} \text{ and } w_o = e^b, \tag{4.10}$$

and we have

$$\delta = 1 - e^{\frac{a}{2}}. \tag{4.11}$$

To apply such an approach, it is necessary to define a fitness proxy. In this paper, we use plant size as a fitness proxy because plant biomass is one of the most widely used proxies and is generally associated positively with fecundity (Younginger et al., 2017). We estimate inbreeding load in two separate situations.

Whole population inbreeding load estimate. We first estimate inbreeding load using the whole population as sample, that is, the inbreeding coefficient and the size

of each individual is measured, and Equation (4.9) is fitted to these data using the least-squares method. This method requires that at least some selfing is occuring in the population, otherwise $F_i \approx 0$ for all individuals.

Inbreeding load estimate using juveniles in pedigrees. In strictly outcrossing species, inbred individuals are often generated artificially using pedigrees (Charlesworth and Willis, 2009). We imitate such an approach by constructing 25 pedigrees with three inbreeding classes. For each pedigree, three different outcrossed parents are sampled from the population and four outcrossed F1s are generated from two crosses, so that one of the parents is common to all individuals. These constitute the class with $F = 0$. Then, the progeny of each cross is used to produce two full siblings crosses, so that we get four individuals with $F = 0.25$. Finally, crosses are performed between progenies to generate four half siblings crosses with $F = 0.125$. Hence, we get three inbreeding classes which we use to estimate the inbreeding load using least-squares fitting of Equation (4.9). The crosses are summarised in Figure 4.1. In this approach, individuals of age 1 are produced so that it would be both unrealistic and tedious to consider the size reached by the produced individuals at all ages across the age distribution of the population. Therefore, the fitness proxy used here is the size reached at maturity.

4.4 Results & (some) discussion

All results are presented as a function of life expectancy. Life expectancy, given survival probability S, is

$$E = \frac{1}{1 - S}. \tag{4.12}$$

In the case of mutations affecting survival, initial life expectancy, that is neglecting the effect of mutations, is used.

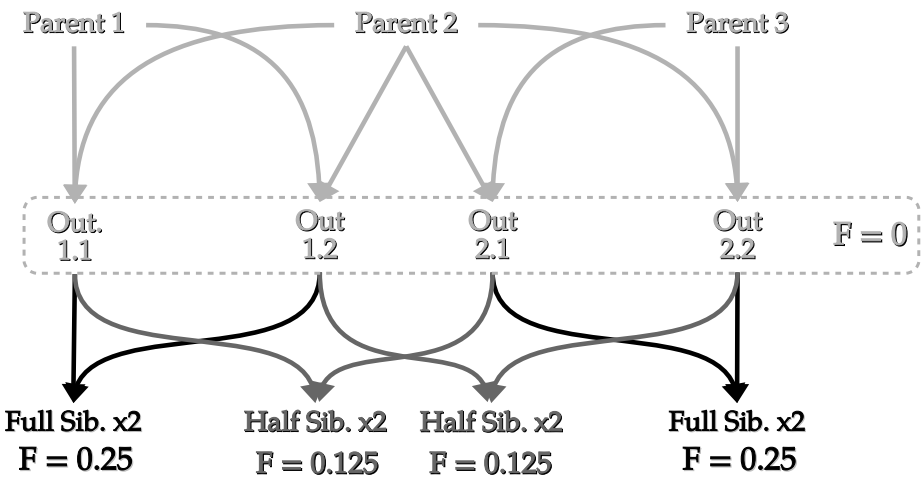

FIGURE 4.1 – Pedigree produced in simulations. Darkening shades of gray indicate increasing levels of inbreeding.

4.4.1 Ritland's estimator

We applied Ritland's estimator to juveniles and to the whole population for a range of life expectancies and selfing rates. Figure 4.2 shows the output of this approach along with true inbreeding depression for life expectancies ranging from 1 to 500 years and three levels of selfing, $\alpha = 0.1$, 0.5 and 0.9. Each row represents a mutation type.

Mutations affecting growth. In the case of mutations affecting growth (top two rows), one would expect Ritland's estimator to predict no inbreeding depression since no selection is occurring on juvenile survival. Yet, none zero inbreeding depression is predicted in every case both on juveniles and using the whole population, and this effect grows more substantial as the selfing rate increases. As shown in Appendix 4.5.2, this pattern is not observed when deleterious mutations are absent ($U = 0$), which indicates that this effect can be attributed to their presence. Under the assumption

that

$$\bar{F}_t \leqslant \frac{1}{2}(1 + F_{t-1}), \tag{4.13}$$

which is always valid when selfing is the only source of inbreeding, as it is the case in our simulations, inbreeding depression should be detected ($\delta > 0$) provided that

$$\bar{F}_t < \frac{\alpha}{2}(1 + \bar{F}_{t-1}). \tag{4.14}$$

Since the term on the right-hand side of equation (4.14) is the expression of $\bar{F}_t$ as a function of $\bar{F}_{t-1}$ in the absence of selection (equation 4.6), it follows that any factor reducing $\bar{F}_t$ should result in Ritland's estimator predicting inbreeding depression. When deleterious mutations affect growth, inbred individuals contribute less to the gamete pool because they tend to be smaller due to their increased homozygosity for said mutations. Thus, homozygosity at neutral markers in gametes is lower than among adults which results in a smaller $\bar{F}_t$ and in inbreeding depression being therefore detected even though no selection occurs on recruitment. This result means that the predictions of Ritland's estimator should not be discussed solely in terms of juvenile survival, as if it were capable of isolating this component of fitness. Instead, one should always keep in mind that any process reducing homozygosity among juveniles compared with parents, including selection against the inbred during reproduction, will interfere with its predictions.

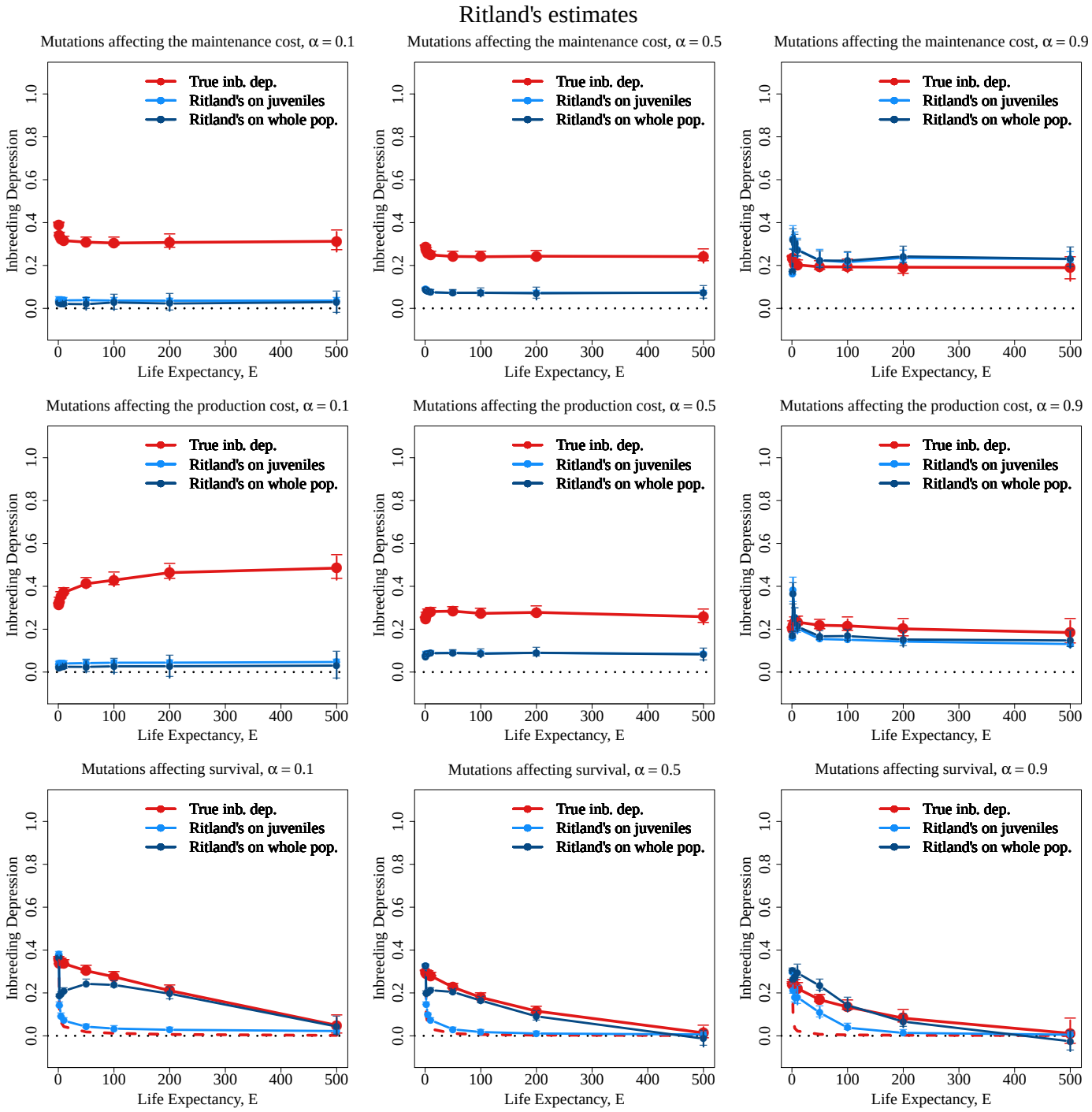

FIGURE 4.2 – Results obtained using Ritland's estimator. The predicted inbreeding depression levels obtained when the estimator is applied to juveniles (light blue) and to the whole population (dark blue) are plotted along with true inbreeding depression (solid red). The expected inbreeding depression affecting juveniles is plotted in dashed red (Appendix 4.5.1). Error bars represent the 95% confidence intervals. Each row corresponds to a mutation type, and each column to an inbreeding level. Other parameter values are $s = 0.05$ for mutations affecting growth and $s = 0.005$ for those affecting survival, $c = \varepsilon = 0.01$, $h = 0.25$, $U = 0.5$, $n_{\text{mark}} = d_{\text{mark}} = 20$, $u_{\text{mark}} = 10^{-3}$, $L = 20$ and $N = 5,000$.

