Université de Lille 1
Thèse pour obtenir le grade de Docteur de l'Université de Lille 1

Discipline
Biologie Évolutive et Ecologie
U.F.R de Biologie

Death and sterility with a side of evolutionary suicide

The interplay of deleterious mutations and population size and the evolution of self-fertilisation

Présentée et soutenue publiquement le 16 décembre 2014 par

Diala Abu Awad

École Doctorale : Science de la Matière, du Rayonnement et de l'Environnement
Laboratoire : Génétique et Évolution des Populations Végétales, UMR CNRS 8198

Composition du jury:

Brian CHARLESWORTH, Professeur, University of Edinburgh	Rapporteur
Sylvain GLÉMIN, Chargé de Recherche CNRS, Université Montpellier II	Rapporteur
Christine DILLMANN, Professeur, Université Paris-Sud	Examinateur
Pauline LAFITTE, Professeur, École Centrale Paris	Examinateur
Emmanuelle PORCHER, Maître de Conférences, Muséum National d'Histoire Naturelle	Examinateur
Xavier VEKEMANS, Professeur, Université de Lille 1	Examinateur
Sylvain BILLIARD, Maître de Conférences, Université de Lille 1	Directeur de thèse
Viet Chi TRAN, Maître de Conférences, Université de Lille 1	Co-directeur de thèse

Remerciements

Ou commencer ? A partir de quelle époque, de quel âge ? Il faudrait tenir compte des expériences et découvertes qui ont influencé mes choix, mes envies, mes passions? Peut être commencer par le tout début? La première molécule capable de se répliquer, le premier organisme multicellulaire, la première espèce avec un système nerveux, puis celle avec une transmission culturelle. Je remercie le premier mâle du genre homo qui, en manifestant une créativité, a pu attirer plus de femelles, et aussi ces femelles qui ont été sensibles à sa créativité et à son intelligence et l'ont choisi comme père pour leurs enfants (donc au final je remercie la sélection sexuelle). Puis les individus qui ont inventé le feu, les outils, les langues, l'écriture, les chiffres... Tout ce qui nous a mené vers cette diversité intellectuelle éblouissante. L'histoire culturelle et les guerres qui ont façonné les esprits de mes parents et ma soeur, les rencontres brèves et moins brèves, qui à leur tour ont façonné le mien. L'amour des livres et la curiosité dont j'ai eu la chance d'hériter, et le milieu dans lequel j'ai grandi qui m'a permis de cultiver ces deux traits, qui les a renforcés. Je remercie donc ma famille et mes amis d'hier et d'aujourd'hui.

Je remercie Sylvain Billiard qui m'a fait découvrir une autre façon de voir le monde du vivant, Chi Tran qui m'a aidé à l'interpréter avec des chiffres et des lettres, Denis Roze pour m'avoir consacré autant de temps, les personnes qui ont supporté mes diatribes contre tout et rien, mes blagues nulles, et mon manque de tact, sans oublier les pompons, ceux qui les ont lancés et/ou été victimes d'au moins une attaque.

Je suis reconnaissante envers le laboratoire qui m'a accueilli depuis mes premiers pas dans le monde scientifique, le GEPV (qui a muté pour devenir le EEP), et aussi le Laboratoire Paul Painlevé. Sans oublier le personnel administratif du GEPV et de l'Ecole Polytechnique de Palaiseau de m'avoir simplifié la vie, ni la Chaire de Modélisation Mathématique et Biodiversité d'avoir suffisamment cru en ce travail pour financer cette thèse.

Contents

Prelude

"His progress through life was hampered by his tremendous sense of his own ignorance, a disability which affects all too few people."

- Terry Pratchett (Maskerade, 1994)

When wading through interdisciplinary waters, which tend be murky at best, having some indication of what is beneath your feet can be of some comfort. With this in mind, I have attempted to provide the members of the present committee who are of a mathematical persuasion with basic information and definitions that the average evolutionary biologist knows like the back of their hand. I hope that I have included all the necessary information, making the water slightly less murky for those who need it, without encumbering those who are equipped with the latest goggles and can see the tropical fish as clear as day.

Act 1

Introduction

> *"The presence of those seeking the truth is infinitely to be preferred to the presence of those who think they've found it."*
>
> - Terry Pratchett (Monstrous Regiment, 2004)

In this first chapter we present the motivations for this thesis and a brief summary of the results obtained. The works undertaken are put into perspective and can be found, presented in more detail, in Chapters 2, 3 and 4.

1.1 Mother Nature is a gambler

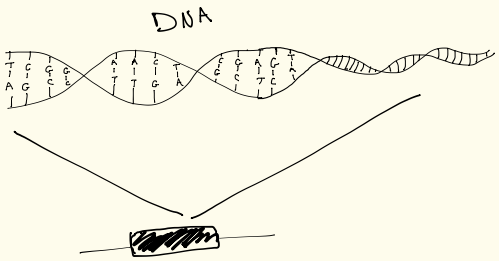
> *"God does not play dice with the universe; He plays an ineffable game of his own devising, which might be compared, from the perspective of any of the other players, to being involved in an obscure and complex version of poker in a pitch dark room, with blank cards, for infinite stakes, with a dealer who won't tell you the rules, and who smiles all the time."*
>
> - Terry Pratchett and Neil Gaiman (Good Omens, 1990)

The ultimate source of genetic variation is mutation, a random event along a strain of DNA (or RNA) that is then subjected to, among other things, selection, which shapes the evolution of species (see Boxes 1.1 and 1.2 for more information on the function of DNA). Of all of the possible changes that can be made in a genome (between substitutions, deletions, insertions, reversals) it is only intuitive that but a few mutations lead to a beneficial change (Agrawal and Whitlock 2011 and see Box 1.3 on Mutation). In spite of the fact that most new mutations have a negative effect on fitness (Keightley and Lynch, 2003), mutation has not been completely eradicated for two main reasons, first the potential introduction of advantageous alleles, and second the metabolic costs of reparation and conservation mechanisms (or the cost of fidelity) are high (Sniegowski et al, 2000). By allowing mutations to happen, organisms gamble with their fitness, there can be good gains, but there is also non-negligible risk. Organisms have therefore evolved mechanisms to deal with the constant introduction of deleterious mutations. For example, it has been suggested that diploidy has evolved in order to decrease the effects of somatic mutations (Orr, 1995) and sexual reproduction and recombination (Keightley and Otto, 2006) as a means to eliminate deleterious alleles present in the genome (Sniegowski et al, 2000).

These deleterious alleles have several consequences, as not only do they reduce individual fitness, they are also one of the causes of inbreeding depression (Charlesworth and Charlesworth 1987 and see Box 1.6). It has been proposed that on average, humans carry up to a thousand deleterious mutations in their genome (Agrawal and Whitlock, 2012). A stranger to population genetics (the study of the evolution of allelic frequencies within populations) may find it odd that deleterious mutations are not immediately eliminated by natural selection, why would such mutations persist? First of all, the majority of mutations are in fact of very small effect (Agrawal and Whitlock, 2011), hence they do not provide a sufficient disadvantage in order to be completely eradicated. Second as they are continuously introduced by

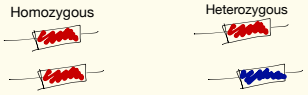
Box 1.1 - The basics

The genetic information necessary for the functioning and development of organisms is stored in structures known as chromosomes as DNA. DNA is mostly made up of nucleotides Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). DNA is double stranded and the two strands are complementary, each type of nucleotide pairs up only one other type (A with T and C with G).



A locus or a gene is a segment of a chromosome containing the information for the production of a protein. Different variants of a gene are known as alleles.

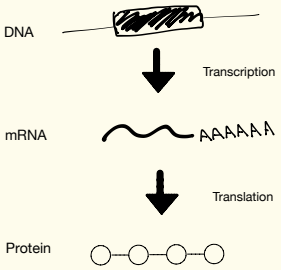
We consider diploid individuals (each individual carries two copies of each chromosome). An individual can therefore be homozygous or heterozygous at a given locus.



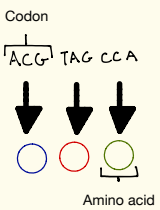
Mutation is the source of new alleles (see Box 1.3).

Box 1.2 - From DNA to proteins

As mentioned in Box 1.1, DNA is made up of two complementary strands of nucleotides. Getting from nucleotides to proteins occur in two main steps: 1) Transcription from DNA on to a template known as messenger RNA (or mRNA), and 2) Translation from mRNA to chains of amino acids, the building blocks of proteins.



The order of nucleotides in the DNA codes for a specific order of amino acids. Three consecutive nucleotides, known as a codon, representing a given amino acids. There are 64 possible codons but only 20 amino acids, giving rise to the possibility of several codons being used for a single amino acid.



Box 1.3 - Mutation

Mutations are a random event that can occur at any point in the genome leading to a change in the DNA sequence. Mutations occur due to errors during replication or due to exposure to radiation and groups several possible events on the genetic level. They can occur in somatic cells (any bodily cells that are neither gametes nor used for producing them) and in the germline (cells that produce gametes). Several examples of mutations are given below, showing the impact on the chain of nucleotides and on the final chain of amino acids (see Box 1.1 for more information).

Original sequence:

Nucleotides	TGT	CAT	ATT	TGT	AAA	GAG	AAT
Amino acids	C	H	I	C	K	E	N

Deletion: When one or more nucleotides are deleted

TGT	CAT	TGT	AAA	GAG	AAT
C	H	C	K	E	N

Insertion: When one or more nucleotides are inserted.

TGT	CAT	ATT	TGT	AAA	GCG	AGA	AT?
C	H	I	C	K	A	S	?

Substitution: When a single nucleotide is replaced by another. There are two kinds of substitutions, those that change the sequence of amino acids (non-synonymous) and those that do not (synonymous). Below is an example of a synonymous substitution:

TGT	CAT	ATT	TGT	AAA	GAA	AAT
C	H	I	C	K	E	N

Reversal: When the sequence of amino acids is reversed

TGT	TTA	TAC	TGT	AAA	GAG	AAT
C	L	Y	C	K	E	N

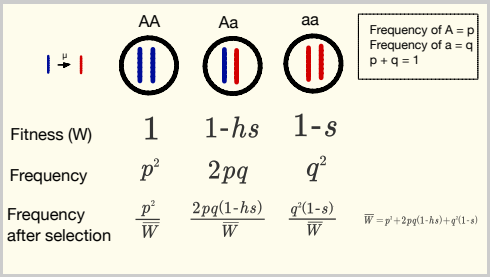
mutation, selection against them may not act fast enough to eliminate all of the mutations. At mutation-selection balance, when the frequency of deleterious mutations is at an equilibrium as the number of mutations introduced becomes compensated by those removed by selection (Gillespie, 1998), the reduction of fitness due to deleterious mutations is known as the mutation or genetic load, usually noted L (see Box 1.4). Understanding the mutational load and its potential consequences has played an important role in the evolution of population genetics models and this work is no exception. Throughout this thesis we will address the role of deleterious mutations in shaping both population demography (*i.e.* reproductive rates, population size, etc.) and genetic structure (*i.e.* allelic frequencies).

1.2 Spice things up: Adding some interplay between population size and selection

Before going further it seems appropriate to define what "fitness" is, as this is what is affected by the mutation load. Fitness is a complex trait, made up of several components, such as an individual's fecundity, it's capacity to

Box 1.4 - The mutation load

In the presence of deleterious mutations, mean population fitness $\overline{W} = 1 - L$, where L the mutation load at mutation selection balance, when the introduction of a deleterious mutation via mutation is countered by its elimination by selection. First let us consider a single locus with two alleles A (wild-type) and a (deleterious).



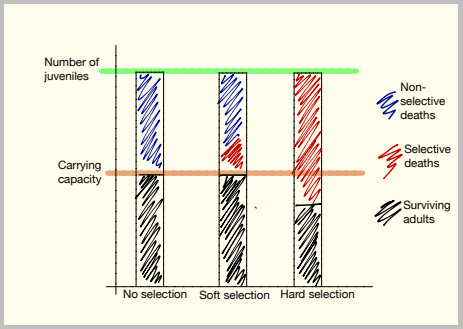
If now we consider that A mutates to a at a rate μ, then the frequency of A after mutation is $p(1-\mu)$. When the population is at mutation selection balance, hence the frequencies of p and q are fixed, then $p = \frac{(p^2 W^{AA} + pq W^{Aa})(1-\mu)}{\overline{W}}$ (Crow and Kimura, 1970, Chapter 6). When a is recessive ($h = 0$, this gives an expected $\overline{W} = 1 - \mu$.

obtain resources, size, growth rate, and the list goes on. However, a complex definition cannot be integrated into a simple model. Initially, fitness was defined as an individual's capacity to produce offspring, be it due to better survival, or actual fecundity (Haldane, 1924). The mutational load reduces fitness and fitness defines how many offspring are created, either in the absolute sense (the actual number of offspring) or relative sense (compared to the maximal number of offspring that can be produced). It seems plausible that as the mutation load decreases fitness and decreasing fitness reduces the reproductive output (or population growth rate as shown in Crow and Kimura, 1970, Chapter 1), a greater mutation load could lead to a smaller population size.

There does however seem to be controversy surrounding this statement (Agrawal and Whitlock, 2012). Wallace (1970) defined two terms to define the demographic consequences of the genetic load: hard and soft selection (see Box 1.5). Hard selection, as defined by (Wallace, 1970, p.89) is "selection resulting from conditions that an organism must meet to function as a breeding individual", i.e. lowered reproduction resulting from mutations that cause sterility or physical disability reducing survival. Soft selection on the other hand does not necessarily involve selection against deleterious mutations but rather a quantity of available resources supporting only a given number of individuals. If more offspring than can be supported by the environment are produced, then a proportion of the offspring will not survive, independently of their fitness. This therefore implies that if there is an increase in the genetic load (given that it does not reduce the reproductive capacity below one descendant per individual in which case the population would go extinct, (Wallace, 1970, p.87), the increase in the number of deaths due to it are absorbed by the deaths that would have occurred due to a lack

Box 1.5 -Hard and soft selection

If we consider that in a population the number of births (juveniles) is greater than the environment can support, death due to a lack of resources can occur, regulating the population so that its size remains at carrying capacity. When we take into account the possible effects of deleterious mutations (deaths due to mutation load in red), then there is either soft selection (which leaves the population size at carrying capacity) or hard selection, which decreases population size to below the carrying capacity.



of resources, leaving population size unaffected. As Agrawal and Whitlock (2012) pointed out, external factors regulating population density (or size), such as limited resources, are ever-present in nature, making the genetic load negligible when considering population size. However, although models have explored the potential effects of the genetic load on population size (Clarke, 1973; Agrawal and Whitlock, 2012, and references within), few have done so by allowing both population size and the mutational load to co-evolve (Bernardes, 1995; Abu Awad et al, 2014).

The separation between demographic models and genetic models seems to be an inherited trait. When pioneering population geneticists Fisher, Wright and Haldane began their work on understanding how allelic frequencies evolve, they made the hypothesis that populations were of infinite or of a fixed size. Hence as the evolution of allelic frequencies was considered to suffice for the understanding of the evolution of population size, in the models that stemmed from these works, population size itself was **independent** of population fitness and considered to be a parameter. Although the explicit interaction between demography and genetics is not taken into account, population genetics models have undertaken introducing demographical constraints. For example, following the evolution of populations with different population sizes (Bataillon and Kirkpatrick, 2000), after a bottleneck (severe reduction of population size Kirkpatrick and Jarne, 2000), population structure (Roze and Rousset, 2004) and changes in the environment that lead to changes in selective pressures (Glémin and Ronfort, 2013), among others. Whether such approaches are indeed sufficient remains to be seen. So far the only models to have allowed the interplay of population size and deleterious mutations are models following the evolution of the mutational meltdown (Lande, 1994; Lynch et al, 1995; Coron et al, 2013). Although models studying the mutational meltdown allow for population size to be a consequence of population fitness (population size is not a parameter), their aim is not to

understand the mutational load in a general context, but the fixation of deleterious mutations in small populations. The recurrent fixations eventually lead to lower and lower fitnesses, eventually leading to their extinction. It seems surprising therefore that in a single scientific community the validity of the mutational meltdown is widely accepted, but that the mutation load may be sufficiently large so as to decrease population size is not (Agrawal and Whitlock, 2012) .

1.2.1 First there was one (locus)

Wallace's work (1970) was but one of several contemporary theoretical works with an interest in the effects of density dependence (the regulation of population growth by population density) on population evolution (for example (Charlesworth, 1971; Clarke, 1972, 1973)). Clarke's (1973) work focused more specifically on the potential effects of a mutation load at a single locus influencing given ecological traits on population size (his work is discussed in more detail in Chapter 2.1). He concluded that whether density-independent selection took place before or after the density regulating factors, there would be a reduction in population size. There are however several shortcomings to Clarke's (1973) work (discussed in further detail in Chapter 2.1), the most important being the introduction of a pre-calculated genetic load, it is not an emergent property of his model but a parameter. Clarke (1973) therefore made the hypothesis that the mutation load itself evolves independently of population demography, a hypothesis we tested in Chapter 2.1. Inspired by Clarke's work, we propose a model in which population size is a consequence of the mutation load and where both the frequencies of a deleterious mutation and population size co-evolve. We have chosen to explore the potential genetic and demographic consequences of the different definitions of fitness used in population genetics models. Depending on the author, when the mutational load is calculated fitness can be defined differently for a same given model. For example, it can either be zygote (or juvenile) survival (Gillespie, 1998) or reproductive capacity (Haldane, 1937). Though both kinds of selection do reduce the genotype's contribution to the next generation, are they equivalent? By clearly defining when selection and census take place, we find that the mutational load can vary. Selection occurring at the adult stage leads to the same mutational load as expected from models that do not explicitly consider the life cycle (Crow and Kimura, 1970, Chapter 6), whereas selection at the zygote stage does not. In the latter case there is a non-monotonic relationship between the mutation load and the deleterious effect (the coefficient of selection and dominance) of the deleterious alleles, a previously unobserved result.

As ways of taking the interaction between selection and demography further, we took on the task of formulating a simple model in which mutations are considered to take place at different times during the life cycle (Chapter 2.2). It is worth mentioning that as population geneticists focus on following the evolution of genetic frequencies over time (generations), the mutations of interest are only those that can be inherited. With this reasoning, such models rely on new mutations being introduced during gamete formation. First of all, the assumption that somatic mutations are not transmitted does not hold for all taxa, as in plants, for example, there is no separation between the germ-line and somatic cells. Any parent acquiring a mutation can transmit it to its offspring. Second, even if they are not hereditary, somatic mutations occur throughout the life cycle (Lynch, 2010) and can decrease fitness (Ally et al, 2010), hence potentially influencing population size. We find that there can be both a mutational load and a decrease in population size from such mutations, be they hereditary or not. We also find that when mutations are introduced at the adult stage, increasing the introduction of new mutation free individuals (or increasing the birth rate) will decrease the observed mutational load as well as its effect on population size.

1.2.2 And then there were many (loci)

In the conclusion of his paper Clarke (1973) verbally suggested that although the mutation load had very little effect on population size, as his model was based on a single locus (as is the case in our model in Section 2.1), the load on several loci should have a non-negligible effect. In (Gervais et al, 2014, included in Supplementary Information) I participated in the development of a model exploring the transition from an outcrossing to a self-fertilising reproductive regime. This work was my first experience in multi-locus populations genetics models, where population size is kept fixed. Fixed population size implies that a population may remain very large in spite of a high mutation load or very small fitness. Results of the simulations run led me to be sceptic on the validity of such an approach as it implied that a population's relative fitness could sink from 1 to 10^{-3} and still have the same size. It seemed necessary to test whether by effectively introducing an interaction between the genetic load and the actual number of offspring this hypothesis still held.

In Chapter 3 we modified existing multi-locus population genetics models in order to allow for a population size that fluctuates with population fitness. Population size is kept in check using a logistic form of density dependence. As the reproductive rate depends on the mutation load, the higher the load the lower the reproductive rate. When the population's mean reproduc-

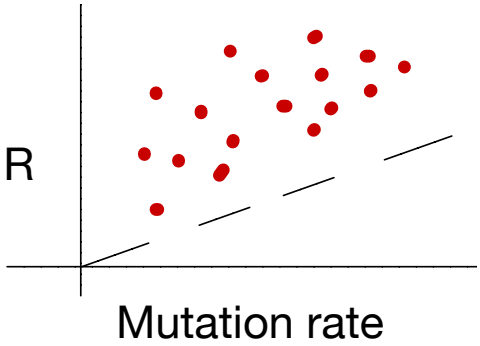


Figure 1.1: The predicted relationship between the genomic mutation rate and the reproductive rate R. The red dots indicate potential qualitative relationships between the mutation rate and the reproductive rate. If our prediction is correct, there should be no points beneath the dashed line.

tive rate drops below 1, then the population goes extinct deterministically. Although the mutation loads of the simulated populations are not greatly influenced in this model compared to models with fixed population sizes, their population size and demographic viability are. Though population sizes in this model are always below the carrying capacity they reach a stable size, which itself depends on the genetic load as verbally predicted by Clarke (1973). As in some cases the mutation load was too high to be demographically viable, and the mutation load is highly dependent on the mutation rate (see Box 1.4), there seems to exist a link between population viability and the mutation rate. Taking this reasoning one step further, if populations with higher intrinsic reproductive rates can be viable with higher mutation loads (as is the case in this model, see Chapter 3), then there may be a relationship between the mutation rate and the reproductive rate in natural populations (see Figure 1.1).

1.3 Reproduction: A team effort or going solo?

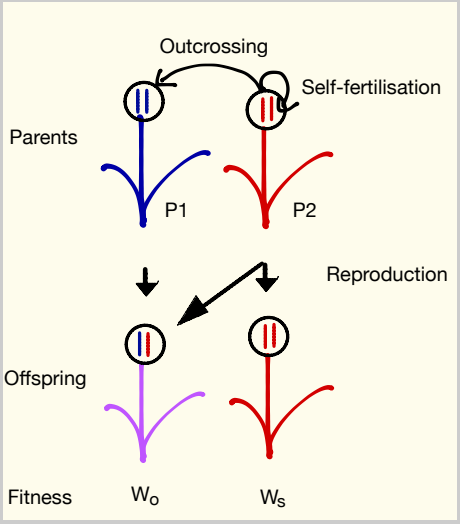
> *"Because this is also a story about sex, although probably not in the athletic, tumbling, count-the-legs-and-divide-by-two sense unless the characters get totally beyond the author's control. They might."*
>
> - Terry Pratchett (Equal rites, 1987)

Literature is littered with theoretical and empirical approaches whose

aims are to understand the role of deleterious mutations on the evolution of reproductive systems. The evolutionary transitions between reproductive systems are frequent and not constrained to specific species or even taxa. In plant families the most common transition in hermaphroditic plants is of the loss of Self-Incompatibility (the ability to avoid self-fertilisation and reproduction with related individuals due), hence leading to the evolution of self-fertilisation from initially outcrossing populations. From a genetic point of view, the evolution of selfing is driven by what is known as Fisher's automatic transmission advantage Fisher 1941, see Box 1.6). The advantage of an outcrossing reproductive system is the avoidance of the cost of inbreeding (Charlesworth and Charlesworth, 1987; Porcher and Lande, 2005), giving fitter offspring (Brennan et al 2005; Randle et al 2009, see Box). However, despite these numerous empirical and theoretical evolutionary studies, we are far from understanding the mechanisms behind the maintenance of outcrossing in a great number of species as in natural populations the levels of inbreeding depression are more often than not below the theoretical limit of the level of inbreeding depression necessary to maintain outcrossing in the presence of a self-fertilising mutant (Winn et al, 2011). The efficiency of the purge (or elimination) of deleterious alleles is also greater in the presence of self-fertilisation (Glémin, 2003), making the necessary levels of inbreeding depression to deter an invading self-fertilising mutant even higher (Porcher and Lande, 2005).

Recent work on the phylogeny of the *Solanaceae* has brought to light the differences in extinction rates between self-fertilisers and outcrossers (Goldberg et al, 2010). The latter seem to be more long-lived species, supporting the hypothesis that self-fertilisation is potentially an evolutionary dead-end (Takebayashi and Morrell, 2001). There are two main arguments to support this hypothesis. The first was presented by Stebbins (1957) wherein the lack of genetic variation found in self-fertilising populations should be to their disadvantage as their adaptive potential would be reduced (see Box 1.7). However, though standing variation does play a significant role in adaptation, a population's adaptive potential also depends on the rapidity with which a beneficial allele is fixed, something that occurs faster in self-fertilising populations (Glémin and Ronfort, 2013). The second argument relies on the negative effects of deleterious mutations, either due to their fixation or the inbreeding depression they engender. As mentioned above, increased self-fertilisation is expected to decrease genetic variation, hence decreasing the effective population size (Wright et al, 2008). Smaller effective size implies greater genetic drift (the change in genetic frequencies due to random events) leading to higher probabilities of the fixation of deleterious alleles and eventually extinction by mutational meltdown (Lande, 1994; Lynch et al, 1995;

Box 1.6 - Inbreeding depression and Fisher's automatic advantage



Inbreeding depression δ is the reduced fitness of individuals due to non-random mating (either with closely related individuals or through selfing). The fitnesses of offspring produced through outcrossing and selfing are noted W_o and W_s respectively. The level of inbreeding depression is measured using the following equation

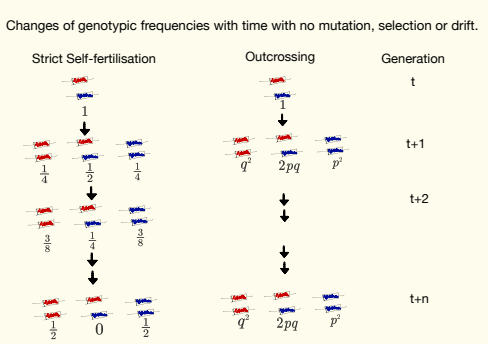
$$\delta = 1 - \frac{W_s}{W_o}$$

Fisher's automatic advantage the transmission of two copies of an individual's genome through selfing compared to a single copy through outcrossing. An outcrossing individual (blue) will transmit only one copy of its genome to each of its descendants. By self-fertilising, as an individual's male gametes fertilise its own ovules, offspring will carry two copies of the parent's genome. Self-fertilisation therefore increases the rate of transmission of genetic material, giving it an advantage over outcrossing.

Coron et al, 2013). Inbreeding depression is considered one of the major genetic factors influencing population extinction (Frankham, 2005). However, both the fixation of deleterious mutations and inbreeding depression are countered by the potential purge of deleterious mutations (see Box 1.7). With the evolution of self-fertilisation, the number of deleterious mutations and inbreeding depression are both expected to decrease (Charlesworth et al, 1990), hence weakening the hypothesis on the importance of deleterious mutations in the probability of population extinction. At this point, as suggested by Glémin and Ronfort (2013), it is adaptive potential (and not deleterious mutations) that accounts for the differences in extinction rates between selfers and outcrossers.

Models studying the evolution and the consequences of self-fertilisation have done so in infinitely large or fixed population sizes. In the model presented in Chapter 3 we also examined how self-fertilisation influences population viability if population size is a consequence of selection. We find a general positive increase in population size and decrease in the probability of extinction when self-fertilisation increases. However, we find that when mutations are completely recessive, intermediate rates of self-fertilisation fare badly compared to strictly outcrossing and high rates of self-fertilisation. We also find that when mutations are of very small effect, they induce a mutational meltdown in strictly self-fertilising populations.

Box 1.7 -Effective population size and purging deleterious mutations



Effective population size Ne depends on the genetic variation present in a population as it represents the expected variation in an "ideal" population (a panmictic population in which all individuals reproduce at the same rate). When there is self-fertilisation, there is loss of genetic variability as homozygous genotypes are fixed and others can be lost through drift.

Purge of deleterious alleles: If we consider the red allele to be deleterious, as it is most often found at the homozygous state in selfing populations, it will be eliminated by selection at a faster rate that in an outcrossing population where it is not as often homozygous.

The results of this first approach confirm that deleterious mutations on their own do not suffice as an the explanation as to why self-fertilising populations do not persist compared to outcrossing ones, as our results generally confirm the higher fitness and better purge of self-fertilising populations. This therefore points to another possibility, what if it is not self-fertilisation in itself but the genetic and demographic consequences of the transition from outcrossing to self-fertilisation that lead to extinction? There are several differences in the genetic backgrounds of selfers and outcrossers (Charlesworth and Wright, 2001), with expected consequences on the genetic structure of a population during a change in the reproductive system. For instance, as outcrossing populations maintain a considerable amount of genetic variability, the evolution of self-fertilisation, if it does not efficiently purge (or eliminate) the deleterious mutations, it could fix them leading to population extinction. Though this has not been observed in previous models, it could be a consequence of the interaction between selection and demography. In Chapter 4, we propose a modified version of the model presented in Chapter 3. This new version allows for the free evolution of self-fertilisation in an initially outcrossing population. From this model there is one main conclusion. As mentioned above, populations that evolve to strict self-fertilisation in the presence of deleterious mutations of very small effect go to extinction. The results presented in Chapter 4 show that most populations evolved to have high self-fertilising rates, but none of the viable populations were strictly self-fertilising. All populations that had evolved strict self-fertilisation went to extinction. Depending on the reproductive rate, the ensuing population sizes rendered the fixation of an allele leading to strict self-fertilisation more

or less probable (the higher the reproductive rate the lower the probability of fixation). None of the populations with mutations of stronger effect had a self-fertilisation rate equal to 1, going against the expectations of the evolution of strict self-fertilisation if there are no ecological constraints (e.g. pollen discounting, Porcher and Lande 2013). There are therefore populations with a genetic structure allowing for the fixation of a strictly self-fertilising genotype but the ensuing consequences do not allow for such population to be viable, making the evolution to strict self-fertilisation an evolutionary suicide.

Lynch et al's (1995) work on the mutational meltdown of initially small populations with strict self-fertilisation had already pointed to the accelerated rate at which such populations went extinct compared to simulated outcrossing populations. In both of the works presented here (with and without the evolution of the self-fertilisation rate), population sizes were not small and have still led to the same conclusions. As mildly deleterious mutations are extremely common (Agrawal and Whitlock, 2011) it is a potential explanation as to why no plants existing in natural populations reproduce solely by self-fertilisation. Our results further support the hypothesis that self-fertilisation is a dead-end.

Act 2

The interaction between demography and selection: A one-locus model

Within this chapter the evolution of the mutation load and its consequence on population size is studied deterministically at a single bi-allelic locus. The first section of this chapter focuses on the importance of timing of selection during the life cycle in determining the potential demographic and genetic consequences of selection. In the second section we present preliminary work on the influence of somatic mutations on population size and the mutation load.

2.1 The timing of selection

The effect of the timing of selection on the mutation load, inbreeding depression and population size

Diala Abu Awad[1,*], Sylvain Billiard[1] Viet Chi Tran[2],
1 UMR-CNRS 8198, Laboratoire Génétique et Évolution des Populations Végétales, Université Lille 1, Villeneuve d'Ascq, France
2 Laboratoire Paul Painlevé, UFR de Mathématiques, UMR CNRS 8524, Université Lille 1, 59655 Villeneuve d'Ascq Cedex, France
∗ **E-mail: Corresponding diala.abu-awad@univ-lille1.fr**

2.1.1 Abstract

If and how the genetic load affects population size has been long debated as on one hand it has been suggested that a high genetic load leads to extinction, and on the other hand, if selection during the life-cycle takes place before density dependence has acted, then population size should not be affected by its genetic load. Explicitly considering the life cycle and the timing of selection is therefore a factor that cannot be ignored when quantifying the effect of deleterious mutations on population size. In addition, population genetics models calculating the expected genetic loads and levels of inbreeding depression ignore the potential effects of demography on these variables. Here we propose a deterministic model in continuous time where deleterious mutations affect individual fitness in one of four ways: by decreasing mating success, fecundity and adult or zygote survival. The genetic load, inbreeding depression and population size are emergent properties. Our results are compared to the predictions from two population genetics models. We find that changing the timing of selection mostly affects population size, but also leads to genetic loads and inbreeding depression that can diverge from the predictions of consensual population genetics models. Our results emphasise the importance of integrating both population demography and genetics in order to study the demographic impact and, more generally, the fate of deleterious mutations.

2.1.2 Introduction

Of the new mutations introduced into populations, ranging from 0.1 to 2.2 new mutations per generation per genome (Sniegowski et al, 2000; Keightley, 2012), most of them are deleterious (Keightley and Lynch, 2003). In spite of their deleterious effects, these alleles are not always immediately eliminated by selection from the genetic pool but can persist for several generations and, in the case of stochasticity, can even go to fixation. These mutations decrease mean population fitness by engendering a mutational load, and are in part responsible for inbreeding depression (Charlesworth and Charlesworth, 1987). In the field of population genetics, the evolution of load and inbreeding depression as a function of population size and structure has been greatly explored (Bataillon and Kirkpatrick, 2000; Roze and Rousset, 2004). However, these models do not consider an explicit interaction between the two, as population size is considered to be a parameter. If and how these mutations influence population size remains unclear; while inbreeding depression is a major concern in conservation biology, whether the mutational load of populations affects population size and viability is debatable. In the case of very small populations and deleterious mutations of very small effects it is widely accepted that there is a risk of mutational meltdown (extinction due to the fixation of deleterious alleles at an accelerating rate; Lande, 1994; Lynch et al, 1995; Coron et al, 2013). Nevertheless, how the mutational load affects the demography of populations that are not at risk of extinction remains widely debated. In the discussion of a recent review (Agrawal and Whitlock, 2012) we find two conflicting statements: "load has no direct relationship to population abundance or persistence" followed by "however, load can reduce population sizes (even with density-dependent regulation) and possibly cause extinction". These contradictory statements just a few paragraphs apart, sum up the complexity of the relationship between the mutational load and population size.

There are two main views in literature (reviewed in Agrawal and Whitlock, 2012); some authors argue that populations cannot persist with high mutational or genetic loads (Kondrashov, 1995), whereas others have insisted that load has little or no ecological consequences (Turner and Williams, 1968; Wallace, 1970). In the latter case, authors have argued that due to density dependence, deaths of individuals due to selection simply replace the unavoidable deaths due to a lack of resources (soft selection), whereas in the former case, the genetic load is expected to directly decrease population size, independently of density dependent factors (hard selection). The main difference between these two types of selection is the timing of the elimination of individuals via selection, either before resource consumption (soft selec-

tion) or after these resources have been used and rendered inaccessible (hard selection). In a simple demographic model proposed by Agrawal and Whitlock (2012), where load was considered a parameter, the authors came to the conclusion that if individuals are eliminated by selective death before having consumed any resources (i.e. at the zygote stage), then the genetic load would not affect population size, as the loss of juveniles would be "masked by ecological compensation". Even though their results are intuitive, this model is limited in that the genetic mechanisms behind the load are not taken into account, therefore ignoring any potential influence that the interaction between the genetic and demographic factors could have on population size.

Clarke (1973) dubbed the decrease of population size due to the presence of deleterious mutations the "numerical load", a term that we will be using throughout this paper. In his model, the genetic frequencies are explicitly taken into account and individuals are subjected to both density-independent and density-dependent selection. His results agree with those of Agrawal and Whitlock (2012), not only can the mutational load lead to a numerical load, but the amplitude of the decrease in population size also depends on when selection is considered to take place, or the timing of selection. However, several points need to be clarified. In this model the mutational load is calculated independently of the demographic context, with selective values of the two phenotypes (homo-zygote for the deleterious allele or carrying at least one wild-type allele) defined in such a way as to keep the mutational load constant. That the genetic load is calculated assuming the deleterious allele reduces zygote survival then extrapolating this as the mutational load observed for selection occurring at a different demographic trait is questionable. Furthermore, the selective values with expressions that change depending on the trait can lead to confusion, as not only do the expressions for the selective values change, but the relation between them changes (i.e. the fitness of the deleterious phenotype is not always $(1 - s)$ the fitness of the wild-type phenotype). This makes it difficult to quantify how a mutation with a given deleterious effect will influence population size depending on the trait it modifies. It is therefore necessary to provide a clearer framework in which the change in the number of deleterious alleles in the population is followed by modelling the selection process and its demographic consequences explicitly. This will also allow the detection of a potential deviation of genetic frequencies from those expected from population genetics models that ignore population demography.

It is generally accepted that when mutations are very rare the mutational load is approximately equal to 2μ (where μ is the mutation rate from the wild-type to the deleterious allele) based on Haldane's (1937) classical theoretical paper. In this work, the author defines fitness as the reproductive

capacity, whereas, while using the same model, other authors (e.g. Gillespie, 1998) define fitness as zygote survival. This possible double-interpretation comes from the method used to calculate the frequency q of the deleterious allele at mutation-selection balance, noted $\widehat{q}$. At mutation-selection balance the change in q due to mutation, Δq_μ, is considered to be exactly equal to the change in q due to the elimination of the deleterious allele by selection Δq_s. This method therefore implies that selection and mutation occur at the same time. However, as shown by (Crow and Kimura, 1970, Chapter 6), for selection that reduces the reproductive capacity, $\widehat{q}$ should be obtained first by reducing the frequency q by Δq_s, then increasing it by Δq_μ. The difference between Crow and Kimura's (1970) model and Haldane's (1937) model is numerically negligible when the mutation rate is small, but this nonetheless highlights the importance of properly defining when selection takes place, as depending on the hypotheses of what fitness is and when it is measured, the mutational load can vary. Explicitly considering the life cycle could therefore modify how selection acts, leading to a genetic load and an amplitude of inbreeding depression that depend on the timing of selection. That the amplitude of inbreeding depression varies at a given trait has been observed empirically (Frankham et al 2010, Chapter 13; Angeloni et al 2014), however, whether this could be due to differences in how selection affects the frequencies of deleterious mutations at given traits has not been explored theoretically.

In this present work we address the validity of the hypothesis made in previous models that the mutational load at a given fitness trait is independent of this trait's demographic consequences. We also provide expressions that allow for predictions of the mutational load, the numerical load and inbreeding depression for mutations of any effect and mutation rate based on a deterministic model where the number of individuals carrying each genotype is considered explicitly. In the presence of selection, genotypes differ in their selective values at a single trait, and there is no density or frequency-dependent effect on fitness. We consider selection at four traits considered to represent fitness: mating success, fecundity, zygote survival and adult survival (Agrawal and Whitlock, 2012). Population size, the mutational load and inbreeding depression are all emerging properties of the model. The genetic properties of our models will be compared to expectations from Haldane's (1937) and Crow and Kimura's (1970) models.

2.1.3 Analytical Model

We consider the evolution of a population with a varying population size and a single bi-allelic locus, where A is the wild type and a the mutant allele. The

population is panmictic and made up of sexually reproducing hermaphrodite individuals. The environment is stable, and the population is isolated and spatially unstructured. Three genotypes can be found in the population, aa, Aa and AA, which, from here onwards, are denoted X, Y, and Z respectively. At a given time t, the population is made up of three kinds of individuals, X_t, Y_t and Z_t representing the number of individuals carrying the respective genotype. We denote the population size $N_t = X_t + Y_t + Z_t$. In a large population setting, these quantities can be considered as continuous, and the evolution of the number of individuals of each genotype is described in continuous time using ordinary differential equations. Three processes affect the change in the number of individuals of each genotype, births (occurring with rate R_t^V, where V can be either X, Y or Z), deaths (at a rate M_t^V) and mutation. Selection and density dependence are introduced in these processes. We consider that the mutation from A to a is unidirectional and occurs with a probability μ at the gamete stage.

We first introduce the demographic and mutational properties of the model without considering selection and show that this model respects the genotypic frequencies predicted by the Hardy-Weinberg model for neutral alleles (and no mutation). Selection is then introduced during different moments of the life cycle and we define the variables measured in order to estimate the effect of the recurrent introduction of deleterious mutations on the numerical load, the mutational load and inbreeding depression. In order to facilitate the reading of the following sections, the notation used throughout the text has been summarised in Table 2.1.

Model without selection

As we consider mutations occurring during gamete formation, the proportions of a gametes produced per genotype are 1, $\frac{1+\mu}{2}$ and μ for X, Y and Z individuals respectively. Mutational events are therefore integrated into the birth rate R_t^V. For example, as Z individuals produce a proportion μ of a gametes and $(1-\mu)$ of A gametes. When two Z individuals are crossed, they produce X, Y and Z offspring with proportions μ^2, $2\mu(1-\mu)$ and $(1-\mu)^2$ respectively. For each reproductive pair, the parents contribute both via the male and the female functions. We consider that the total number of female gametes produced by all individuals in the population is limited and, when there is no selection, depends only on the number of individuals, whereas male gametes are produced in very large quantities and are subject to competition. The probability that an individual reproduces via the male function depends on the proportion of male gametes contributed compared to the total amount of male gametes available. For example, when the X and Y indi-

Table 2.1: **Notation.**

Λ, L, δ	The numerical load, the mutational load and inbreeding depression.
N_t, N_{eq}	Population size at time t with or with selection, and population size at equilibrium when there is no selection.
$V_t, \widetilde{V}_t$	The total number of either X, Y or Z individuals (with genotypes aa, Aa and AA respectively) at time t, and the number of individuals at time t that contribute to the genetic pool.
N_{mut}, V_{mut}	Population size and number of V individuals at population equilibrium.
R_t, M_t	The total birth and death rates of the population at time t.
R_t^V, M_t^V	The birth and death rates of individuals of genotype V at time t.
b, d	The intrinsic birth and death rates of individuals.
s, h	The coefficient of selection and dominance of allele a. The relative fitnesses of X, Y and Z individuals at a given trait are $(1-s)$, $(1-hs)$ and 1 respectively.
μ, μ_{fix}	The mutation rate from A to a and the threshold value of μ for which there is deterministic fixation of a.

viduals cross to give X individuals, X individuals contribute X_t ovules and a proportion of $\frac{X_t}{N_t}$ male gametes, while Y individuals contribute $(1+\mu)\frac{Y_t}{2}$ ovules and a proportion of $(1+\mu)\frac{Y_t}{2N_t}$ male gametes (as only $\frac{(1+\mu)}{2}$ of the gametes produced carry an a allele).

Generally, the equation describing the change in the number of individuals for each genotype is given by

$$\frac{dV_t}{dt} = R_t^V - M_t^V. \tag{2.1}$$

For each of the genotypes, when there is mutation and no selection, the

birth rates R_t^V are given by

$$R_t^X = \frac{b}{N_t}\Bigg(X_t^2 + 2X_tZ_t\mu + Z_t^2\mu^2 + X_tY_t(1+\mu) + Y_tZ_t\mu(1+\mu) + \frac{1}{4}Y_t^2(1+\mu)^2\Bigg)$$

$$R_t^Y = \frac{b}{N_t}\Bigg(X_tY_t(1-\mu) + 2X_tZ_t(1-\mu) + 2Z_t^2(1-\mu)\mu + \frac{1}{2}Y_t^2(1-\mu^2) + Y_tZ_t\left(1+\mu-2\mu^2\right)\Bigg)$$

$$R_t^Z = \frac{b}{N_t}\left(\frac{1}{4}Y_t^2(1-\mu)^2 + Y_tZ_t(1-\mu)^2 + Z_t^2(1-\mu)^2\right).$$

The birth rate depends on an intrinsic birth rate b, which, by default, holds the same value for all genotypes, on the reproductive events that lead to the production of new individuals with genotype V and on the mutation rate μ. The death rate M_t^V depends on an intrinsic death rate d and is density dependent (we consider a carrying capacity K). The equation for M_t^V is given by

$$M_t^V = d\frac{N_t}{K}V_t. \tag{2.2}$$

When solving $\frac{dN}{dt} = \frac{dX}{dt} = \frac{dY}{dt} = \frac{dZ}{dt} = 0$ we find the optimal population size is given by (see Supplementary Material S4 for the proof)

$$N_{eq} = \frac{bK}{d}. \tag{2.3}$$

If we consider that there is neither selection nor mutation ($\mu = 0$), then we find that the frequencies of X, Y and Z are at Hardy-Weinberg equilibrium (see Supplementary Material S3). Explicitly considering the demography of a population leads to the same genotypic frequencies at a neutral locus as those predicted by deterministic population genetics models. This implies that, once selection is introduced, any differences observed between our model and the two population genetics models (Haldane, 1937; Crow and Kimura, 1970, Chapter 6) are due to the interaction between the timing of selection and demography.

Timing of selection

Fitness can be defined as being an individual's relative mating success, fecundity or survival either at the zygote or adult stages. We consider all

four definitions of fitness. Selection can occur at different times during the life cycle, affecting either reproduction or survival. As a is deleterious, Z individuals always have the maximal fitness. The relative fitness of each genotype at a given trait (i.e. its reproductive rate or survival) is equal to $(1-s)$, $(1-hs)$ and 1 for genotypes X, Y and Z respectively, where s is the selection coefficient and h the dominance of the mutant allele a. If we consider that a affects the inherent birth rate and if the inherent birth rate of Z individuals is b^Z and the inherent birth rate of X individuals is b^X, then the relative fitness of X individuals is $\frac{b^X}{b^Z} = (1-s)$. The full equations for the change in the number of individuals of each genotype for these models can be found in Supplementary Material S2.

Selection on reproduction: In order to model the effect of the deleterious allele a on the reproductive success of individuals, we introduce a new term $\widetilde{V}_t$ instead of V_t in the R_t^V function. This term represents the contribution of V individuals to the genetic pool, which is proportional to their fitness and can reduce their reproductive success (i.e. $\widetilde{X}_t = (1-s)X_t$). There are two ways in which carrying a can affect reproductive success; it can reduce the mating success of individuals (i.e. for X individuals only a proportion of $(1-s)$ matings are successful or lead to fertilisation) or by reducing the fecundity of individuals (i.e. the proportion of gametes produced by X individuals is $(1-s)$ that produced by Z individuals).

Mating success: When mating success is reduced, all individuals produce the same quantity of gametes and the proportion of male gametes an individual V contributes to the next generation is $\frac{\widetilde{V}_t}{N_t}$. The probability of a successful reproductive event is proportional to the parental fitnesses. For example, R_t^X for this model of selection is given by

$$R_t^X = \frac{b}{N_t}\Bigg(\widetilde{X}_t^2 + 2\widetilde{X}_t\widetilde{Z}_t\mu + \widetilde{Z}_t^2\mu^2 + \widetilde{X}_t\widetilde{Y}_t(1+\mu) + \widetilde{Y}_t\widetilde{Z}_t\mu(1+\mu) + \frac{1}{4}\widetilde{Y}_t^2(1+\mu)^2\Bigg). \tag{2.4}$$

Fecundity: When fecundity is affected by selection, an individual V contributes $\widetilde{V}_t$ female gametes and a proportion of $\frac{\widetilde{V}_t}{\widetilde{X}_t+\widetilde{Y}_t+\widetilde{Z}_t}$ male gametes to the next generation (the proportion of male gametes produced by V depends on the total amount of male gametes produced and not on the number of

individuals in the population). For example

$$R_t^X = \frac{b}{\widetilde{X}_t + \widetilde{Y}_t + \widetilde{Z}_t}\Big(\widetilde{X}_t^2 + 2\widetilde{X}_t\widetilde{Z}_t\mu + \widetilde{Z}_t^2\mu^2 + \widetilde{X}_t\widetilde{Y}_t(1+\mu) + \widetilde{Y}_t\widetilde{Z}_t\mu(1+\mu) + \frac{1}{4}\widetilde{Y}_t^2(1+\mu)^2\Big). \tag{2.5}$$

Note that in both models with selection on reproduction the probability of reproduction via the female function remains unaffected as we consider that there is no competition between the female gametes.

Selection on survival: Selection can also occur during the life cycle, independently of reproductive success, affecting either zygote or adult survival. Zygote survival can be translated as the proportion of germinating seeds, or, more generally, viable offspring. Selection on adult survival is considered to occur before reproduction.

Zygote survival: The probability of zygote survival is decreased by considering a birth rate R_t^V that is genotype dependent. This can be done by introducing a term b^V, an intrinsic birth rate that is proportional to the genotype's fitness. For example, $b^X = (1-s)b$ and

$$R_t^X = \frac{b(1-s)}{N_t}\left(X_t^2 + 2X_tZ_t\mu + Z_t^2\mu^2 + X_tY_t(1+\mu) + Y_tZ_t\mu(1+\mu) + \frac{1}{4}Y_t^2(1+\mu)^2\right). \tag{2.6}$$

Adult survival: We consider that the number of adults that survive selection before reproduction of genotype type V is $\widetilde{V}$, hence proportional to their fitness. As only surviving individuals reproduce and compete for resources, V is replaced by $\widetilde{V}$ in the birth rate R_t^V and in the death rate M_t^V. Therefore we obtain the same expression for R_t^V as for selection on fecundity and M_t^V is given by

$$M_t^V = d\frac{\widetilde{X}_t + \widetilde{Y}_t + \widetilde{Z}_t}{K}V_t. \tag{2.7}$$

Population equilibrium

In order to understand how the interaction between selection and population demography impacts population size and the frequency of a recurrent deleterious mutation, we derive the deterministic equilibrium values for each of the

models of selection described above (mating success, fecundity and zygote and adult survival) by solving $\frac{dX_t}{dt} = \frac{dY_t}{dt} = \frac{dZ_t}{dt} = 0$. This allows us to obtain the number of individuals carrying each genotype at equilibrium (X_{mut}, Y_{mut} and Z_{mut}), the sum of which gives us the population size at equilibrium N_{mut}. Using N_{mut} we obtain the expression for the numerical load Λ (the decrease of population size due to the presence of deleterious mutations), a term defined by Clarke (1973) and given by

$$\Lambda = \frac{N_{eq} - N_{mut}}{N_{eq}}, \tag{2.8}$$

where N_{eq} is the population size at equilibrium when there is no selection ($s = 0$, see equation 2.3). We also use the expressions for X_{mut}, Y_{mut} and Z_{mut} to derive the expressions for the mutational load L and inbreeding depression δ. The mutational load L is defined as the decrease in population fitness due to the presence of deleterious mutations and is given by (Gillespie, 1998, p.61):

$$L = 1 - \frac{(1-s)X_{mut} + (1-hs)Y_{mut} + Z_{mut}}{N_{mut}} \tag{2.9}$$

Inbreeding depression δ is defined as the difference in fitness between offspring produced via selfing and via outcrossing. We calculate it using equation 3 in Roze and Rousset (2004):

$$\delta = 1 - \frac{(1-s)X_{mut} + \left(\frac{1}{4} + \frac{1-hs}{2} + \frac{1-s}{4}\right) Y_{mut} + Z_{mut}}{(1-s)X_{mut} + (1-hs)Y_{mut} + Z_{mut}}. \tag{2.10}$$

We then compare L and δ to expectations from population genetics models. In order to compare our results to these models, we replace X_{mut}, Y_{mut} and Z_{mut} with q^2, $2q(1-q)$ and $(1-q)^2$ respectively, where q is the frequency of the deleterious mutant a at mutation-selection balance. We will compare our models to the explicit expressions for q (for any value of μ) from Haldane's model (1937), obtained using the equations given by (Gillespie, 1998, p.71)

$$q^H = \frac{2\mu}{hs(1+2\mu) + \sqrt{s\left(4\mu(1+\mu) - 8h\mu(1+\mu) + hs(1+2\mu)^2\right)}}, \tag{2.11}$$

as well as the expression from the model by (Crow and Kimura, 1970, Chapter 6), where

$$q^{CK} = \frac{2\mu}{hs(1+\mu) + \sqrt{s\left(4\mu - 8h\mu + h^2 s(1+\mu)^2\right)}}. \tag{2.12}$$

For all four models of selection, there exists a solution where the population is made entirely of X individuals. There is therefore a threshold value of the mutation rate μ, as a function of the selection coefficient s and the dominance h, which leads to the deterministic fixation of a. This threshold value is noted μ_{fix} and is calculated by solving the equations for μ when considering that N_{mut} is equal to X_{mut}. In order to calculate the fixation threshold for the population genetics models, we solve for μ when q (equation 2.11 and 2.12) is equal to 1.

2.1.4 Results

Model	h	N_{mut}	Λ	L	δ
Mating success	0	$N_{eq}(1-\mu)^2$	$2\mu-\mu^2$	μ	$\frac{\sqrt{\mu s}-\mu}{2(1-\mu)}$
	0.5	$N_{eq}\frac{(1-\mu)^2}{(1+\mu)^2}$	$\frac{4\mu}{(1+\mu)^2}$	$\frac{2\mu}{1+\mu}$	0
Fecundity	0	$N_{eq}(1-\mu)$	μ	μ	$\frac{\sqrt{\mu s}-\mu}{2(1-\mu)}$
	0.5	$N_{eq}\frac{1-\mu}{1+\mu}$	$\frac{2\mu}{1+\mu}$	$\frac{2\mu}{1+\mu}$	0
Adult survival	0	N_{eq}	0	μ	$\frac{\sqrt{\mu s}-\mu}{2(1-\mu)}$
	0.5	N_{eq}	0	$\frac{2\mu}{1+\mu}$	0
Zygote survival	0	$N_{eq}(1-\mu)$	μ	$\frac{\mu(1-s)}{1-\mu}$	$\frac{\sqrt{\mu s}-\mu}{2-2\mu(2-s)}$
	0.5	$N_{eq}\frac{1-\mu}{1+\mu}$	$\frac{2\mu}{1+\mu}$	$\frac{\mu(2-\mu s-s)}{1+\mu^2}$	0
Haldane's model	0	–	–	$\frac{\mu}{1+\mu}$	$\frac{\sqrt{(1+\mu)\mu s}-\mu}{2}$
	0.5	–	–	$\frac{2\mu}{1+2\mu}$	0

Table 2.2: Expressions for population size N_{mut}, numerical load Λ, mutational load L, inbreeding depression δ and the threshold value of the mutation rate for deterministic fixation μ_{fix} at population equilibrium for $h = 0$ and 0.5 for selection on mating success, fecundity and zygote and adult survival, as well as Haldane's (1937) model. Crow and Kimura's (1970) model gives the same expressions as our models, with the exception of zygote survival. Exact expressions or any $h \neq 0.5$ obtained using Wolfram's Mathematica 9 (Wolfram Research, 2012) are given in Table S1.

By solving the equations given above, we have found explicit solutions for population size N_{mut}, the numerical load Λ, the mutational load L and

inbreeding depression δ at population equilibrium for all four models of selection (mating success, fecundity and zygote and adult survival). The full expressions (valid for all parameter values when $h \neq 0.5$) can be found in Table S1 of the Supporting Information, as for the sake of legibility, we present only the expressions for recessive and co-dominant mutations ($h = 0$ and $h = 0.5$ respectively) in the main text in Table 2.2. For $h \neq 0$, the expressions were found by using Wolfram's Mathematica 9.0 (Wolfram Research, 2012) , whereas the proofs for population size for $h = 0$ can be found in Supplementary Material S4. The expressions for the frequencies of each genotype at population equilibrium are also in the Supplementary Material S3. The equations are valid given that there is no fixation of the deleterious mutation, which occurs at the mutation rate $\mu_{fix} = \frac{(1-h)s}{1-hs}$. If the conditions for fixation are met, there is only one valid solution, where $N_{mut} = X_{mut}$.

General results

We can distinguish two expressions when considering the genetic properties, L and δ, of the populations at equilibrium. Selection on mating success, fecundity and adult survival all lead to the same L and δ, while selection on zygote survival leads to different equilibrium values (see Table 2.2 and Table S1 in Supplementary Material). Concerning the numerical load Λ we find three different equilibria, one for mating success, one for fecundity and zygote survival, and one for adult survival (see Table 2.2 and Table S1). In the latter model, we find that selection has no effect on population size. As adult survival shares the same genetic equilibria as selection on mating success and fecundity and there is no numerical load, we will not graphically represent this model in the results.

Concerning the expectations from the population genetics models, we find that our results for mating success, fecundity and adult survival are the same as Crow and Kimura's (1970) model. The assumption that mutation and selection occur at the same time lead to a slight deviation in the case of Haldane's (1937) model, with a very slight numerical difference for low mutation rates, but a large difference for higher mutation rates. It is important to note that we do not make the simplifying assumption that $\mu << s$, therefore the mutational load L does not take on its simplified form 2μ as given by Haldane (1937). The equations for the predictions of these models are valid for any and all rates of mutation and coefficients of selection, while dominance must be between 0 and 1.

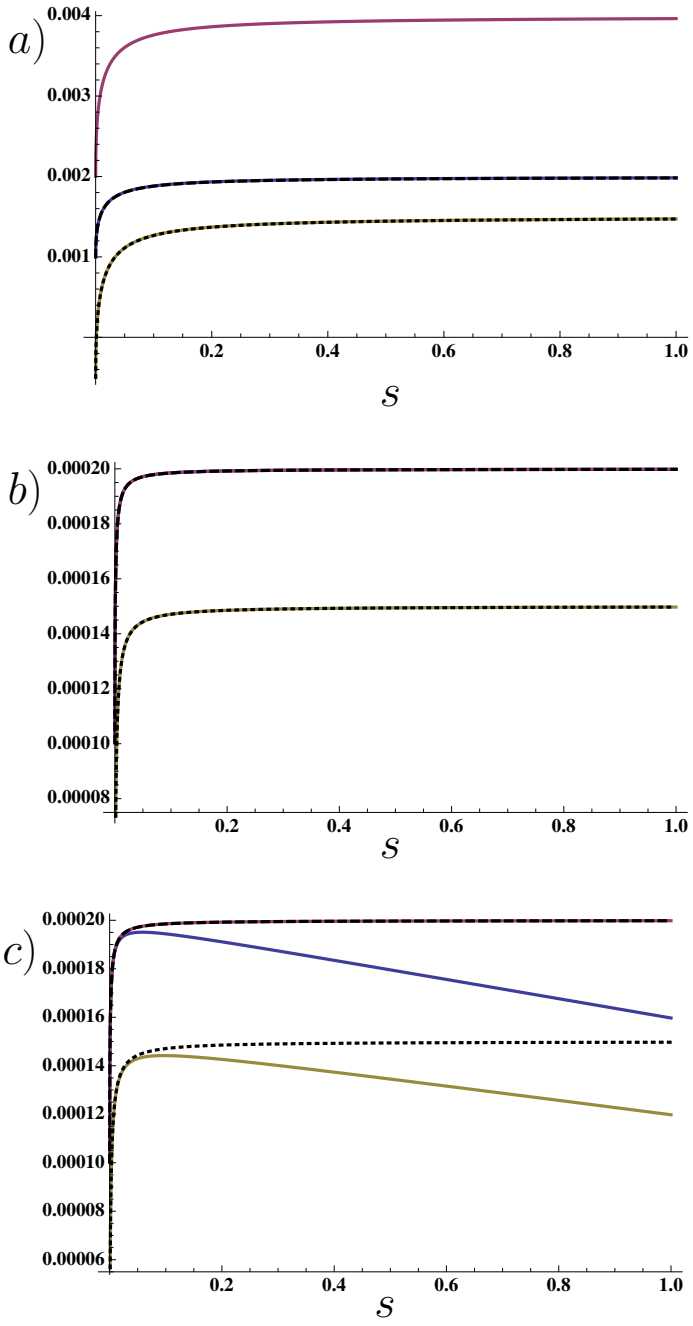


Figure 2.1: The numerical load Λ (red), mutational load L (blue) and inbreeding depression δ (yellow) at population equilibrium as a function of the coefficient of selection s for selection on a) mating success, b) fecundity and c) zygote survival. The black dashed and dotted lines represent L and δ respectively as predicted from Haldane's model. Results from Crow and Kimura's model are not represented as they are equal to those with selection on mating success and fecundity. Lines are plotted using the expressions in Table S1, for $h = 0.2$ and $\mu = 10^{-4}$.

The mutational load and inbreeding depression

Selection on mating success, fecundity and adult survival all lead to the same genetic properties at equilibrium. The equations obtained are exactly the same as those from Crow and Kimura's (1970), but deviate from Haldane's (1937) model (see Table 2.2 and Table S1). When the mutation rate μ is small, these models are numerically close, as can be seen in Figures $1a$ and $1b$, where the red line representing the mutational load L and the yellow line representing inbreeding depression δ from our models are superposed by the black dashed and dotted lines representing L and δ respectively from Haldane's (1937) model. However, this holds only for small values of μ, increasing the mutation rates leads to higher values of L and δ than those predicted by Haldane's (1937) model. For selection on mating success, fecundity and adult survival, as well as both population genetics models, L increases with the coefficient of selection s (Figure 1) and the dominance (h, results not shown). Inbreeding depression also increases with s, but decreases with h.

The genetic properties at equilibrium when selection is on zygote survival differs both quantitatively and qualitatively from the other models. These differences are due to a greater effect of the coefficient of selection s on both L and δ. This can clearly be seen in Table 2.2 when mutations are completely recessive ($h = 0$), where s is present in the equation for L for zygote survival, but completely absent in the other equations for L. L and δ are both lower for zygote survival than for the other models and both have a non-monotonic relationship with s, for low values of s the mutational load and inbreeding depression increase. Once they have reached a maximum, further increasing s leads to a decrease in both variables. This non-monotonic effect is also observed for L when s is constant and h is increased (results not shown).

The numerical load

In Figure 1, the numerical load Λ is represented in blue lines for selection on mating success, fecundity and zygote survival (Figures $1a$, $1b$ and $1c$ respectively). For these three models, increasing the mutation rate μ, coefficient of selection s and the dominance h all lead to an increased Λ. Selection on zygote survival and on fecundity both have the same numerical load Λ, with $\Lambda = \mu$ when mutations are completely recessive (Table 2.2). In the case of selection on fecundity, $\Lambda = L$, whereas the relationship between L and Λ is more complex for the other models. Selection on mating success leads to the highest numerical load, almost double that observed for the other two models.

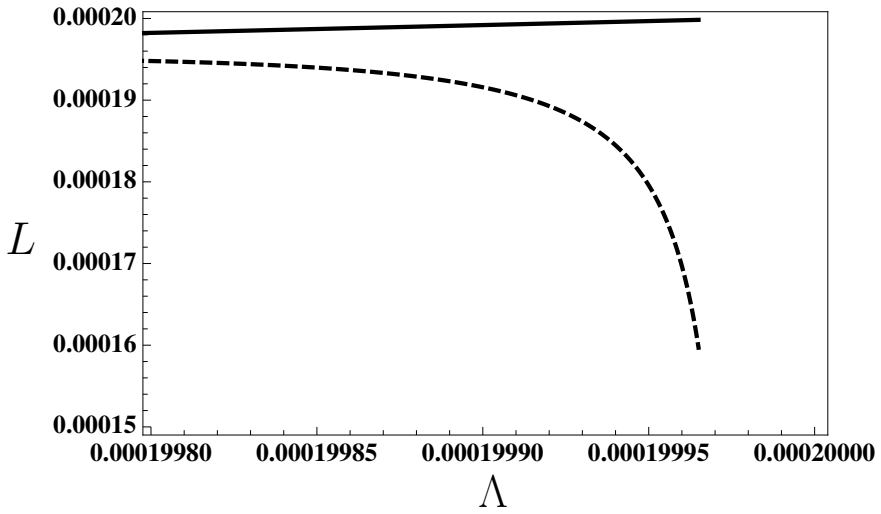


Figure 2.2: The genetic load L as a function of the numerical load Λ for selection on fecundity (full line) and zygote survival (dashed line) at population equilibrium for a coefficient of selection s between 0.01 and 1. Inbreeding depression as a function of the numerical load gives a qualitatively equal result. Lines are plotted using the expressions in Table S1, for $h = 0.2$ and $\mu = 10^{-4}$.

When selection is on fecundity and mating success, a higher numerical load is associated with a higher genetic load (see Figure 2.2). The same holds for the relationship between inbreeding depression and the numerical load. As increasing selection lowers the genetic load when selection is on zygote survival, a higher Λ is associated with a lower L (see Figure 2.2) and δ.

2.1.5 Discussion

In this article we have presented models in which the timing of selection is explicitly taken into account in order to explore the link between the genetic load and population size. We have also tested whether the different definitions of fitness proposed in population genetics models are indeed interchangeable. We find that population size, the mutational load and inbreeding depression depend on the fitness trait affected, with, in the case of selection on zygote survival, a non-monotonic relationship between the mutational load and the strength of selection and dominance of the deleterious allele. Unlike the proposed population genetics models in literature, we have avoided making approximations concerning the values of the mutation rate and the coefficient of selection. This has allowed us to define mathematical

conditions under which there is a deterministic fixation of deleterious mutations, a property that has been overlooked in the past. This implies that the mutational meltdown is potentially possible even in large populations provided that mutations of a small enough deleterious effect are continuously introduced at a large number of loci.

Population size as a consequence of the mutational load

Previous models have found that the genetic load can affect population size and does so differently depending on the trait under selection (Clarke, 1973; Agrawal and Whitlock, 2012). In such models the genetic load is introduced as the population's mean decrease in performance at a given trait and the mutational load is a fixed parameter. In the model proposed by Agrawal and Whitlock (2012) the decrease in demographic performance is considered on the population and not the genotypic level. In this model resources are consumed and regenerated at different rates. The authors allow for selection to occur at different life stages (zygote and adult survival) by considering that individuals at each age consume different amounts of resources. Lower reproductive efficiency and survival engender a numerical load, with zygote survival leading to a smaller numerical load than lowered adult survival. In our model genotypic mutational loads are taken into account and resources (which could be described as available patches) become immediately available upon death. We also observe a numerical load when the efficiency of reproduction is diminished, however, depending on whether the mutational load influences mating success or fecundity (the quality or the quantity of gametes), the numerical load is not the same. Selection on mating success induces a greater numerical load because there is competition between gametes of differing quality, i.e. gametes that do or do not lead to a successful reproductive attempt. Lower mating success implies that "good" gametes are wasted when they meet gametes of lower quality, in which case reproduction is not successful. Concerning selection on adult vs. zygote survival, our results contradict those of Agrawal and Whitlock (2012) as reducing zygote survival generates a numerical load whereas reducing adult survival does not. This is because when zygotes do not survive all of the resources invested in producing them are lost and the low reproducing adults continue to occupy patches, whereas upon the death of an adult's resources are freed and immediately used by new individuals. In our demographic model, reducing zygote survival equates to reducing fecundity, as can be deduced from the equations for the numerical load which are equal for these two models of selection.

The timing of selection and genetic properties at equilibrium

First we would like to point out that though Haldane's (1937) and Crow and Kimura's (1970) models are equivalent when the mutation rate is smaller than the coefficient of selection (with $L \approx 2\mu$), higher mutation rates and multiple loci should amplify the difference between them. Excepting selection on zygote survival, explicitly modelling selection and mutation in our models results in the same genotypic frequencies as those predicted by Crow and Kimura's model. Clarke (1973) assumed that selection occurred at zygote survival and used the same genetic load to determine the numerical load, independently of the demographic trait affected. The potential effect of the interaction between demography and selection on the mutational load at equilibrium is therefore neglected. In our models we find that 1) the timing of selection can influence the genetic load and 2) selection on zygote survival leads to a mutational load different than that calculated by Clarke (1973), which he gave as being equal to μ for recessive mutations.

Selection at the zygote stage engenders a non-monotonic relationship between the coefficients of selection and dominance and the mutational load. The mutational load L is also lower for this model than for the other proposed models of selection. These discrepancies are in part due to having defined the mutational load as being due to the genotypic frequencies among the individuals in the adult population. Deleterious mutations are introduced into the population through the zygotes. If selection immediately eliminates individuals carrying a before they are a part of the adult population, this automatically leads to lower observed frequencies of a. As the deleterious effect of a increases (higher values of s and h), it is more efficiently selected against. This leads to lower frequencies of adult individuals carrying a, decreasing the number of a alleles introduced by reproduction and making mutation the main source of new a alleles. The maximum of the genetic load observed in Figure 1c represents the turning point where mutation becomes a more important source of individuals carrying a than reproduction. As of this point it is the value of s that defines the efficiency of selection. In the other models, the deleterious mutation rate μ continues to play an important role in the evolution of the genetic load as offspring carrying a will arrive to the adult stage. This reasoning is also applicable to explaining the differences in levels of inbreeding depression between selection on zygote survival and the other models. Inbreeding depression related to a trait early in life may therefore be lower than that of a trait observed in later stages. Interestingly, segregating mutations that affect selection on zygote survival that will be observed within the population will not have any coefficient of selection associated with any dominance, as contrary to the other models

of selection and population genetics models, the coefficient of selection and the dominance both play an important role in selection. For a mutation of strong deleterious effect on zygote survival to make it to the adult population it must be recessive or nearly so, whereas mutations of small effect can be associated with any dominance. This association between s and h has been documented empirically (Simmons and Crow, 1977; Steinmetz et al, 2002; Agrawal and Whitlock, 2011). It is not impossible that a fraction of the observed associations between the coefficient of selection and dominance of deleterious mutations may be an artefact due to the timing of selection.

The interaction between demography and genetics

As predicted by other models, a given mutational load can lead to different numerical loads depending on the timing of selection. This model, however, is the first to consider both genetic frequencies and population size as properties and not as parameters of the model. How the numerical load and the genetic properties co-evolve is represented in Figure 2.2 for selection on fecundity and zygote survival. For selection on mating success and fecundity, as the mutational load increases so does the numerical load. This is not the case when selection is on zygote survival, where the numerical load effectively reflects the consequences of selection on demography, but does not follow the non-monotonic nature of the mutational load. Selection on resource acquisition has been shown to lead to lower numerical loads in populations with higher mutational loads (Clarke, 1973), but none have shown that, depending on the properties of the deleterious allele, the numerical load can be higher for a lower mutational load. There can therefore exist a demographic cost to selection that is not necessarily reflected in the population's genetic traits.