Mutations affecting survival. In the case of mutations affecting survival, the same kind of bias occurs, although less conspicuously. Indeed, inbred individuals tend to be younger and hence smaller, which results in a smaller inbreeding coefficient among gametes. For low selfing rates, the bias is small and Ritland's estimator of inbreeding depression affecting juvenile survival agrees with analytical predictions, but this is no longer the case for $\alpha = 0.9$. We also applied Ritland's estimator to the whole population following recruitment, so that selection against the inbred among adults would be accounted for. The predictions of this second method agree relatively well with the true inbreeding depression, at least for high life expectancies. Hence, it may constitute a meaningful way of quantifying inbreeding depression in perennials. Its application to actual populations seems, however, rather unrealistic considering that variations in the inbreeding coefficient have to be captured across all age-classes in the population for the prediction to be accurate, which is feasible in a simulation where all individuals can be sampled, but is unreasonable in practice.

4.4.2 Inbreeding load

We applied an inbreeding load approach to the population in two ways. The considered parameter values are the same as for Ritland's estimator results, and the results are presented in a similar way in Figure 4.3.

Pedigrees. In the first approach, we mimicked the construction of artificial pedigrees in order to obtain three classes of inbreeding through controlled crosses, and size at maturity was used as a fitness proxy. As expected of such an approach no inbreeding depression is detected when mutations affect survival, as all individuals reach maturity at the same age and therefore do not differ in size. On the other hand, some inbreeding depression is predicted to occur when mutations affect growth, but results differ widely between mutation types. Indeed, the predicted inbreeding depression is generally lower

than the true inbreeding depression in the case of mutations affecting the maintenance cost, except in the annual case ($E = 1$), while it is much larger in every case for mutations affecting the production cost. Hence, depending on the way mutations affect growth, the use of size at maturity as a fitness proxy yields completely different results which are not consistent with the actual magnitude of inbreeding depression. This result shows that focusing on early ages to quantify inbreeding depression may sometimes lead to misleading results.

Whole population. In the second approach, we measured the size and inbreeding coefficient of each individual at each timestep to compute the inbreeding load. In the case of mutations affecting survival, this approach predicts almost no inbreeding depression because selfed individuals only differ in size from the outcrossed by the fact that they tend to be younger. As for mutations affecting growth, this approach yields somewhat better estimates of inbreeding depression than pedigrees as expected because the whole population, and not only juveniles, is accounted for here. These estimates remain, however, in poor agreement with true inbreeding depression, highlighting again the fact that fitness proxies measured at a given age may not always be used to extrapolate over the whole lifespan.

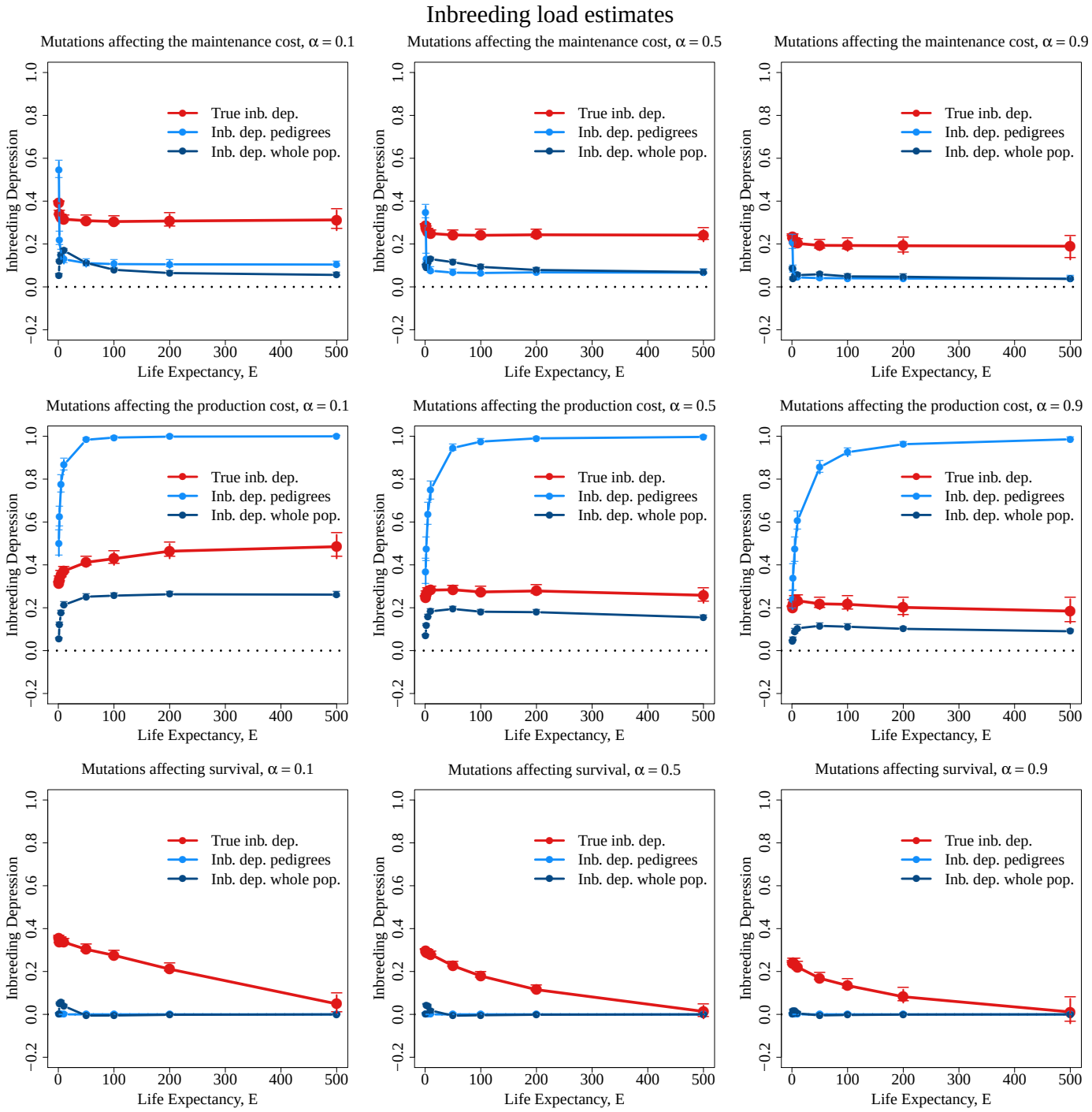

FIGURE 4.3 – Results obtained using inbreeding load approaches. The predicted inbreeding depression levels obtained using pedigrees (light blue) and to the whole population (dark blue) are plotted along with true inbreeding depression (solid red). Error bars represent the 95% confidence intervals. Each row corresponds to a mutation type, and each column to an inbreeding level. Other parameter values are $s = 0.05$ for mutations affecting growth and $s = 0.005$ for those affecting survival, $c = \varepsilon = 0.01$, $h = 0.25$, $U = 0.5$, $n_{\text{mark}} = d_{\text{mark}} = 20$, $u_{\text{mark}} = 10^{-3}$, $L = 20$ and $N = 5,000$.

Acknowledgements

This work was funded by the European Research Council (NOVEL project, grant #648321). The authors also thank the Région Hauts-de-France, and the Ministère de l'Enseignement Supérieur et de la Recherche (CPER Climibio), and the European Fund for Regional Economic Development for their financial support.

4.5 Appendices

4.5.1 Analytical prediction of juvenile inbreeding depression

Let us define indicator variables X_i and $\overset{*}{X}_i$, which are worth one when a deleterious allele is present at the ith locus on the paternally and maternally inherited chromosomes, respectively, and zero otherwise. Using these variables, the fitness of a juvenile can be written as

$$w = \prod_i \left[1 - sh\left(X_i + \overset{*}{X}_i\right) - s(1-2h)X_i\overset{*}{X}_i\right]. \tag{A1}$$

These indicator variables are such that $\mathbb{E}\left[X_i\right] = \mathbb{E}\left[\overset{*}{X}_i\right] = p_i$ where p_i is the frequency of the deleterious allele at the ith locus. They can be used to define centered variables $\zeta_i = X_i - p_i$ and $\overset{*}{\zeta}_i = \overset{*}{X}_i - p_i$ which can in turn be used to quantify genetic associations within and between loci (Kirkpatrick et al., 2002). Injecting centered variables into equation (A1) and taking its natural log, we have

$$\ln w \approx -s\sum_i \left[2hp_i + h\left(\zeta_i + \overset{*}{\zeta}_i\right) + (1-2h)\left(\zeta_{i,i} + p_i\left(\zeta_i + \overset{*}{\zeta}_i\right) + p_i^2\right)\right], \tag{A2}$$

where $\zeta_{i,i} = \zeta_i \times \overset{*}{\zeta}_i$. The expectation of this product, $D_{i,i} = \mathbb{E}\left[\zeta_{i,i}\right]$, quantifies the excess in homozygotes at the ith locus compared with panmixia. The mean log-fitness is given by

$$\overline{\ln w} = -s\sum_i \left[2hp_i + (1-2h)\left(D_{i,i} + p_i^2\right)\right], \tag{A3}$$

and to leading order in $\Delta \ln w = \ln w - \overline{\ln w}$, fitness can be approximated by

$$w = e^{\ln w} \approx e^{\overline{\ln w}}\left(1 + \Delta \ln w\right). \tag{A4}$$

Since $\mathbb{E}\left[\Delta \ln w\right] = 0$, the mean fitness is given by

$$\bar{w} \approx e^{\overline{\ln w}}. \tag{A5}$$

Among selfed individuals following reproduction, the excess in homozygotes is given by

$$D_{i,i}^{\text{self}} = \frac{1}{2}\left(p_i q_i + D_{i,i}\right) \approx \frac{1}{2}\left(1 + F\right) p_i, \tag{A6}$$

while it is worth zero among the selfed. Hence, inbreeding depression among juveniles is given by

$$\delta = 1 - \frac{\bar{w}_{\text{self}}}{\bar{w}_{\text{out}}} \approx 1 - e^{-s(1-2h)\frac{1+F}{2}\sum_i p_i}. \tag{A7}$$

A more detailed explanation of this result can be found in Roze (2015).