Another point we wish to bring forward is the relationship between the numerical load and inbreeding depression, which is qualitatively equal to that between the mutational load and the numerical load. In population genetics models where population size is a fixed parameter, inbreeding depression is expected to decrease with population size as in small populations polymorphism is lost with random drift (Bataillon and Kirkpatrick, 2000). In our model, population size decreases due to selection leading to a higher inbreeding depression associated with smaller population sizes when selection influences mating success and fecundity. This implies that if a population were small because of its numerical load and not because of external constraints, we would expect to observe higher inbreeding depression in traits that affect reproductive success. Due to the non-monotonic relationship between inbreeding depression and selection, traits that influence zygote survival will portray lower levels of inbreeding depression in small populations.

Acknowledgements

We would like to thank Sylvain Glémin, Emmanuelle Porcher, François Massol, Camille Coron and Xavier Vekemans for their input and helpful comments. This work benefited from the support of the Chair "Modélisation Mathémathique et Biodiversité" of Veolia Environnement - École Polytechnique - Muséum National d'Histoire Naturelle - Fondation X, the French Agence Nationale de la Recherche Modéles Aléatoires en Écologie, Génétique et Évolution (MANEGE; ANR-09-BLAN-0215) and the French Agence Nationale de la Recherche ANR-11-BSV7- 013-03. Viet Chi Tran was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01).

2.2 Somatic mutations

2.2.1 Introduction

Somatic mutations are constantly being introduced into populations, and at mutation rates that are usually higher than those observed in germ-line cells (Lynch, 2010). In spite of their frequency, how such mutations can influence the genetic and demographic evolution of populations is not clear. In current literature, existing population genetic models only address mutations that occur in the germ-line as only they are considered to be heritable. This therefore excludes how genetic frequencies evolve in plants, protists and fungi, where the germ-line and the somatic cells are not separated (*i.e.* mutations acquired during an individual's lifetime can be transmitted to their offspring). This aside, a recent work on the frequency of neutral alleles in which both germ-line mutations and somatic (non-heritable) mutations were taken into account, the authors found that there was a potential over-estimation of the heritable genetic variation (Ezawa and Innan, 2013). Though the goal of this work was to examine how using somatic cells for polymorphism data engenders biased results, the authors have pointed out the importance of also accounting for somatic mutations, be they heritable or not. With these results in mind, one can intuitively suggest that if such mutations can have an impact on individual fitness, then they could contribute to the observed genetic load. It is generally accepted that deleterious somatic mutations do occur and are thought to be one of the causes of senescence (the decrease of fitness with age, Ally et al 2010) and to have contributed to the evolution of diploidy (Newman and Pilson, 1997). It has also been suggested that there exists a relationship between somatic mutations and demography, as the number of genetic abnormalities is correlated with population size in voles (Cheprakov and Rakitin, 2012). In light of the available empirical data and lack of theoretical works, there is clearly a need for a theoretical framework exploring the role somatic mutations (hereditary or not) play in decreasing population fitness and potentially population size has yet to be examined. Here we propose a simple model based on the models presented in Chapter 2.1 modified so as to introduce deleterious somatic mutations that are either hereditary or not and that may appear at either the adult or zygote stage. The results from this model are compared to those with mutation occurring during gamete formation (Chapter 2.1).

2.2.2 Model

We consider a single bi-allelic locus affecting fecundity at which mutations occur at a rate μ from the wild-type allele A to the deleterious mutant a (there are no back mutations). The number of individuals carrying each genotype at a given time t are noted X_t, Y_t and Z_t for genotypes aa, Aa and AA respectively. We test two kinds of somatic mutations, those that are hereditary and those that simply decrease adult fitness but are not transmitted. X, Y and Z individuals have fitnesses of $(1-s)$, $(1-hs)$ and 1 respectively, where s is the coefficient of selection and h the dominance.

Hereditary mutations

When mutations are hereditary, the change in the number of individuals carrying each genotype is given by

$$\frac{dV_t}{dt} = R_t^V - M_t^V, \tag{2.13}$$

where V represents X, Y or Z, R_t^V is the birth rate and MV_t the mortality of individual's with genotype V. We consider that reproduction is sexual and all the individuals are hermaphrodites, hence contributing both via the male and the female function. When there is no selection or mutation and a is hereditary, the birth rates are

$$\begin{aligned}
R_t^X &= \frac{b}{N_t}\left(X_t^2 + X_tY_t + \frac{1}{4}Y_t^2\right)\\
R_t^Y &= \frac{b}{N_t}\left(X_tY_t + 2X_tZ_t + \frac{1}{2}Y_t^2 + Y_tZ_t\right)\\
R_t^Z &= \frac{b}{N_t}\left(\frac{1}{4}Y_t^2 + Y_tZ_t + Z_t^2\right).
\end{aligned}$$

The birth rate depends on an intrinsic birth rate b, which, by default, holds the same value for all genotypes, and on the reproductive events that lead to the production of new individuals with each genotype. The death rate M_t^V depends on an intrinsic death rate d and is density dependent (we consider a carrying capacity K). The equation for M_t^V is given by equation 2.2. As shown in Chapter 2.1, when solving $\frac{dN}{dt} = \frac{dX}{dt} = \frac{dY}{dt} = \frac{dZ}{dt} = 0$ we find the optimal population size is $N_{eq} = \frac{bK}{d}$ and the genotypic frequencies are at Hardy-Weinberg equilibrium.

The number of available female gametes is limited and depends on the number of individuals present in the population and their respective fitnesses

whereas male gametes are produced in very large quantities and are subject to competition. The probability that an individual reproduces via the male function depends on the proportion of male gametes contributed compared to the total amount of male gametes available. We consider that selection decreases gamete production, lowering fecundity. Below we use R_t^X to illustrate how the equations are changed when selection is introduced. For example, when a decreases fecundity in general (fewer ovules and sperm) R_t^X becomes

$$R_t^X = \frac{b}{((1-s)X_t + (1-hs)Y_t + Z_t)} \left((1-s)^2 X_t^2 + (1-s)X_t(1-hs)Y_t + (1-hs)^2 \frac{1}{4} Y_t^2\right). \tag{2.14}$$

Mutations occur at a rate μ from A to a at either the adult stage, where the change in the number of individuals is

$$\begin{aligned} \frac{dX_t}{dt} &= R_t^X + \mu Y_t - M_t^X \\ \frac{dY_t}{dt} &= R_t^Y + 2\mu Z_t - \mu Y_t - M_t^Y \\ \frac{dZ_t}{dt} &= R_t^Z - 2\mu Z_t - M_t^Z, \end{aligned} \tag{2.15}$$

or exclusively at the zygote stage, with

$$\begin{aligned} \frac{dX_t}{dt} &= R_t^X + \mu R_t^Y - M_t^X \\ \frac{dY_t}{dt} &= (1-\mu) R_t^Y + 2\mu R_t^Z - M_t^Y \\ \frac{dZ_t}{dt} &= (1-2\mu) R_t^Z - M_t^Z. \end{aligned} \tag{2.16}$$

Note that in the latter case mutation only introduces Y but not X individuals. This is because we consider the zygote stage as being too short-lived for two mutations to occur simultaneously.

2.2.3 Non-hereditary mutations

If we consider that the mutations introduced into the population lower fitness but are not in any way hereditary, then reproduction always leads to the

introduction of new Z individuals. In this case, when there is no selection, the reproductive rate $R_t^U = bN_t$, as it depends only on the intrinsic birth rate and the number of individuals in the population. Once selection is introduced and a decreases fecundity, R_t^U becomes

$$R_t^U = \frac{b}{((1-s)X_t + (1-hs)Y_t + Z_t)} \left((1-s)^2 X_t^2 + 2(1-s)(1-hs)X_t Y_t + 2(1-s)X_t Z_t + (1-hs)^2 Y_t^2 + 2(1-hs)^2 Y_t Z_t + Z_t^2 \right). \tag{2.17}$$

In the case of mutations occurring at the adult stage the change in the number of individuals becomes

$$\begin{aligned} \frac{dX_t}{dt} &= \mu Y_t - M_t^X \\ \frac{dY_t}{dt} &= 2\mu Z_t - \mu Y_t - M_t^Y \\ \frac{dZ_t}{dt} &= R_t^U - 2\mu Z_t - M_t^Z. \end{aligned} \tag{2.18}$$

When mutations are considered to occur exclusively at the zygote stage, only Y individuals can be introduced via mutation, simplifying the change in the number of individuals to

$$\begin{aligned} \frac{dY_t}{dt} &= 2\mu R_t^Z - M_t^Y \\ \frac{dZ_t}{dt} &= (1-2\mu) R_t^U - M_t^Z. \end{aligned} \tag{2.19}$$

Population equilibrium

We derive the deterministic expectations at population equilibrium for each of the models of mutation (adult or zygote), selection (general fecundity, male fecundity or female fecundity) and heritability described above by solving $\frac{dX_t}{dt} = \frac{dY_t}{dt} = \frac{dZ_t}{dt} = 0$. This allows us to obtain the number of individuals carrying each genotype at equilibrium (X_{mut}, Y_{mut} and Z_{mut}), the sum of which gives us the population size at equilibrium N_{mut}. Using N_{mut} we obtain the expression for the numerical load Λ (the decrease of population size due to the presence of deleterious mutations), a term defined by Clarke

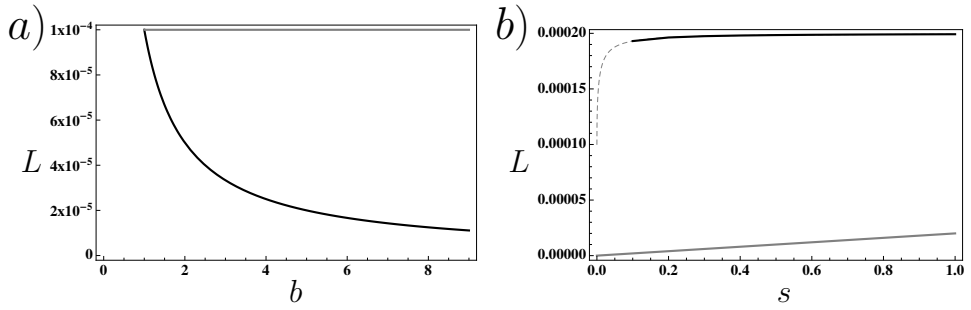


Figure 2.3: *a*) L as a function of the birth rate for hereditary somatic mutations occurring at the adult stage (in black) and the zygote stage (in grey). $s = 0.1$, $h = 0$ and $\mu = 10^{-4}$. b) L as a function of the coefficient of selection for hereditary (in black) and non-hereditary (in grey) somatic mutations occurring at the zygote stage. The dashed line is the genetic load for the model with selection on fecundity from section 2.1. $b = 1$, $h = 0.1$ and $\mu = 10^{-3}$). For these parameter values, somatic mutations at the adult stage give numerically equivalent values.

(1973) is given in equation 2.8. We also use the expressions for X_{mut}, Y_{mut} and Z_{mut} to derive the expressions for the mutational load L (equation 2.9).

2.2.4 Preliminary Results and Conclusions

As in the model presented in Chapter 2.1, selection on fecundity leads to a numerical load Λ equal to the genetic load L for both the hereditary and non-hereditary mutations, independently of the timing of mutation. Hence from the point onwards, as $\Lambda = L$, we will only be referring to the genetic load. We find that somatic mutations occurring at the adult stage, be they hereditary or not, lead to loads that are dependent on the intrinsic birth rate b (see Figure 2.3a). When mutations are recessive ($h = 0$) and hereditary, $L = \frac{\mu}{b}$ for mutations at the adult stage, increasing the birth rate therefore leads to a decrease in the observed loads. The genetic load due mutations introduced exclusively at the zygote stage is completely independent of b, with (when $h = 0$) $L = \mu$. If b is set at 1, then the timing of mutation has no effect on the observed loads and when mutations are hereditary, and for $h \neq 0$, L is in fact numerically equivalent to that obtained when mutations occur during gamete formation (see Chapter 2.1). In figure 2.3b we present the genetic load L as a function of the coefficient of selection for mutations that occur at the zygote stage. When the mutations are non-hereditary, both the numerical and the genetic loads are simply $2h\mu s$. This result is intuitive as mutations can only be found in the heterozygote state and can only be introduced by mutation. Similarly, when mutations are not hereditary, regardless of the

timing of mutation, the dominance h has no effect on the frequencies of each genotype.

In light of these exploratory results, as well as the results of Ezawa and Innan (2013) on the polymorphism of neutral mutations, somatic mutations influence the observed allelic frequencies and may play a role in decreasing population size and increasing the genetic load. That mutation at the adult stage lead to a load that is dependent on the birth rate is intuitive. The more new mutations free individuals introduced, the lower the actual frequency of "older" individuals carrying the deleterious allele a. In their paper, Cheprakov and Rakitin (2012) had tested for structural chromosomal aberrations (among other genetic abnormalities) in voles. The authors found that the frequency of this particular abnormality (induced only in somatic cells) decreased with an increase in population size. Their interpretation was that this was due to a reduced somatic mutation rate that given year of sampling. In light of our results however, the lower frequencies of detected genetic abnormalities mays simply be due to higher birth rate that year.

Non-hereditary mutations could be compared to the onset of cancer (one of the consequences of somatic mutations, Lynch 2010). No models to our knowledge have studied how the onset of cancer within populations influences demographic variables. As the mutation rate is predicted to play a role in an individual's susceptibility of developing cancer (Nunney, 1999), a perspective to this preliminary work would be to consider a heritable "mutation rate".

Interlude

In the previous chapter we examined the potential influence of the interplay between demography and selection on both population size and the genetic load at a single locus. Clarke (1973) verbally predicted that taking multiple loci into account could lead to a non-negligible effect of the genetic load on population size. In the following chapters we develop stochastic individual-based models with a potentially infinite number of loci testing his prediction. We also take these models one step further by introducing self-fertilisation, either as a fixed (Chapter 3) or an evolving trait (Chapter 4).

Act 3

Population viability and self-fertilisation: a multi-locus model

Published work: PLoS ONE 9(1): e86125. DOI:10.1371/journal.pone.0086125

The interaction between selection, demography and selfing and how it affects population viability

Diala Abu Awad[1,*], Sophie Gallina[1], Cyrille Bonamy[2], Sylvain Billiard[1]
1 UMR-CNRS 8198, Laboratoire Génétique et Évolution des Populations Végétales, Université Lille 1, Villeneuve d'Ascq, France
2 Centre de Ressources Informatiques (CRI), Université Lille 1, Villeneuve d'Ascq, France.
∗ **E-mail: Corresponding diala.abu-awad@ed.univ-lille1.fr**

Abstract

Population extinction due to the accumulation of deleterious mutations has only been considered to occur at small population sizes, large sexual populations being expected to efficiently purge these mutations. However, little is known about how the mutation load generated by segregating mutations affects population size and, eventually, population extinction. We propose a simple analytical model that takes into account both the demographic and genetic evolution of populations, linking population size, density dependence, the mutation load, and self-fertilisation. Analytical predictions were found to be relatively good predictors of population size and probability of population viability when verified using an explicit individual based stochastic model. We show that initially large populations do not always reach mutation-selection balance and can go extinct due to the accumulation of segregating deleterious mutations. Population survival depends not only on the relative fitness and demographic stochasticity, but also on the interaction between the two. When deleterious mutations are recessive, self-fertilisation affects viability non-monotonically and genomic cold-spots could favour the viability of outcrossing populations.

3.1 Introduction

Population size and viability are both affected by extrinsic (e.g. environmental change and interspecific interactions) and intrinsic factors (e.g. genetic and demographic components). The genetic factors most frequently considered as contributing to population decline are the lack of adaptive potential in a changing environment, inbreeding depression and the reduction of fitness due to the accumulation of deleterious mutations (reviewed in Frankham

2005). The accumulation of deleterious mutations has often been put forth as an explanation for species extinction, especially concerning the differences in extinction rates between sexual and asexual species, or selfers and outcrossers. The relevance of the accumulation of deleterious mutations on population extinction, however, remains unclear.

Both empirical and theoretical works have insisted on the importance of deleterious mutation fixation on the decline and extinction of populations. Some models have shown that small populations can go extinct due to the acceleration of recurrent fixation of deleterious mutations with a small effect, the so-called mutational meltdown (Lynch and Gabriel, 1990; Lande, 1994; Lynch et al, 1995; Coron et al, 2013). Several empirical works have also supported this hypothesis. The fitness of experimental populations has been shown to decrease after several generations during which new deleterious mutations are fixed (Newman and Pilson, 1997; Frankham et al, 2001; Zeyl et al, 2001; Vassilieva and Lynch, 1999; Baer et al, 2005), and data from small highly inbred natural populations follow the same trend (Packer et al, 1991; Westemeier et al, 1998; Gelatt et al, 2010). However, in these theoretical and empirical works, populations are considered to be small and isolated or, because of successive growth-dilution cycles, are subjected to recurrent and strong bottlenecks. When these conditions are not met (i.e. when populations are larger, or are not isolated or are not subject to strong recurrent bottlenecks) populations are more likely to go extinct because of other genetic and demographic factors before the fixation of deleterious mutations has an effect on population decline (Frankham, 2005).

What about large populations, can they decline in size due to recurrent deleterious mutations? It has been suggested that the mutation load due to segregating mutations might be important in population decline (Bernardes, 1995; Lynch et al, 1995). This however remains controversial as it is generally thought that in large sexual populations deleterious mutations should be efficiently purged (Hedrick, 2004; Whitlock M.C., 2004). First, segregating deleterious mutations are expected to have no consequence on demography (Agrawal and Whitlock, 2012), especially in the presence of density-dependence, where there is compensation of the death of individuals due to selection by those that would have been lost from the population due to the lack of resources (soft selection). Empirical evidence on the other hand supports more "hard selection" (Saccheri and Hanski, 2006), in which density-independent deaths due to the genetic load are not completely compensated. It is crucial to determine whether segregating mutations are important or not for determining population viability as this has direct empirical implications, especially when considering the genetic rescue of populations. Second, many aspects of population survival and extinction in presence of a high mutation

load still remain unclear. When taking into account empirical estimations of genomic mutation rates (between 0.1 and more than 1 for higher Eukaryotes per genome and per generation, (Sniegowski et al, 2000)) and the effects of deleterious mutations, population genetics theories (Haldane, 1937) imply that we should expect extremely high mutation loads. Population genetics models consider population size as a fixed variable, and their stochastic estimations of the mutation load, even in finite populations, also agree with the existence of high mutation loads (Bataillon and Kirkpatrick, 2000; Haag and Roze, 2007). Can these predictions still hold if we allow selection to influence population size, or would the mutational load evolve differently? One of the earliest models, to our knowledge, that has taken into account the effect of the mutation load on mean fitness, and the latter's effect on population size is a single-locus model proposed by Clarke (1973). In this model, the mutation load had a very small, almost negligible effect on population size, however Clarke (1973) verbally suggested that extending the model to a whole genome could possibly lead to a more important effect.

In order to better understand the interaction between the mutation load and demography, we propose a model that combines simple deterministic population genetics and demographic models. We consider sexual reproduction, with a mean reproductive rate that depends both on population density and on the population's mutation load, and recurrent mutations segregating at an infinite number of loci. Using this model we predict the relationship between population size and the mutational load at mutation-selection equilibrium. We also predict the threshold fitness value depending on both the genetic and demographic parameters under which the population is not expected to be viable. These predictions are then verified using an individual-based simulation model in which we explicitly model the introduction of new mutations in the genome and the effect of the mean genomic recombination rate. Population size varies from one generation to the next, as it depends on individual fitness and competition. This simulation model allows for the better understanding of the mechanisms leading to population extinction, more specifically the relative importance of the polymorphism of deleterious mutations, the fixation of these mutations, and the mutational meltdown.

To illustrate the importance of such an approach (i.e. combining genetic and demographic factors) in ecology and evolution, we will address the question of the effect of self-fertilisation on population size and viability. This is indeed a long running question, as the prevalence of outcrossing species is still puzzling both in animals and plants (Jarne and Auld, 2006). From an evolutionary standpoint, self-fertilisation should be greatly advantaged because of Fisher's transmission advantage (Fisher, 1941), a more efficient purge of deleterious mutations (Glémin, 2003), and also because of their reproductive

assurance (e.g. Baker's law, Baker 1955; Stebbins 1957). This is correlated with the empirical estimation of high transition rates from outcrossing to selfing reproductive systems, for instance in the *Solanaceae* (Goldberg et al, 2010). Despite this transition rate, and high speciation rates in selfers compared to outcrossers, outcrossers still represent more than 40% of species in this family. Other studies in other plant families also come to this conclusion (Schoen et al, 1997; Perez-Barrales et al, 2006; Sakai and Wright, 2008). Goldberg et al (2010) show that this puzzling prevalence of outcrossers is due to higher extinction rates in selfing species than in outcrossing ones. One hypothesis to explain this difference in extinction rates is that selfers are more prone to mutational meltdown than outcrossers (Lynch et al, 1995). Empirical work on fungi, more specifically *Neurospora*, strongly supports this hypothesis, as selfing lineages accumulate more deleterious mutations and are less fit than outcrossing ones (Gioti et al, 2013). We will therefore extend our model in order to include different rates of self-fertilisation and test this hypothesis.

3.2 Model

3.2.1 Deterministic model and expectations

We consider a population in a constant environment, with discrete, non-overlapping generations. At generation t, the population is made up of N_t hermaphrodite individuals, where

$$N_t = N_{t-1} R_{t-1}. \tag{3.1}$$

R_t is the absolute multiplicative fitness of a population at a given generation t, with trait value (or relative fitness) W_t, in a population of density N_t (Chevin and Lande, 2010) is given by

$$R_t = r_0^{1-\frac{N_t}{K}} W_t. \tag{3.2}$$

The density-dependent component of fitness depends on K, the carrying capacity of a population with all individuals having the optimal relative fitness, and on r_0, the intrinsic reproductive rate of such a population (Chevin and Lande, 2010). The second factor, W_t, corresponds to the mean relative fitness of individuals in the population, as a function of their mutation load, and so on the number of segregating or fixed deleterious mutations in the population. In this model, we consider that density dependence affects all individuals in the same way, independently of their relative fitness (or genotype).

If the mean relative fitness is at equilibrium (when the population is at mutation-selection balance), and there is no demographic stochasticity, the equilibrium of population size denoted N_{det}, can be expressed as a function of the relative fitness at equilibrium W_{eq}, r_0 and K, giving

$$N_{det} = K\left(1 + \frac{\mathrm{Ln}(W_{eq})}{\mathrm{Ln}(r_0)}\right). \tag{3.3}$$

By solving $N_{det} = 0$ from equation 3.3, we can determine the relative fitness threshold

$$W_{min} = \frac{1}{r_0}, \tag{3.4}$$

under which the population is not expected to be viable.

A deterministic value of W_{eq}, noted W_{det}, can be calculated at mutation-selection balance for a large population of diploid individuals with a large number of independent bi-allelic loci, where deleterious mutations with selection coefficient s and dominance h can segregate. This is done using equations for the mutation load L derived from Wright's equations for allele frequencies at equilibrium at a single locus (Caballero and Hill, 1992; Roze and Rousset, 2004), and gives

$$L = u \tag{3.5a}$$

for recessive mutations (the dominance coefficient $h = 0$), and

$$L = u\frac{4h + \alpha(1-4h)}{2h + \alpha(1-2h)} \tag{3.5b}$$

when $h >> 0$, where u is the deleterious mutation rate at a single locus and α is the proportion of offspring produced via self-fertilisation. If we consider that there is no epistasis and no linkage disequilibrium, then the relative population fitness when considering a genome-wide mutation load is given by $W_{det} = e^{-L}$ (Haldane, 1937), where u is replaced by U, the haploid genomic mutation rate, when calculating L.

Finally we can calculate N_{det}, using W_{det} instead of W_{eq} in Equation 3.3 as an estimation of population size at equilibrium N_{eq}. Populations are not expected to be viable ($N_{det} \leq 0$) when for $h = 0$

$$r_0 \leq e^{U} \tag{3.6a}$$

and for $h >> 0$

$$r_0 \leq e^{\left(U\frac{4h+\alpha(1-4h)}{2h+\alpha(1-2h)}\right)}. \tag{3.6b}$$

High mutation rates and low reproductive rates are both expected to contribute to the decrease of population viability.

3.2.2 Simulation Model

An individual-centred model with discrete non-overlapping generations was used to follow the evolution of an isolated population of variable size, made up of diploid hermaphrodite individuals.

Genomic assumptions

The genetic properties of this model, mutation and recombination, are those described in Roze's (2009) model. Each individual is represented by two homologous chromosomes of length $2D$ with a potentially infinite number of loci. The map length is considered to be D from the centre of the chromosome to the edge, hence representing a chromosome with a defined centromere. The life cycle is as follows: mutation, selection, meiosis and reproduction.

The number of new deleterious mutations occurring per chromosome per generation, is sampled from a Poisson distribution with mean U, where $2U$ is the genomic mutation rate. Their position on the chromosome is sampled from a uniform distribution in $[-D, D]$. The effect of deleterious mutations on the fitness of individual i living at time t, W_t^i, is multiplicative and depends on the number of homozygous, n_{ho}, and heterozygous, n_{he}, deleterious mutations each individual carries

$$W_t^i = (1 - hs)^{n_{he}}(1 - s)^{n_{ho}}, \tag{3.7}$$

where s and h are the selective and dominance coefficients respectively and are fixed parameters. All mutations are deleterious and have the same values of s and h. The deleterious effect of these mutations is independent of population density. Recombination occurs during gamete production and is considered to be uniform along the chromosome. New individuals are a combination of two gametes, either from two different individuals for reproduction via outcrossing, or the same individual via selfing.

Demography and selection

At a given time t, population size N_t is given by

$$N_t = \sum_{i=1}^{N_{t-1}} X_{t-1}^i \tag{3.8}$$

where X_t^i is the number of viable offspring an individual i at time t contributes to the next generation via the female function (we consider that there is no limitation in the number of offspring an individual i can sire).

X_t^i is sampled from a Poisson distribution with mean $R_t^i = W_t^i r_0^{1-\frac{N_t}{K}}$ (the individual reproductive rate).

Self-fertilisation occurs at a probability α_t^i for individual i, at time t given by

$$\alpha_t^i = \frac{\alpha_0 W_t^i}{\alpha_0 W_t^i + (1-\alpha_0)\frac{\sum_{j\neq i} W_t^j}{N_t - 1}}. \tag{3.9}$$

The proportion of selfed offspring depends on α_0, the proportion of an individual's male gametes that are available for self-fertilization, and on the individual's fitness W_t^i compared to the average relative fitness of the other possible fathers in the population ($\frac{\sum_{j\neq i} W_t^i}{N_t - 1}$). The lower an individual's relative fitness as a father, the lower the proportion of offspring produced via selfing. We consider that there is no limitation in the availability of male gametes. The proportion of an individual's offspring produced by self-fertilisation is sampled from a binomial distribution with parameters α_t^i and X_t^i. When $\alpha_0 = 0$ the population is strictly outcrossing and the population is automatically considered non-viable if $N_t < 2$.

In order to facilitate the reading of the following sections, the notation used through the text has been summarized in Table 3.1.

Initial conditions and simulations run

At the beginning of each simulation, we consider population size to be equal to K and that there are no deleterious mutations present in the population. The simulations are run until the population reaches equilibrium or goes extinct. We define equilibrium as the stabilisation of the mean population fitness, denoted $\overline{W}_t$; the average $\overline{W}_t$ over one thousand generations is calculated and compared to the average $\overline{W}_t$ of the previous thousand generations. If the difference between the two is lower than 1 per cent the population is considered to be at equilibrium. Population size N_{eq}, mean fitness $\overline{W}_{eq}$ and the mean number of mutations per chromosome were measured at equilibrium. Throughout the results, we will mostly be addressing the average of the mean value of the population's relative fitness across all simulations for each set of parameter values, which we note $\widehat{W}_{eq}$. If the population goes extinct, then the time to extinction is measured.

The mutational meltdown is defined as the acceleration of the decrease of population size due to the accumulation of deleterious mutations (Lande,

1994; Lynch et al, 1995; Coron et al, 2013). In order to evaluate this acceleration, once non-viable populations reach a population size of 250 individuals, the best fitting quadratic polynomial regression ($a + yt + zt^2$) is calculated to fit the decrease in $\overline{W}_t$, N_t and $\overline{R}_t$ independently of one another. When these variables are decreasing, the first order coefficient is negative. As these variables decrease with time, if they decrease in a linear fashion, then there is no mutational meltdown and the second order coefficient z is equal to 0. In the case of acceleration of the decrease of these variables with time, as expected in the case of a mutational meltdown, then the second order coefficient z, like the first order coefficient y, is negative. This second order coefficient z is calculated for each simulation run that results in population extinction. We also measure the mean population size at the fixation of the first deleterious mutation in order to detect whether population decline is associated to mutation fixation.

A wide range of values, from 0 to 1, were run for the parameter α_0, with $r_0 = 2, 4$ and 10. For $r_0 = 2$, simulations were run for U between 0.1 and 0.6, mutations were mildly deleterious, moderately deleterious or lethal ($s = 0.02$, 0.2 and 1), that were completely recessive, almost recessive, or moderately recessive ($h = 0$, 0.02 and 0.2). The recombination rates taken into account reflect conditions where mutations were highly linked, moderately linked or very slightly linked ($D = 0.1$, 1 and 10). Aside from the general effect of recombination rates over a whole genome, these genomic recombination rates can also reflect how mutation loads evolve within a genome in specific genomic hot and cold spots. Increasing D over 10 has very little effect on the results, which allows us to make the assumption that the mutations act as though they were independent (Roze, 2012). However, it is possible, when there is selfing, that there is some linkage due to the genomic consequences of self-fertilisation. One thousand replicates were run for each group of parameter values, coming to a total of 1458000 simulations run for $r_0 = 2$. For $r_0 = 4$ and 10, one hundred replicas were run for $U = 0.2$ to 1, $s = 0.02$ and 0.2, $h = 0$ and 0.02 and $D = 10$, leading to a total of 54000 simulations run.

We compare the expected deterministic values of the mean fitness at equilibrium W_{det}, as well as the expected deterministic fitness threshold value W_{min} under which populations should not be viable, with our simulation results.

3.2.3 Estimating the stochastic fluctuations of population size

The stochastic fluctuation of population size from one generation to the next can be due to two mechanisms: demographic stochasticity alone or the interaction between demography and genetic selection. In order to estimate the importance of each of these two sources of stochasticity, we first estimated the fluctuations that would be observed with only demographic stochasticity and no mutations (the relative fitness $\overline{W}_t$ is a constant) and then compared these estimations with the fluctuations observed in our simulations. We use the standard deviation of population size over time σ_N as a measure of these fluctuations. We compare σ_N calculated from simulations run for 100000 generations for different constant values of $\overline{W}_t$ to the standard deviation of N_t over 100 generations when the populations in our simulations were at equilibrium. If demographic stochasticity alone can explain the fluctuation of population size, then the stochastic fluctuations calculated from the simulations with a dynamic component of relative fitness (denoted $\widehat{\sigma}_N$) should not be very different than those run with $\overline{W}_t$ as a constant.

In our simulations, population extinction is inevitable, as when time is very long, all populations go extinct due to demographic stochasticity with a probability of 1. In order to test whether populations that are expected to be viable, in other words with an expected relative fitness W_{det} greater than the threshold value W_{min} (Equation 3.3), go extinct because of demographic stochasticity alone, we ran stochastic simulations of Equation 3.8. We assume that all individuals have the same constant relative fitness W_{det} (equation 3.5) and the initial population size is N_{det} (Equation 3.3). We calculate the probability of population extinction before one order of magnitude higher than the highest time to extinction observed in our simulations with a dynamic component of relative fitness.

Table 3.1: **Notation.**

$V, \overline{V}, \widehat{V}$	No superscript indicates deterministic values (except in the case of population size N, where it is mentioned clearly in the text), a bar indicates that the variable is the intra-population mean for one simulation run ($\overline{W}_{eq}$ is the mean relative fitness for one population) and a hat indicates that the variable is the mean across all simulations ($\widehat{W}_{eq}$ is the mean relative fitness across all simulations, conditional to survival).
$N_t, \widehat{N}_{eq}$	Population sizes at generation t and mean population size at population equilibrium across all simulations
$\sigma_N, \widehat{\sigma}_N$	Standard deviation of population size over time as a measure of the fluctuation of population size for one population and across simulations.
N_{det}	Expected population size without demographic or genetic stochasticity
r_0	The intrinsic reproductive rate.
K	The carrying capacity.
R_t	Expected number of offspring per individual at generation t.
X_t^i, R_t^i	Respectively the number of offspring produced by individual i at generation t and its expectation.
$\overline{W}_t, \overline{W}_{eq}, \widehat{W}_{eq}$	Means of the population's relative fitness at generation t and at population equilibrium for one or across simulations conditional to population survival.
W_t^i	Relative fitness of individual i at generation t.
W_{det}	Expected value of the population's component of fitness calculated using the Wright-Fisher model.
W_{min}	The threshold value of the population's mean relative fitness, under which populations are not viable.
U, D	Genomic properties: the haploid mutation rate and the recombination rate or map length.
s, h	Mutational effects: the selection coefficient and the dominance.
n_{he}, n_{ho}	The number of heterozygote and homozygote mutations on an individual's genome.
a_0	The proportion of male gametes available for selfing.
$\alpha_t^i, \overline{\alpha}_t$	Proportion of offspring produced via self-fertilisation by individual i and the population's mean proportion of offspring produced via self-fertilisation at generation t.

3.3 Results

3.3.1 Demographic and genetic evolution to equilibrium and extinction

Populations either evolve both demographically and genetically to a quasi-stationary equilibrium of population size and the relative fitness, noted respectively N_{eq} and $\overline{W}_{eq}$, or go extinct (illustrated in Figure 3.1A).

For all values of the intrinsic reproductive rate r_0, the simulation results agree with deterministic expectations of the interaction between W_{eq} and N_{eq} from Equation 3.3 (see Figure 3.1B) for viable populations. When $r_0 = 10$ and W_{eq} values are greater than 0.74, the deterministic equilibrium is oscillatory and unstable, as confirmed by a stability analysis of equation 3.3. This is clearly seen in the simulation results, where the populations oscillate between two states.

Very few populations are viable with a relative fitness at equilibrium $\overline{W}_{eq}$ that is lower than the estimated value of the relative fitness threshold W_{min}, and as expected from Equation 3.4, W_{min} decreases with an increasing intrinsic reproductive rate r_0. This strong relationship between $\overline{W}_{eq}$ and N_{eq} is not enough, however, to explain population extinction. When taking a closer look at the relationship between the probability of population extinction and mean population fitness, we do not have a positive linear relationship between the two variables, but a more bimodal distribution (Figure 3.1C). We observe a great range of values of $\widehat{W}_{eq}$ (i.e. the average $\overline{W}_{eq}$ of viable populations) across simulations for which all populations survive, a similar range of $\widehat{W}_{eq}$ for parameter sets for which very few populations were viable and intermediate-low values of $\widehat{W}_{eq}$ for populations with an intermediate probability of extinction.