4.5.2 Results in the absence of mutations ($U = 0$)

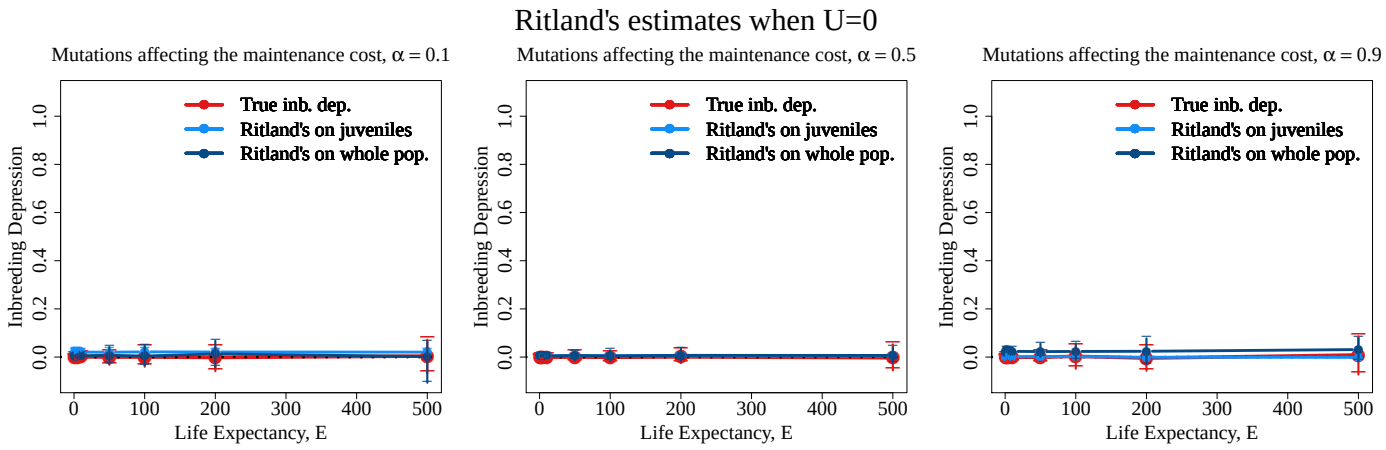

FIGURE S1 – Results obtained using Ritland's estimator with no deleterious mutations ($U = 0$). The predicted inbreeding depression levels obtained when the estimator is applied to juveniles (light blue) and to the whole population (dark blue) are plotted along with true inbreeding depression (solid red). The expected inbreeding depression affecting juveniles is plotted in dashed red (Appendix 4.5.1). Error bars represent the 95% confidence intervals.

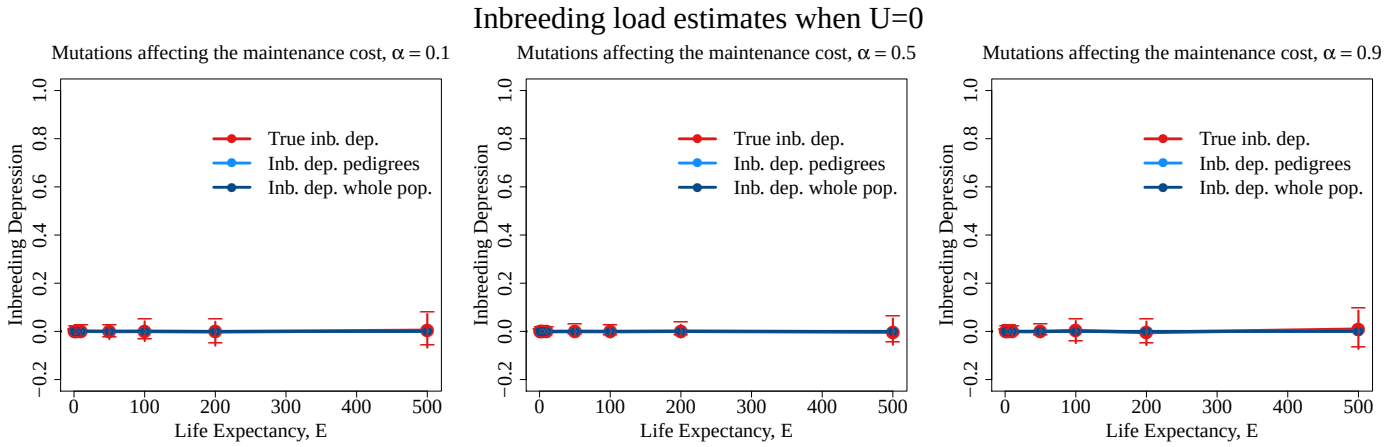

FIGURE S2 – Results obtained using the inbreeding load approach with no deleterious mutations ($U = 0$). The predicted inbreeding depression levels obtained when the estimator is applied to juveniles (light blue) and to the whole population (dark blue) are plotted along with true inbreeding depression (solid red). The expected inbreeding depression affecting juveniles is plotted in dashed red (Appendix 4.5.1). Error bars represent the 95% confidence intervals.

Chapitre 5

Conclusion

5.1 Retour à la question initiale

La question initiale du travail présenté dans cette thèse était celle des mécanismes évolutifs assurant le maintien de l'association empiriquement observée entre annualité et autofécondation d'une part, et pérennialité et allofécondation d'autre part (Figure 2.4, Barrett and Harder, 1996; Munoz et al., 2016). Dans le premier chapitre, nous avons étudié la coévolution entre autofécondation et longévité et nous avons identifié la dépression de consanguinité affectant la survie des individus entre les saisons de floraison comme le principal mécanisme capable de maintenir cette association, en faisant l'hypothèse d'une dépression de consanguinité fixée par un paramètre. Nous avons montré, en accord avec Morgan et al. (1997), que l'action répétée d'une telle dépression de consanguinité réduit drastiquement le succès reproducteur total des individus issus d'autofécondation et ce d'autant plus qu'une espèce plus longévive est considérée, aboutissant ainsi au maintien de l'allofécondation pour une plus large gamme de paramètres dans ces espèces. De plus, lorsque cette dépression de consanguinité n'est pas suffisante pour empêcher l'évolution de l'autofécondation, nous avons montré qu'elle induit une réduction de la longévité.

Dans le deuxième chapitre, nous avons levé l'hypothèse d'une dépression de consanguinité fixée par un paramètre, et nous avons montré que sous l'hypothèse la mieux empiriquement et théoriquement supportée quant aux bases génétiques de la dépression de consanguinité (Charlesworth and Charlesworth, 1987; Charlesworth and Willis, 2009), c'est-à-dire sous l'hypothèse qu'elle est générée par un fardeau de mutations délétères partiellement récessives maintenues en faibles fréquences dans la population, la dépression de consanguinité générée par des mutations affectant la survie des individus d'une saison à une autre, loin de croître avec la pérennialité, diminue à l'équilibre mutation-sélection chez les espèces plus longévives du fait de l'intensification de la sélection endurée par les mutations chez ces dernières. Ce résultat fait voler en éclat les conclusions que nous avons obtenues au premier chapitre. Pourtant, l'association entre annualité et autofécondation est bien observée et les données empiriques suggèrent que la dépression de consanguinité augmente fortement chez les espèces pérennes (Petit and Hampe, 2006; Duminil et al., 2009; Angeloni et al., 2011). La question que nous nous sommes posée est donc la suivante. Si l'effet répété de la dépression de consanguinité sur la survie n'est pas responsable de son augmentation avec la pérennialité lorsqu'elle est mesurée à l'échelle de l'existence complète des individus, quels sont les mécanismes responsables de cette augmentation ?

Dans le chapitre deux, nous avons exploré l'idée que des variations de l'effet sur la fitness des mutations avec la pérennialité, dont des mutations affectant la survie ne sont qu'un cas particulier, pourraient en être l'origine. En nous servant d'un modèle physiologique de croissance (West et al., 2001), nous avons montré que seules des mutations ayant un effet fort sur la fitness diminuant nettement avec la pérennialité auraient en principe la capacité de générer une augmentation de la dépression de consanguinité chez les espèces plus longévives. Dans le chapitre trois, nous nous sommes intéressés à la seule autre hypothèse proposée dans la littérature pour expliquer cette augmentation, l'accumulation de mutations somatiques (Scofield and Schultz, 2006). Nous avons

étudié l'évolution du taux de mutation dans une population où les individus accumulent des mutations héritables à la fois pendant leur croissance et pendant la méiose, et nous avons montré que même si le taux de mutation par événement mutagène diminue avec la pérennialité à l'ESS, l'accumulation de mutations délétères pendant la croissance induit quand même une augmentation de la dépression de consanguinité à l'équilibre mutation-sélection lorsque des mutations affectant la fécondité des modules qui composent les individus sont considérées. L'intensité de l'augmentation observée dépend toutefois fortement de la mutagénicité relative de la méiose par rapport aux mitoses qui se produisent durant la croissance. En effet, plus la mutagénicité relative de la méiose est élevée, et plus l'augmentation de la dépression de consanguinité avec la pérennialité est faible.

Le modèle présenté dans le chapitre trois suggère donc que l'accumulation de mutations somatiques pourrait générer l'augmentation de la dépression de consanguinité observée chez les espèces pérennes, en accord avec l'argument verbal de Scofield and Schultz (2006). Cependant, ce modèle a été construit de la manière la plus simple possible et ignore de ce fait plusieurs mécanismes potentiellement importants pour la dynamique suivie par les mutations à la fois au sein des organismes et au niveau de la population complète. Trois éléments ressortent particulièrement.

Premièrement, un modèle de croissance linéaire est supposé ici, ce qui ne correspond pas à la dynamique de croissance typiquement observée chez les végétaux. En effet, l'accroissement de la biomasse suit en général une dynamique exponentielle dans un premier temps avant de ralentir du fait de contraintes hydrauliques par exemple. Ainsi, la vitesse à laquelle les individus accumulent des mutations devrait varier avec l'âge et la taille des individus. De plus, le modèle linéaire supposé dans ce chapitre crée une structure d'apparentement et une structuration en âges des modules au sein de l'individu qui ne correspondent probablement pas à celles observées dans la réalité, le processus de croissance suivi par les végétaux étant intrinsèquement branchant (section

1.2.1 du premier chapitre de l'introduction générale).

Deuxièmement, du fait de ces hypothèses simplistes sur le développement des individus et de la manière dont les mutations délétères affectent la fitness, aucune sélection intra-organisme ne se produit dans le modèle bien que ce type de processus ait été mis en évidence entre modules (e.g. Edwards et al., 1990) et soit susceptible d'affecter significativement l'accumulation de mutations délétères en conduisant à l'élimination par compétition des modules défectueux au cours de la croissance (Otto and Orive, 1995).

Troisièmement, la mutagénicité relative de la méiose et des mitoses qui se produisent pendant la croissance, qui joue un rôle déterminant pour l'évolution du taux de mutation et le lien entre dépression de consanguinité et pérennialité, est contrôlée ici par un paramètre. Cette quantité, en dépit de son rôle important, a d'ailleurs une signification relativement vague, dans la mesure où elle englobe à la fois de possibles différences de mutagénicité des divisions cellulaires méiotiques et mitotiques, du fait par exemple de cassures double brins, mais aussi des différences de nombre de divisions cellulaires.

Les trois limites que nous venons d'énumérer montrent la relation intime qui existe entre l'accumulation de mutations somatiques et les processus physiologiques qui permettent la croissance végétale, que le modèle présenté dans le chapitre trois capture mal. Ainsi, il n'est pas possible de conclure rigoureusement sur le rôle des mutations somatiques dans l'association entre magnitude de la dépression de consanguinité - et donc taux d'autofécondation - et pérennialité simplement à partir des résultats de ce modèle.

5.2 Vers une une théorie physiologique et génétique de l'évolution végétale

Pour aller plus loin, il est nécessaire de modéliser la croissance végétale d'une manière plus fidèle à nos connaissances de la physiologie. C'est ce travail que j'ai voulu initier à travers le premier chapitre de l'introduction générale et sur lequel je reviens dans cette ultime section, en essayant ce faisant d'en identifier les carences et zones d'ombre qui demeurent pour moi, au moins au niveau intuitif. Il n'est bien sûr pas envisageable de construire un modèle manipulable - et par manipulable j'entends « dont on peut tirer des résultats analytiques » -, incorporant toute la complexité du fonctionnement physiologique des végétaux. À mon sens, le rôle du théoricien n'est d'ailleurs pas de produire un modèle le plus détaillé possible mais de discerner dans le réel les mécanismes qui importent afin de produire, ultimement, des prédictions testables. En cohérence avec cette vision du travail théorique, la section dédiée à la physiologie végétale dans le premier chapitre de l'introduction générale fait d'ailleurs plusieurs raccourcis, ne mentionnant pas tel ou tel type particulier de méristème, de dormance induite ou de dédifférenciation, et ne s'attardant pas non plus sur les processus chimiques qui induisent les mécanismes qu'elle décrit.

5.2.1 Modéliser la croissance végétale

Dans le vaste programme que constitue l'intégration de processus physiologiques dans une théorie de génétique des populations végétales, concentrons-nous d'abord sur l'établissement d'un modèle minimal de la croissance végétale auquel nous ajouterons ensuite différents processus, comme la mutation ou la sélection, suivant la question envisagée. Je propose à cet effet de considérer le méristème comme l'unité de la croissance, c'est-à-dire comme son plus petit élément constitutif insécable, sa particule élémentaire. Cette hypothèse est évidemment une approximation, un méristème étant

en fait le lieu de la production des cellules, qui sont les véritables « atomes » de la croissance. Dans la mesure où c'est le type du méristème qui détermine le destin des cellules qu'il produit, je veux croire à sa pertinence. Sous cette hypothèse, une plante peut être décrite comme une population de méristèmes dont il s'agit de comprendre la dynamique des naissances et morts pour en décrire la croissance. Bien qu'il en existe d'autres types plus compliqués, comme les méristèmes d'inflorescence, ou plus spécifiques à un type de plantes, comme le cambium, je propose de simplement supposer ici l'existence de deux types de méristèmes, les méristèmes végétatifs et floraux.

Un méristème végétatif produit de nouveaux méristèmes végétatifs, ce qui correspond aux naissances dans la population de méristèmes, à une cadence qui dépend des ressources qui lui parviennent. Incorporer la dynamique de répartition des ressources, c'est-à-dire de compétition, entre les méristèmes revient à considérer un processus de branchement dans lequel le taux de division dépend de la taille totale de l'arbre produit, ce qui est analytiquement difficile à prendre en compte. Une approximation possible est de simplement ignorer cette dynamique en supposant que le taux de division est constant, mais la taille de la population de méristèmes risque alors d'exploser. Une solution de repli est de considérer que le taux de division décroît avec le temps ou (peut-être) d'incorporer une régulation densité-dépendante sur la mort des méristèmes.

Lors de la floraison, une partie des méristèmes végétatifs exécute une transition vers l'état floral pour produire l'ensemble des organes reproducteurs et des gamètes. Même si de rares cas de réversion ont été documentés dans la littérature chez certaines lignées mutantes, la règle semble être que cette transition est irréversible (Friedman, 2020). Deux zones d'ombre importantes persistent pour moi autour d'elle. La première se situe au niveau de son timing : un méristème végétatif actif peut-il se transformer en cours d'activité en un méristème floral, ou son destin est-il déterminé pendant la dormance ? La seconde concerne la durée de vie d'un méristème floral : une fois

la saison de floraison terminée, ce méristème persiste-t-il pour les floraisons futures ou devient-il inactif ? Dans tous les cas, sous l'hypothèse que la transition florale est irréversible, cette dernière constitue une première manière de « mourir » du point de vue de la population de méristèmes. Remarquons que le trade-off entre croissance et reproduction, abondamment documenté chez les plantes (Obeso, 2002), est ainsi inclus dans le modèle. Une seconde manière de mourir pour un méristème végétatif est simplement de cesser son activité, ce qui peut se produire par sénescence ou lorsque le méristème ne reçoit plus suffisamment de ressources, du fait de la compétition avec toutes les branches qui ont émergé des méristèmes axillaires qu'il a produits et de contraintes hydrauliques.

5.2.2 Questions théoriques associées

À partir d'un modèle minimal de croissance tel que celui esquissé dans la section précédente, il est possible de s'intéresser à la fois à des questions de théorie des traits d'histoire de vie et de génétique des populations. Par exemple, la notion d'effort reproducteur prend un sens nouveau sous ce modèle, puisqu'il peut à présent être quantifié comme une proportion de méristèmes convertis plutôt qu'une proportion de ressources dépensées. Ses liens avec la croissance et la survie sont intrinsèques au modèle puisque la reproduction implique la conversion de méristèmes végétatifs, et plus une population (de méristèmes) est petite, plus elle a de chances de s'éteindre du fait de la stochasticité démographique. Un cas extrême est celui où tous les méristèmes sont dédiés à la reproduction lorsque cette dernière est déclenchée, comme c'est le cas chez les espèces annuelles obligatoires. Ici, la taille de la population est brutalement réduite à zéro. Nous pourrions donc envisager d'étudier l'évolution des histoires de vie non plus à travers de trade-offs arbitrairement supposés mais qui émergent du modèle de croissance considéré, en revisitant les questions classiques de l'évolution de l'effort reproducteur ou de l'âge à maturité. Une complexification possible du modèle serait de considérer

le rôle de la dormance méristèmatique dans le maintien de l'appareil végétatif aérien entre les saisons de floraison. En effet, nous avons vu dans l'introduction générale que nombre d'espèces herbacées produisent un nouvel appareil végétatif tous les ans après que le précédent ait été détruit pendant la mauvaise saison.