When the deterministic value of mean fitness W_{det} is greater than the fitness threshold W_{min} but populations are not viable, extinction can be attributed to either demographic stochasticity alone, or to an interaction between both demographic and genetic stochastic processes. When we consider the mean relative fitness $\overline{W}_t$ to be a constant, so that population extinction is due only to demographic stochasticity, we find that the probability of extinction is extremely low compared to what is observed in our results. For example for $r_0 = 2$ and $\overline{W}_t = 0.52$, the probability of extinction within a time equivalent to one order of magnitude greater than our highest time to extinction observed is almost null. In other words, these populations are expected to survive as their relative fitness is above the threshold value W_{min}, and the expected time to stochastic demographic extinction is extremely long. This

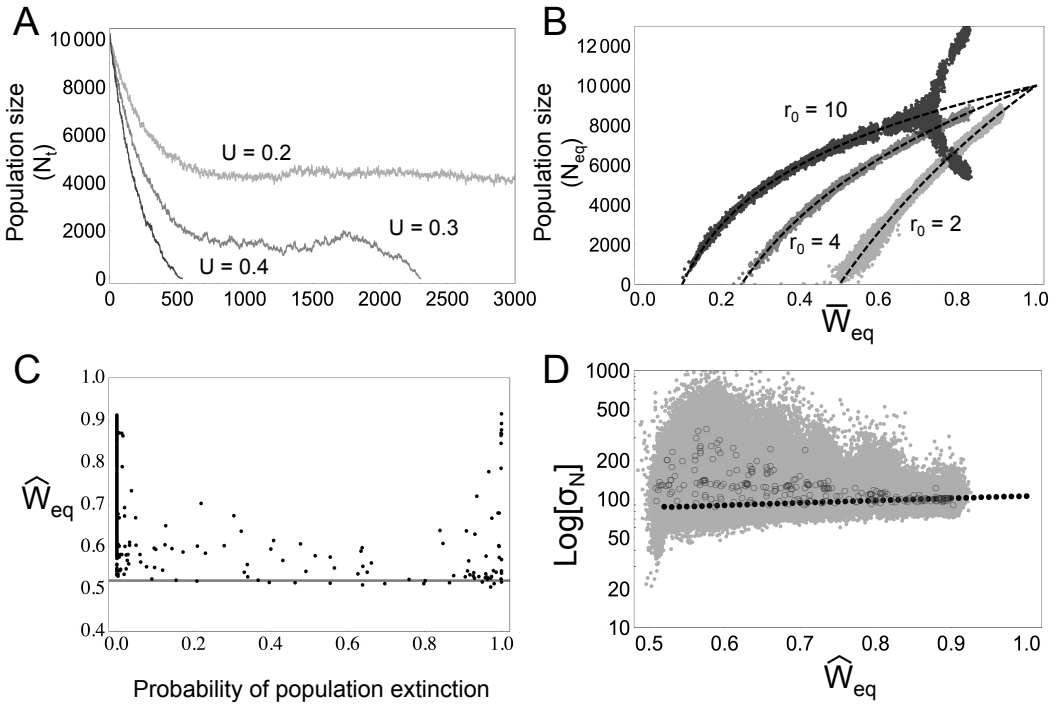


Figure 3.1: Demographic and genetic evolution of populations. A) Typical evolution of population size N_t with time for a viable population ($U = 0.2$), one that reaches equilibrium but goes extinct due to stochasticity ($U = 0.3$), and one that is not viable ($U = 0.4$). $\alpha_0 = 0$, $s = 0.02$, $h = 0.2$, $D = 10$, $K = 10000$ and $r_0 = 2$. $\overline{W}_t$ follows the same pattern. B) Population size at equilibrium, N_{eq} as a function of mean population fitness $\overline{W}_{eq}$ for different values of r_0. The dashed line represents the expected population size N_{det} from Equation 3.3, and the points represent simulation results for all viable populations for all parameter sets with $D = 10$ and $K = 10000$. C) Probability of population extinction from simulations run for all sets of parameter values for $r_0 = 2$ and $K = 10000$ as a function of the average population's mean relative fitness $\hat{W}_{eq}$. The grey line represents the population fitness above which the probability of population extinction in less than 10^5 generations due to demographic stochasticity alone is almost null. D) Standard deviation of population size over time at population equilibrium σ_N as a function of the relative fitness $\overline{W}_{eq}$ from simulations run for all sets of parameter values for $r_0 = 2$ and $K = 10000$. The light grey points each represent σ_N of a single simulation, the open circles represent the mean standard deviation across simulations per group of parameter values $\hat{\sigma}_N$ and the full circles represent results from simulations run that take into account only demographic stochasticity.

however is not the case in simulations with a dynamic component of fitness, suggesting that it is the interaction between demography and genetics that leads to population extinction in such a relatively short time scale.

The importance of this interaction is observed when taking into account the standard deviation of population size over time σ_N (Figure 3.1D). We observe that the values of σ_N for each simulation run (grey points) fluctuate around the expected standard deviation if the change in population size from one generation to the next were due only to demographic stochasticity (full circles). Fluctuations are even higher when the relative fitness $\overline{W}_{eq}$ decreases. We note that the mean value of the standard deviation of population size over time σ_N for any group of parameter values $\widehat{\sigma}_N$ (open circles) is always greater than the expected values of standard deviation of this variable when only demographic stochasticity affects population size. This suggests that the interaction between demography, genetics and selection highly increases the stochastic fluctuations of population size. Below, we show that this increased stochasticity is important when estimating the probability of population extinction.

From this point onwards, we will mostly be addressing results with an intrinsic reproductive rate $r_0 = 2$ since that we observe the same patterns for higher values of r_0 when considering population extinction and mutational loads. For simplicity, in order to compare the effect of selfing between populations, we have chosen to compare population characteristics for different values of the proportion of male gametes available for selfing α_0 and not $\overline{\alpha}_t$, the effective mean proportion of offspring produced via self-fertilisation of the population. The difference between the two is slight (unless the population is very near extinction), with $\alpha_0 \leq \overline{\alpha}_t \leq \alpha_0(1 + 0.1)$, except for $\alpha_0 = 0$ and 1, where it remains fixed.

3.3.2 Probability of population extinction

As expected from deterministic approximations (see section "Deterministic model and expectations", Equation 3.6), when mutations are only slightly linked ($D = 10$) we find that for a particular intrinsic reproductive rate r_0, increasing values of the haploid mutation rate U leads to higher probabilities of extinction (Figure 3.2A) and lower $\widehat{W}_{eq}$ (Figure 3.2B). Contrary to deterministic expectations, the coefficient of selection s affects both population extinction and mean relative fitness. The effect of the coefficient of the selection and the proportion of selfed offspring both depend greatly on the dominance h of the deleterious mutations. Generally, increasing the coefficient of selection decreases the probability of extinction at higher mutation

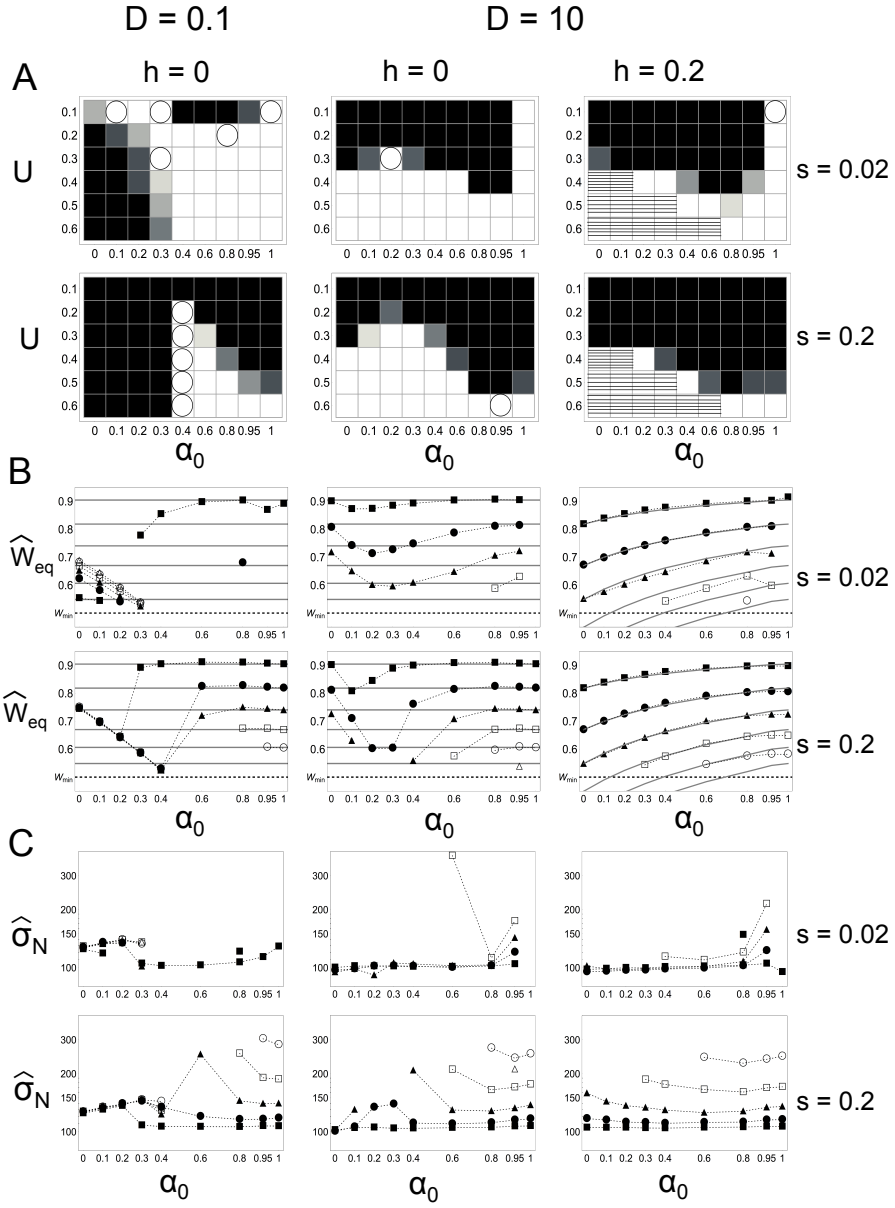


Figure 3.2: Population extinction and equilibrium. A) Probability of population extinction calculated from the 1000 simulations run ($r_0 = 2$ and $K = 10000$), from black (0% extinction) to white (100% extinction), as a function of α_0. The circles indicate that more than 95% of populations went extinct. The horizontal lines indicate that deterministic extinction was predicted (Equation 3.3). B) Mean values of the observed population fitness at equilibrium across simulations $\tilde{W}_{eq}$ of viable populations as a function of α_0 for $r_0 = 2$ and $K = 10000$. The grey lines represent the expected mean fitness W_{det} for (from top to bottom) $U = 0.1$, 0.2, 0.3, 0.4, 0.5 and 0.6. Top row $s = 0.02$ and below $s = 0.2$. Missing points indicate parameter values for which none of the 1000 simulations run were viable. C) Standard deviation of population size over time at population equilibrium across simulations $\widehat{\sigma}_N$ with a logarithmic scale as a function of α_0, with $r_0 = 2$ and $K = 10000$. Note that this variable is underestimated for parameter sets with less than 100% population survival, as the standard deviation of extinct populations is not taken into account. For B) and C) : ■ $U = 0.1$, ● $U = 0.2$, ▲ $U = 0.3$, □ $U = 0.4$, ○ $U = 0.5$, △ $U = 0.6$

rates and increases mean relative fitness $\widehat{W}_{eq}$.

From the deterministic equations, the proportion of selfed offspring should not affect either of these variables when mutations are completely recessive and almost recessive ($h = 0$). However, self-fertilisation has a non-monotonic effect on both of these variables. Between $\alpha_0 = 0$ and 0.2 the probability of extinction increases, while $\widehat{W}_{eq}$ decreases. For α_0 between 0.2 and 0.95, the opposite tendency is generally observed, the greater α_0, the lower the probability of extinction and the higher $\widehat{W}_{eq}$. When a strictly selfing regime is imposed, no simulated populations survive when mutations are almost neutral ($s = 0.02$) and when mutations are mildly deleterious ($s = 0.2$) the probability of extinction (respectively the mean relative fitness) is the same, or slightly higher (respectively lower), as what is observed for $\alpha_0 = 0.95$. The same patterns are observed when mutations are almost recessive ($h = 0.02$). When mutations are moderately recessive ($h = 0.2$), we find that the deterministic expectations are more accurate. There is generally a monotonic effect of α_0 on both the probability of extinction viability and $\widehat{W}_{eq}$, the former decreases and the latter increases with increasing α_0. At higher mutation rates, we observe that the probability of extinction increases and $\widehat{W}_{eq}$ decreases at very high values of α_0 (> 0.8). Once again, when mutations are almost neutral ($s = 0.02$), no populations survive when $\alpha_0 = 1$. For all values of h, increasing U accentuates the effect of α_0 on the decrease of $\widehat{W}_{eq}$ and increase of the probability of extinction.

At lower recombination rates ($D = 1$ and 0.1), $\widehat{W}_{eq}$ is generally not affected by s and h, but the probability of population extinction is generally greater with increasing U than what is observed for high recombination rates (for $D = 10$). This is especially true for lower proportions of selfed offspring. The effect of recombination on viability decreases with increasing selection. We observe one particular case, when mutations are very tightly linked (the recombination rate $D = 0.1$), of small effect (the coefficient of selection $s \leq 0.2$) and completely recessive (dominance $h = 0$), where increasing the haploid mutation rate U can, for low rates of α_0, decrease the probability of extinction and increase the mean relative fitness or have no effect on either (Figures 3.2A and 3.2B). This could be due to more efficient selection against deleterious mutations at higher mutation rates, as increasing the mutation rate could increase the probability that tightly linked groups of recessive mutations are found at a homozygote state and eliminated.

As mentioned above, population fitness and demographic stochasticity alone cannot fully explain population viability. Figure 3.2C represents the mean standard deviation of population size over time at equilibrium conditional to population survival $\widehat{\sigma}_N$. We observe that, when considering pa-

rameter sets with the same mutation rate U, $\widehat{\sigma}_N$ increases with increasing probability of population extinction. It is important to note that $\widehat{\sigma}_N$ is most probably underestimated for parameter sets for which not all populations are viable, as the standard deviation of extinct populations is not taken into account during its calculation.

3.3.3 Accumulation and fixation of deleterious mutations

Our results show that when deleterious mutations are fixed, it is only in populations that are very small, in other words already on their way to extinction. Even though the mean number of mutations is smaller at higher proportions of selfed offspring (Figure 3.3A), when populations are not viable, increasing α_0 leads to the fixation of deleterious mutations at greater population sizes (Figure 3.3B) and with a higher probability (results not shown). The smaller the coefficient of selection and the greater the mutation rate, the greater the population size at first fixation. Therefore, if mutations are indeed fixed, they are fixed at larger population sizes when they are mildly deleterious, frequently introduced and in populations with greater proportions of selfed offspring. Mean population size at first fixation is generally relatively small compared to the carrying capacity K (see Figure 3.3B), except for one group of parameter values ($U = 0.1$, $s = 0.02$, $h = 0$ and $\alpha_0 = 1$, results not shown) where fixation can occur at a population size of 8000 individuals. Mean population size at first fixation is greater for high-intermediate values of α_0 when there is little recombination than when recombination rates ($D = 10$) are high, whereas outcrossing populations are not affected.

The mean number of mutations per chromosome at population equilibrium across all simulations (conditional to population survival) decreases with increasing coefficients of selection s and/or the dominance h and for increasing proportions of male gametes available for selfing α_0 (Figure 3.3A shows typical results). Increasing the haploid mutation rate U increases the mean number of mutations per chromosome. The lower number of mutations per chromosome for mutations with stronger effects (either at the homo- or heterozygous state) and for higher proportions of selfed offspring can be explained by a more efficient purging (Glémin, 2003).

Lower recombination rates generally do not affect the mean number of mutations per chromosome, except when mutations are completely recessive ($h = 0$) and there is almost no recombination ($D = 0.1$). In this case the mean number of mutations per chromosome can be more than doubled for low values of α_0 independently of the coefficient of selection s, but not greatly

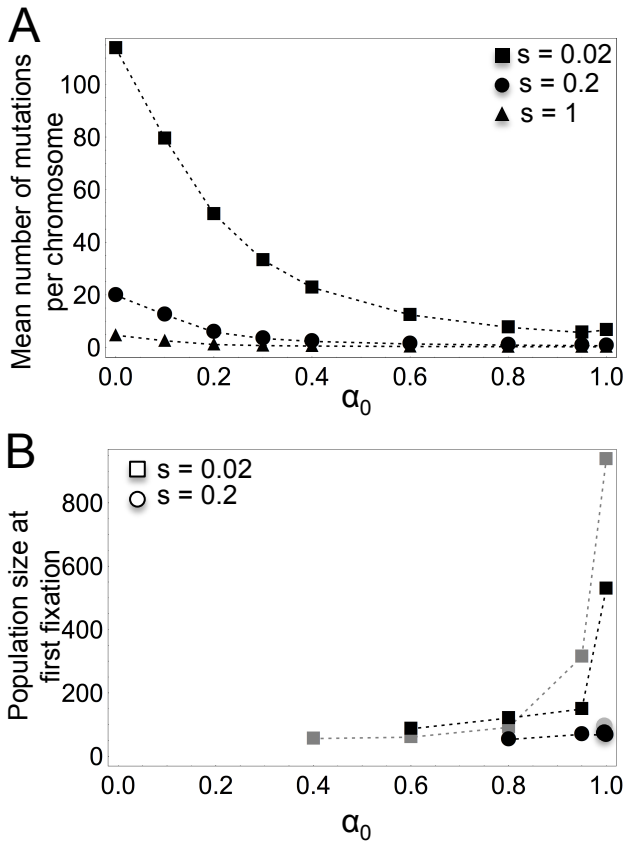


Figure 3.3: Accumulation and fixation of deleterious mutations. A) Mean number of mutations per chromosome equilibrium across simulations as a function of α_0. $U = 0.1$, $h = 0.02$, $D = 10$ and $r_0 = 2$. Missing points present parameter values for which no populations were viable. B) Mean population size at first fixation of deleterious mutations for populations greater than 50 individuals as a function of α_0. $U = 0.5$ in grey, $U = 0.6$ in black. $h = 0.2$, $D = 10$ and $r_0 = 2$.

changed for higher values of α_0 when selection is strong $s \geq 0.2$.

3.3.4 Mutational meltdown and time to extinction

When a population is on its way to extinction, there is a weak but significant acceleration in the decrease of the mean relative fitness $\overline{W}_t$ and in the mean reproductive rate $\overline{R}_t$, but there is a deceleration in the decrease of population size N_t for high recombination rates ($D = 10$, see Figure 3.4A). Higher mutation rates U, lower dominance of mutations h and lower proportions of selfed offspring α_0 contribute to the acceleration (respectively the deceleration) of the decrease of $\overline{W}_t$ and $\overline{R}_t$ (respectively N_t). Even though both the mean relative fitness and the reproductive rate show an overall tendency to decrease at an accelerating rate, the low population density allows for the deceleration of the decrease of population size as the smaller the population size, the more resources available per individual for reproduction, as $\overline{R}_t$ is density dependent (Equation 3.2).At low recombination rates, there is neither an acceleration or a deceleration in the decrease of $\overline{W}_t$, $\overline{R}_t$ and N_t.

The mean time to extinction for non-viable populations has a complex relationship with the proportion of male gametes available for selfing α_0 (Figure 3.4B). The general patterns that are observed, however, are that the effect of α_0 on the time to extinction is attenuated with increasing values of U. Increasing selection (s and h) reduces the mean time to extinction for populations with low proportions of offspring produced via self-fertilisation, but increases it for higher α_0 (Figure 3.4B). The contrary is observed when mutations are very mildly deleterious ($s = 0.02$) and recessive ($h = 0$ and 0.02), with longer times to extinction for outcrossing populations than for selfing populations. The pattern of the effect of α_0 and selection on time to extinction remains unchanged when recombination rates are low.

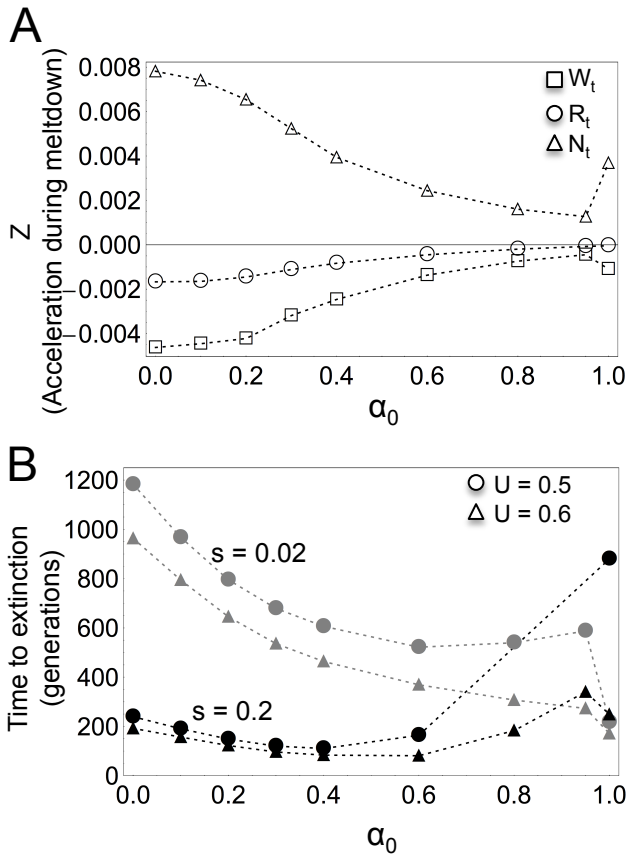


Figure 3.4: A) Median value of acceleration in the rate of decrease of $\overline{W}_t$, $\overline{R}_t$ and N_t, noted z, for non-viable populations from $N_t = 250$ until extinction as a function of α_0. $U = 0.6$, $s = 0.02$, $h = 0.02$, $D = 10$, $r_0 = 2$ and $K = 10000$. B) Mean time to extinction as a function of α_0. $D = 10$, $r_0 = 2$ and $K = 10000$.

3.4 Discussion

It is generally accepted that selection is less effective in small populations, which could lead to their extinction due to mutational meltdown (Lynch and Gabriel, 1990; Lande, 1994; Lynch et al, 1995; Coron et al, 2013), whereas large populations are able to purge recurrent deleterious mutations and remain at mutation-selection balance (Wright, 1931; Hedrick, 2004). Our results suggest that there are values of genetic parameters for which even large populations cannot purge deleterious mutations fast enough to reach mutation-selection balance and go to extinction due to the increasing number of segregating mutations, which in turn increase the mutational load. This implies two things: 1) Mutation selection balance is not attainable for all genetic parameters as hypothesized by population genetics models and 2) Populations can go to rapid extinction due to segregating deleterious mutations. Self-fertilisation, while expected to allow for more efficient purging of deleterious mutations (Glémin, 2003), does not always allow for lower probability of extinction. Population fitness as well as the amplitude of the fluctuation of population size both contribute to the eventual fate of a population, with lower population fitness and greater fluctuations leading to higher probabilities of extinction. More specifically, our results show that there is a synergistic interaction between genetics and demography, which increases the stochastic fluctuations of population size.

Empirical estimations of the genetic parameters used in our model are now becoming available. The idea that there is a strong correlation between h and s has often been shown in empirical works (Simmons and Crow 1977; Steinmetz et al 2002; Agrawal and Whitlock 2011; also see Phadnis and Fry 2005). As the dominance coefficient has been estimated as being around 0.25 (Manna et al, 2011) and most of the deleterious mutations that make up an individual's mutational load are of small effect, this implies that the most realistic parameters we have run are for $h = 0.2$ and $s = 0.02$. However, a new approach using a phenotypic landscape model has shed doubt on this interpretation; the dominance and the coefficient of selection of mutations might well be independent of one another (Manna et al, 2011). This justifies our choice to study several values and combinations of these values of s and h values. We have chosen to consider mutations with constant and only deleterious effects, which is one of the limits of this model. Mutations found in natural populations have variable coefficients of selection s and dominance h and the distributions of these variables are still under debate as they can vary not only between species (Eyre-Walker and Keightley, 2007), but also between populations (Keightley and Halligan, 2009). How the variability of mutation affects the evolution of populations is still unclear and yet to be

fully taken into account in theoretical models.

The genomic mutation rates (equivalent to $2U$ in our model), estimated empirically in various organisms range from 0.01 to 1 (Sniegowski et al, 2000) and even greater (Agrawal and Whitlock, 2012) in eukaryotes. We therefore explore realistic mutation rates, even though in our case we consider that all new mutations are deleterious.

3.4.1 Population size, viability and the mutational meltdown

Generally, if we are able to predict demographic factors (such as the intrinsic reproductive rate, the carrying capacity and density dependence) and the mean relative fitness of populations (or the mutational load), we are able to predict population size and a threshold value of mean relative fitness, below which a population is not viable (Figure 3.1A). However, predicting a relatively large (or non-null) population size is not enough to guarantee population survival within a relatively long time scale. In our model we find that population extinction is not due to demographic stochasticity alone, but to increased levels of stochasticity that result from the interaction between demography and genetics.

The importance of this interaction on population extinction has already been suggested in literature (Lande, 1988). Our results find that it is indeed non-negligible as shown by the fluctuation of population size over time in our model, which is a result of demographic stochasticity and a dynamic mutation load or relative fitness, which are not independent of one another (Figure 3.1D). The mean fluctuation of the population size for a given set of parameter values (open circles) are too great to be due to demographic stochasticity alone (full circles). Even though stochastic processes affect the reproductive rate, contemporary stochasticity alone does not account for this variance. Past stochastic events, or the mutational history of the population (where the deleterious mutations are in the genome, and at what frequencies), can also influence this variable, as observed in Figure 3.1D, where the standard deviation of population size σ_N (grey points) varies greatly around the mean standard deviation observed $\widehat{\sigma}_N$ (open circles). For parameter values where the probability of extinction is different than 0 and 1, the importance of past stochastic events, is even more evident as the fate of a population is not sealed. The importance of past mutational events has also been observed in experimental mutation accumulation lines, where replicate populations with the same genetic origin do not all go to extinction (e.g. for yeast populations (Zeyl et al, 2001)). We propose, that in order to predict the probability of

extinction of populations, it is not only important to predict the mean fitness, but also the fluctuation of population size. Further exploration of this model is required to estimate how the genetic and demographic parameters affect the amplitude of the fluctuations of population size. No theoretical work, to our knowledge, has taken on this question from a demo-genetic point of view.

Clarke's (1973) work highlights the importance of the effect of selection on demography, and his prediction that taking into account the accumulation of deleterious mutations throughout the genome would allow a significant decrease of population size is confirmed by our model, where in some cases populations go to extinction. It has been suggested that the timing of selection is crucial in order to assess the effect of the mutation load on population size (Clarke, 1973; Agrawal and Whitlock, 2012). In our model we have chosen relative fitness to affect only the reproductive rate. However, it is possible that selection that affects the consumption of resources (K) could lead to different results (Clarke, 1973). It is often considered that selection has no effect on demography (e.g. Agrawal and Whitlock, 2012), but if the mutation load has a direct effect on an individual's reproductive capacity, as is the case in our model, then the effect of selection on population size cannot be ignored. This has important implications on how data on population size from natural populations should be interpreted (see below).

The importance of the reproductive rate concerning population extinction has already been suggested by other models, where populations with higher intrinsic reproductive rates have longer times to extinction (Lynch et al, 1995; Robert et al, 2002) or lower probabilities of extinction (Bernardes, 1995). These predictions are an inherent property of our model, as populations with high intrinsic reproductive rates r_0 are expected to be viable at higher mutational loads (Equation 3.6), and are supported by our stochastic simulations (Figure 3.1B). In our model, we consider a stable environment, which is an unrealistic hypothesis. It is therefore probable that we overestimate population viability, as shown by Higgins and Lynch's model (Higgins and Lynch, 2001), which, upon taking environmental stochasticity into account, suggests that it increases the rate of accumulation of deleterious mutations.

To what extent are we capable of estimating the mean relative fitness (or the mutational load)? In spite of explicitly considering demography, we find that the simplified deterministic predictions of the mutation load are reliable when mutations have a strong effect ($s = 0.2$ and $h = 0.2$). However, when mutations are almost neutral, mean fitness is overestimated, especially when the mutations are recessive ($h = 0$) and the genomic recombination rate is low as the number of mutations per chromosome is increased (Kondrashov, 1982), affecting the purging process. The deterministic expectations of mean

fitness when comparing them to simulated results have already been shown to be reliable by population genetics models (Bataillon and Kirkpatrick, 2000; Haag and Roze, 2007), however none of these models explicitly include the effect of demography. This interaction, between genetics and demography, could in fact be important, as shown by the unexpected non-monotonic relationship between population fitness and the proportion of selfed offspring in our model (Figure 3.2B), which we discuss below.

It is important to note that a high mutation rate and a large number of segregating deleterious mutations do not necessarily lead to a higher mutation load. In the case of very little recombination ($D = 0.1$), increasing the mutation rate increases mean fitness when mutations are recessive and of small effect, and hence decreases the mutation load. Therefore, the mutation rate in itself is not sufficient, and depending on the effects of deleterious mutations and the recombination rates, increasing the genomic mutation rate does not always lead to an increase in the mutational load. The extreme case of very little recombination could also be translated as the existence of genomic regions with low recombination rates known as cold-spots (reviewed in Petes, 2001), allowing for the accumulation of deleterious mutations (Charlesworth and Charlesworth, 2010, p. 555). The existence of such genomic regions could in fact have an important influence on the survival of populations. Low rates of recombination are expected to reduce population fitness and increase the rate of fixation of deleterious mutations (Charlesworth et al, 1993). Though our results confirm this for mutations that are moderately recessive ($h = 0.2$), it is not the case when considering recessive mutations, where the contrary is observed. At such low recombination rates, the high levels of linkage-disequilibrium lead to the formation of tightly linked groups of mutations. These groups of mutations act as a single "super locus". When the mutations are recessive, this load remains silent at the heterozygous state, but once at the homozygous state, the multiplicative effects of these small mutations are expressed and lead to a very deleterious effect. The relative fitness of individuals that become homozygous for only one of these different super loci is extremely low. In this case, the advantage of outcrossing is much higher, as outcrossed offspring have a higher probability of being heterozygous at these loci than selfed offspring.

From our simulated results, we conclude that when populations are on their way to extinction, whether we observe a mutational meltdown depends not only on the values of the genetic parameters, but also on the variable measured. Due to the nature of the density dependence in this model, the decrease in population size decelerates when population density is low: the smaller the population size, the more resources available to the few remaining individuals. Even though both the reproductive rate and the mean relative

fitness do show an acceleration in their decrease when populations are on their way to extinction, we find that the existence of this mutational meltdown depends greatly on the effect of the deleterious mutations, the mutation rate, and the proportion of offspring produced via self-fertilisation.

The importance of segregating mutations has already been suggested by Lynch et al (1995), but this work concluded that fixation had a greater effect on the meltdown. Our model suggests that the fixation of deleterious mutations is a consequence rather than a cause of decline towards population extinction, this however could be due to the fact that in Lynch et al's (1995) model the genetic load affected offspring survival, whereas in our model there is a direct link between the reproductive rate and the mutational load. The effect of the accumulation of segregating deleterious mutations has been considered to be negligible when considering large populations (Hedrick, 2004; Whitlock M.C., 2004), even more so when considering their eventual extinction over a short time scale because of this process (but see Bernardes, 1995). This does not seem to be the case when considering the interaction between demography and genetics. A feed-back loop between these two properties seems to lead to a "cost of purging" : Unfit individuals do not reproduce, decreasing population size at the next generation, which in turn increases the effect of drift, leading to a lower efficiency of purging and more unfit individuals. This continual increase of the number of segregating deleterious mutations eventually leads to a demo-genetic extinction. Mutation-selection balance is therefore not the automatic fate of initially large populations, and the cost of purging can lead to a very rapid extinction (Bernardes, 1995).

3.4.2 How does selfing affect population size and viability?

Our results indicate that selfing has an effect both on population size and viability. We often observe that selfing populations have lower probability of extinction than outcrossing populations at higher mutation rates (see Figure 3.2A), especially when there is strong selection, in which case selfers are always expected to have larger population sizes. However, when selection is weak, we find that strict-selfing and low levels of selfing (but not strictly outcrossing) hinder both size and viability. As this has not been noted in other models, even when genetic drift is taken into account (Bataillon and Kirkpatrick, 2000; Roze and Rousset, 2004), it is possibly a consequence of the interaction between genetics and demography. A possible hypothesis to explain this observation is that the viability of populations concerning the accumulation of deleterious mutations depends on two opposing properties

1) The purge of these mutations and 2) The avoidance of expressing them. Outcrossers accumulate deleterious mutations (Glémin, 2003), but avoid the cost of inbreeding depression (Charlesworth and Charlesworth, 1987), with most of their mutations being at a heterozygous state. Selfers purge these mutations (Glémin, 2003) and even though they do not accumulate as many as outcrossers do, many are at a homozygous state (Charlesworth, 2003). We therefore propose that populations with low proportions of offspring produced via self-fertilisation suffer from both the inconveniences of outcrossers and selfers, not only do they accumulate deleterious mutations, as purging is not as efficient as for high proportions of selfed offspring, they also express them, and suffer from the demographic cost of purging.

The effect of self-fertilisation on the extinction of populations due to the accumulation of deleterious mutations has long been debated (Stebbins, 1957; Lynch et al, 1995; Takebayashi and Morrell, 2001). Our results suggest that the accumulation, hence fixation, of deleterious mutations is probably an insufficient explanation for higher extinction rates. In spite of this model's limitations, we find that even though self-fertilisation does affect population extinction due to genetic deterioration, the effects of the genetic parameters are complex and do not result in a simple pattern of the effect of selfing on the time to extinction. When selection is weak, strict outcrossers are less likely to go extinct than strict selfers, agreeing with Lynch et al's (1995) results. However, allowing for a small amount of outcrossing (e.g. a proportion of male gametes available for selfing $\alpha_0 = 0.95$) is enough to greatly decrease the probability of extinction, even allowing for a higher probability of population survival at higher mutation rates than for strict outcrossing (see Figure 3.2A). Strong selection reverses this observation, with strict outcrossers being more prone to a mutational meltdown than selfers, in accordance with Bernardes's (1995) results. What rate of selfing is more likely to cause extinction or lead to population vulnerability is not clear and greatly depends on both the genetic (mutation rates, genomic recombination rates, deleterious effects of mutations) and demographic (intrinsic reproductive rate) parameters. It has already been suggested that selfing has a greater effect on population extinction when considering the fate of beneficial mutations and their effects after environmental change. For instance Glémin and Ronfort (2013) showed that if adaptation is due to standing variation, then outcrossers are less prone to extinction than selfers. Their model, however, is not demographically explicit.