Par ailleurs, l'évolution des histoires de vie est également discutée en lien avec la sénescence au travers de la théorie de la lignée somatique jetable, qui propose que la sénescence est due à une accumulation inexorable de mutations délétères au fil des divisions cellulaires que l'investissement dans le maintien de la lignée somatique au détriment de la reproduction permet de ralentir (Medawar, 1952; Kirkwood, 1977). Cette théorie est basée sur une conception unitaire des organismes, dans laquelle la défaillance d'un organe implique celle de l'individu. Chez les organismes modulaires, dont les plantes sont un formidable exemple, le flétrissement d'un organe pourrait ne pas nécessairement porter à conséquence car ils sont constitués d'un ensemble d'organes (modules) physiologiquement équivalents. En effet, ceux qui acquièrent une mutation délétère pourraient être éliminés soit parce que cette mutation arrête leur développement, soit parce qu'elle leur inflige un désavantage compétitif par rapport aux autres modules. Cette hypothèse verbale proposée Finch (1990) pourrait également être testée théoriquement à partir du modèle minimal de croissance proposé dans la section précédente. Intuitivement, nous pouvons considérer l'analogie suivante. Si une plante est une population de méristèmes végétatifs, elle s'apparente alors à une population d'« individus » clonaux et pérennes, et chaque nouveau clone dans cette population accumule des mutations qui affectent ses capacités de survie ou de reproduction clonale. Cette population est forcément petite au moment de la germination et pourrait être sujette au Cliquet de Muller (Felsenstein, 1974), c'est-à-dire à l'accumulation inexorable de mutations délétères jusqu'à son effondrement, du fait de son incapacité à produire des recombinants. Il s'en suit que l'expression d'une sénescence à l'échelle de l'organisme complet, c'est-à-dire l'augmentation dans le temps de sa probabilité

d'extinction, devrait finir par être observée tôt ou tard suivant la taille que l'individu atteint, son taux de mutation et l'effet sur la fitness des mutations qu'il accumule, particulièrement lorsqu'elles sont à l'état hétérozygote. Le fait qu'aucune sénescence n'est en général observée chez les plantes (Baudisch et al., 2013) suggère qu'elles parviennent à repousser l'effondrement mutationnel (*mutational meltdown*) suffisamment loin pour mourir de causes indépendantes de la sénescence, c'est-à-dire que la population de méristèmes se situe dans un régime dans lequel l'accumulation de mutations est suffisamment lente. L'intensité de la reproduction, c'est-à-dire de la transition de méristèmes à l'état floral, pourrait en outre jouer un rôle important dans cette dynamique car elle implique une surmortalité du point de vue de la population de méristèmes, établissant ainsi un lien naturel entre sénescence, effort reproducteur, et accumulation de mutations somatiques à travers la croissance. L'augmentation de la dépression de consanguinité avec la pérennialité et l'association entre système de reproduction et histoire de vie qui en résulte probablement, point de départ de cette thèse que nous avons un peu perdu de vue dans cette section finale, pourrait être la conséquence de cette fascinante interaction.

Bibliographie

Aarssen, L., Schamp, B., and Pither, J. (2006). Why are there so many small plants ? implications for species coexistence. *Journal of Ecology*, 94 :569–580.

Abrams, P. and Ludwig, D. (1995). Optimality theory, Gompertz'law, and the disposable soma theory of senescence. *Evolution*, 49 :1055–1066.

Abu Awad, D., Billiard, S., and Tran, V. C. (2016). Perenniality induces high inbreeding depression in self-fertilising species. *Theoretical Population Biology*, 112 :43–51.

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

Abu Awad, D. and Roze, D. (2019). Epistasis, inbreeding depression and the evolution of self-fertilization. *BioRXiv*.

Agrawal, A. (2011). Current trends in the evolutionary ecology of plant defence. *Functional Ecology*, 25 :420–432.

Agren, J. and Schemske, D. (1993). The cost of defense against herbivores : an experimental study of trichome production in brassica rapa. *The American Naturalist*, 141 :338–350.

Aidoo, M.and Terlouw, D., Kolczak, M., McElroy, P., ter Kuile, F., Kariuki, S., Nahlen, B., Lal, A., and Udhayakumar, V. (2002). Protective effects of the sickle cell gene against malaria morbidity and mortality. *The Lancet*, 359 :1311–1312.

Angeloni, F., Ouborg, N., and Leimu, R. (2011). Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation*, 144 :35–43.

Antolin, M. and Strobeck, C. (1985). The population genetics of somatic mutation in plants. *The American Naturalist*, 126(1) :52–62.

Arbel-Eden, A. and Simchen, G. (2019). Ele-

vated mutagenicity in meiosis and its mechanism. *BioEssays*, 41(4).

Auld, J. and Rubio de Casas, R. (2012). The correlated evolution of dispersal and mating-system traits. *Evolutionary Biology*, 40 :185–193.

Barrett, S. (2010). Darwin's legacy : the forms, function and sexual diversity of flowers. *Philosophical Transactions of the Royal Society B*, 365 :351–368.

Barrett, S. C. (2002). The evolution of plant sexual diversity. *Nature Reviews Genetics*, 3 :274–284.

Barrett, S. C. and Harder, L. D. (1996). The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London B*, 351 :1271–1280.

Barrett, S. C. and Harder, L. D. (2017). The ecology of mating and its evolutionary consequences in seed plants. *Annual Review of Ecology, Evolution, and Systematics*, 48 :135–157.

Barton, N. and Turelli, M. (1991). Natural and sexual selection on many loci. *Genetics*, 127 :229–255.

Bataillon, T. and Kirkpatrick, M. (2000). Inbreeding depression due to mildly deleterious mutations in finite populations : size does matter. *Genetics Research*, 75 :75–81.

Baucher, M., El Jaziri, M., and Vandeputte, O. (2007). From primary to secondary growth : origin and development of the vascular system. *Journal of Experimental Botany*, 58 :3485–3501.

Baudisch, A., Salguero-Gómez, R., O.R., J., Wrycza, T., Mbeau-Ache, C., Franco, M., and Colchero, F. (2013). The pace and shape of senescence in angiosperms. *Journal of Ecology*, 101 :596–606.

Bell, G. (1980). The costs of reproduction and their consequences. *The American Naturalist*, 116 :45–76.

Benton, T. and Grant, A. (1999). Optimal reproductive effort in stochastic, density-dependent environments. *Evolution*, 53 :677–688.

Bernard, C., Compagnoni, A., and Salguero-Gómez, R. (2020). Testing finch's hypothesis : The role of organismal modularity on the escape from actuarial senescence. *Functional Ecology*, 34 :88–106.

Bobiwash, K., Schultz, S., and Schoen, D. (2013). Somatic deleterious mutation rate in a woody plant : estimation from phenotypic data. *Heredity*, 111 :338–344.

Bonser, S. and Aarssen, L. (1996). Meristem allocation : a new classification theory for adaptive strategies in herbaceous plants. *Oikos*, pages 347–352.

Bonser, S. and Aarssen, L. (2006). Meristem allocation and life-history evolution in herbaceous plants. *Botany*, 84 :143–150.

Brommer, J. (2000). The evolution of fitness in life-history theory. *Biological Reviews*, 75 :377–404.

Brown, K. and Kelly, J. (2020). Severe inbreeding depression is predicted by the "rare allele load" in *Mimulus guttatus*. *Evolution*, 74 :587–596.

Brunet, J. (1992). Sex allocation in hermaphroditic plants. *Trends in Ecology & Evolution*, 7 :79–84.

Burian, A., Barbier de Reuille, P., and Kuhlemeier, C. (2016). Patterns of stem cell divisions contribute to plant longevity. *Current Biology*, 26 :1385–1394.

Charlesworth, B. (1980). *Evolution in age-structured populations*. Cambridge Studies in Mathematical Biology, first edition.

Charlesworth, B. (1990). Optimization models, quantitative genetics, and mutation. *Evolution*, 44 :520–538.

Charlesworth, B. (1994). *Evolution in age-structured populations*. Cambridge Studies in Mathematical Biology, second edition.

Charlesworth, B. and Leon, J. (1976). The relation of reproductive effort to age. *The American Naturalist*, 110 :449–459.

Charlesworth, B., Morgan, M., and Charlesworth, D. (1991). Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. *Genetics*, 57 :177–194.

Charlesworth, D. and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, 18 :237–268.

Charlesworth, D., Morgan, M., and Charlesworth, B. (1990). Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution*, 44 :1469–1489.

Charlesworth, D. and Willis, J. (2009). The genetics of inbreeding depression. *Nature reviews genetics*, 10 :783–796.

Charnov, E. (2005). Reproductive effort is inversely proportional to average adult life span. *Evolutionary Ecology Research*, 7 :1221–1222.

Charnov, E. and Schaffer, W. (1973). Life-history consequences of natural selection : Cole's result revisited. *The American Naturalist*, 107 :791–793.

Cichoń, M. (1997). Evolution of longevity through optimal resource allocation. *Proceedings of the Royal Society of London B*, 264 :1383–1388.

Clauss, M. and Mitchell-Olds, T. (2006). Population genetic structure of *arabidopsis lyrata* in europe. *Molecular Ecology*, 15 :2753–2766.

Cody, M. (1966). A general theory of clutch size. *Evolution*, pages 174–184.

Cole, L. (1954). The population consequences of life history phenomena. *The Quarterly review of biology*, 29 :103–137.

Crow, J. (1958). Some possibilities for measuring selection intensities in man. *Human Biology*, 61 :763–775.

Crow, J. (1999). Dominance and overdominance. *Genetics and exploitation of heterosis in crops*, pages 49–58.

Crow, J. and Kimura, M. (1970). *An introduction to population genetics theory*. Black Burn Press.

Damgaard, C., Weiner, J., and Nagashima, H. (2002). Modelling individual growth and competition in plant populations : growth curves of *Chenopodium album* at two densities. *Journal of Ecology*, 90 :666–671.

Darwin, C. (1876). *The effects of cross and self fertilisation in the vegetable kingdom*. London John Murray.

DeHaan, L. and Van Tassel, D. (2014). Useful insights from evolutionary biology for developing perennial grain crops. *American Journal of Botany*, 101 :1801–1819.

Delmas, C., Cheptou, P.-., Escaravage, N., and Pornon, A. (2014). High lifetime inbreeding depression counteracts the reproductive assurance benefit of selfing in a mass-flowering shrub. *BMC Evolutionary Biology*, 14 :243.

Duminil, J., Hardy, O., and Petit, R. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology*, 9 :177.

Duputié, A. and Massol, F. (2013). An empiricist's guide to theoretical predictions on the evolution of dispersal. *Interface Focus*, 40 :20130028.

Durand, E., Chantreau, M., Le Veve, A., Stetsenko, R., Dubin, M., Genete, M., Llaurens, V., Poux, C., Roux, C., Billiard, S., Vekemans, X., and Castric, V. (2020). Evolution of self-incompatibility in the brassicaceae : Lessons from a textbook example of natural selection. *Evolutionary Applications*, 13 :1279–1297.

Eckert, C., Samis, K., and Dart, S. (2006). Reproductive assurance and the evolution of uniparental reproduction in flowering plants. In Harder, L. and Barrett, S., editors, *Ecology and Evolution of Flowers*, pages 183–203. Oxford University Press.

Edwards, P., Wanjura, W., Brown, W., and Dearn, J. (1990). Mosaic resistance in plants. *Nature*, 347(6292) :434–434.

Ehrlén, J. and Lehtilä, K. (2002). How perennial are perennial plants ? *Oikos*, 98 :1070–1072.

Enquist, B. (2002). Universal scaling in tree and vascular plant allometry : toward a general quantitative theory linking plant form and function from cells to ecosystems. *Tree physiology*, 22 :1045–1064.

Enquist, B., Brown, J., and West, G. (1998). Allometric scaling of plant energetics and population density. *Nature*, 395 :163–165.

Epinat, G. and Lenormand, T. (2009). The evolution of assortative mating and selfing with in- and outbreeding depression. *Evolution*, 63 :2047–2060.

Eyre-Walker, A. and Keightley, P. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8 :610–618.

Fagen, R. (1972). An optimal life-history strategy in which reproductive effort decreases with age. *The American Naturalist*, 106 :258–261.

Felsenstein, J. (1971). Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics*, 68 :581–597.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, 78(2) :737–756.

Finch, C. (1990). *Longevity, senescence, and the genome*. University of Chicago Press.

Fisher, R. (1930). *The Genetical Theory of Natural Selection*. Oxford University Press.

Fisher, R. (1941). Average excess and average effect of a gene substitution. *Annals of Human Genetics*, 11 :53–63.

Franco, M. and Silvertown, J. (1996). Life-history variation in plants : an exploration of the fast-slow continuum hypothesis. *Philosophical Transactions of the Royal Society of London B*, 351 :1341–1348.

Friedman, J. (2020). The evolution of annual and perennial plant life histories : ecological correlates and genetic mechanisms. *Annual Review of Ecology, Evolution, and Systematics*, 51 :461–480.

Friedman, J. and Rubin, M. (2015). All in good time : understanding annual and perennial strategies in plants. *American Journal of Botany*, 102 :497–499.

Gadgil, M. and Bossert, W. (1970). Life historical consequences of natural selection. *The American Naturalist*, 104 :1–24.

Galipaud, M. and Kokko, H. (2020). Sex, senescence, sources and sinks. *Functional Ecology*, 34 :55–64.

Gardner, S. and Mangel, M. (1997). When can a clonal organism escape senescence ? *The American Naturalist*, 150 :462–490.

Geber, M. (1990). The cost of meristem limitation in polygonum arenastrum : negative genetic correlations between fecundity and growth. *Evolution*, 44 :799–819.

Gemmell, N. and Slate, J. (2006). Heterozygote advantage for fecundity. *PLoS One*, 1 :e125.

Gervais, C. and Roze, D. (2017). Mutation rate evolution in partially selfing and partially asexual organisms. *Genetics*, 207 :1561–1575.

Goodwillie, C., Kalisz, S., and Eckert, C. (2005). The evolutionary enigma of mixed mating systems in plants : Occurence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, 36 :47–79.

Groover, A., Mansfield, S., DiFazio, S., Dupper, G., Fontana, J., Millar, R., and Wang, Y. (2006). The populus homeobox gene arborknox1 reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. *Plant molecular biology*, 61 :917–932.

Gros, P.-A., Le Nagard, H., and Tenaillon, O. (2009). The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation. *Genetics*, 182 :277–293.

Grossenbacher, D., Briscoe Runquist, R., Goldberg, E., and Brandvain, Y. (2015). Geographic range size is predicted by plant mating system. *Ecology Letters*, 18 :706–713.

Hanlon, V., Otto, S., and Aitken, S. (2019). Somatic mutations substantially increase the per-generation mutation rate in the conifer *Picea sitchensis*. *Evolution letters*, 3 :348–358.

Harper, J. (1977). *Population biology of plants*. Academic Press.

Hartl, D. and Clark, A. (1997). *Principles of population genetics*. Sinauer associates Sunderland, MA.

Hedrick, P. and Garcia-Dorado, A. (2016). Understanding inbreeding depressio, purging and genetic rescue. *Trends in Ecology & Evolution*, 31 :940–952.

Hedrick, P. W. and Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual review of ecology and systematics*, 31(1) :139–162.

Herms, D. and Mattson, W. (1992). The dilemma of plants : to grow or defend. *The quarterly review of biology*, 67 :283–335.

Hirshfield, M. and Tinkle, D. (1975). Natural selection and the evolution of reproductive effort. *Proceedings of the national academy of sciences*, 72 :2227–2231.

Hofmeister, B., Denkena, J., Colomé-Tatché, M., Shahryary, Y., Hazarika, R., Grimwood, J., Mamidi, S., Jenkins, J., Grabowski, P., Sreedasyam, A., Shu, S., Barry, K., Lail, K., Adam, C., Lipzen, A., Sorek, R., Judrna, D., Talag, J., Wing, R., Hall, D., Tuskan, G., Schmutz, J., Johannes, F., and Schmitz, R. (2019). The somatic genetic and epigenetic mutation rate in a wild long-lived perennial *Populus trichocarpa*. *BioRxiv*.

Horvath, D., Anderson, J., Chao, W., and Fo-

ley, M. (2003). Knowing when to grow : signals regulating bud dormancy. *Trends in plant science*, 8 :534–540.

Husband, B. and Schemske, D. W. (1996). Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution*, 50 :54–70.

Igic, B., Lande, R., and Kohn, J. (2008). Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences*, 169 :93–104.

Johnson, M., Smith, S., and Rausher, M. (2009). Effects of plant sex on range distributions and allocation to reproduction. *New Phytologist*, 186 :769–779.

Jösson, K. and Tuomi, J. (1994). Costs of reproduction in a historical perspective. *Trends in ecology & evolution*, 9 :304–307.

Käfer, J., Marais, G., and Pannell, J. (2017). On the rarity of dioecy in flowering plants. *Molecular Ecology*, 26 :1225–1241.

Kiær, L., Weisbach, A., and Weiner, J. (2013). Root and shoot competition : a meta-analysis. *Journal of Ecology*, 101 :1298–1312.

Kimura, M. (1967). On the evolutionary adjustment of spontaneous mutation rates. *Genetics Research*, 9 :23–34.

Kimura, M. and Maruyama, T. (1966). The mutation load with epistatic gene interactions in fitness. *Genetics*, 54 :1337–1351.

Kirkpatrick, M., Johnson, T., and Barton, N. (2002). General models of multilocus evolution. *Genetics*, 161 :1727–1750.

Kirkwood, T. (1977). Evolution of ageing. *Nature*, 270 :301–304.

Kirkwood, T. (2017). The disposable soma theory. *The evolution of senescence in the tree of life*, pages 23–39.

Kirkwood, T. and Rose, M. (1991). Evolution of senescence : late survival sacrificed for reproduction. *Philosophical Transactions of the Royal Society of London. Series B : Biological Sciences*, 332 :15–24.

Klekowski, E. (2003). Plant clonality, mutation, diplontic selection and mutational meltdown. *Biological Journal of the Linnean Society*, 79 :61–67.

Klinkhamer, P., Meelis, E., De Jong, T., and Weiner, J. (1985). On the analysis of size-

dependent reproductive output in plants. *Functional Ecology*, 6 :308–316.

Kondrashov, A. (1985). Deleterious mutations as an evolutionary factor. II. facultative apomixis and selfing. *Genetics*, 111 :635–653.

Kozłowski, J. (1992). Optimal allocation of resources to growth and reproduction : implications for age and size at maturity. *Trends in Ecology & Evolution*, 7 :15–19.

Kozłowski, J. and Wiegert, R. (1987). Optimal age and size at maturity in annuals and perennials with determinate growth. *Evolutionary Ecology*, 1 :231–244.

Kozłowski, J. (1993). Measuring fitness in life-history studies. *Trends in Ecology & Evolution*, 8 :84–85.

Lande, R. and Schemske, D. (1985). The evolution of self-fertilization and inbreeding depression in plants. I. genetic models. *Evolution*, 39 :24–40.

Lande, R., Schemske, D., and Schultz, S. (1994). High inbreeding depression, selective interference among loci, and the threshold selfing rate for purging recessive lethal mutations. *Evolution*, 48 :965–978.

Lanfear, R. (2018). Do plants have a segregated germline ? *PloS Biology*, 16.

Lanfear, R., Ho, S., Davies, T., Moles, A., Aarssen, L., Swenson, N., Warman, L., Zanne, A., and Allen, A. (2013). Taller plants have lower rates of molecular evolution. *Nature Communications*, 4.