When considering genomic cold-spots with low recombination rates, outcrossers are greatly advantaged when mutations are recessive, as they do not express these accumulated mutations. The lower viability of selfing populations in our results for such low recombination rates supports the observation

that selfing species could be more likely to evolve higher recombination rates, in order to avoid the hitchhiking of deleterious mutations (Roze and Lenormand, 2005). The difference in extinction rates between outcrossers and selfers observed empirically (Goldberg et al, 2010) could perhaps be due to such genomic regions where mutations of small effect segregate.

The non-monotonic effect of selfing on the probability of extinction in our results, could offer a possible explanation to the differences in extinction rates between selfers and outcrossers within the same family. For a transition to be successful, the transition in the reproductive mode has to be of large effect, going from complete outcrossing to high proportions of selfed offspring, for in some cases the mutational load that an outcrossing population could put up with could prove lethal for a reproductive mode with low proportions of offspring produced via self-fertilisation. The observed high extinction rates could therefore be related to the transition process and not the reproductive mode in itself.

3.4.3 Empirical implications

How can the correlation between population size and population fitness be interpreted? In most empirical works, a positive correlation between the two is translated as the negative effect of a small population size on population fitness due to inbreeding, the fixation of deleterious mutations or a lack of reproductive assurance (for example Fischer and Matthies, 1998; Dostalek et al, 2010). Another possible interpretation which is not often considered is simply that population size is a consequence and not the initial cause of a high mutation load, just as in some cases small population size does not seem to lead to a decline in fitness (Costin et al, 2001).

Generally, small populations are considered to be most at risk of extinction within a relatively short time frame due to inbreeding depression, mutational meltdown and demographic stochasticity (Lande, 1988; Hedrick, 2004; Frankham, 2005). Empirical experiments have therefore concentrated on the extinction of small populations, through the accumulation of deleterious mutations (Newman and Pilson, 1997; Frankham et al, 2001; Zeyl et al, 2001; Vassilieva and Lynch, 1999; Baer et al, 2005). Even though the fixation of deleterious mutations can lead to the mutational meltdown of small populations (Lande, 1994; Lynch et al, 1995; Coron et al, 2013), our results suggest that the interaction between demography and genetics can lead to the extinction in large populations due to segregating mutations alone and at relatively fast rates. In initially large populations, once the "mutational meltdown" is underway, fixation is rare and is a consequence rather than a cause of population decline. The importance of segregating mutations in population decline

could have implications in conservation biology, as in most empirical studies, it is automatically assumed that the load leading to population decline is fixed or almost fixed (Fredrickson et al, 2007; Heber et al, 2013). If population decline is mostly due to segregating mutations, then there exists a real potential of purging the deleterious load through conservation efforts. Small populations are also expected to be more prone to demographic stochasticity, which should act more rapidly on population extinction than genetic factors (Lande, 1988; Frankham, 2005), as has been shown through empirical experiments (Wootton and Pfister, 2013). The higher levels of stochasticity in the variation of population size observed due to this interaction compared to the expected effect of demographic stochasticity alone in our results (Figure 3.1D) indicate that stochastic events (that are not due to external factors such as environmental stochasticity) are not only detrimental in small populations, but can also be so in large populations.

Our results suggest that measuring the decline in population size could be misleading when attempting to asses whether a population is going into a mutational meltdown or not. Depending on the density dependence, a population on the way to extinction could seem to be demographically stable, as the decrease of population size could potentially decelerate with time, becoming barely detectable. As the mutational load is not accessible, measuring the acceleration of the decline of the mean relative fitness is not possible in natural populations. However, measuring the acceleration in the decline of the mean reproductive rate over several generations could be a more informative measure concerning population extinction, or the mutational meltdown, and is empirically more accessible. The lack or rarity of a mutational vortex when our simulated populations are on the decline could indicate that, if the segregating deleterious mutations can be purged at any time, then, as there is very little or no increase in the rate of reduction of population size, conservation efforts could be applied successfully even when populations reach relatively small sizes.

In conservation biology, Population Viability Analyses (PVA) are the most frequently used tool for estimating the probability of population extinction. PVAs take mostly demographic data and parameters into consideration and do not take into account genetics explicitly. They have proved useful and accurate when considering external pressures (i.e. over-fishing, fragmentation, etc.) that affect population demography. However, the effectiveness of PVA's remains ambiguous, as even though they can be relatively accurate predictors of the evolution of population demography (Schoedelbauerova et al, 2010), in other cases the population growth rates can be over-estimated (Bell et al, 2013). Could the overestimation of growth rates of growth rates be due to the omission of the genetic effects on population

demography?

The direct relationship between higher intrinsic reproductive rates and greater population sizes for the same relative fitness in our model (Equation 3.3, Figure 3.1B) leads to a relationship between the reproductive and mutational rates in viable populations (Equation 3.6). This could mean that for a population with a high genomic mutation rate to be viable, it must have a large enough intrinsic reproductive rate. This relationship has not been studied either theoretically or empirically, and it could be interesting to perform a comparative analysis between species to test if there is such a relationship. We suggest that there could be a correlation between the genomic mutation rates of a species and the number of gametes produced, which could represent the intrinsic reproductive capacity.

Acknowledgments

Numerical results presented in this paper were carried out using the regional computational cluster supported by Université Lille 1, CPER Nord-Pas-de-Calais/FEDER, France Grille, CNRS and on the Biomed virtual organization of the European Grid Infrastructure (http://www.egi.eu) via DIRAC (http://diracgrid.org). We thank the technical staff of the CRI-Lille 1 centre, the European Grid Infrastructure and the supporting National Grid Initiatives for providing the technical support and infrastructure. We would also like to thank Denis Roze, Sylvain Glémin, Xavier Vekemans and Luis-Miguel Chevin for their input and helpful comments.

Act 4

The evolution of self-fertilisation

The demographic consequences of the transition from outcrossing to self-fertilisation.

Diala Abu Awad[1,*], Sylvain Billiard[1],
1 UMR-CNRS 8198, Laboratoire Génétique et Évolution des Populations Végétales, Université Lille 1, Villeneuve d'Ascq, France
∗ **E-mail: Corresponding diala.abu-awad@univ-lille1.fr**

4.1 Abstract

The prevalence of outcrossing in the plant kingdom in spite of numerous transitions from strictly outcrossing to self-fertilising reproductive systems remains poorly understood. It has been suggested that self-fertilising species have higher extinction rates, thus giving outcrossers a long-term advantage. Whether higher extinction rates are due to self-fertilisation being an evolutionary dead-end or whether they are a consequence of the already marginal nature of the populations in which transitions tend to be common is unknown. Using an individual-centred model in which both population size and the genetic load co-evolve we test the conditions for which the transition to self-fertilisation is responsible for population extinction due to the accumulation of deleterious mutations. We find that when mutations have a small deleterious effect, populations sometimes evolve to reproduce strictly by self-fertilisation, leading to the accumulation of deleterious mutations and extinction.

4.2 Introduction

One of the most frequently observed evolutionary transitions is that of the evolution of self-fertilisation (selfing) in primarily outcrossing plant families (Stebbins, 1957). Self-fertilization provides two main advantages 1) The demographic advantage of reproductive assurance which facilitates colonisation as pollen limitation is eliminated (Baker, 1955) and 2) The genetic advantage of Fisher's automatic advantage or the transmission of two copies of an individual's genome to its offspring, against one copy if the individual does not self-fertilize (Fisher, 1941). Theoretical models studying the question have shown that Fisher's automatic advantage can be countered by high levels of inbreeding depression, where outcrossing is expected to be maintained if selfed progeny are at least 50% less fit than those produced by outcrossing (Lande and Schemske, 1985; Charlesworth and Charlesworth, 1987; Charlesworth et al, 1990; Porcher and Lande, 2005). However, there is recent empirical evidence that levels of inbreeding depression in outcrossed populations are below the limit for the maintenance of outcrossing (Winn et al, 2011), and as the transition is frequent and seemingly uni-directional (Igic et al. 2004), it is paradoxical that only 11% of flowering plants are predominantly selfing (Wright et al, 2013).

In order to explain the prevalence of outcrossing (in spite of the facility with which selfing can evolve) it has been suggested that selfing lineages within outcrossing families present higher extinction rates (Goldberg et al, 2010), supporting that self-fertilization is an evolutionary dead-end (the "dead-end hypothesis", reviewed in Takebayashi and Morrell, 2001). These higher rates of extinction have been attributed to two phenomena; firstly selfing-species have lower standing genetic variation and hence are expected to have lower adaptive potential Stebbins (1957). Secondly selfing populations accumulate deleterious mutations at the homozygous state that can eventually go to fixation, leading to population extinction via mutational meltdown Lynch et al (1995). Both of these hypotheses can be disputed, as even though selfing does decrease genetic variation, it does however increase the probability of fixing new advantageous mutations and have the same, if not in some cases better, adaptive potential as expected in outcrossing populations (Glémin and Ronfort, 2013). Concerning deleterious mutations, not only can selfing lead to (in general) a more efficient purge of the genetic load (Glémin, 2003), it can also purge inbreeding load during the transition from outcrossing (Lande and Schemske, 1985; Charlesworth et al, 1990; Porcher and Lande, 2005; Gervais et al, 2014) and even lead to lower genetic loads (Charlesworth et al, 1990). As the evolution of selfing has been greatly observed in marginal populations (Barrett, 2010), the already deteriorat-

ing genetic and demographic background in populations where selfing has evolved could be the reason why newly evolved selfing lineages are not long-lived (Wright and Barrett, 2010). Though this explanation is plausible, it would concern only certain populations in given conditions, making it insufficient. There is therefore no clear explanation for the higher extinction rates observed in selfing lineages. Wright and Barrett's (2010) hypothesis does however lead to an important question: what are the potential demographic consequences of a transition from strictly outcrossing to self-fertilising and how do they influence population viability?

Models that have studied the transition from outcrossing to selfing have done so by considering populations of fixed size (Charlesworth et al, 1990; Porcher and Lande, 2005). Although several works have integrated ecological or environmental factors in transition models (Cheptou and Dieckmann, 2002; Porcher and Lande, 2005; Porcher et al, 2009), to our knowledge none have taken into account the interaction between selection and demography. It has been suggested that depending on the intrinsic demographic and genetic properties of a population, when the interaction between genetics and demography is considered explicitly, different rates of self-fertilisation can indeed affect the probability of extinction (Abu Awad et al, 2014). One could therefore argue that the genetic changes that could ensue (Charlesworth and Wright, 2001) also lead to demographic consequences, hence leading to extinction during or after the transition. Here we present an individual based model in which population size evolves as a consequence of the genetic load, and self-fertilizing mutants are introduced into an initially outcrossing population at equilibrium in order to test this hypothesis.

4.3 Model

4.3.1 Deterministic model and expectations

As in (Abu Awad et al, 2014, presented in Chapter 3), we consider a population of N_t hermaphroditic individuals at generation t, where

$$N_t = N_{t-1} R_{t-1}. \tag{4.1}$$

Population size therefore depends on R_t, the absolute multiplicative fitness of a population at a given generation t, with trait value (or relative fitness) W_t, in a population of density N_t (Chevin and Lande, 2010) given by

$$R_t = r_0^{1-\frac{N_t}{K}} W_t. \tag{4.2}$$

The environment is considered to have a constant carrying capacity of K. The intrinsic reproductive rate of such a population is $r_0 \geq 1$ (Chevin and Lande, 2010), and W_t is the mean relative fitness of individuals in the population as a function of their mutation load. Therefore the number of segregating or fixed deleterious mutations in the population can decrease the reproductive rate and eventually population size. In this model, we consider that density dependence affects all individuals in the same way, independently of their relative fitness (or genotype).

At equilibrium (in this case, mutation selection balance) when there is no demographic stochasticity, population size N_{det} is given by

$$N_{det} = K\left(1 + \frac{\mathrm{Ln}(W_{eq})}{\mathrm{Ln}(r_0)}\right). \tag{4.3}$$

This implies that for a population to be viable ($N_{det} > 0$), the mean relative fitness at mutation-selection balance W_{eq} must be greater than the threshold value W_{min} (obtained by solving for $N_{det} = 0$ from equation 4.3), where

$$W_{min} = \frac{1}{r_0}. \tag{4.4}$$

Initially, the population is made of strictly outcrossing individuals. The deterministic mean relative fitness at a very large number of independent loci of the population is approximately e^{-L} (Haldane, 1937), where L is the mutation load. We consider that at $t = 0$ there is a sudden transition of all of the individuals in the population from strictly outcrossing to outcrossing at a rate $1 - \alpha$ and self-fertilising at a rate α. As of this point, we consider that all offspring produced by outcrossing have a fitness $W_o = W_{eq}$, whereas those produced by self-fertilisation will automatically have a mean relative fitness W_s that is decreased by a factor δ, representing the effect of inbreeding depression with

$$\delta = 1 - \frac{W_s}{W_o}. \tag{4.5}$$

W_s can therefore be expressed as a function of W_o, giving $W_s = W_o(1 - \delta)$. The population's new mean fitness is now $W_\alpha = (1 - \alpha)W_{eq} + \alpha W_{eq}(1 - \delta)$ which simplifies to $W_\alpha = W_{eq}(1 - \alpha\delta)$. Demographically speaking, in order for there to be a successful transition from outcrossing to selfing at a given rate α, N_{det} must remain greater than 0, hence W_α should remain greater than W_{min} (equation 4.4). Therefore for a population to successfully evolve a rate of self-fertilisation α, it must have a level of inbreeding depression lower

than δ_{lim} (obtained by solving for $W_\alpha = W_{min}$) given by

$$\delta_{lim} = \frac{r_0 W_{eq} - 1}{r_0 W_{eq} \alpha}. \tag{4.6}$$

There exist two possible scenarios for the evolution of self-fertilisation, either the purge of deleterious mutations is relatively slow compared to the evolution of α (sudden transition) or there is sufficient time for deleterious mutations to be purged during the transition. In the first scenario, W_{eq} in equation δ_{lim} does not change with α, hence the fitness of the population remains fixed as it was before $t = 0$. In the second scenario, a rapid purge implies that the mean fitness of the population changes with the rate of self-fertilisation, increasing as α increases (Roze and Rousset, 2004). In Abu Awad et al (2014), we find that the equations of population fitness provided in Roze and Rousset (2004) are accurate in estimating the mean fitness of simulated populations. If there is no purge, we replace W_{eq} in equation 4.6 with e^{-U} for completely recessive mutations (dominance $h = 0$) and e^{-2U} for all other values of $h >> 0$, where U is the haploid genomic mutation rate. In the second scenario where deleterious mutations are purged, W_{eq} is considered to change with the selfing rate α, but this occurs only when $h \neq 0$ in which case it is replaced by $W_{det} = e^{-U\frac{4h+\alpha(1-4h)}{2h+\alpha(1-2h)}}$ (Roze and Rousset 2004). As fitness increases with α, the value of δ_{lim} greater in the second scenario than in the first, implying that if the deleterious mutations are purged then the evolution of selfing is facilitated demographically speaking as the population will be less likely to go extinct.

4.3.2 Simulation Model

In order to follow the evolution of the selfing rate, we used an individual-centred model with discrete non-overlapping generations. We consider a population of variable size in a stable environment, made of diploid hermaphroditic individuals. As in Abu Awad et al. (2014), at a given time t, population size N_t is given by

$$N_t = \sum_{i=1}^{N_{t-1}} X_{t-1}^i \tag{4.7}$$

where X_t^i is the number of viable offspring an individual i at time t contributes via the female function to the next generation. X_t^i is sampled from a Poisson distribution with mean $R_t^i = W_t^i r_0^{1-\frac{N_t}{K}}$ (the individual reproductive rate), where W_t^i is the individual's relative fitness (see below for more

details) and r_0 is the intrinsic reproductive rate (which is the same for all individuals).

The genetic properties of this model, mutation and recombination are modelled as in (Roze, 2009). We consider that each individual is represented by two homologous chromosomes of length $2D$ with a potentially infinite number of loci. The map length is considered to be D from the centre of the chromosome to the edge, hence representing a chromosome with a defined centromere. Near the centromere, we have included a modifier locus that affects the rate of self-fertilization. The life cycle is as follows: mutation, selection, meiosis and reproduction. Recombination occurs during gamete production and is considered to be uniform along the chromosome. New individuals are a combination of two gametes, either from two different individuals for reproduction via outcrossing, or the same individual via selfing.

Self-fertilisation

If there is evolution to self-fertilisation, then self-fertilisation occurs at a probability α_t^i and is given by

$$\alpha_t^i = \frac{\alpha_0^i W_t^i}{\alpha_0^i W_t^i + \frac{\sum_{j \neq i} (1-\alpha_0^j) W_t^j}{N_t - 1}}, \tag{4.8}$$

where α_0^i is the proportion of an individual i's male gametes that are available for self-fertilization. As the probability of self-fertilization depends on the individual's relative fitness W_t^i compared to the average relative fitness of the other possible fathers in the population. The lower an individual's relative fitness as a father, the lower the proportion of offspring produced via selfing. There is no limitation in the availability of male gametes. The proportion of an individual's offspring produced by self-fertilisation is sampled from a binomial distribution with parameters α_t^i and X_t^i. At the beginning of all simulations, $\alpha_0^i = 0$ for all individuals in the population as the population is strictly outcrossing (the population is automatically considered non-viable if $N_t < 2$).

Mutations

In this model we introduce two types of mutations: mutations at the modifier locus that influence the rate of self-fertilisation by modifying α_0^i and deleterious mutations that appear randomly throughout the chromosome but not on the modifier locus. The number of new deleterious mutations occurring

per chromosome per generation, is sampled from a Poisson distribution with mean U, where $2U$ is the genomic mutation rate. Their position on the chromosome is sampled from a uniform distribution in $[-D, D]$.

As little is known of the mutations causing self-fertilisation arising in natural populations we explore two types of mutations at the modifier locus. The first type of mutation is the introduction of mutants with a fixed rate of self-fertilisation d defined at the beginning of the simulation, *i.e.* there is only one type of self-fertilising mutant and if it invades the population then the population's mean self-fertilisation rate is d. The second type of mutation involves the continuous evolution of α_0, with at every generation both self-fertilizing and outcrossing individuals can mutate to have a new value of α_0 that is sampled from a uniform distribution between $-d$ and d around the initial value, but the new value is bounded between 0 and 1. Therefore the second type of mutation leads to a gradual evolution of selfing with mutations of either large or small effects defining the rate of selfing (depending on the value of d) until the stabilization of the population's mean selfing rate, which can go as high as 1. We consider that the alleles at the modifier locus are co-dominant.

In order to facilitate the reading of the following sections, the notation used through the text has been summarized in Table 4.3.1.

Initial conditions and simulations run

At the beginning of each simulation, we consider that the population is strictly outcrossing ($\alpha_0 = 0$) with size K, no deleterious mutations present and all individuals have the same identity at the modifier locus. At first, only deleterious mutations are introduced. The simulations are run until the population's mean relative fitness $\overline{W}_t$ reaches equilibrium or goes extinct. We define equilibrium as the stabilization of $\overline{W}_t$, when the average $\overline{W}_t$ over one thousand generations varies les than 1 per cent compared to the average W_t of the previous thousand generations. Once at equilibrium, mutation at the modifier locus is introduced and the simulations are run until the stabilization of the mean rate of self-fertilization (detected using the same method as for the relative fitness) or the population goes extinct. We keep track of the frequency of all of the self-fertilising alleles in the population, independently of the rate of self-fertilisation they induce. The level of inbreeding depression at generation t is given by $\delta_t = 1 - \frac{W_s}{W_o}$, where W_s is the fitness of offspring produced by self-fertilization and W_o is the fitness of offspring produced though outcrossing. Every generation, 200 individuals are sampled and the population's $\overline{\delta}_t$ is estimated by producing hypothetical offspring through selfing and outcrossing. If population size N_t is smaller

than 200 individuals, then the entire population is sampled.

Mean fitness $\overline{W}_{eq}$ and the level of inbreeding depression $\overline{\delta}_{eq}$ are measured at the fitness equilibrium before the introduction of mutations at the modifier locus. After the introduction of mutations at the modifier locus, once the mean rate of self-fertilization $\overline{\alpha_0}$ is stabilized, the population's new mean fitness $\overline{W}_{tr}$ and level of inbreeding depression $\overline{\delta}_{tr}$ are measured and compared to those obtained at equilibrium before the introduction of mutations at the modifier locus.

We ran simulations for an intrinsic reproductive rate $r_0 = 2$, 4 and 10, with a haploid genomic rate U between 0.1 and 0.3 for $r_0 = 2$ (as for higher values of U all populations went extinct) and U between 0.1 and 0.6 for the other values of r_0. Mutations were either mildly deleterious or moderately deleterious (coefficient of selection $s = 0.02$ and 0.2), that were completely recessive or moderately recessive (dominance $h = 0$ and 0.2). We considered a recombination rate $D = 10$ as D over 10 has very little effect on the results, which allows us to make the assumption that the mutations act as though they were independent (Roze, 2012). However, it is possible, when there is selfing, that there is some linkage due to the genomic consequences of self-fertilization. Mutations at the modifier locus were either of fixed effect $d = 0.3$, 0.6 and 0.9, or sampled around the initial value for $d = 0.1$ and 0.5. One thousand replicates were run for each group of parameter values, coming to a total of 300000 simulations run.

We compare the values of the mean levels of inbreeding depression $\widehat{\delta}_{eq}$ and $\widehat{\delta}_{tr}$ across all the simulations run in order to determine how the level of inbreeding depression evolves after the transition to selfing. We also compare these values of inbreeding from our simulations to our deterministic limit of inbreeding depression δ_{lim} to test our deterministic predictions for the maximal rate of self-fertilisation. The notation used throughout the Results and Discussion is resumed in Table 4.3.1.

Table 4.1: **Notation.**

$V, \overline{V}, \widehat{V}$	No superscript indicates deterministic values (except in the case of population size N, where it is mentioned clearly in the text), a bar indicates that the variable is the intra-population mean for one simulation run ($\overline{W}_{eq}$ is the mean genetic component of fitness for one population) and a hat indicates that the variable is the mean across all simulations ($\widehat{W}_{eq}$ is the mean genetic component of fitness across all simulations, conditional to survival).
$\widehat{N}_t$	Mean population size at generation t across all simulations conditional to survival.
r_0	The intrinsic reproductive rate.
$\widehat{W}_{eq}, \widehat{W}_{tr}$	Mean population fitness at population equilibrium before the introduction of mutants at the modifier locus and at equilibrium after their introduction across all simulations conditional to survival.
W_{min}	The threshold value of the population's genetic component of fitness, under which populations are not viable.
$\widehat{\delta}_{eq}, \widehat{\delta}_{tr}$	Mean level of inbreeding depression at population equilibrium before the introduction of mutants at the modifier locus and at equilibrium after across all simulations conditional to survival.
δ_{lim}	The threshold value of inbreeding depression (with or without the purge of deleterious mutations affecting fitness) providing the limit of the rate of self-fertilisation above which populations are not viable.
α_0	The proportion of male gametes available for selfing. At the beginning of each simulation it is set at 0.
$\widehat{\alpha_{tr}}$	The proportion of male gametes available for selfing at equilibrium after the introduction of mutations at the modifier locus across all simulations conditional to survival.
U, D	Genomic properties: the haploid mutation rate and the recombination rate or map length.
s, h	Mutational effects of deleterious mutations: the selection coefficient and the dominance.
d	The effect of the mutations at the modifier locus on the proportion of available mate gametes for self-fertilisation. We consider two kinds of mutations at the modifier locus, either new alleles at the modifier locus will have a value d of α_0 or new values of α_0 are sampled from a uniform distribution between $-d$ and d around the initial value.

4.4 Results

4.4.1 Initial levels of inbreeding depression and allelic frequencies at the modifier locus

At equilibrium, before the introduction of mutations at the modifier locus, the observed levels of inbreeding depression $\widehat{\delta}_{eq}$ increase with the mutation rate U and the coefficient of selection s, but decrease with the dominance h (Figure 4.1a). The intrinsic reproductive rate r_0 influences $\widehat{\delta}_{eq}$ only when there are very large differences in population sizes between the different values of r_0. This occurs when $s = 0.02$, $h = 0.2$ for $U = 0.3$ when $r_0 = 2$ (none of the simulations run for $r_0 = 2$ with $U > 0.3$ were viable) and for $U = 0.6$ when $r_0 = 4$. In these two cases, increasing r_0 leads to a higher $\widehat{\delta}_{eq}$ (results not shown). As of this point, these populations will be referred to as "marginal" as they have fitnesses close to W_{min} (equation 4.4) and are hence at risk of going extinct.

The initial levels of inbreeding depression provide some information on the expected frequencies of mutant alleles at the modifier locus (Figure 4.1b). For very high levels of initial inbreeding depression (≥ 0.8), which occurs when deleterious mutations are completely recessive ($h = 0$), the frequencies of mutant alleles at the modifier locus remain low. This is true for both types of mutations at the modifier locus and independently of the value of d (results not shown). The highest $\widehat{\delta}_{eq}$ for which there is fixation is 0.547, however having a $\widehat{\delta}_{eq}$ lower than this value does not guarantee that the population will be invaded by alleles at the modifier locus increasing self-fertilisation. Whether fixation is possible or not and is associated with the "marginal" nature of the population, and hence depends greatly on the coefficient of selection s. Although most simulations run with "marginal" populations end, when viable, with low to intermediate frequencies of the mutant modifier alleles, we find that within a similar time frame there are a few cases of fixation. This could be attributed to the stochastic nature of the mutational history (and allele associations within the genome) within populations. In these conditions, more often than not, a mutant increasing the rate of self-fertilisation is associated with a genome carrying deleterious mutations that are difficult to purge (discussed below), making the fixation of this mutant a rare occurrence.

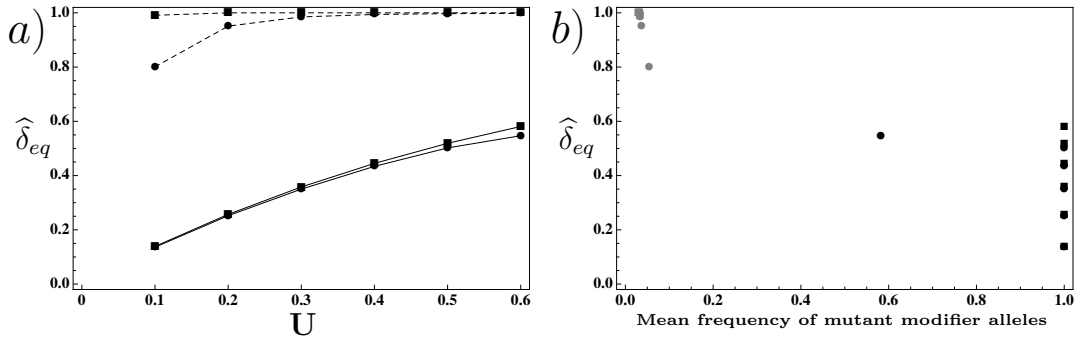


Figure 4.1: *a)* $\widehat{\delta}_{eq}$ (conditional to population survival) as a function of the mutation rate U for $r_0 = 4$. The full lines represent $h = 0.2$ and the dashed lines $h = 0$. The circles are for $s = 0.02$ and the squares for $s = 0.2$. *b)* The mean frequency of mutant modifier alleles at equilibrium conditional to survival as a function of δ_{eq} for mutations at the modifier locus with a fixed value of $d = 0.3$ for $r_0 = 4$. The black plot markers represent $h = 0.2$ and the grey ones $h = 0$. The circles are for $s = 0.02$ and the squares for $s = 0.2$.

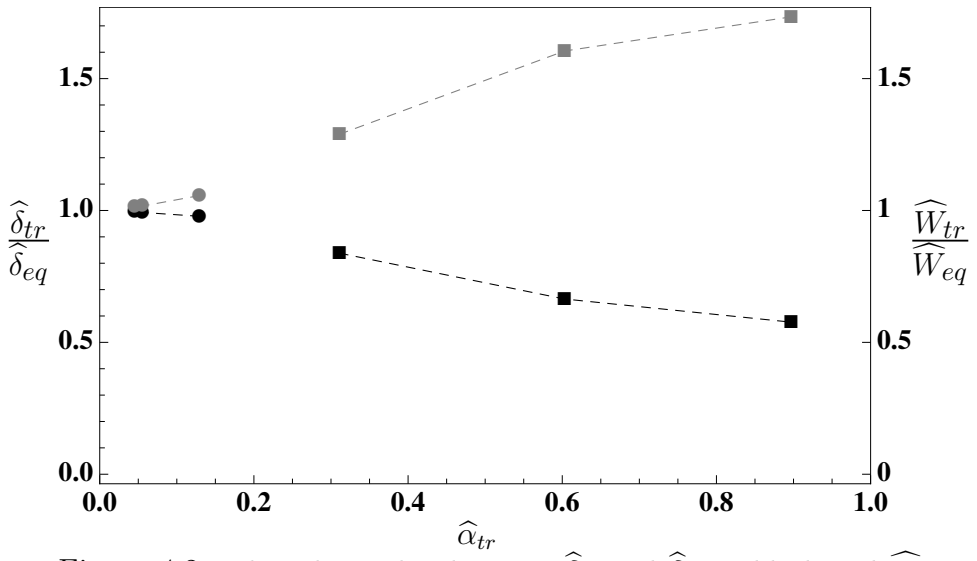


Figure 4.2: The relationship between $\widehat{\delta}_{tr}$ and $\widehat{\delta}_{eq}$ in black and $\widehat{W}_{tr}$ and $\widehat{W}_{eq}$ in grey as a function of $\widehat{\alpha}_{tr}$ for $U = 0.6$, $h = 0.2$, $r_0 = 4$ and the transition regime with fixed modifier effects $d = 0.3$, 0.6 and 0.9. $\widehat{\alpha}_{tr}$ differs for $s = 0.02$ and $s = 0.2$, but the values of d are always in ascending order from left to right for each value of s. The circles are for $s = 0.02$ and the squares for $s = 0.2$.

4.4.2 Purging deleterious mutations and the evolution of self-fertilisation in viable populations

When populations are viable once self-fertilisation evolves we find that in most cases the mean levels of inbreeding depression decrease and mean population fitnesses $\widehat{W}_{eq}$ increase (see Figure 4.2). These changes are due to the purge of deleterious mutations brought on by the increased rates of self-fertilisation. In the cases where deleterious mutations cannot be purged, as in Figure 4.2 where the ratios of $\widehat{\delta}_{tr}$ and $\widehat{\delta}_{eq}$, as well as those between $\widehat{W}_{tr}$ and $\widehat{W}_{eq}$, are almost equal to 1, the mean rate of self-fertilisation remains relatively small. In Figure 4.3 for parameter sets that allow little or no purge (circles) the rates of self-fertilisation are well below those predicted by δ_{lim} (Equation 4.3). This implies that although demographically the population could survive the evolution of higher rates of self-fertilisation (as N_{det} would remain greater than 0, see equation 4.3), the increase in the frequency of self-fertilising mutants is constrained by the cost of inbreeding depression. When considering that the purge of deleterious mutations occurs rapidly relative to the evolution of self-fertilisation we find that the demographic constraints are small compared to when there is no purge (see full black line in Figure 4.3) and always predict the evolution of high self-fertilisation rates for the observed levels of inbreeding depression.

When the mutant alleles at the modifier locus invade the population, we find that the rate of self-fertilisation is equivalent to the proportion of self-pollen available, the mean α_0 of the population. Therefore when the mutants at the modifier locus have a fixed d, the rate of self-fertilisation is equal to d (see Figure 4.3). In simulations where α_0 is allowed to evolve in small steps (as the effect of the alleles at the modifier locus is sampled between $-d$ and d), populations present very high self-fertilisation rates, very close, but never equal, to 1.

4.4.3 Extinction

In the previous sections we addressed the evolution of self-fertilisation in viable populations. When populations go extinct in our simulations it is for one of two reasons, either the initial parameter values were those resulting in "marginal" populations, or populations went extinct via mutational meltdown. In the case of marginal populations, it is not clear whether the introduction of the mutant modifier alleles influences the probability of extinction within a given time frame. There is however a slight indication that this is the case. For marginal populations when $r_0 = 2$ (but not for $r_0 = 4$) there is a clear increase of the proportion of extinct populations with increas-

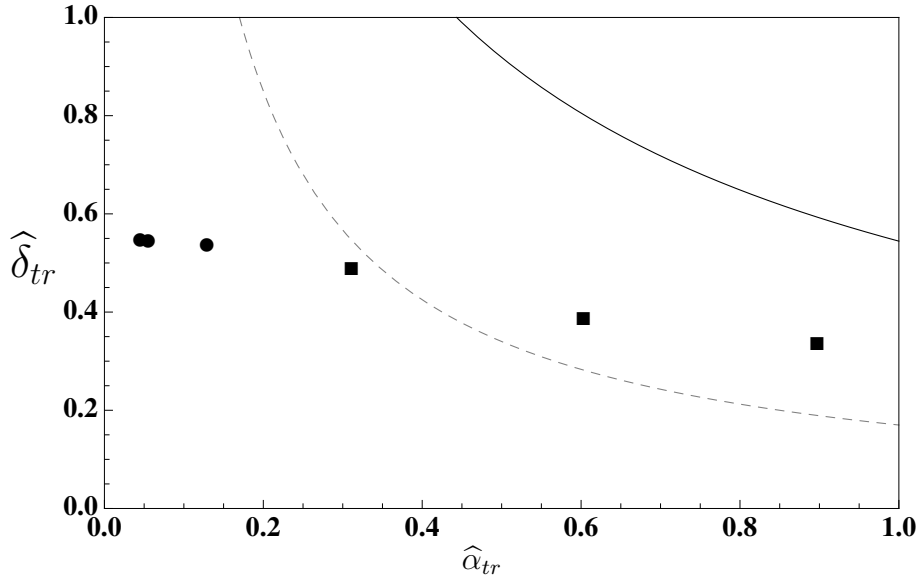


Figure 4.3: δ_{tr} as a function of $\widehat{\alpha}_{tr}$ for $U = 0.6$, $h = 0.2$, $r_0 = 4$ and the transition regime with fixed modifier effects $d = 0.3$, 0.6 and 0.9. The $\widehat{\alpha}_{tr}$ differ for $s = 0.02$ and $s = 0.2$, but the values of d are always in ascending order from left to right for each value of s. The circles are for $s = 0.02$ and the squares for $s = 0.2$. The full line represents δ_{lim} when the purge of deleterious mutations is taken into account, whereas the dashed line represents δ_{lim} with a fixed value of W_{eq} (there is no purge).