Lang, G. (1987). Dormancy - a new universal terminology. *HortScience*, 22 :817–820.

Lanner, R. and Connor, K. (2001). Does bristlecone pine senesce ? *Experimental Gerontology*, 36 :675–685.

Lehtilä, K. and Larsson, A. (2005). Meristem allocation as a means of assessing reproductive allocation. In *Reproductive allocation in plants*, pages 51–75. Academic Press.

Lesaffre, T. and Billiard, S. (2020). On deleterious mutations in perennials : inbreeding depression, mutation load and life-history evolution. *bioRxiv*.

Levins, R. (1968). *Evolution in changing environments*. Princeton University Press.

Lloyd, D. (1979). Some reproductive factors affecting the selection of self-fertilization in

plants. *The American Naturalist*, 113 :67–79.

Lloyd, D. (1992). Self- and cross-fertilization in plants. II. the selection of self-fertilization. *International Journal of Plant Sciences*, 153 :370–380.

Lynch, M. (2011). The lower bound to the evolution of mutation rates. *Genome Biology and Evolution*, 3 :1107–1118.

Magni, G. and Von Borstel, R. (1962). Different rates of spontaneous mutation during mitosis and meiosis in yeast. *Genetics*, 47(8).

Manna, F., Martin, G., and Lenormand, T. (2011). Fitness landscapes : An alternative theory for the dominance of mutation. *Genetics*, 189 :923–937.

Marbà, N., Duarte, C., and Agustí, S. (2007). Allometric scaling of plant life history. *Proceedings of the National Academy of Sciences*, 104 :15777–15780.

Maynard Smith, J. and Price, G. (1973). The logic of animal conflict. *Nature*, 246 :15–18.

Medawar, P. (1952). An unsolved problem of biology.

Michalski, S. and Durka, W. (2007). High selfing and high inbreeding depression in peripheral populations of *Juncus atratus*. *Molecular Ecology*, 16 :4715–4727.

Mithöfer, A. and Boland, W. (2012). Plant defense against herbivores : chemical aspects. *Annual review of plant biology*, 63 :431–450.

Morgan, M. (2001). Consequences of life history for inbreeding depression and mating system evolution in plants. *Proceedings of the Royal Society of London B*, 268 :1817–1824.

Morgan, M., Schoen, D., and Bataillon, T. (1997). The evolution of self-fertilization in perennials. *The American Naturalist*, 150 :618–638.

Morton, N., Crow, J., and Muller, H. (1956). An estimate of the mutational damage in man from data on consanguineous marriages. *Proceedings of the National Academy of Sciences of the United States of America*, 42 :855–863.

Munné-Bosch, S. (2015). Senescence : is it universal or not ? *Trends in plant science*, 20 :713–720.

Munné-Bosch, S. (2018). Limits to tree growth and longevity. *Trends in plant science*, 23 :985–993.

Munoz, F., Violle, C., and Cheptou, P.-O. (2016). CSR ecological strategies and plant mating systems : outcrossing increases with competitiveness but stress-tolerance is related to mixed mating. *Oikos*, 125 :1296–1303.

Mylius, S. and Diekmann, O. (1995). On evolutionarily stable life histories, optimization and the need to be specific about density dependence. *Oikos*, pages 218–224.

Nietlisbach, P., Muff, S., Reid, J., Whitlock, M., and Keller, L. (2019). Nonequivalent lethal equivalents : Models and inbreeding metrics for unbiased estimation of inbreeding load. *Evolutionary Applications*.

Niklas, K. and Enquist, B. (2001). Invariant scaling relationships for interspecific plant biomass production rates and body size. *Proceedings of the National Academy of Sciences*, 98 :2922–2927.

Nord-Larsen, T., Damgaard, C., and Weiner, J. (2006). Quantifying size-asymmetric growth among individual beech trees. *Canadian Journal of Forest Research*, 36 :418–425.

Obeso, J. R. (2002). The costs of reproduction in plants. *New Phytologist*, 155 :321–348.

Orr, A., Padovan, A., Kainer, D., Külheim, C., Bromham, L., Bustos-Segura, C., Foley, W., Haff, T., Hsieh, J.-F., Morales-Suarez, A., Cartwright, R., and Lanfear, R. (2020). A phylogenomic approach reveals a low somatic mutation rate in a long-lived plant. *Proceedings of the Royal Society of London B*, 287.

Ott, J., Klimešová, J., and Hartnett, D. (2019). The ecology and significance of below-ground bud banks in plants. *Annals of botany*, 123 :1099–1118.

Otto, S. and Day, T. (2011). *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press.

Otto, S. and Orive, M. (1995). Evolutionary consequences of mutation and selection within an individual. *Genetics*, 141 :1173–1187.

Partridge, L. and Sibly, R. (1991). Constraints in the evolution of life

histories. *Philosophical Transactions of the Royal Society of London. Series B : Biological Sciences*, 332 :3–13.

Perrin, N. and Sibly, R. (1993). Dynamic models of energy allocation and investment. *Annual Review of Ecology and Systematics*, 24 :379–410.

Peters, R. (1983). *The ecological implications of body size*. Cambridge studies in ecology.

Petit, R. and Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, 37 :187–214.

Phillips, P. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics*, pages 855–867.

Pineda-Krch, M. and Lehtilä, K. (2002). Cell lineage dynamics in stratified shoot apical meristems. *Journal of Theoretical Biology*, 291 :495–505.

Plomion, C., Aury, J.-M., Amselem, J., Leroy, T., Murat, F., Duplessis, S., Faye, S., Francillonne, N., Labadie, K., Le Provost, G., Lesur, I., Bartholomé, J., Faivre-Rampant, P., Kohler, A., Leplé, J.-C., Chantret, N., Chen, J., Diévrat, A., Alaeitabar, T., Barbe, V., Belser, C., Bergès, H., Bodénès, C., Bogeat-Triboulot, M.-B., Bouffaud, M.-L., Brachi, B., Chancerel, E., Cohen, D., Couloux, A., Da Silva, C., Dossat, C., Ehrenmann, F. Gaspin, C., Grima-Pettenati, J., Guichoux, E., Hecker, A., Herrmann, S., Hugueney, P., Hummel, I., Klopp, C., Lalanne, C., Lascoux, M., Lasserre, E., Lemainque, A. Desprez-Loustau, M.-L., Luyten, I., Madoui, M.-A., Mangenot, S., Marchal, C., Maumus, F., Mercier, J., Michotey, C., Panaud, O., Picault, N., Rouhier, N., Rué, O., Rustenholz, C., Salin, F., Soler, M., Tarkka, M., Velt, A., Zanne, A., Martin, F., Wincker, P., Quesneville, H., Kremer, A., and Salse, J. (2018). Oak genome reveals facets of long lifespan. *Nature Plants*, 4 :440–452.

Price, C., Gilooly, J., Allen, A., Weitz, J., and Niklas, K. (2010). The metabolic theory of ecology : prospects and challenges for plant biology. *New Phytologist*, 188 :696–710.

Primack, R. (1979). Reproductive effort in annual and perennial species of plantago (plantaginaceae). *The American Naturalist*, 114 :51–62.

Pugliese, A. and Kozlowski, J. (1990). Op-

timal patterns of growth and reproduction for perennial plants with persisting or not persisting vegetative parts. *Evolutionary Ecology*, 4 :75–89.

Ågren, J. and Schemske, D. (1993). Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata*. *Evolution*, 47 :125–135.

Renner, S. (2014). The relative and absolute frequencies of angiosperm sexual systems : dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of botany*, 101 :1588–1596.

Reznick, D. (1985). Costs of reproduction : an evaluation of the empirical evidence. *Oikos*, 44 :257–267.

Ricker, W. (1954). Stock and recruitment. *Journal of the Fisheries Board of Canada*, 11 :559–623.

Ritland, K. (1990). Inferences about inbreeding depression based on changes of the inbreeding coefficient. *Evolution*, 44 :1230–1241.

Ritland, K. (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetics Research*, 67 :175–185.

Roze, D. (2015). Effects of interference between selected loci on the mutation load, inbreeding depression, and heterosis. *Genetics*, 201 :745–757.

Roze, D. and Blanckaert, A. (2014). Epistasis, pleiotropy, and the mutation load in sexual and asexual populations. *Evolution*, 86 :137–149.

Roze, D. and Michod, R. (2010). Deleterious mutations and selection for sex in finite, diploid populations. *Genetics*, 184 :1095–1112.

Roze, D. and Rousset, F. (2005). Inbreeding depression and the evolution of dispersal rates : a multilocus model. *The Amrican Naturalist*, 166 :708–721.

Salguero-Gómez, R. (2017). Applications of the fast–slow continuum and reproductive strategy framework of plant life histories. *New Phytologist*, 213 :1618–1624.

Salguero-Gómez, R., Jones, O. R., Jongejans, E., Blomberg, S. P., Hodgson, D. J., Mbeau-Ache, C., Zuidema, P. A., de Kroon, H., and Buckley, Y. M. (2016). Fast–slow continuum and reproductive strategies structure plant life-history varia-

tion worldwide. *Proceedings of the National Academy of Sciences*, 113(1) :230–235.

Schaffer, W. (1974a). Optimal reproductive effort in fluctuating environments. *The American Naturalist*, 108 :783–790.

Schaffer, W. (1974b). Selection for optimal life histories : the effects of age structure. *Ecology*, 55 :291–303.

Schaffer, W. M. and Schaffer, M. V. (1977). *The adaptive significance of variations in reproductive habit in the Agavaceae*, pages 261–276. Macmillan Education UK.

Schmid-Siegert, E., Sarkar, N., Iseli, C., Calderon, S., Gouhier-Darimont, C., Chrast, J., Cattaneo, P., Schütz, F., Farinelli, L., Pagni, M., Schneider, M., Voumard, J., Jaboyedoff, L., Fankhauser, C., Hardtke, C., Keller, L., Pannell, J., Reymond, A., Robinson-Rechavi, M., Xenarios, I., and Reymond, P. (2017). Low number of fixed somatic mutations in a long-lived oak tree. *Nature Plants*, 12 :926–929.

Schoen, D. and Schultz, S. (2019). Somatic mutation and evolution in plants. *Annual Review of Ecology, Evolution and Systematics*, 50 :2.1–2.25.

Schultz, S. and Scofield, D. (2009). Mutation accumulation in real branches : fitness assays for genomic deleterious mutation rate and effect in large-statured plants. *The American Naturalist*, 174 :163–175.

Scofield, D. and Schultz, S. (2006). Mitosis, stature and evolution of plant mating systems : low-ϕ and high-ϕ plants. *Proceedings of the Royal Society of London B*, 273 :275–282.

Sicard, A. and Lenhard, M. (2011). The selfing syndrome : a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany*, 107 :1433–1443.

Silvertown, J., Franco, M., and Perez-Ishiwara, R. (2001). Evolution of senescence in iteroparous perennial plants. *Evolutionary Ecology Research*, 3 :393–412.

Simberloff, D. and Leppanen, C. (2019). Plant somatic mutations in nature conferring insect and herbicide resistance. *Pest Management Science*, 75(1) :14–17.

Stearns, S. (1989). Trade-offs in life-history evolution. *Functional Ecology*, 3 :259–268.

Stearns, S. (1992). *The Evolution of Life Histories*. Oxford University Press.

Stearns, S. and Crandall, R. (1981). Quantitative predictions of delayed maturity. *Evolution*, 35 :455–463.

Stebbins, G. (1950). *Variation and evolution in plants*. Columbia University Press.

Thomas, H. (2013). Senescence, ageing and death of the whole plant. *New Phytologist*, 197 :696–711.

Thomas, H., Thomas, H., and Ougham, H. (2000). Annuality, perenniality and cell death. *Journal of Experimental Botany*, 51 :1781–1788.

Trinder, C., Brooker, R., Davidson, H., and Robinson, D. (2012). Dynamic trajectories of growth and nitrogen capture by competing plants. *New Phytologist*, 193 :948–958.

Van Etten, M., Tate, J., Anderson, S., Kelly, D., Ladley, J., Merrett, M., Peterson, P., and Robertson, A. (2015). The compounding effects of high pollen limitation, selfing rates and inbreeding depression leave a new zealand tree with few viable offspring. *Annals of Botany*, 116 :833–843.

Vayssières, A., Mishra, P., Roggen, A., Neumann, U., Ljung, K., and Albani, M. (2020). Vernalization shapes shoot architecture and ensures the maintenance of dormant buds in the perennial arabis alpina. *New Phytologist*, 227 :99–115.

Wang, L., Ji, Y., Hu, Y., Hu, H., Jia, X., Jiang, M., Zhang, X., Zhao, L., Zhang, Y., Jia, Y., Hurst, L., and Tian, D. (2019). The architecture of intra-organism mutation rate variation in plants. *PLoS biology*, 17(4).

Watkinson, A. and White, J. (1986). Some life-history consequences of modular construction in plants. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 313 :31–51.

Watson, J., Platzer, A., Kazda, A., Akimcheva, S., Valuchova, S., Nizhynska, V., Nordborg, M., and Riha, K. (2016). Germline replications and somatic mutation accumulation are independent of vegetative life span in arabidopsis. *Proceedings of the National Academy of Sciences*, 113 :12226–12231.

Weiner, J. (1990). Asymmetric competition in plant populations. *Trends in ecology & evolution*, 5 :360–364.

Weiner, J., Campbell, L., Pino, J., and

Echarte, L. (2009). The allometry of reproduction within plant populations. *Journal of Ecology*, 97 :1220–1233.

West, G., Brown, J., and Enquist, B. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276 :122–126.

West, G., Brown, J., and Enquist, B. (1999). The fourth dimension of life : fractal geometry and allometric scaling of organisms. *Science*, 284 :1677–1679.

West, G., Brown, J., and Enquist, B. (2001). A general model for ontogenetic growth. *Nature*, 413 :628–631.

Williams, G. (1957). Pleiotropy, natural selection, and the evolution of senescence. *Evolution*, 11 :398–411.

Williams, G. (1966a). *Adaptation and Natural Selection : A Critique of Some Current Evolutionary Thought*. Princeton University Press.

Williams, G. (1966b). Natural selection, the costs of reproduction, and a refinement of lack's principle. *The American Naturalist*, 100 :687–690.

Winn, A., Elle, E., Kalisz, S., Cheptou, P.-O., Eckert, C., Goodwillie, C., Johnston, M., Moeller, D., Sargent, R., and Vallejo-Marìn, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution*, 65 :3339–3359.

Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist*.

Younginger, B., Sirová, D., Cruzan, M., and Ballhorn, D. (2017). Is biomass a reliable estimate of plant fitness ? *Applications in plant sciences*, 5(2) :1600094.

Zhang, D. Y. (2000). Resource allocation and the evolution of self-fertilization in plants. *The American Naturalist*, 155 :187–199.

Züst, T. and Agrawal, A. (2017). Trade-offs between plant growth and defense against insect herbivory : an emerging mechanistic synthesis. *Annual review of plant biology*, 68 :513–534.

Contribution à une théorie physiologique et génétique de l'évolution végétale
Fardeau génétique, systèmes de reproduction et évolution du taux de mutation
dans les populations structurées en classes

RÉSUMÉ. Chez les Angiospermes, il existe une forte association entre histoire de vie et système de reproduction. En effet, la plupart des espèces autofécondantes sont annuelles alors que la majorité des espèces pérennes sont allofécondantes. Cette association est le point de départ du travail réalisé dans cette thèse. Dans le premier chapitre, je m'intéresse à la coévolution entre longévité et autofécondation sous l'hypothèse que la dépression de consanguinité affecte la survie des individus entre les saisons de floraison et qu'elle est fixée par un paramètre. Sous ces hypothèses, l'évolution de l'autofécondation est d'autant plus empêchée que la longévité de l'espèce considérée augmente, car la dépression de consanguinité subie à l'échelle de l'existence complète des individus devient plus forte. Lorsqu'elle se produit néanmoins, elle induit une diminution de la longévité à l'équilibre évolutif. Ces résultats sont en accord avec le patron empirique observé. Dans le chapitre deux, je lève l'hypothèse d'une dépression de consanguinité fixée par un paramètre en donnant une base génétique explicite à cette dernière. Je montre que des mutations affectant la survie, loin de générer une augmentation de la dépression de consanguinité avec la longévité, aboutissent à une diminution de cette dernière. Partant du constat qu'une augmentation de la dépression de consanguinité est pourtant bien observée chez les espèces pérennes, j'explore deux hypothèses pour expliquer cette augmentation. Dans le chapitre deux, j'étudie l'idée plus générale selon laquelle des variations des effets sélectifs des mutations avec la longévité, dont des mutations affectant la survie ne sont qu'un cas particulier, pourraient générer ce patron. Dans le chapitre trois, je m'intéresse aux conséquences de l'accumulation de mutations somatiques pour l'évolution du taux de mutation et la dépression de consanguinité résultante chez les plantes. Le chapitre quatre est quant à lui consacré à une évaluation théorique des méthodes d'estimation de la dépression de consanguinité. Je termine ce manuscrit en proposant des pistes pour l'élaboration d'une théorie physiologique et génétique de l'évolution végétale.

Mots-clefs : Populations structurées en classes, Systèmes de reproduction, Fardeau génétique, Évolution du taux de mutation, Génétique des Populations théorique

---

Contribution to a physiological and genetic theory of plant evolution
Mutation load, mating systems and mutation rate evolution in class-structured populations

ABSTRACT. In Angiosperms, there exists a strong association between life-history and mating system. Indeed, most predominantly selfing species are annual while the majority of perennials are outcrossers. This association is the starting point of the work presented in this thesis. In the first chapter, I study the joint evolution of lifespan and selfing assuming that inbreeding depression affects survival between flowering seasons and is fixed by a parameter. Under these assumptions, lifetime inbreeding depression increases as lifespan increases therefore preventing the evolution of self-fertilisation in a wider parameter range. When it occurs, it induces the evolution of shorter lifespans. These results are in agreement with the empirically observed pattern. In chapter two, I relax the assumption that inbreeding depression is fixed by a parameter, by assuming an explicit genetic basis. Far from generating higher inbreeding depression in more long-lived species, deleterious mutations affecting survival result in a decrease of inbreeding depression with longevity at mutation-selection equilibrium. Yet, increased inbreeding depression is empirically observed in long-lived species. In the following chapters, I explore two hypotheses to explain this increase. In chapter two, I study the more general idea that variations in the fitness effects of mutations with longevity, of which mutations affecting survival are a mere special case, may generate increased inbreeding depression in longer-lived species. In chapter three, I model the consequences of inheritable somatic mutations accumulating during growth for the evolution of the mutation rate and the resulting inbreeding depression in plants. As for chapter four, it is devoted to a theoretical evaluation of indirect inbreeding depression estimation methods. I end this manuscrit by proposing leads and ideas for the development of a physiological and genetic theory of plant evolution.

Keywords : Class-structured populations, Mating systems, Mutation load, Mutation rate evolution, Theoretical population genetics

La pensée n'est qu'un éclair au milieu d'une longue nuit, mais c'est cet éclair qui est tout.

- Henri Poincaré, *La valeur de la science* (1905).