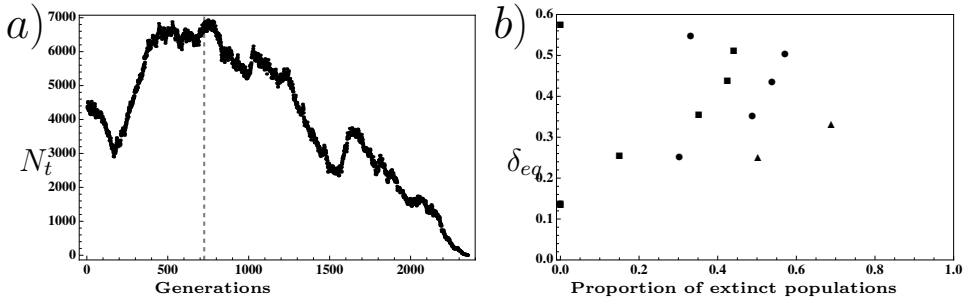


Figure 4.4: *a*) The evolution of population size N_t with time after the introduction of mutations at the modifier locus (second transition regime, $d = 0.5$) for one simulation run with $r_0 = 2$, $U = 0.2$, $s = 0.02$ and $h = 0.2$. The dashed line represents the moment in time when the population becomes strictly self-fertilising. *b*) δ_{eq} (conditional to population survival) as a function of the proportion of extinct populations post the introduction of mutations sampled between $-d$ and d ($d = 0.1$) at the modifier locus for $h = 0.2$ and $s = 0.02$. The triangles, circles and squares represent $r_0 = 2$, 4 and 10 respectively.

ing d when d is a fixed value, rising from 10% to 98% extinction rates for $d = 0.3$ and 0.9 respectively.

The observed mutational meltdowns occur only for $s = 0.02$ and when the effects of mutations at the modifier locus are sampled between $-d$ and d. In these simulations, first there is loss of the initial modifier allele (initially $\alpha_0 = 0$) and then, by chance, the population becomes strictly self-fertilising. This then leads to the fixation and accumulation of deleterious alleles and to the extinction of the population (see Figure 4.4a). Increasing the intrinsic reproductive rate leads to lower rates of extinction due to the transition to a strictly self-fertilising reproductive regime (Figure 4.4b)).

4.5 Discussion

4.5.1 The importance of inbreeding depression and demography in the evolution of self-fertilisation

In this work we have explored how taking into account the interaction between selection and demography influences the evolution of self-fertilisation in an initially outcrossing population. As predicted by previous models, increasing the rate of self-fertilisation leads to a decrease in inbreeding depression due to the purge of deleterious mutations, with greater coefficients of selection s and dominance h leading to a more efficient purge (Charlesworth

et al, 1990; Porcher and Lande, 2013). In Charlesworth et al (1990), the authors proposed that for deleterious mutations of weak effect, if the level of inbreeding depression is greater than 0.5 there would always be selection against a modifier allele only slightly increasing the rate of self-fertilisation, but that a modifier allele greatly increasing self-fertilisation would be selected for. In our results, for $\widehat{\delta}_{eq} \geq 0.5$ the types of mutations introduced seem to have no effect on the probability of the evolution of self-fertilisation. This could be due to the fact that in our model we introduce recurrent mutations at the modifier locus, whereas in Charlesworth et al (1990) mutant alleles at the modifier locus are introduced at a given frequency into the population only once. The recurrent introduction of these mutants could therefore create a gradual purge, eventually leading to favourable conditions for them to increase in frequency.

The introduction of an interaction between population fitness and population size has not greatly influenced the genetic evolution of populations (given that they are viable). We find that our analytical predictions of the rates of self-fertilisation by considering that there are solely demographic constraints to the evolution of selfing (i.e. decreased number of offspring, hence population decline and extinction) do not suffice. When there is no purge of deleterious mutations, the predicted maximal rate of self-fertilisation in Figure 4.3 is at around 0.3. However, none of the types of mutations at the modifier locus lead to a level of self-fertilisation that is that high, generally remaining below 0.1. The importance of introducing demography therefore is not in furthering our understanding of the evolution of the rates of self-fertilisation, but the viability of populations that do.

4.5.2 Evolutionary suicide and extinction of marginal populations

That strictly selfing populations go extinct when mutations are mildly deleterious ($s = 0.02$) was not surprising as this had been predicted in a previous model (Abu Awad et al, 2014). However, as this occurs in simulations where the effect of mutations at the modifier locus that can either increase or decrease self-fertilisation, it is surprising that once the population was on the decline, mutant alleles at the modifier locus decreasing the selfing rate were not selected for. This can be explained by the fact that the levels of inbreeding depression at such high selfing rates are low (see Figure 4.2,Charlesworth et al 1990), and inbreeding depression also continues to decrease with population size (Bataillon and Kirkpatrick, 2000). In this scenario, in spite of the decrease in population fitness due to the accumulation of deleterious mu-

tations, the automatic advantage of self-fertilisation outweighs any potential advantage of outcrossing. The role played by the intrinsic reproductive rate R_0 in decreasing the probability of these populations going extinct is mostly due to higher population sizes for otherwise equal parameter sets. Increased population size decreases the probability that a strictly self-fertilising modifier allele goes to fixation due to genetic drift.

Whether what we have termed to be "marginal" populations have a higher probability of going extinct because of the introduction of mutations at the modifier locus remains unclear and should be tested further. It does however seem plausible that perturbations in the reproductive system could engender a greater genetic (hence demographic) stochasticity, eventually leading to a decrease in fitness below the predicted threshold value W_{min} and extinction.

4.5.3 Self-fertilisation: an evolutionary dead-end?

In a recent work on the phylogeny of the *Solanaceae*, Goldberg et al (2010) came to the conclusion that the net diversification rates of self-fertilising lineages were smaller than those of self-incompatible lineages due to higher extinction rates. They concluded that this observation was due to species selection, self-incompatible species presenting a long-term advantage. Natural selection is short-sighted, a long-term advantage must therefore stem from a short-term advantage. Models on the evolution of self-fertilisation have all come to the same conclusion: in order to maintain outcrossing in the face of recurrent introductions of self-fertilising mutants, a high level of inbreeding depression is necessary. However, these conditions are rarely met in natural populations. What then can explain the prevalence of outcrossing?

The hypothesis that self-fertilisation is a dead end has received considerable attention, but has yet to be proven, as although self-fertilisation can be disadvantageous (by decreasing genetic variability) it also presents both genetic (purging of deleterious alleles) and demographic advantages (reproductive assurance). Our results point to the possibility of an evolutionary suicide that is independent of ecological constraints and due solely to the interaction between population size and deleterious mutations. As most novel deleterious mutations are of small effect (Agrawal and Whitlock 2011) and it is only in cases where mutations were only slightly deleterious ($s = 0.02$) that the mutational meltdown occurs, our results seem biologically plausible.

Act 5

Conclusions and Perspectives

"They both savoured the strange warm glow of being much more ignorant than ordinary people, who were only ignorant of ordinary things."

- Terry Pratchett (Equal Rites, 1987)

5.1 Some demography can go a short (but not negligible) way

Introducing demography into population genetics models is not necessarily an easy task. The number of hours spent eating away at pen and staring at a blank piece of paper, cursing the program you are using for not being able to resolve these equations (even if they are rather complicated) should be enough to put anyone off. However, the results are worth the time spent. With this work we have in some cases validated the hypotheses of existing population genetics models and conditions for which taking demography in to account is not necessary. The work presented in Chapter 2.1 has clarified when both selection and census take place for a give genetic load, something as a masters student had never quite been clear to me. We also have presented comprehensible relationships between the genetic load at a given locus and its effect on population size. Clarke (1973) had also presented such relationships, however in his model fitnesses were density dependent and his genetic load was imposed. I aim to take this model a step further by first integrating several timings of selection, for example for a given genotype there may be better survival at the zygote stage by individuals are less fit when it comes to mating success. The pleiotropy of mutations is well known, and if we consider a mutation that decreases a metabolic function, it is intuitive that its effects may not be constrained to a given moment in the life cycle. Second of all, the preliminary work presented in Chapter 2.2 has aroused my curiosity and I am preparing to integrate both somatic and germ-line mutations in order to test my personal predictions of whether the simultaneous effects of these mutations will be additive or have a more complex relationship when it comes to the genetic load. I will also be expanding this model further to test the timing of selection as in Chapter 2.1.

The multi-locus models presented in Chapters 3 and 4 were a result of endless hours of programming, running simulations, re-programming and re-running all million simulations all over again because of a bug. This work however has comforted me in my initial reaction to the unintuitive existence of populations with extremely low fitnesses. Though one of the drawbacks, in my opinion, is that we have only tested a single demographic scenario, at the same time, it was necessary to keep the simulations as simple as possible so as to avoid the drowning out of the separate parameter effects. The finding of higher mutation rates actually increasing fitness and viability in the case of highly linked mutations giving an advantage to outcrossers was one of the surprising results of this work. From my point of view, this result seems to support why highly self-fertilising species have such small chromo-

somes. If only the chromosomal regions with higher rates of recombination are maintained then extinction due to tightly linked loci is eliminated. Although effective recombination rates are very low in selfers (Charlesworth, 2003), they may be high enough to eliminate a newly introduced deleterious mutation.

The non-monotonic relationship between the rate of self-fertilisation and population fitness observed for recessive mutations was at first quite surprising and unexpected. Though we have attributed this observation to being due to the possible interaction between selection and demography, this does not seem to be the case. During a discussion with Denis Roze he confirmed to me that this observation occurs also in simulations where population size is kept fixed. It is due to the fact that there exists some linkage between the mutations that results from self-fertilisation. However, the observation that populations with intermediate self-fertilisation rates can go extinct more often than either extreme does provide a possible explanation for the observed outcrossing and self-fertilising rates in natural populations (Jarne and Auld, 2006).

After having obtained the results from the model in Chapter 3, I was convinced that during a transition from outcrossing to self-fertilisation, as the self-fertilisation rate evolved, its non-monotonic effect on fitness would lead to it extinction. This however was not the case, a result that to me was rather disappointing. That extinction was mainly due to the off-chance (in some cases it was actually very frequent) self-fertilisation would evolve to 1 and go extinct by mutational meltdown was also surprising as the recurrent mutations at the modifier locus could have prevented the maintenance of such high self-fertilisation rates.

Introducing demography into population genetics models has provided a possible means to evaluate what genetic parameters, such as the effects and rates of introduction of deleterious mutations, are indeed reasonable and how a decrease in the fitness of a given trait could potentially lead to a decrease in population size. Although in most cases presented here, the genetic properties of populations were not very different than those obtained from models with fixed sizes, the exception is the case of selection on zygote survival in Chapter 2.1. The non-monotonic fitness with increasing selection has pointed to a potential relationship between the timing of selection and the observed relationships between dominance and the coefficient of selection. This finding further encourages taking the demographic components into consideration if only to provide a clearer idea of how, for example, the timing of selection can influence the interpretation of results obtained empirically.

Epilogue

"Light thinks it travels faster than anything but it is wrong. No matter how fast light travels it finds the darkness has always got there first, and is waiting for it."

- Terry Pratchett (Reaperman, 1999)

Something that I come across regularly is the look of horror or sheer boredom (in some cases both) of an audience made up of evolutionary biologists and/or ecologists when the current slide being shown is one of a mathematical equation. It has been suggested that this aversion to mathematics is widespread, as works flaunting their mathematical innards (or equations) are shunned (Fawcett and Higginson, 2012). The irony is that most of what we do as evolutionary biologists and ecologists has stemmed from mathematical works, from the Hardy-Weinberg law, to population dynamics and even to statistically verifying whether two measurements of a given trait are significantly different. The importance of mathematics in all biological disciplines is undeniable (Karlin, 1984; May, 2004), yet so few seem conscious of this fact. My goal in this short essay is not give a full history of how mathematics has influenced the biological sciences (there are many reviews on the subject, such as Gunawardena 2014), but a brief account of how this writer's view of mathematical modelling has evolved over time.

When I first decided to go into biology I knew nothing of mathematical models in biology. From the outside, as a molecular biologist, mathematics didn't seem to have its place in my discipline other than for statistical purposes or measuring enzyme reactions and so forth. Everything that could be known about evolution was already known. Everything that had to do with mathematics was already known. The reader can therefore imagine my surprise when I realised that biology and mathematics are living, breathing, -dare I say it?- evolving disciplines. There seems to be some lack of information somewhere along the line. How is it that while within the educational system we are not conscious of all that we do not yet understand on a fun-

damental level? Biology is presented as an empirical science in which theory has nothing to offer.

When I stumbled upon modelling population dynamics during my master's degree, I was in awe. What was this side of biology that had been kept hidden all of these years? I had discovered the cave of wonders. Theoretical and mathematical approaches in biology offer (almost) endless possibilities. My current thesis has focused on understanding the impact of deleterious mutations both on a demographic and genetic level. Due to the difficulty of accessing the mutational history and the frequencies of deleterious mutations of both individuals and populations, tackling these questions empirically is something we don't have the technologies for. Theory can and already has provided a basis for future empirical works, explanations for current observations and predictive models (*i.e.* in predicting the evolution of life-history traits, Stearns and Koella 1986).

My first experience in modelling was writing a seemingly over-simplified multi-locus program, which at the time I thought too biologically unrealistic. As I progressed to writing analytical models, I started to understand the satisfaction in obtaining clear results, knowing which parameter did what. The simplicity of mathematical models is key, as are the assumptions made. I am now steering further and further away from complex models and have become somewhat critical of computer simulations using digital organisms (Elena et al, 2007). However, when scanning Postdoctoral offers, most theoretical works aim to develop such "realistic" hyper-parametered simulations. This leaves me to question whether the availability of such approaches is not linked to the sense that more fundamental models are too simplistic or perhaps to biologists' lack of mathematical know-how. In both cases, it is clear that biologists need to be re-introduced to mathematics for a better understanding of their own theoretical approaches.

In their introductory paragraph, Crow and Kimura (1970) stated that they preferred "generality and realism to precision and rigour". Seeing how complex biological systems are, it is only natural that some rigour be sacrificed. Mathematical models are all about assumptions, which is why it is necessary, in my opinion, to test these assumptions. As Gunawardena (2014) stated, "Reproducibility improves credibility." If by using a separate mathematical approach we find the same response, then that serves to validate a previous model. A perfect example of this is Coron et al's (2013) more rigorous mathematical proof of what Lande (1994) had found intuitively, the existence of the mutational meltdown. Coron et al (2013) also provided clearer limitations and more detailed consequences of both the demographic and genetic parameters taken into account. Just because something has been proven, that should not be a barrier to going back with new points of view

and new approaches in order to test them again. The main and very personal goal of my thesis was just that: testing the assumptions we have made about population demography in population genetics models. We have both confirmed their validity and refuted them, these assumptions do not always hold. But these results are "true" if and only if, we accept the assumptions and limits of the models presented in this thesis. Testing these assumptions will be a next step.

Bibliography

Abu Awad D, Gallina S, Bonamy C, Billiard S (2014) The Interaction between Selection, Demography and Selfing and How It Affects Population Viability. Plos One 9(1), DOI 10.1371/journal.pone.0086125

Agrawal A, Whitlock M (2011) Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 187:553–U274

Agrawal AF, Whitlock MC (2012) Mutation load: The fitness of individuals in populations where deleterious alleles are abundant. Annual Review of Ecology, Evolution, and Systematics 43:115–135, http://www.annualreviews.org/doi/pdf/10.1146/annurev-ecolsys-110411-160257

Ally D, Ritland K, Otto SP (2010) Aging in a long-lived clonal tree. Plos Biology 8(8), DOI 10.1371/journal.pbio.1000454

Angeloni F, Vergeer P, Wagemaker C, Ouborg N (2014) Within and between population variation in inbreeding depression in the locally threatened perennial scabiosa columbaria. Conserv Genet 15:331–342

Baer C, Shaw F, Steding C, Baurngartner M, Hawkins A, Houppert A, Mason N, Reed M, Sinnonelic K, Woodard W, Lynch M (2005) Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. Proceedings of the National Academy of Sciences of the United States of America 102(16):5785–5790, DOI 10.1073/pnas.0406056102

Baker H (1955) Self compatibilty and establishment after long distance dispersal. Evolution 9:347–349, DOI 10.2307/2405656

Barrett SCH (2010) Understanding plant reproductive diversity. Philosophical transactions of the Royal Society of London Series B-Biological Sciences 365(1537):99–109, DOI 10.1098/rstb.2009.0199

Bataillon T, Kirkpatrick M (2000) Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Genetical Research 75:75–81, DOI 10.1017/S0016672399004048

Bell T, Powell K, Bowles M (2013) Viability model choice affects projection accuracy and reintroduction decisions. Journal of Wildlife Management 77(6):1104–1113, DOI 10.1002/jwmg.525

Bernardes A (1995) Mutational meltdown in large sexual populations. Journal de Physique I France 5:1501–1515

Brennan A, Harris S, Hiscock S (2005) Modes and rates of selfing and associated inbreeding depression in the self-incompatible plant Senecio squalidus (Asteraceae): a successful colonizing species in the British Isles. New Phytologist 168(2):475–486, DOI 10.1111/j.1469-8137.2005.01517.x

Caballero A, Hill W (1992) Effects of partial inbreeding on fixation rates and variation of mutant-genes. Genetics 131:493–507

Charlesworth B (1971) Selection in density-regulated populations. Ecology 52:469–474, DOI 10.2307/1937629

Charlesworth B, Charlesworth D (2010) Elements of evolutionary genetics. Roberts and Company

Charlesworth D (2003) Effects of inbreeding on the genetic diversity of populations. Philosophical transactions of the Royal Society of London Series B-Biological Sciences 358:1051–1070, DOI 10.1098/rstb.2003.1296

Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. Annual Review of Ecology and Systematics 18:237–268, DOI 10.1146/annurev.ecolsys.18.1.237

Charlesworth D, Wright S (2001) Breeding systems and genome evolution. Current Opinion In Genetics & Development 11(6):685–690, DOI 10.1016/S0959-437X(00)00254-9

Charlesworth D, Morgan MT, Charlesworth B (1990) Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. Evolution 44(6):1469–1489, DOI 10.2307/2409330

Charlesworth D, Morgan M, Charlesworth B (1993) Mutation accumulation in finite outbreeding and inbreeding populations. Genetical Research 61:39–56

Cheprakov MI, Rakitin SB (2012) Interrelationship between the level of chromosomal aberrations and demographic parameters. Zhurnal Obshchei Biologii 73(4):253–258

Cheptou P, Dieckmann U (2002) The evolution of self-fertilization in density-regulated populations. Proceedings of the Royal Society B-Biological Sciences 269(1496):1177–1186, DOI 10.1098/rspb.2002.1997

Chevin L, Lande R (2010) When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? Evolution 64:1143–1150, DOI 10.1111/j.1558-5646.2009.00875.x

Clarke B (1972) Density-dependent selection. American Naturalist 106(947):1–&, DOI 10.1086/282747

Clarke B (1973) Mutation and population size. Heredity 31:367–379, DOI 10.1038/hdy.1973.92

Coron C, Meleard S, Porcher E, Robert A (2013) Quantifying the mutational meltdown in diploid populations. American Naturalist 181(5):623–636, DOI 10.1086/670022

Costin B, Morgan J, Young A (2001) Reproductive success does not decline in fragmented populations of Leucochrysum albicans subsp albicans var. tricolor (Asteraceae). Biological Conservation 98(3):273–284, DOI 10.1016/S0006-3207(00)00165-8

Crow JF, Kimura M (1970) An introduction to population genetic theory. Harper and Row, New York

Dostalek T, Munzbergova Z, Plackova I (2010) Genetic diversity and its effect on fitness in an endangered plant species, Dracocephalum austriacum L. Conservation Genetics 11(3):773–783, DOI 10.1007/s10592-009-9879-z

Elena SF, Wilke CO, Ofria C, Lenski RE (2007) Effects of population size and mutation rate on the evolution of mutational robustness. Evolution 61:666–674, DOI 10.1111/j.1558-5646.2007.00064.x

Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. Nature Reviews Genetics 8(8):610–618, DOI 10.1038/nrg2146

Ezawa K, Innan H (2013) Theoretical framework of population genetics with somatic mutations taken into account: application to copy number variations in humans. Heredity 111(5):364–374, DOI 10.1038/hdy.2013.59

Fawcett TW, Higginson AD (2012) Heavy use of equations impedes communication among biologists. Proceedings of the National Academy of Sciences of the United States of America 109(29):11,735–11,739, DOI 10.1073/pnas.1205259109

Fischer M, Matthies D (1998) Experimental demography of the rare Gentianella germanica: seed bank formation and microsite effects on seedling establishment. Ecography 21(3):269–278, DOI 10.1111/j.1600-0587.1998.tb00564.x

Fisher R (1941) Average excess and average effect of a gene substitution. Annals of EugenicsA 11:53–63

Frankham R (2005) Genetics and extinction. Biological Conservation 126:131–140, DOI 10.1016/j.biocon.2005.05.002

Frankham R, Ballou JD, Briscoe DA (2010) Introduction to Conservation Genetics, 2nd edn. Cambridge University Press, Cambridge, UK

Frankham R, Gilligan D, Morris D, Briscoe D (2001) Inbreeding and extinction: Effects of purging. Conservation Genetics 2(3):279–285, DOI 10.1023/A:1012299230482

Fredrickson R, Siminski P, Woolf M, Hedrick P (2007) Genetic rescue and inbreeding depression in Mexican wolves. Proceedings of the Royal Society B-Biological Sciences 274(1623):2365–2371, DOI 10.1098/rspb.2007.0785

Gelatt T, Davis C, Stirling I, Siniff D, Strobeck C, Delisle I (2010) History and fate of a small isolated population of Weddell seals at White Island, Antarctica. Conservation Genetics 11(3):721–735, DOI 10.1007/s10592-009-9856-6

Gervais C, Awad DA, Roze D, Castric V, Billiard S (2014) Genetic architecture of inbreeding depression and the maintenance of gametophytic self-incompatibility. Evolution DOI 10.1111/evo.12495, URL http://dx.doi.org/10.1111/evo.12495

Gillespie JH (1998) Population genetics: A concise guide, 2nd edn. The John Hopkins University Press, Baltimore

Gioti A, Stajich J, Johannesson H (2013) Neurospora and the dead-end hypothesis: Genomic consequences of selfing in the model genus. Evolution Accepted, DOI 10.1111/evo.12206

Glémin S (2003) How are deleterious mutations purged? Drift versus nonrandom mating. Evolution 57:2678–2687, DOI 10.1554/03-406

Glémin S, Ronfort J (2013) Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. Evolution 67:225–240

Goldberg E, Kohn J, Lande R, Robertson K, Smith S, Igic B (2010) Species selection maintains self-incompatibility. Science 330:493–495, DOI 10.1126/science.1194513

Gunawardena J (2014) Models in biology: 'accurate descriptions of our pathetic thinking'. BMC Biology 12, DOI 10.1186/1741-7007-12-29

Haag C, Roze D (2007) Genetic load in sexual and asexual diploids: Segregation, dominance and genetic drift. Genetics 176:1663–1678, DOI 10.1534/genetics.107.073080

Haldane J (1924) A mathematical-theory of natural and artificial selection .1. Transactions of the Cambridge Philosophical Society 23(1-2):19–41

Haldane J (1937) The effect of variation on fitness. American Naturalist 71:337–349, DOI 10.1086/280722

Heber S, Varsani A, Kuhn S, Girg A, Kempenaers B, Briskie J (2013) The genetic rescue of two bottlenecked South Island robin populations using translocations of inbred donors. Proceedings of the Royal Society B-Biological Sciences 280(1752), DOI 10.1098/rspb.2012.2228

Hedrick P (2004) Recent developments in conservation genetics. Forest Ecology and Management 197(1-3):3–19, DOI 10.1016/j.foreco.2004.05.002

Higgins K, Lynch M (2001) Metapopulation extinction caused by mutation accumulation. Proceedings of the National Academy of Sciences of the United States of America 98:2928–2933, DOI 10.1073/pnas.031358898

Jarne P, Auld J (2006) Animals mix it up too: The distribution of self-fertilization among hermaphroditic animals. Evolution 60:1816–1824, DOI 10.1554/06-246.1

Karlin S (1984) Mathematical-models, problems, and controversies of evolutionary-theory. Bulletin of the American Mathematical Society 10(2):221–273, DOI 10.1090/S0273-0979-1984-15239-X

Keightley P (2012) Rates and fitness consequences of new mutations in humans. Genetics 190(2):295–304, DOI 10.1534/genetics.111.134668

Keightley P, Lynch M (2003) Toward a realistic model of mutations affecting fitness. Evolution 57:683–685

Keightley PD, Halligan DL (2009) Analysis and implications of mutational variation. Genetica 136(2):359–369, DOI 10.1007/s10709-008-9304-4

Keightley PD, Otto SP (2006) Interference among deleterious mutations favours sex and recombination in finite populations. NATURE 443(7107):89–92, DOI 10.1038/nature05049

Kirkpatrick M, Jarne P (2000) The effects of a bottleneck on inbreeding depression and the genetic load. American Naturalist 155(2):154–167, DOI 10.1086/303312

Kondrashov A (1982) Selection against harmful mutations in large sexual and asexual populations. Genetical Research 40:325–332

Kondrashov A (1995) Contamination of the genome by very slightly deleterious mutations - why have we not died a hundred times over. J Theor Biol 175:583–594, DOI 10.1006/jtbi.1995.0167

Lande R (1988) Genetics and demography in biological conservation. Science 241(4872):1455–1460, DOI 10.1126/science.3420403

Lande R (1994) Risk of population extinction from fixation of new deleterious mutations. Evolution 48:1460–1469, DOI 10.2307/2410240

Lande R, Schemske D (1985) The evolution of self-fertilization and inbreeding depression in plats. i. genetic models. Evolution 39(1):24–40, DOI 10.2307/2408514

Lynch M (2010) Evolution of the mutation rate. Trends in Genetics 26(8):345–352, DOI 10.1016/j.tig.2010.05.003

Lynch M, Gabriel W (1990) Mutation load and the survival of small populations. Evolution 44:1725–1737, DOI 10.2307/2409502

Lynch M, Conery J, Burger R (1995) Mutational meltdowns in sexual populations. Evolution 49:1067–1080, DOI 10.2307/2410432

Manna F, Martin G, Lenormand T (2011) Fitness landscapes: An alternative theory for the dominance of mutation. Genetics 189:923–U303, DOI 10.1534/genetics.111.132944

May R (2004) Uses and abuses of mathematics in biology. Science 303(5659):790–793, DOI 10.1126/science.1094442

Newman D, Pilson D (1997) Increased probability of extinction due to decreased genetic effective population size: Experimental populations of Clarkia pulchella. Evolution 51(2):354–362, DOI 10.2307/2411107

Nunney L (1999) Lineage selection and the evolution of multistage carcinogenesis. Proceedings of the Royal Society B-Biological Sciences 266(1418):493–498, DOI 10.1098/rspb.1999.0664

Orr H (1995) Somatic mutation favors the evolution of diploidy. Genetics 139(3):1441–1447

Packer C, Pusey A, Rowley H, Gilbert D, Martenson J, Obrien S (1991) Case-study of a population bottleneck - lions of the ngorongoro crater. Conservation Biology 5(2):219–230, DOI 10.1111/j.1523-1739.1991.tb00127.x

Perez-Barrales R, Vargas P, Arroyo J (2006) New evidence for the Darwinian hypothesis of heterostyly: breeding systems and pollinators in Narcissus sect. Apodanthi. New Phytologist 171:553–567, DOI 10.1111/j.1469-8137.2006.01819.x

Petes T (2001) Meiotic recombination hot spots and cold spots. Nature Reviews Genetics 2:360–369, DOI 10.1038/35072078

Phadnis N, Fry J (2005) Widespread correlations between dominance and homozygous effects of mutations: Implications for theories of dominance. Genetics 171(1):385–392, DOI 10.1534/genetics.104.039016

Porcher E, Lande R (2005) Loss of gametophytic self-incompatibility with evolution of inbreeding depression. Evolution 59:46–60, DOI 10.1111/j.0014-3820.2005.tb00893.x

Porcher E, Lande R (2013) Evaluating a simple approximation to modeling the joint evolution of self-fertilization and inbreeding depression. Evolution 67(12):3628–3635, DOI 10.1111/evo.12216

Porcher E, Kelly JK, Cheptou PO, Eckert CG, Johnston MO, Kalisz S (2009) The genetic consequences of fluctuating inbreeding depression and the evolution of plant selfing rates. Journal of Evolutionary Biology 22(4):708–717, DOI 10.1111/j.1420-9101.2009.01705.x

Randle AM, Slyder JB, Kalisz S (2009) Can differences in autonomous selfing ability explain differences in range size among sister-taxa pairs of Collinsia (Plantaginaceae)? An extension of Baker's Law. New Phytologist 183(3):618–629, DOI 10.1111/j.1469-8137.2009.02946.x

Robert A, Couvet D, Sarrazin F (2002) Fitness heterogeneity and viability of restored populations. Animal Conservation 5:153–161, DOI 10.1017/S1367943002002214

Roze D (2009) Diploidy, population structure, and the evolution of recombination. American Naturalist 174:S79–S94, DOI 10.1086/599083

Roze D (2012) Spatial heterogeneity in the strength of selection against deleterious alleles and the mutation load. Heredity 109(3):137–145, DOI 10.1038/hdy.2012.24

Roze D, Lenormand T (2005) Self-fertilization and the evolution of recombination. Genetics 170:841–857, DOI 10.1534/genetics.104.036384

Roze D, Rousset F (2004) Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. Genetics 167:1001–1015, DOI 10.1534/genetics.103.025148

Saccheri I, Hanski I (2006) Natural selection and population dynamics. Trends in Ecology and Evolution 21(6):341–347, DOI 10.1016/j.tree.2006.03.018

Sakai S, Wright SJ (2008) Reproductive ecology of 21 coexisting Psychotria species (Rubiaceae): when is heterostyly lost? Biological Journal of the Linnean Society 93:125–134

Schoedelbauerova I, Tremblay R, Kindlmann P (2010) Prediction vs. reality: Can a PVA model predict population persistence 13 years later? Biodiversity and Conservation 19(3):637–650, DOI 10.1007/s10531-009-9724-1

Schoen D, Johnston M, LHeureux A, Marsolais J (1997) Evolutionary history of the mating system in Amsinckia (Boraginaceae). Evolution 51:1090–1099, DOI 10.2307/2411038

Simmons M, Crow J (1977) Mutations affecting fitness in drosophilia populations. Annual Review of Genetics 11:49–78, DOI 10.1146/annurev.ge.11.120177.000405

Sniegowski P, Gerrish P, Johnson T, Shaver A (2000) The evolution of mutation rates: separating causes from consequences. Bioessays 22(12):1057–1066

Stearns S, Koella J (1986) The evolution of phenotypic plasticity in life-history traits - predictions of reaction norms for age and size at maturity. Evolution 40(5):893–913, DOI 10.2307/2408752

Stebbins G (1957) Self fertilization and population variability in the higher plants. American Naturalist (861):337–354, DOI 10.1086/281999

Steinmetz L, Scharfe C, Deutschbauer A, Mokranjac D, Herman Z, Jones T, Chu A, Giaever G, Prokisch H, Oefner P, Davis R (2002) Systematic screen for human disease genes in yeast. Nature Genetics 31(4):400–404, DOI 10.1038/ng929

Takebayashi N, Morrell PL (2001) Is self-fertilization an evolutionary dead end? revisiting an old hypothesis with genetic theories and a macroevolutionary approach. American Journal of Botany 88:1143–1150, http://www.amjbot.org/content/88/7/1143.full.pdf+html

Turner JRG, Williams MH (1968) Population size, natural selection and genetic load. Nature 218:700, DOI 10.1038/218700a0

Vassilieva L, Lynch M (1999) The rate of spontaneous mutation for life-history traits in Caenorhabditis elegans. Genetics 151(1):119–129

Wallace B (1970) Genetic load. Its biological and conceptual aspects. Prentice-Hal, Englewood Cliffs, New Jersey

Westemeier R, Brawn J, Simpson S, Esker T, Jansen R, Walk J, Kershner E, Bouzat J, Paige K (1998) Tracking the long-term decline and recovery of an isolated population. Science 282(5394):1695–1698, DOI 10.1126/science.282.5394.1695

Whitlock MC BR (2004) Fixation of new mutations in small populations. In: Ferrière R CD Dieckmann U (ed) Evolutionary conservation biology, Cambridge University Press, chap 9, pp 155–177

Winn AA, Elle E, Kalisz S, Cheptou PO, Eckert CG, Goodwillie C, Johnston MO, Moeller DA, Ree RH, Sargent RD, Vallejo-Marin M (2011) Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. Evolution 65(12):3339–3359, DOI 10.1111/j.1558-5646.2011.01462.x

Wolfram Research I (2012) Mathematica. Wolfram Research, Inc., Champaign, Illinois, version 9.0 edn

Wootton J, Pfister C (2013) Experimental separation of genetic and demographic factors on extinction risk in wild populations. Ecology 94(10):2117–2123, DOI 10.1890/12-1828.1

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Wright SI, Barrett SCH (2010) The long-term benefits of self-rejection. Science 330:459–460, DOI 10.1126/science.1198063

Wright SI, Ness RW, Foxe JP, Barrett SCH (2008) Genomic consequences of outcrossing and selfing in plants. International Journal of Plant Sciences 169:105–118, DOI 10.1086/523366

Wright SI, Kalisz S, Slotte T (2013) Evolutionary consequences of self-fertilization in plants. Proceedings of the Royal Society B-Biological Sciences 280(1760), DOI 10.1098/rspb.2013.0133

Zeyl C, Mizesko M, de Visser J (2001) Mutational meltdown in laboratory yeast populations. Evolution 55:909–917

Supplementary Information

Gervais et al. : Genetic architecture of inbreeding depression and the maintenance of gametophytic self-incompatibility
Table S1
File S1: Domain of validity for the expressions for population size, numerical and genetic loads and inbreeding depression
File S2: Equations for the change in the number of individuals
File S3: Genotypic frequencies at equilibrium
File S4: Proofs for population size at equilibrium

BRIEF COMMUNICATION

doi:10.1111/evo.12495

GENETIC ARCHITECTURE OF INBREEDING DEPRESSION AND THE MAINTENANCE OF GAMETOPHYTIC SELF-INCOMPATIBILITY

Camille Gervais,[1,2] Diala Abu Awad,[3] Denis Roze,[1,2] Vincent Castric,[3] and Sylvain Billiard[3,4]

[1]*UMI 3614, Evolutionary Biology and Ecology of Algae, CNRS, 29680 Roscoff, France*

[2]*Sorbonne Universités, UPMC University Paris 06, 29680 Roscoff, France*

[3]*Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université Lille 1 – Sciences et Technologies, 59655 Villeneuve d'Ascq, France*

[4]*E-mail: sylvain.billiard@univ-lille1.fr*

Received March 16, 2014
Accepted July 2, 2014

Gametophytic self-incompatibility (GSI) is a widespread genetic system, which enables hermaphroditic plants to avoid self-fertilization and mating with close relatives. Inbreeding depression is thought to be the major force maintaining SI; however, inbreeding depression is a dynamical variable that depends in particular on the mating system. In this article we use multilocus, individual-based simulations to examine the coevolution of SI and inbreeding depression within finite populations. We focus on the conditions for the maintenance of SI when self-compatible (SC) mutants are introduced in the population by recurrent mutation, and compare simulation results with predictions from an analytical model treating inbreeding depression as a fixed parameter (thereby neglecting effects of purging within the SC subpopulation). In agreement with previous models, we observe that the maintenance of SI is associated with high inbreeding depression and is facilitated by high rates of self-pollination. Purging of deleterious mutations by SC mutants has little effect on the spread of those mutants as long as most deleterious alleles have weak fitness effects: in this case, the genetic architecture of inbreeding depression has little effect on the maintenance of SI. By contrast, purging may greatly enhance the spread of SC mutants when deleterious alleles have strong fitness effects.

KEY WORDS: **Deleterious mutation, inbreeding depression, purging, self-incompatibility.**

Self-incompatibility (SI) is a genetic system that enables hermaphroditic plants to avoid self-fertilization and limit mating with close relatives by recognition and rejection by the pistil of pollen expressing cognate specificities. In many species, SI specificity is controlled by a single multiallelic locus, the S-locus. SI is widespread, found in more than 100 families of angiosperms (Igic et al. 2008) despite the fact that it entails a transmission disadvantage. Indeed, a self-compatible (SC) mutant occurring in an SI population should benefit from a higher number of potential mates, and from a transmission advantage through self-fertilization. This last advantage is equal to two-thirds (Charlesworth and Charlesworth 1979), and is thus higher than the 50% advantage of an allele coding for selfing in an outcrossing population (Fisher 1941), because only the SC pollen contributes to the selfed offspring in heterozygous individuals for SC. Recently, it has been argued that SI may be maintained by selection acting at the species level, due to the fact that SI species diversify at higher rates (Goldberg et al. 2010). This form of selection only acts in the long term, however, and implies that SC mutations should occur very rarely. The main advantage of SI that could possibly explain its maintenance in the short term is the avoidance of inbreeding depression (e.g., Charlesworth and Charlesworth 1987). However, Charlesworth and Charlesworth (1979) showed that for SI to be maintained in the presence of SC mutants, inbreeding depression has to be high; its minimal value depends on the number of SI alleles segregating in the population and on the selfing rate of SC mutants, but is often close to Two-thirds when the number of SI alleles is large—this minimal

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

value is higher and may even reach 1 when the number of SI alleles is not large and the selfing rate of SC mutants is small to moderate, as SC mutants also benefit from a higher number of potential mates. Furthermore, the minimal value of inbreeding depression needed to maintain SI may be underestimated by Charlesworth and Charlesworth's (1979) model because inbreeding depression was treated as a fixed variable, therefore neglecting the effect of purging of deleterious alleles by the self-fertilizing SC individuals. Purging is expected to facilitate the spread of SC mutants, as these mutants tend to better eliminate partially recessive deleterious alleles, thereby reducing the magnitude of inbreeding depression experienced by their selfed offspring (Uyenoyama and Waller 1991; Glémin 2003). This was confirmed by a deterministic simulation model by Porcher and Lande (2005), assuming that a given proportion of self-pollen lands on the stigma ("mass-action" pollination model, Holsinger 1991) and that inbreeding depression results from recessive lethal mutations segregating at a very large (effectively infinite) number of loci (Kondrashov 1985). In this model, invasion of an SI population by an SC mutant (i.e., the breakdown of SI) is easier than expected from results of Charlesworth and Charlesworth (1979), in particular when the selfing rate of SC mutants is moderate to high (so that purging can occur). Porcher and Lande (2005) also observed that in a small region of parameter space (namely, for high values of inbreeding depression and low selfing rates), the population may reach a stable, polymorphic equilibrium in which both SI and SC individuals are present. This observation is important in the context of how new SI specificities may arise through SC intermediates, assuming that compensatory mutations may secondarily restore a novel SI functionality (Uyenoyama et al. 2001; Gervais et al. 2011).

Although Porcher and Lande (2005) model showed that considering the joint dynamics of SI and inbreeding depression may strongly affect predictions concerning the maintenance of SI, the fact that they considered lethal mutations may overestimate the importance of purging in situations in which inbreeding depression is (at least partly) generated by mildly recessive deleterious alleles. Indeed, analytical models have shown that the advantage of a selfing modifier through purging is expected to increase with the strength of selection against deleterious alleles (Uyenoyama and Waller 1991; Epinat and Lenormand 2009), which was also observed in deterministic simulations representing the spread of a mutation affecting the selfing rate (Charlesworth et al. 1990). More recently, Porcher and Lande (2013) showed that the effect of purging on the spread of a mutation affecting selfing is much lower when the strength of selection against deleterious alleles is mild ($s = 0.05$) than when it is strong ($s = 1$). However, Porcher and Lande (2013) only considered weak-effect selfing modifiers (changing the selfing rate by 10^{-6}), and the effect of purging may be stronger in the case of a mutation having a large effect on the selfing rate (Charlesworth et al. 1990) such as a mutation disrupting SI. Because the results of Porcher and Lande (2005) showed that purging may strongly limit conditions for the maintenance of SI in the presence of lethal recessive mutations, it is important to assess the generality of this result (in particular to situations in which a substantial proportion of mutations are mildly deleterious) to better understand how SI can be maintained in natural populations (note that Porcher and Lande (2005) also considered situations in which inbreeding depression is partly due to mildly deleterious mutations, but this component of inbreeding depression was not dynamic in their model).

All previous simulation models explicitly considering the genetic basis of inbreeding depression (Charlesworth et al. 1990; Porcher and Lande 2005, 2013) were based on Kondrashov (1985) model, representing the dynamics of deleterious alleles at an infinite number of unlinked loci, in an infinite population. In principle, genetic linkage should increase the effect of purging, by increasing the association between SC alleles and purged genetic backgrounds; however, whether the effect of linkage is substantial for realistic values of genomic recombination rates is unclear. Furthermore, genetic linkage to the S-locus may affect the dynamics of deleterious alleles ("sheltered load," Glémin et al. 2001; Porcher and Lande 2005), in turn affecting the conditions for the maintenance of SI. Finite population size is expected to affect the number of SI alleles maintained in the population—which depends both on population size and on the mutation rate toward new SI alleles (Yokoyama and Hetherington 1982). Because conditions for the maintenance of SI depend on the number of SI alleles segregating (Charlesworth and Charlesworth 1979; Porcher and Lande 2005; Gervais et al. 2011), they should thus be affected by population size. Furthermore, the number of SI alleles may also change as SC mutants increase in frequency, which may in turn affect the conditions needed to maintain SI and SC individuals at a polymorphic equilibrium (Gervais et al. 2011).

In this article, we use a multilocus individual-based simulation program to explore the conditions for the maintenance of SI when inbreeding depression is generated by deleterious alleles segregating at a large number of partially linked loci in a finite population. Our model also differs from the previous models cited above by assuming that inbreeding depression affects both seed and pollen production, which in turn affects the selfing rate of SC individuals (as it depends on the quantity of self-pollen relative to the quantity of pollen received from other individuals). We show that in many cases, conditions for the maintenance of SI are similar to those obtained from an analytical model assuming fixed inbreeding depression (i.e., the effect of purging on the spread of SC mutants remains small), unless a substantial proportion of deleterious mutations have strong fitness effects. Linkage has only little effect for the parameter values tested, as long as the mean number of crossovers per genome (i.e., genetic map length)

is not too small. Finally, we almost never observe a polymorphic equilibrium, which is probably due to the fact that the number of SI alleles decreases as SC individuals increase in frequency, further enhancing the advantage of self-compatibility.

Methods

We consider a population with a GSI system, that is, fertilization is possible only if the specificity expressed by the pollen is different from the two specificities codominantly expressed in the style. Specificities are subject to negative frequency-dependent selection, because pollen bearing rare specificities can fertilize more individuals than pollen bearing more common specificities (Wright 1939). We assume that SI is coded by a single S-locus with many alleles (denoted S_i): a plant with genotype S_iS_j ($i \neq j$) is self-incompatible and can be fertilized by pollen of genotype S_l, with $l \neq i$ and $l \neq j$. We also consider a mutant, SC allele S_C segregating at the same locus: pollen carrying S_C can fertilize all plants, whereas a plant with genotype S_iS_C is partially SC (through its S_C pollen) and can be fertilized by any pollen whose genotype is different from S_i. Finally, S_CS_C individuals are fully SC and can be fertilized by any pollen. A parameter α measures the proportion of pollen produced by a plant that stays on the same plant ("self-pollen"), leading to self-fertilization if it carries the S_C allele. We assume that selfed offspring suffer from inbreeding depression, generated by partially recessive deleterious alleles segregating at a number of different loci. Several selective forces may affect the frequency of the S_C allele: (1) automatic transmission advantage through selfing, when $\alpha > 0$; (2) transmission advantage through outcrossing, as S_C pollen can fertilize all plants; (3) negative consequences of increased homozygosity due to selfing on the mean fitness of offspring (inbreeding depression); and (4) indirect benefits stemming from a better elimination of deleterious alleles as a consequence of increased homozygosity (purging). This last effect occurs more rapidly when deleterious alleles have stronger effects (e.g., Charlesworth and Willis 2009). S_C is thus favored by effects (1), (2), and (4), and disfavored by effect (3); but note that purging also tends to reduce inbreeding depression. In the following, we use a multilocus, individual-based simulation model to represent explicitly the genetic causes of effects (3) and (4) (inbreeding depression and purging). Results on the spread of S_C are compared to the predictions of a simple analytical model in which effect (4) is ignored by treating inbreeding depression as a fixed parameter.

ANALYTICAL MODEL

Our analytical model represents an infinite population with discrete, nonoverlapping generations. We assume that $n + 1$ alleles segregate at the S-locus: n self-incompatible alleles and the SC S_C allele. We denote x_1 and x_2 the frequencies of S_CS_C individuals produced by selfing (x_1) and by outcrossing (x_2), and x_3 and x_4 the frequencies of S_iS_C individuals produced by selfing (x_3) and by outcrossing (x_4), where S_i can be any SI allele. We suppose that selfed individuals produce fewer gametes (inbreeding depression), the number of male and female gametes produced by outcrossed individuals being proportional to $W_2 = W_4 = 1$, whereas the number of gametes produced by selfed individuals is proportional to $W_1 = W_3 = 1 - \delta$; mean fecundity $\overline{W}$ is thus given by $1 - \delta(x_1 + x_3)$. The selfing rate a_i (proportion of selfed seeds) of an individual of type i (1, 2, 3 or 4) is given by

$$a_i = \frac{\alpha\,\theta_i\,W_i}{\alpha W_i + (1-\alpha)\overline{W}}, \tag{1}$$

where $\theta_i = 1$ for $i = 1, 2$ and $\theta_i = 1/2$ for $i = 3, 4$. The numerator of equation (1) represents the quantity of compatible self-pollen (bearing allele S_C), whereas the denominator is the total quantity of pollen received by the individual. The frequency of allele S_C among gametes is given by $q = \sum_i \theta_i\,W_i\,x_i/\overline{W}$, whereas, by symmetry, each SI allele is present in frequency $p = (1 - q)/n$. From this, frequencies at the next generation are given by the following:

$$\begin{aligned}
\overline{W}x_1' &= W_1 a_1 x_1 + W_2 a_2 x_2 + \frac{1}{2}W_3 a_3 x_3 + \frac{1}{2}W_4 a_4 x_4,\\
\overline{W}x_2' &= W_1(1-a_1)q\,x_1 + W_2(1-a_2)q\,x_2\\
&\quad + \frac{[W_3(1-a_3)x_3 + W_4(1-a_4)x_4]\,q}{2(1-p)},\\
\overline{W}x_3' &= \frac{1}{2}W_3 a_3 x_3 + \frac{1}{2}W_4 a_4 x_4,\\
\overline{W}x_4' &= W_1(1-a_1)(1-q)\,x_1 + W_2(1-a_2)(1-q)\,x_2\\
&\quad + \frac{1}{2}W_3(1-a_3)x_3 \quad + \frac{1}{2}W_4(1-a_4)x_4\\
&\quad + (1 - x_1 - x_2 - x_3 - x_4)\,\frac{q}{2(1-p)}\,.
\end{aligned} \tag{2}$$

In the Supplementary Material, we use a local stability analysis to determine the values of α, n, and δ for which S_C increases in frequency when rare (which involves solving a fourth-order equation in δ numerically).

MULTILOCUS SIMULATIONS

Individual-based, multilocus simulations were used to explore the conditions for the maintenance of SI when deleterious alleles segregate at a large number of partly linked loci, and when the number of SI alleles evolves freely by mutation, selection, and drift. The simulation program (written in C++, and available from Dryad) corresponds to a modified version of the program described in Roze and Michod (2010), representing a population of N diploid individuals (the parameters used in the simulation

Table 1. Parameters and default values used in simulations.

Description	Symbol	Default Value
Population size	N	2000
Mean number of crossovers per genome per generation	L	10
Proportion of self-pollen	α	
Selection coefficient of deleterious mutations	s	0.05
Dominance coefficient of deleterious mutations	h	0.2
Rate of deleterious mutation per haploid genome	U	
Rate of mutation from S_i to any S_j (with $j \neq i$)	U_{SI}	10^{-5}
Rate of mutation from S_i to S_C	U_{SC}	10^{-4}

model are summarized in Table 1). Each individual possesses two copies of a linear chromosome, and an S-locus located at the midpoint of the chromosome. We assume that a maximum of $k + 1$ alleles can segregate at the S-locus: k self-incompatible alleles S_i ($1 \leq i \leq k$) and an SC allele S_C; k is fixed at 100 in all simulations. Every generation, each S_i allele mutates to any of the $k - 1$ other SI alleles with probability U_{SI} and to allele S_C with probability U_{SC} (loss of SI). There is no reverse mutation from S_C to S_i, that is, loss of SI is irreversible. Deleterious mutations occur at a rate U per haploid genome, that is, the number of new deleterious mutations on each chromosome is drawn from a Poisson distribution with parameter U, whereas the position of each new mutation along the chromosome is sampled from a (continuous) uniform distribution—the number of sites at which deleterious alleles may segregate is thus effectively infinite. All deleterious mutations have the same selection and dominance coefficients (s and h respectively), although we also consider cases in which a given proportion of mutations are lethal (as detailed below). Reproduction occurs as follows: for each of the N individuals of the next generation the maternal parent is sampled randomly among all individuals of the previous generation, the probability of sampling parent i being proportional to its fecundity $W_i = (1 - hs)^{N_{\text{he}}}(1 - s)^{N_{\text{ho}}}$, where N_{he} and N_{ho} are the number of mutations in the heterozygous and homozygous state within its genome. If the maternal plant carries at least one SC allele S_C, its selfing rate a_i is calculated as follows:

$$a_i = \frac{\alpha\, \gamma_{ii}\, W_i}{\alpha\, \gamma_{ii}\, W_i + \frac{1-\alpha}{N-1} \sum_{j \neq i} \gamma_{ij}\, W_j}, \qquad (3)$$

where γ_{ij} is 0, 1, or 2 and represents the number of individual j's S-alleles that are compatible with individual i. In the case of selfing, the offspring's genome is formed from two of i's recombinant gametes; otherwise an individual j is sampled randomly with a probability proportional to W_j, and contributes as a father only if one of its S-alleles is compatible with those of the mother (if not, another individual is sampled until a compatible partner is found). To form a recombinant gamete, the number of crossovers occurring along the chromosome is drawn from a Poisson distribution with parameter L (genome map length, in Morgans) and the position of each crossover is sampled from a uniform distribution.

At the beginning of each simulation, individuals are free of deleterious mutations and heterozygous for randomly sampled SI alleles. During the first 2000 generations, the number of SI alleles segregating in the population is allowed to reach equilibrium by considering only mutation between SI alleles ($U_{SC} = 0$, $U = 0$). Deleterious mutations are then introduced and allowed to reach mutation-selection balance over the next 2000 generations. Finally, during the next 500,000 generations, SC mutants are also introduced. Because loss of SI is irreversible, simulations are stopped after 50,000 generations if S_C is fixed in the population. Every 100 generations, different variables are measured from the population: the frequency of allele S_C, the effective number of SI alleles present and the level of inbreeding depression. The effective number of SI alleles n_e (measured before allele S_C is introduced) corresponds to the number of alleles that would yield the same genetic diversity at the S-locus if all alleles were present in frequency $1/n_e$; it is calculated as $n_e = 1/\sum_{i=1}^{k} p_i^2$, where p_i is the frequency of allele S_i. Inbreeding depression is measured as $\delta = 1 - W_s/W_o$, where W_s and W_o are the mean fecundities of selfed and outcrossed offspring, respectively (estimated by creating 100 selfed and 100 outcrossed offspring from randomly sampled parents, without taking into account the compatibility between their S-alleles). For each simulation run, δ and n_e are averaged over the last 50 samples before introduction of SC mutants, whereas the frequency of S_C is averaged over the last 300 samples of the simulation (last 30,000 generations). The minimal value of δ necessary to maintain SI in the population is determined by running simulations with increasing values of U for each set of parameters values (U_{SI}, U_{SC}, s, h, N, L, and α). When the frequency of S_C stays lower than 0.05 throughout the simulation, SI is considered maintained.

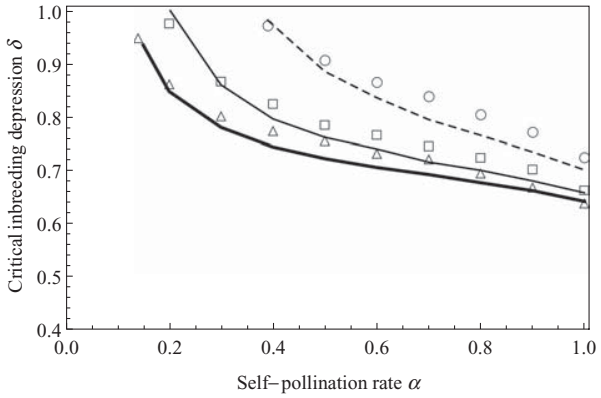


Figure 1. Minimal inbreeding depression needed to prevent invasion of an SI population by an SC mutant, when the genetic load depends only on small-effect deleterious alleles, and for different population sizes *N*. The points correspond to multilocus simulation results and the curves to analytical predictions. Circles, dashed curve: $N = 500$; squares, solid curve: $N = 2000$; triangles, thick curve: $N = 5000$. Other parameter values: $L = 10$, $s = 0.05$, $h = 0.2$.

Results

MILDLY DELETERIOUS MUTATIONS

Figure 1 shows the minimal value of inbreeding depression δ necessary to maintain SI in the population as a function of the rate of self-pollination α, for different values of population size N. In all figures, the highest value of inbreeding depression (in the absence of S_C) for which we observed that S_C invades the population (i.e., reaches frequency 0.05) is just slightly below the points (see Fig. S3 for more detailed results). In almost all simulations the frequency of S_C at equilibrium was either close to zero or equal to 1. Note that because the loss of SI is irreversible in our model (no back mutation from SC to SI), one expects that allele S_C should necessarily become fixed after a sufficiently long time. However, changing the number of generations with $U_{SC} > 0$ to 10^5 or to 10^6 did not lead to significant differences in the threshold values of δ shown in Figure 1 (results not shown), suggesting that as one enters the area above the points, the expected fixation time of S_C quickly reaches extremely high values (i.e., SI is stably maintained). In the same vein, increasing the mutation rate towards S_C (U_{SC}) from 10^{-5} to 10^{-4} has very little effect on the results (not shown). In a few cases, S_C was still polymorphic at the end of the simulation (at frequency >0.05), or the time to fixation of S_C was higher than the average time to fixation of a neutral allele ($4N$ generations), suggesting that selection may maintain polymorphism for these parameter values. However, this was only observed for narrow parameter ranges at the leftmost limit of the region in which SI is maintained (see Fig. S3).

Curves on Figure 1 correspond to the predictions derived from the analytical model (neglecting the effect of genetic associations between the S-locus and selected loci), in which the number of SI alleles n is set to the average effective number of alleles n_e measured in simulations corresponding to the critical δ. Overall, simulation results are qualitatively and quantitatively consistent with these analytical expectations: maintenance of SI is always observed when inbreeding depression is high ($\delta > 2/3$), the critical δ being lower when the self-pollination rate α is higher, in agreement with previous results neglecting effects of purging (Charlesworth and Charlesworth 1979; Uyenoyama et al. 2001). The effect of population size N on the critical δ is due to the fact that the effective number of SI alleles is reduced in smaller populations through the loss of low-frequency alleles by drift (e.g., Yokoyama and Hetherington 1982): on average n_e at the critical δ equals 11.9, 23.0, and 36.4 for $N = 500$, 2000, and 5000, respectively. Lower effective numbers of SI alleles favor the spread of S_C by increasing the transmission advantage of S_C through outcrossing, as S_C pollen never encounters incompatible pistils and can fertilize every potential mate in the population (Charlesworth and Charlesworth 1979; Porcher and Lande 2005; Gervais et al. 2011). Simulations for different values of U_{SI} (rate of mutation toward new SI alleles) show that increasing U_{SI} (with $NU_{SI} = 0$, 0.02, and 0.2) has similar effects as increasing N (Fig. S1).

Overall, the good match between the simulation results (with purging) and predictions from the analytical model (without purging) indicates that purging has little effect on the spread of S_C for these parameter values. Note that the analytical model systematically underestimates the critical δ for SI to be maintained, which suggests that some purging may still be taking place. In most cases however, the discrepancy remains slight, suggesting that the magnitude of this effect is low, except when α increases and hence the selfing rate of SC individuals becomes more important. The effect of purging on the spread of S_C should be more important when selection against deleterious alleles is stronger (e.g., Charlesworth et al. 1990), as deleterious alleles are eliminated more rapidly when present in homozygotes, and as the benefit of being associated with chromosomes carrying fewer deleterious alleles is stronger. In agreement with this prediction, Figure 2 shows that discrepancies between analytical and simulation results becomes more important for higher values of s and h. One can also see from Figure 2 that, at the critical δ, s has little effect on the effective number of SI alleles n_e maintained in the population (the curves on Fig. 2 A are almost superposed), whereas higher values of h lead to lower values of n_e. This is likely due to the fact that n_e is affected by background selection (reduction in diversity due to selection at linked sites, e.g., Charlesworth 1993). Indeed, background selection is stronger when the deleterious mutation rate U is higher (e.g., Hudson and Kaplan 1995), and higher values of U are needed to reach the critical δ when h is increased,

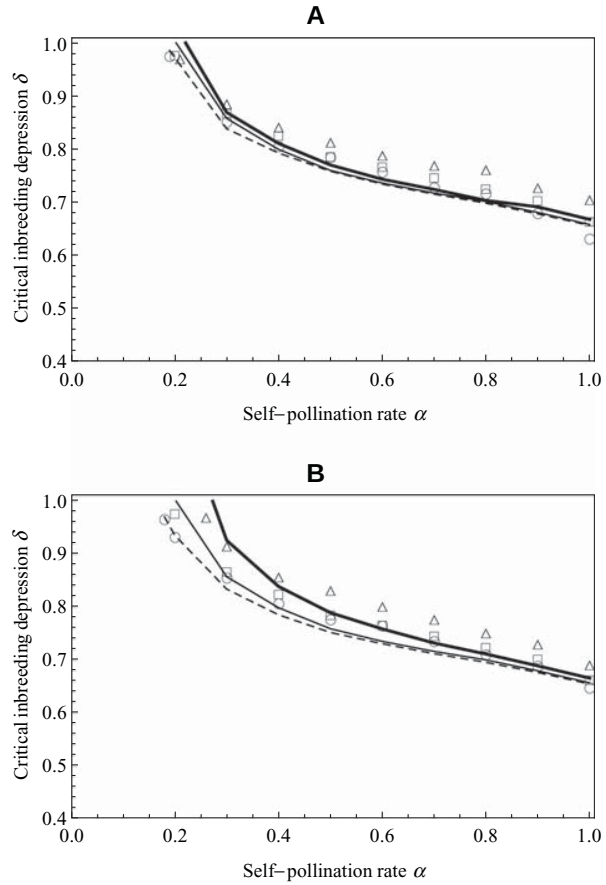


Figure 2. Minimal inbreeding depression needed to prevent invasion of an SI population by an SC mutant, when the genetic load depends only on small-effect deleterious alleles, and for different selection (A) and dominance (B) coefficients of deleterious alleles. The points correspond to multilocus simulation results and the curves to analytical predictions. (A) circles, dashed curve: $s = 0.02$; squares, solid curve: $s = 0.05$; triangles, thick curve: $s = 0.1$. (B) Circles, dashed curve: $h = 0.1$; squares, solid curve: $h = 0.2$; triangles, thick curve: $h = 0.3$. Other parameter values: $N = 2000$, $L = 10$, $h = 0.2$ (in A), $s = 0.05$ (in B).

leading to stronger background selection effects at the critical δ. By contrast, δ is only weakly dependent on s (as long as population size is sufficiently large, e.g., Bataillon and Kirkpatrick 2000), and s has thus little effect on the strength of background selection at the critical δ. Finally, we find that varying map length L has relatively little effect on the results as long as it is sufficiently large (roughly, $L > 5$ using our default parameter values shown in Table 1). Figure S2 shows that reducing L decreases the effective number of SI alleles (background selection) and increases selection for S_C through purging (higher discrepancy between analytical and simulation results when L is lower), both effects reducing the parameter range in which SI is maintained.

ADDING A PROPORTION OF NEARLY RECESSIVE LETHAL MUTATIONS

Using a deterministic model in which mutations generating inbreeding depression are lethal and nearly recessive, Porcher and Lande (2005) found much more stringent conditions for the maintenance of SI than those shown on Figures 1 and 2. We modified our simulation program so that a proportion λ of deleterious mutations are nearly recessive lethals (selection and dominance coefficients: $s_l = 1$ and $h_l = 0.02$, respectively) whereas the other mutations (in proportion $1 - \lambda$) are mildly deleterious ($s = 0.05$ and $h = 0.2$). As shown by Figure 3, the discrepancy between analytical and simulation results becomes stronger as λ increases, and the parameter range in which SI is maintained is considerably reduced. In particular, allele S_C invades the population for much larger values of inbreeding depression when the rate of self-pollination α is high, so that individuals carrying S_C frequently self. The proportion of lethal mutations has almost no effect on the effective number of SI alleles maintained in the absence of S_C, and therefore the analytical predictions are nearly unaffected by λ (curves on Fig. 3 are nearly superposed).

Discussion

SI (and especially GSI) is widespread among angiosperms, despite the fact that SC mutants should benefit from a direct transmission advantage. Different models have shown that, provided that it is sufficiently high, inbreeding depression can allow the stability of SI despite the recurrent occurrence of SC mutants (Charlesworth and Charlesworth 1979; Uyenoyama et al. 2001; Porcher and Lande 2005; Gervais et al. 2011). However, these models differ in the way inbreeding depression is introduced: some models treat inbreeding depression as a fixed parameter, thereby neglecting the effects of purging within SC backgrounds (Charlesworth and Charlesworth 1979; Uyenoyama et al. 2001; Gervais et al. 2011), whereas the model by Porcher and Lande (2005) explicitly represents the genetic architecture of inbreeding depression (infinite number of unlinked loci subject to recessive lethal mutations) and shows that purging can dramatically decrease the parameter range in which SI is maintained, in particular when the selfing rate of SC mutants is high. The importance of this effect should, in principle, depend on the genetic basis of inbreeding depression, because purging may occur over just a few generations when deleterious alleles are highly deleterious, but much more slowly when mutations tend to have weak fitness effects. In this article, we compare predictions from an analytical model assuming fixed inbreeding depression to the results of multilocus simulations in which deleterious alleles occur along a linear genetic map, to assess the importance of purging on the spread of SC mutants. When inbreeding depression is mainly due to weak-effect mutations, purging has limited effects. In most

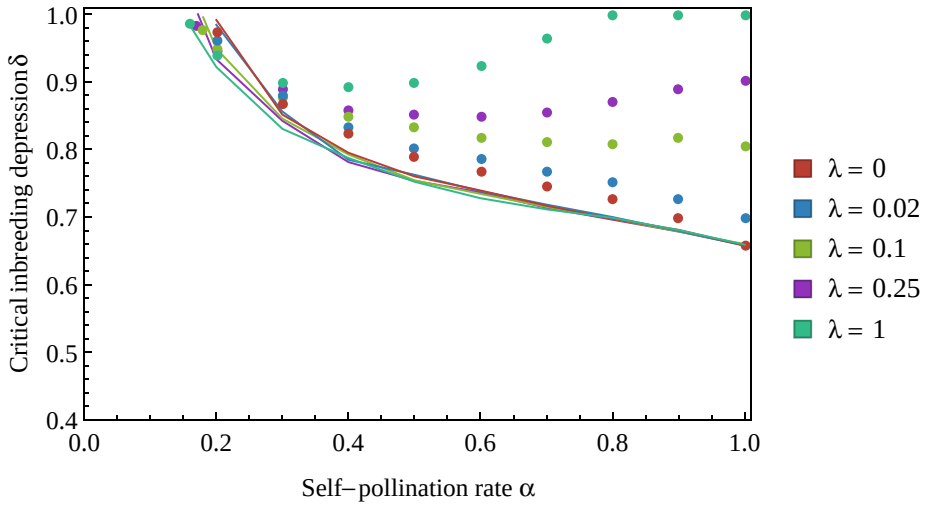


Figure 3. Minimal inbreeding depression needed to prevent invasion of an SI population by an SC mutant, when the genetic load depends on a mix of small-effect and nearly recessive lethal deleterious alleles, and for different proportions λ of lethal mutations. The points correspond to multilocus simulation results and the curves to analytical predictions. Mildly deleterious mutations: $s = 0.05$ and $h = 0.2$. Nearly recessive lethal mutations: $s = 1$ and $h = 0.02$. Other parameter values: $N = 2000$, $L = 10$.

cases, maintenance of SI mainly depends on the number of SI alleles segregating in the population, the rate of self-pollination, and inbreeding depression, independently of population size and the genetic architecture of inbreeding depression—a similar result was obtained recently by Porcher and Lande (2013) on the evolution of self-fertilization through weak-effect modifiers. Note that our model does not incorporate pollen limitation, which would tend to favor the loss of SI (Porcher and Lande 2005). However, our results relative to the effect of purging should remain valid in the presence of pollen limitation: pollen limitation adds direct selection for self-compatibility, but should not modify the indirect effect of deleterious alleles on SC mutants.

Although our current knowledge on the genetic basis of inbreeding depression remains fragmentary, several lines of evidence suggest an important role of mutations of small effects (Carr and Dudash 2003; Charlesworth and Willis 2009), although a study by Fox et al. (2010) showed a rapid reduction of inbreeding depression after several generations of inbreeding, indicating a potentially important effect of strongly deleterious mutations. More experimental work is thus necessary to assess whether purging is susceptible to significantly affect the spread of SC mutants within self-incompatible populations.

Finally, we almost never observe stable polymorphic equilibria involving both SI and SC alleles (except for restricted cases involving very high inbreeding depression and low rates of self-pollination). This stands in contrast to previous models involving infinite populations (Uyenoyama et al. 2001; Porcher and Lande 2005; Gervais et al. 2011), in which wider regions of parameter space allowing polymorphism were observed. A possible explanation for this discrepancy is that, in infinite populations, the number of SI alleles stays constant and is not affected by the frequency of SC, whereas in our model the number of SI alleles decreases as SC increases in frequency (because the size of the SI subpopulation decreases). The decrease in number of SI alleles tends to favor SC, whose frequency can further increase until reaching fixation. This result is consistent with the fact that SC alleles are rarely found in natural SI populations (Stone 2002), although some cases have been reported in which SI appears to be quantitative rather than qualitative, with some partially SC alleles (Mena-Ali and Stephenson 2007; Paape et al. 2011). The lack of polymorphic equilibria should impose restrictions on the evolution of new SI specificities. In the present model we assume that new SI alleles appear in a single mutational step, but in reality the evolution of a new specificity involves at least two mutations: one affecting the protein expressed by the pollen and the other the receptor expressed by the pistil (both genes being part of the S-locus). Most scenarios for the evolution of new specificities rely on an intermediate step involving a SC mutant present at an intermediate frequency in the population (e.g., Uyenoyama et al. 2001; Gervais et al. 2011), which should become more difficult in the absence of polymorphic equilibrium (unless the mutation rate at the S-locus is sufficiently high, so that a compensatory mutation can appear before SC reaches fixation). Furthermore, SC mutations occurring in the receptor part of the S-locus can spread under more restricted conditions than SC mutations in the pollen part, because they do not benefit from a fertilization advantage under outcrossing. Modifying our simulation model to explicitly

represent the pollen and pistil components of the S-locus (to explore conditions for the evolution of new SI specificities with dynamical inbreeding depression) would be an interesting extension of the present work.

ACKNOWLEDGMENTS

We thank two anonymous reviewers for helpful comments. This work was supported by the French Agence Nationale de la Recherche (ANR-11-BSV7-013).

DATA ARCHIVING

The doi for our data is 10.5061/dryad.rv746.

LITERATURE CITED

Bataillon, T., and M. Kirkpatrick. 2000. Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Genet. Res. 75:75–81.

Carr, D. E., and M. R. Dudash. 2003. Recent approaches into the genetic basis of inbreeding depression in plants. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358:1071–1084.

Charlesworth, B. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303.

Charlesworth, D., and B. Charlesworth. 1979. Evolution and breakdown of s-alleles systems. Heredity 43:41–55.

———. 1987. Inbreeding depression and its evolutionary consequences. Ann. Rev. Ecol. Syst. 18:237–268.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10:783–796.

Charlesworth, B., M. T. Morgan, and B. Charlesworth. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. Evolution 44:1469–1489.

Epinat, G., and T. Lenormand. 2009. The evolution of assortative mating and selfing with in- and outbreeding depression. Evolution 63:2047–2060.

Fisher, R. 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11:53–63.

Fox, C. W., K. L. Scheibly, and D. H. Reed. 2010. Experimental evolution of the genetic load and its implications for the genetic basis of inbreeding depression. Evolution 62:2236–2249.

Gervais, C. E., V. Castric, A. Ressayre, and S. Billiard. 2011. Origin and diversification dynamics of self-incompatibility haplotypes. Genetics 188:625–636.

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

Glémin, S., T. Bataillon, J. Ronfort, A. Mignot, and I. Olivieri. 2001. Inbreeding depression in small populations of self-incompatible plants. Genetics 159:1217–1229.

Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith, and B. Igić. 2010. Species selection maintains self-incompatibility. Science 330:493–495.

Holsinger, K. E. 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. Am. Nat. 138:606–622.

Hudson, R. R., and N. L. Kaplan. 1995. Deleterious background selection with recombination. Genetics 141:1605–1617.

Igic, B., R. Lande, and J. R. Kohn. 2008. Loss of self-incompatibility and its evolutionary consequences. Int. J. Plant Sci. 169:93–104.

Kondrashov, A. S. 1985. Deleterious mutations as an evolutionary factor. II. Facultative apomixis and selfing. Genetics 111:635–653.

Mena-Ali, J. I., and A. G. Stephenson. 2007. Segregation analyses of partial self-incompatibility in self and cross progeny of solanum carolinense reveal a leaky s-allele. Genetics 177:501–510.

Paape, T., T. Miyake, N. Takebayashi, D. Wolf, and J. R. Kohn. 2011. Evolutionary genetics of an s-like polymorphism in papaveraceae with putative function in self-incompatibility. PLoS One 6:e23635.

Porcher, E., and R. Lande. 2005. Loss of gametophytic self-incompatibility with evolution of inbreeding depression. Evolution 59:46–60.

———. 2013. Evaluating a simple approximation to modeling the joint evolution of self-fertilization and inbreeding depression. Evolution 67:3628–3635.

Roze, D., and R. E. Michod. 2010. Deleterious mutations and selection for sex in finite, diploid populations. Genetics 184:1095–1112.

Stone, J. 2002. Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. Quart. Rev. Biol. 77:17–32.

Uyenoyama, M. K., and D. M. Waller. 1991. Coevolution of self-fertilization and inbreeding depression. I. Mutation-selection balance at one and two loci. Theoret. Popul. Biol. 40:14–46.

Uyenoyama, M., Y. Zhang, and E. Newbigin. 2001. On the origin of self-incompatibility haplotypes: transition through self-compatible intermediates. Genetics 157:1805–1817.

Wright, S. 1939. The distribution of self-sterility alleles in populations. Genetics 24:538–552.

Yokoyama, S., and L. Hetherington. 1982. The expected number of self-incompatibility alleles in finite plant-populations. Heredity 48:299–303.

Associate Editor: J. Wilkins

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Effect of mutation rate towards new SI alleles U_{SI}.
Figure S2. Effect of map length L.
Figure S3. Full simulation results.
Text S1. Condition for increase of the SC mutant in the absence of purging.

N_{mut}	Mating success	$N_{eq}\frac{(1-\mu)^2\left(2+h^4(1+\mu)^2s^2-2h(4+H_1)+2h^2(4+s+2\mu s+2H_1)-h^3s(4+8\mu+(1+\mu)H_1)\right)}{2(1-2h)^2}$
	Fecundity and Zygote Survival	$N_{eq}\frac{(1-\mu)\left(2+h^2(1+\mu)s-h(4+H_1)\right)}{2(1-2h)}$
	Adult Survival	N_{eq}
Λ	Mating success	$1-\frac{(1-\mu)^2\left(2+h^4(1+\mu)^2s^2-2h(4+H_1)+2h^2(4+s+2\mu s+2H_1)-h^3s(4+8\mu+(1+\mu)H_1)\right)}{2(1-2h)^2}$
	Fecundity and Zygote Survival	$1-\frac{(1)\mu)(2+h^2(1+\mu)s-h(4+H_1))}{2-4h}$
	Adult Survival	0
L	Mating success, Fecundity and Adult Survival	$\frac{h^2\mu^2s+\mu(2-h(H_1+4))-h(hs-H_1)}{2(1-2h)}$
	Haldane's model	$\frac{2(1+\mu)-h\left(4+4\mu-h(s+2s\mu)+\sqrt{s(4\mu(1+\mu)-8h\mu(1+\mu)+s(h+2h\mu)^2)}\right)}{(2-4h)(1+\mu)^2}$
δ	Mating success, Fecundity and Adult Survival	$\frac{(1-h)(H_1-hs)-\mu\left(3h^2s-h(s+4)+2\right)}{4(1-\mu)(1-h(2-hs))}$
	Haldane's model	$\frac{(1-2h)\mu\left(h(s+2s\mu)+\sqrt{s(4\mu(1+\mu)-8h\mu(1+\mu)+s(h+2h\mu)^2}-2\mu\right)}{2\left(2\mu+h\left(h(s+2s\mu)+\sqrt{s(4\mu(1+\mu)-8h\mu(1+\mu)+s(h+2h\mu)^2}-4\mu\right)\right)}$

Table S1: Explicit solutions at mutation-selection balance for population size N_{mut}, the numercial load Λ, the genetic load L, inbreeding depression δ for values of dominance h between 0 and 1 different than 0.5. $H_1 = \sqrt{s\left((4-8h)\mu + h^2(1+\mu)^2 s\right)}$. The equations for L and δ for Zygote Survival are too long to be represented, but can be obtained using the frequencies of each genotype at equilibrium in File S3 and the equations in the main text. The domain of validity of these expressions is addressed in more detail in File S1.

File S1

Domain of validity for the expressions for population size, numerical and genetic loads and inbreeding depression

Full expressions for all measured variables (population size, the numerical and genetic load, and inbreeding depression) can be found in Table SS1. These equations are valid for all values of dominance h between 0 and 1, but not codominant mutations ($h = 0.5$); expressions for co-dominant mutations can be found in Table 2 of the main text. The validity of these equations depends on the threshold value of the mutation rate $\mu_{fix} = \frac{(1-h)s}{1-hs}$ for the deterministic fixation of the deleterious allele a and on the validity of $H_1 = \sqrt{s\left((4-8h)\mu + h^2(1+\mu)^2 s\right)}$. Figure SS1 is a graphical representation of the domains of validity for each of the limitations. For H_1 to be in the domain of real numbers, and hence for the expressions to be biologically realistic, s must be greater than $\frac{(8h-4)\mu}{h^2(1+\mu)^2}$. This limitation does not interfere with the domain of validity for the expressions, as it is always below the domain of validity imposed by the deterministic fixation of a ($\mu > \frac{(1-h)s}{1-hs}$).

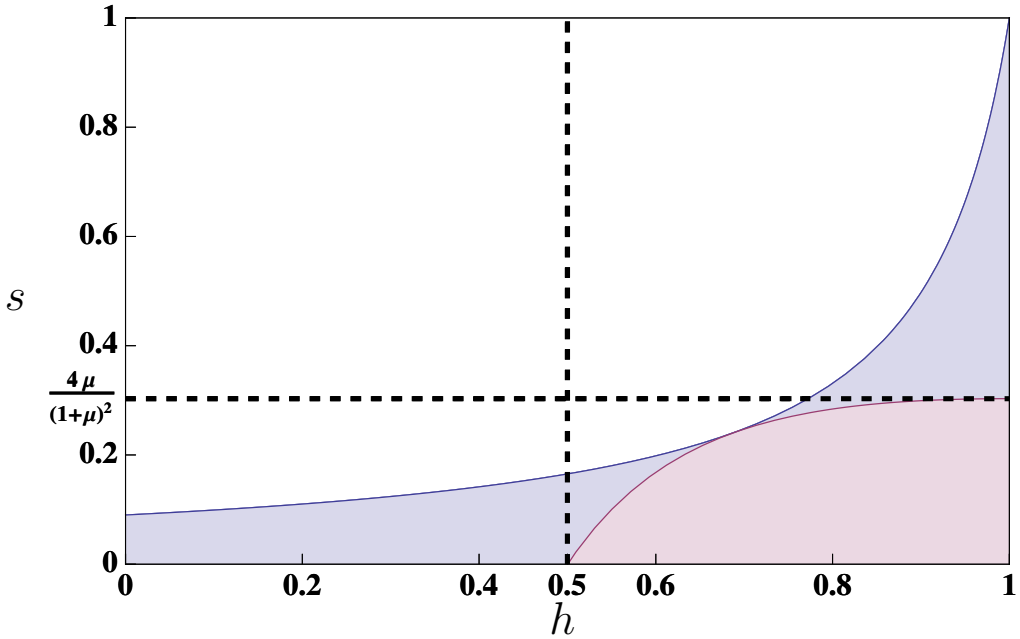


Figure S1: Limitations for the validity of the expressions in Table SS1. The area below the blue (respectively red) curve depicts all values of s and h for which the expressions are not valid as mutations are deterministically fixed (respectively as the expression H_1 is not in the domain of real numbers) for $\mu = 10^{-2}$.

File S2

Equations for the change in the number of individuals

Here we provide the full general equations for the change in the number of individuals of each genotype for all three models of selection (mating success, fecundity and survival). The expressions presented in Table SS1 as well as the genotypic frequencies of X, Y and Z and mutation selection balace presented below are obtained by solving for the steady state of these differential equations ($\frac{dX_t}{N_t} = \frac{dY_t}{N_t} = \frac{dZ_t}{N_t} = 0$). These expressions were obtained using Wolfram's Mathematica 9.0.

Selection on mating success

$$\begin{aligned}
\frac{dX_t}{N_t} =& \frac{b}{N_t}\Bigg((1-s)^2X_t^2 + 2(1-s)X_tZ_t\mu + Z_t^2\mu^2 + (1-s)(1-hs)X_tY_t(1+\mu) \\
&+ (1-hs)Y_tZ_t\mu(1+\mu) + \frac{1}{4}(1-hs)^2Y_t^2(1+\mu)^2\Bigg) - d\frac{N_t}{K}X_t \\
=& \frac{b}{N_t}(r^X)^2 - d\frac{N_t}{K}X_t \qquad \text{(S1a)} \\
\frac{dY_t}{N_t} =& \frac{b}{N_t}\Bigg((1-s)(1-hs)X_tY_t(1-\mu) + 2(1-s)X_tZ_t(1-\mu) + 2Z_t^2(1-\mu)\mu \\
&+ \frac{1}{2}(1-hs)^2Y_t^2(1-\mu)(1+\mu) \\
&+ 2(1-hs)Y_tZ_t\left(\frac{1}{2}(1-\mu)\mu + \frac{1}{2}(1-\mu)(1+\mu)\right)\Bigg) - d\frac{N_t}{K}Y_t \\
=& \frac{2b}{N_t}(r^Xr^Y) - d\frac{N_t}{K}Y_t \qquad \text{(S1b)} \\
\frac{dZ_t}{N_t} =& \frac{b}{N_t}\left(\frac{1}{4}(1-hs)^2Y_t^2(1-\mu)^2 + (1-hs)Y_tZ_t(1-\mu)^2 + Z_t^2(1-\mu)^2\right) - d\frac{N_t}{K}Z_t \\
=& \frac{b}{N_t}(r^Z)^2 - d\frac{N_t}{K}Z_t \qquad \text{(S1c)}
\end{aligned}$$

where

$$r^X = \left((1-s)X_t + (1-hs)Y_t\frac{(1+\mu)}{2} + Z_t\mu\right) \qquad \text{(S2a)}$$

and

$$r^Z = \left(Z_t(1-\mu) + (1-hs)Y_t\frac{(1-\mu)}{2}\right). \qquad \text{(S2b)}$$

Selection on fecundity

$$\begin{aligned}
\frac{dX_t}{N_t} =& \frac{b}{(1-s)X_t + (1-hs)Y_t + Z_t}\Big((1-s)^2X_t^2 + 2(1-s)X_tZ_t\mu + Z_t^2\mu^2 \\
&+ (1-s)(1-hs)X_tY_t(1+\mu) + (1-hs)Y_tZ_t\mu(1+\mu) + \frac{1}{4}(1-hs)^2Y_t^2(1+\mu)^2\Big) \\
&- d\frac{N_t}{K}X_t \\
=& \frac{b}{N_t - sX_t - hsY_t}(r^X)^2 - d\frac{N_t}{K}X_t \qquad \text{(S3a)} \\
\frac{dY_t}{N_t} =& \frac{b}{(1-s)X_t + (1-hs)Y_t + Z_t}\Big((1-s)(1-hs)X_tY_t(1-\mu) + 2(1-s)X_tZ_t(1-\mu) \\
&+ 2Z_t^2(1-\mu)\mu + \frac{1}{2}(1-hs)^2Y_t^2(1-\mu)(1+\mu) \\
&+ 2(1-hs)Y_tZ_t\left(\frac{1}{2}(1-\mu)\mu + \frac{1}{2}(1-\mu)(1+\mu)\right)\Big) - d\frac{N_t}{K}Y_t \\
=& \frac{2b}{N_t - sX_t - hsY_t}(r^Xr^Z) - d\frac{N_t}{K}Y_t \qquad \text{(S3b)} \\
\frac{dZ_t}{N_t} =& \frac{b}{(1-s)X_t + (1-hs)Y_t + Z_t}\Big(\frac{1}{4}(1-hs)^2Y_t^2(1-\mu)^2 + (1-hs)Y_tZ_t(1-\mu)^2 \\
&+ Z_t^2(1-\mu)^2\Big) - d\frac{N_t}{K}Z_t \\
=& \frac{b}{N_t - sX_t - hsY_t}(r^Z)^2 - d\frac{N_t}{K}Z_t \qquad \text{(S3c)}
\end{aligned}$$

where

$$r^X = \left((1-s)X_t + (1-hs)Y_t\frac{(1+\mu)}{2} + Z_t\mu\right) \qquad \text{(S4a)}$$

and

$$r^Z = \left(Z_t(1-\mu) + (1-hs)Y_t\frac{(1-\mu)}{2}\right). \qquad \text{(S4b)}$$

Selection on zygote survival

$$\begin{aligned}
\frac{dX_t}{N_t} =& \frac{b(1-s)}{N_t}\left(X_t^2 + 2X_tZ_t\mu + Z_t^2\mu^2 + X_tY_t(1+\mu) + Y_tZ_t\mu(1+\mu)\right. \\
&\left. + \frac{1}{4}Y_t^2(1+\mu)^2\right) - d\frac{N_t}{K}X_t \\
=& \frac{b(1-s)}{N_t}(r^X)^2 - d\frac{N_t}{K}X_t \qquad \text{(S5a)} \\
\frac{dY_t}{N_t} =& \frac{b(1-hs)}{N_t}\left(X_tY_t(1-\mu) + 2X_tZ_t(1-\mu) + 2Z_t^2(1-\mu)\mu\right. \\
&\left. + \frac{1}{2}Y_t^2(1-\mu^2) + Y_tZ_t\left(1+\mu-2\mu^2\right)\right) - d\frac{N_t}{K}Y_t \\
=& \frac{b(1-hs)}{N_t}(r^Xr^Z) - d\frac{N_t}{K}Y_t \qquad \text{(S5b)} \\
\frac{dZ_t}{N_t} =& \frac{b}{N_t}\left(\frac{1}{4}Y_t^2(1-\mu)^2 + Y_tZ_t(1-\mu)^2 + Z_t^2(1-\mu)^2\right) - d\frac{N_t}{K}Z_t \\
=& \frac{b}{N_t}(r^Z)^2 - d\frac{N_t}{K}Z_t \qquad \text{(S5c)}
\end{aligned}$$

where

$$r^X = \left(X_t + Y_t\frac{(1+\mu)}{2} + Z_t\mu\right) \qquad \text{(S6a)}$$

and

$$r^Z = \left(Z_t(1-\mu) + Y_t\frac{(1-\mu)}{2}\right). \qquad \text{(S6b)}$$

Selection on adult survival

$$\begin{aligned}\frac{dX_t}{N_t} =& \frac{b}{(1-s)X_t+(1-hs)Y_t+Z_t}\Big((1-s)^2X_t^2+2(1-s)X_tZ_t\mu+Z_t^2\mu^2\\ &+(1-s)(1-hs)X_tY_t(1+\mu)+(1-hs)Y_tZ_t\mu(1+\mu)+\frac{1}{4}(1-hs)^2Y_t^2(1+\mu)^2\Big)\\ &-d\frac{(1-s)X_t+(1-hs)Y_t+Z_t}{K}X_t\\ =& \frac{b}{N_t-sX_t-hsY_t}(r^X)^2-d\frac{(1-s)X_t+(1-hs)Y_t+Z_t}{K}X_t \quad \text{(S7a)}\\ \frac{dY_t}{N_t} =& \frac{b}{(1-s)X_t+(1-hs)Y_t+Z_t}\Big((1-s)(1-hs)X_tY_t(1-\mu)+2(1-s)X_tZ_t(1-\mu)\\ &+2Z_t^2(1-\mu)\mu+\frac{1}{2}(1-hs)^2Y_t^2(1-\mu)(1+\mu)\\ &+2(1-hs)Y_tZ_t\left(\frac{1}{2}(1-\mu)\mu+\frac{1}{2}(1-\mu)(1+\mu)\right)\Big)-d\frac{(1-s)X_t+(1-hs)Y_t+Z_t}{K}Y_t\\ =& \frac{2b}{N_t-sX_t-hsY_t}(r^Xr^Z)-d\frac{(1-s)X_t+(1-hs)Y_t+Z_t}{K}Y_t \quad \text{(S7b)}\\ \frac{dZ_t}{N_t} =& \frac{b}{(1-s)X_t+(1-hs)Y_t+Z_t}\Big(\frac{1}{4}(1-hs)^2Y_t^2(1-\mu)^2+(1-hs)Y_tZ_t(1-\mu)^2\\ &+Z_t^2(1-\mu)^2\Big)-d\frac{(1-s)X_t+(1-hs)Y_t+Z_t}{K}Z_t\\ =& \frac{b}{N_t-sX_t-hsY_t}(r^Z)^2-d\frac{(1-s)X_t+(1-hs)Y_t+Z_t}{K}Z_t \quad \text{(S7c)}\end{aligned}$$

where

$$r^X=\left((1-s)X_t+(1-hs)Y_t\frac{(1+\mu)}{2}+Z_t\mu\right) \quad \text{(S8a)}$$

and

$$r^Z=\left(Z_t(1-\mu)+(1-hs)Y_t\frac{(1-\mu)}{2}\right). \quad \text{(S8b)}$$

File S3

Genotypic frequencies at equilibrium

Hardy-Weinberg genotypic frequencies at equilibrium

We will prove that in our model, the expectations at Hardy-Weinberg equilibrium are met when there is no selection or mutation.

Proposition 1. *When there is no selection or mutation (s and $\mu = 0$), the genotypic frequencies in our model at are at Hardy-Weinberg equilibrium.*

Proof of Proposition 1. Let q be the frequency of allele a, defined by

$$q = \frac{2X_{eq} + Y_{eq}}{2N_{eq}}, \tag{S9}$$

and $1 - q$ is the frequency of A.
First, at Hardy-Weinberg equilibrium the frequency of each genotype is q^2, $2q(1-q)$ and $(1-q)^2$ for genotypes aa, Aa and AA respectively. From the expected genotype frequencies, q can be expressed as a function of the frequency of heterozygotes (F^{Aa}).
From $F^{Aa} = 2q(1-q)$ we obtain the expected frequency of a:

$$p = \frac{1 \pm \sqrt{1 - 2FAa}}{2} \tag{S10}$$

Let us consider our model and prove that we have the same frequency of the allele a. Equilibrium is defined by $\frac{dV_t}{d_t} = 0$, where V represents X, Y and Z. When there is no mutation or selection, the number of homozygous individuals in the population (either X_{eq} or Z_{eq}, both noted H_{eq}) can be written as a function of the number of number of Y_{eq} (or heterozygous) individuals:

$$\frac{b}{N_{eq}}\left(H_{eq}^2 + Y_{eq}H_{eq} + \frac{1}{4}Y_{eq}^2\right) - d\frac{N_{eq}}{K}H_{eq} = 0.$$

By using equation 6 from the main text we can simplify the previous equation to:

$$\frac{{Y_{eq}}^2}{4N_{eq}} + H_{eq}\left(\frac{Y_{eq}}{N_{eq}} - 1\right) + \frac{{H_{eq}}^2}{N_{eq}} = 0.$$

This gives us a quadratic equation in H_{eq} with discriminant

$$\Delta = \left(1 - 2\frac{Y_{eq}}{N_{eq}}\right) \geq 0$$

This gives two possible solutions, corresponding to frequencies of homozygous individuals greater or smaller than $\frac{1}{4}N_{eq}$:

$$H_{eq} = \frac{1}{2}\left(N_{eq} - Y_{eq} \pm N_{eq}\sqrt{1 - 2\frac{Y_{eq}}{N_{eq}}}\right) \tag{S12}$$

If we replace X_{eq} with equation S12 in equation S9, then we find that

$$q = \frac{N_{eq} \pm N_{eq}\sqrt{1 - 2\frac{Y_{eq}}{N_{eq}}}}{2N_{eq}}.$$

Which is equation S10 and the proposition is proved. □

Genotypic frequencies at mutation-selection balance

The genotypic frequencies of aa, Aa and AA individuals are given by the expressions $F^{X_{mut}}$, $F^{Y_{mut}}$ and $F^{Z_{mut}}$ respectively. For each of the models, we provide the expressions for all values of the dominance h of the deleterious allele a between 0 and 1 and different than 0.5, as well as the seperate expressions for $h = 0.5$. The domain of validity of these equations depends on the threshold value of the mutation rate for the deterministic fixation of the deleterious allele a $\mu_{fix} = \frac{(1-h)s}{1-hs}$ and on the validity of $H_1 = \sqrt{s\left((4-8h)\mu + h^2(1+\mu)^2 s\right)}$ (see File $S1$ for more details). These expressions were obtained using Wolfram's Mathematica 9.0.

Selection on mating success, fecundity and adult survival

For $h \neq 0.5$:

$$F^{X_{mut}} = \frac{(2-4h)\mu + h^2(1+\mu)^2 s - hH_1 - h\mu H_1}{2(1-2h)^2 s}$$

$$F^{Y_{mut}} = \frac{\mu(h(4-s+H_1)-2) - (1-h)(hs-H_1) - h^2\mu^2 s}{(1-2h)^2 s}$$

$$F^{Z_{mut}} = \frac{(2-h(6-5h))s + h^2\mu^2 s - (2-3h)H_1 + \mu(2-h(4-2(1-h)s+H_1))}{2(1-2h)^2 s}$$

For $h = 0.5$

$$F^{X_{mut}} = \frac{4\mu^2}{(1+\mu)^2 s^2}$$
$$F^{Y_{mut}} = \frac{4\mu(s-\mu(2-s))}{(1+\mu)^2 s^2}$$
$$F^{Z_{mut}} = \frac{(s-\mu(2-s))^2}{(1+\mu)^2 s^2}$$

Selection on zygote survival

For $h \neq 0.5$

$$F^{X_{mut}} = \frac{(1-s)\,(2\mu - 4h\mu + h^2(1+\mu)s - hH_1)}{2(1-2h)(1-m)s(1-h(2-hs))}$$
$$F^{Y_{mut}} = \frac{2(1-hs)((1-h)(H_1-hs) - \mu(2-h(4-s+H_1))) - h^2\mu^2 s}{(1-2h)(1-\mu)s\,(2+h^2(1+\mu)s - h(4+H_1))}$$
$$F^{Z_{mut}} = \frac{(2-h(6-5h))s + h^2\mu^2 s - (2-3h)H_1 + \mu(2-h(4-2(1-h)s+H_1))}{(1-2h)(1-\mu)s\,(2+h^2(1+\mu)s - h(4+H_1))}$$

For $h = 0.5$

$$F^{X_{mut}} = \frac{4\mu^2(1-s)}{(1-\mu^2)s^2}$$
$$F^{Y_{mut}} = \frac{2\mu(2-s)(s-\mu(2-s))}{(1-\mu^2)s^2}$$
$$F^{Z_{mut}} = \frac{(s-\mu(2-s))^2}{(1-\mu^2)\,s^2}$$

File S4

Proofs for population size at equilibrium

General expressions for population size at equilibrium found in Table SS1 are obtained using Wolfram's Mathematica 9.0 by solving for the steady state solutions of the differential equations $\frac{dN_t}{dt} = \frac{dX_t}{dt} = \frac{dY_t}{dt} = \frac{dZ_t}{dt} = 0$. Expressions and the proofs for population size at equilibrium without selection and for recurrent recessive mutations are presented in this section.

No selection

When there is no selection, the expression for population size at equilibrium N_{eq} is given by equation 6 in the main text.

Proposition 2. *At equilibrium and when there is no selection ($s = 0$), the equation $\frac{dN_t}{dt} = \frac{dX_t}{dt} = \frac{dY_t}{dt} = \frac{dZ_t}{dt} = 0$ admits a unique attractive non-trivial solution N_{eq} (equation 6 in the main text).*

Proof of proposition 2. Population size is at equilibrium when

$$R_t^X + R_t^Y + R_t^Z - M_t^X - M_t^Y - M_t^Z = 0. \tag{S13}$$

When this condition is met, then population size is noted N_{eq}, and X_{eq}, Y_{eq} and Z_{eq} are the numbers of individuals carrying each genotype, with $N_{eq} = X_{eq} + Y_{eq} + Z_{eq}$. Equation S13 therefore gives:

$$\frac{b}{N_{eq}}\left(X_{eq}^2 + Y_{eq}^2 + Z_{eq}^2 + 2X_{eq}Y_{eq} + 2Y_{eq}Z_{eq} + 2X_{eq}Z_{eq}\right) - d\frac{N_{eq}}{K}\left(X_{eq} + Y_{eq} + Z_{eq}\right) = 0.$$

leading to equation 6. □

Mutation-selection balance

Population size at mutation-selection balance N_{mut} (s and $\mu \neq 0$) can be presented as a function of population size at equilibrium with no or neutral mutations N_{eq}. For each model of selection (mating success, fecundity and survival), we find two attractive non-trivial solutions, one for a polymorphic population (with X, Y and Z individuals) and one which is monomorphic as the deleterious allele a is fixed (for $\mu \geq \frac{(1-h)s}{1-hs}$). Proofs for N_{mut} for recessive mutations ($h = 0$) can be found in the following section for selection on mating success, fecundity and zygote survival.

Selection on mating success

Proposition 3. *There are two expressions for N_{mut} when mutations are recessive:*

$$N_{mut} = N_{eq}(1-\mu\wedge s)^2 \tag{S15}$$
$$\mu\wedge s = min(\mu, s)$$

Proof of proposition 3. At equilibrium, $\frac{dN_t}{d_t} = \frac{dX_t}{d_t} = \frac{dY_t}{d_t} = \frac{dZ_t}{d_t} = 0$. From equations S1 and S2 in File S2, we find:

$$\begin{aligned}
0 &= \frac{b}{N_{mut}}((r^X)^2 + 2(r^X)(r^Z) + (r^Z)^2) - d(X_{mut}+Y_{mut}+Z_{mut})\frac{N_{mut}}{K} \\
\Leftrightarrow 0 &= \frac{b}{N_{mut}}(r^X + r^Z)^2 - d\frac{N_{mut}^2}{K} \\
\Leftrightarrow 0 &= \frac{b}{N_{mut}}((1-s)X_{mut} + Y_{mut} + Z_{mut})^2 - d\frac{N_{mut}^2}{K} \\
\Leftrightarrow 0 &= \frac{b}{N_{mut}}(N_{mut} - sX_{mut})^2 - d\frac{N_{mut}^2}{K}
\end{aligned}$$

$N_{mut} = 0$ is a trivial solution, else N_{mut} satisfies

$$N_{mut} = \frac{bK}{d}(1-\frac{sX_{mut}}{N_{mut}})^2. \tag{S16}$$

A second equation for N_{mut} can be obtained by finding an expression for the change in the number of A alleles at mutation-selection $\frac{dN^A}{dt} = 0$, where $N^A = 2Z_t + Y_t$. From equations S1 and S2, and for $Z_{mut} + Y_{mut} \neq 0$ we find:

$$\begin{aligned}
0 &= \frac{b}{N_{mut}}2r^Z(r^X + r^Z) - d(2Z_{mut} + Y_{mut})\frac{N_{mut}}{K} \\
\Leftrightarrow 0 &= \frac{b(1-\mu)}{N_{mut}}(2Z_{mut} + Y_{mut})(N_{mut} - sX_{mut}) - d(2Z_{mut} + Y_{mut})\frac{N_{mut}}{K} \\
\Leftrightarrow 0 &= (2Z_{mut} + Y_{mut})(b(1-\mu)(1-\frac{sX_{mut}}{N_{mut}}) - d\frac{N_{mut}}{K}) \\
\Leftrightarrow 0 &= (b(1-\mu)(1-\frac{sX_{mut}}{N_{mut}}) - d\frac{N_{mut}}{K})
\end{aligned}$$

$$\Leftrightarrow N_{mut} = \frac{bK}{d}(1-\mu)(1-\frac{sX_{mut}}{N_{mut}}). \tag{S17}$$

Using equations S16 and S17 we find:

$$b(1-\frac{sX_{mut}}{N_{mut}})^2 = b(1-\mu)(1-\frac{sX_{mut}}{N_{mut}})$$

which implies

$$(1-\frac{sX_{mut}}{N_{mut}}) = (1-\mu).$$

By using this equality in equation S16 or S17, we obtain equation S15

If $N_{mut} = X_{mut}$, we consider that there is fixation of the a allele. Using this information in equation S16, we find that

$$N_{mut} = N_{eq}(1-s)^2 \tag{S19}$$

as proposed in equation S15. □

Selection on fecundity

Proposition 4. *When selection is on fecundity*

$$N_{mut} = N_{eq}(1-\mu \wedge s) \tag{S20}$$

Proof of proposition 4. Using similar calculations as in Proposition 3, we find an expression for N_{mut} using the expression for N^A. From equations S3 and S8 we find:

$$\begin{aligned}
0 &= 2(R^Z_{mut} - dZ_{mut}\frac{N_{mut}}{K}) + (R^Y_{mut} - dY_{mut}\frac{N_{mut}}{K}) \\
\Leftrightarrow 0 &= \frac{b(1-\mu)}{N_{mut}-sX_{mut}}(2Z_{mut}+Y_{mut})(N_{mut}-sX_{mut}) \\
&\quad - d(2Z_{mut}+Y_{mut})\frac{N_{mut}}{K} \\
\Leftrightarrow 0 &= (2Z_{mut}+Y_{mut})(b(1-\mu) - d\frac{N_{mut}}{K})
\end{aligned}$$

which implies for $Z_{mut} + Y_{mut} \neq 0$

$$N_{mut} = N_{eq}(1-\mu). \tag{S21}$$

If $N_{mut} = X_{mut}$, then the mutant allele a has gone to fixation, implying that $Z_{mut} = Y_{mut} = 0$. This gives

$$\begin{aligned}
0 &= \frac{b}{(1-s)X_{mut}}((1-s)X_{mut})^2 - d\frac{X^2_{mut}}{K} \\
\Leftrightarrow 0 &= b(1-s) - d\frac{X_{mut}}{K} \\
N_{mut} &= N_{eq}(1-s)
\end{aligned} \tag{S22}$$

as proposed in equation S20. □

Selection on zygote survival

Proposition 5. *When mutations are recessive*

$$N_{mut} = N_{eq}(1 - \mu \wedge s) \tag{S23}$$

Proof of proposition 5. The number of A alleles is noted N^A and as in Proposition 3, from equations S5 and S6 we find:

$$\begin{aligned} 0 &= \frac{b}{N_{mut}} 2r^Z(r^X + r^Z) - d(2Z_{mut} + Y_{mut})\frac{N_{mut}}{K} \\ \Leftrightarrow 0 &= b(1-\mu)(2Z_{mut} + Y_{mut}) - d(2Z_{mut} + Y_{mut})\frac{N_{mut}}{K} \\ \Leftrightarrow b(1-\mu) &= d\frac{N_{mut}}{K} \end{aligned}$$

providing $N_{mut} = N_{eq}(1 - \mu)$.

If we consider that there is fixation of the a allele, then the entire population is made of X individuals, then Y_{mut} and $Z_{mut} = 0$,

$$0 = \frac{b}{N_{mut}}(1-s)N_{mut}^2 - d\frac{N_{mut}^2}{K}$$

gives

$$N_{mut} = N_{eq}(1 - s) \tag{S24}$$

if $N_{mut} \neq 0$. □

Abstract

As the ultimate source of genetic variation, mutation has the inconvenience of introducing deleterious mutations. These mutations shape the evolution of species, from genetic mechanisms on the cellular level to reproductive systems, which lessen their effects on fitness. In this thesis we explore how these mutations influence population size by allowing the interaction between population size and selection, which has been little explored in conventional population genetics models. In a deterministic context with a single locus, germ-line and somatic mutations influence population size and the mutation load, both of which depend on the timing of the expression of these mutations. Multi-locus individual based models show that population viability depends on the demographic properties, the rate of introduction and impact of mutations. Though self-fertilisation generally increases population viability, strictly self-fertilising populations go extinct due to mutational meltdown when mutations are of small effect. When selfing is allowed to evolve from an outcrossing reproductive regime, there are cases of evolutionary suicide where strict selfing evolves and leads to extinction. We predict that the genetic properties of populations may not be a consequence but a cause of population size. We emphasize the importance of taking the demographic consequences of deleterious mutations into account when studying the evolution of populations, as in the case of the evolution of self-fertilisation where evolutionary suicide was observed. This result may explain the observed higher extinction rates in selfing compared to outcrossing species.

***K*eywords : Genetic load, population size, self-fertilisation, evolutionary suicide, extinction, somatic mutations.**

Résumé

La présence des mutations délétères a favorisé l'évolution de mécanismes, au niveau cellulaire et au niveau des organismes (ex. les régimes de reproduction), permettant de diminuer leurs effets négatifs. Au cours de cette thèse nous avons étudié leur impact sur la taille des populations à travers des modèles tenant compte de l'interaction entre la démographie et la sélection, cette interaction étant souvent mise de coté dans les modèles conventionnels de génétique des populations. Dans un contexte déterministe à un seul locus des mutations somatiques et gamétiques influencent la taille et le fardeau génétique des populations (ces derniers étant dépendants du moment d'expression des mutations dans le cycle de vie). Nos modèles stochastiques avec un grand nombre de locus indiquent que la viabilité des populations dépend des paramètres démographiques et génétiques (taux de mutation, effet délétère des mutations). L'autofécondation est généralement avantageuse, augmentant la taille et la viabilité des populations, mais lorsque les mutations sont de faible effet un régime d'autogamie stricte mène à l'extinction par fonte mutationnelle. En permettant l'évolution de l'autofécondation à partir d'une population allogame nous observons des cas de suicide évolutif où les populations évoluent vers l'autogamie stricte et s'éteignent, ce qui pourrait expliquer les taux d'extinctions élevés des espèces auto-fécondantes comparées aux allo-fécondantes. Ces modèles prédisent que la taille des populations pourrait être une conséquence et non une cause de leurs propriétés génétiques, soulignant l'importance de la prise en compte leur interaction dans l'étude de l'évolution des populations.

***Mots clés* : Fardeau génétique, taille de populations, auto-fécondation, suicide evolutif, extinction, mutations somatiques.**