Évolution des cycles de vie : modélisation et évolution expérimentale sur la levure Saccharomyces cerevisiae

Marie Rescan

THÈSE DE DOCTORAT DE
l'UNIVERSITÉ PIERRE ET MARIE CURIE

École doctorale 227 Sciences de la Nature et de l'Homme : évolution et écologie
UMI 3614 Evolutionary Biology and Ecology of Algae

Evolution des cycles de vie: Modélisation et évolution expérimentale sur la levure *Saccharomyces cerevisiae*

Présentée par
Marie Rescan

Dirigée par Myriam Valero et Denis Roze

Pour obtenir le grade de
DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Jury:

Mme. Simone Immler	Rapporteur
M. Sylvain Billiard	Rapporteur
M. Nicolas Loeuille	Examinateur
M. Olivier Tenaillon	Examinateur
M. Denis Roze	Directeur de thèse
Mme. Myriam Valero	Directeur de thèse

Contents

Introduction

Contents

1 Hofmeister-Strasburger alternation of generations

Alternation of meiosis and syngamy in sexual organisms results in the alternation of haploid and diploid generations. Haploid individuals carry one unique copy of each gene, while diploid individuals have pairs of homologous chromosomes and carry two copies. Alternation between haploid and diploid nuclear phases was first described by Hofmeister (1851) and Strasburger (1894). From a theoretical perspective, the evolutionary forces affecting the evolution of life

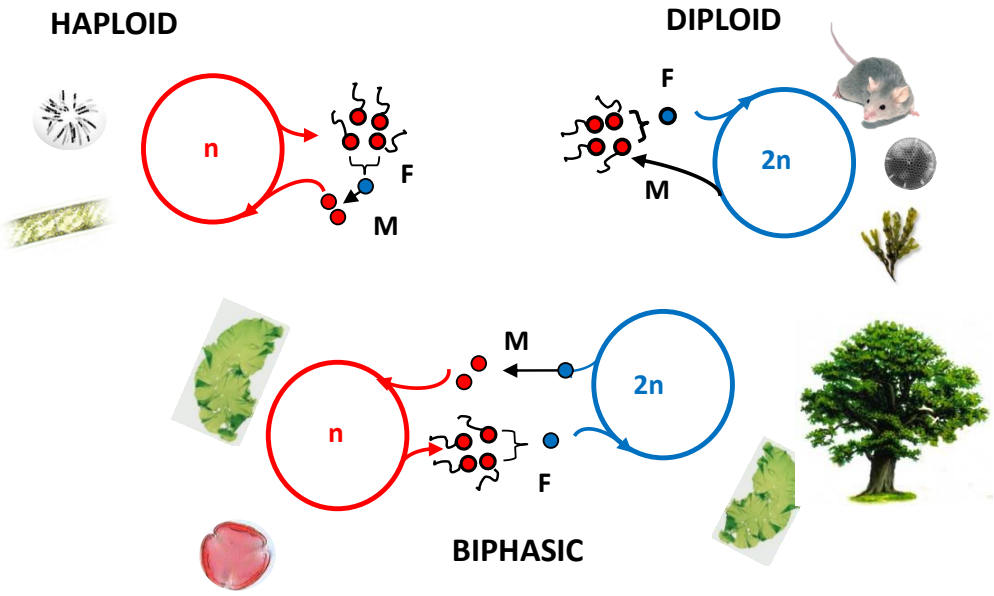

Figure 1: Haploid, diploid and biphasic life cycles. Blue and red arrows represent mitotic development of the diploid and the haploid phase, respectively. Red dots: haploid spores. Red dots with flagella: haploid gametes. Blue dots: diploid zygote. Abbreviations: f, fecondation; m, meiosis.

cycles remained little explored before the end of the 20th century, when several scientists began to study life cycle diversity, and to propose hypotheses explaining the evolution of diploidy, and the maintenance of a haploid phase.

1.1 Structure of haploid diploid life cycles

Length and development of both phases vary largely between taxa (Valero et al. 1992, Mable and Otto 1998, Otto and Gerstein 2008). Some species, including diatoms, oocmycetes and most animals have a diploid life cycle: the haploid phase is reduced to the gametes, which undergo syngamy before any mitotic division. In haploid species (ascomycetes, charophytes, dinoflagellates), the diploid phase is reduced to one single cell, the zygote. Finally, in many species, vegetative growth occurs both in the haploid and the diploid phases, delimiting two generations (haploid-diploid cycle). Gametophytic generations are generally haploid and produce gametes, while sporophytic generations are generally diploid and produce spores. However, alternation of generation and sexual reproduction do not always correspond: asexual organisms can reproduce without alternation of meiosis and syngamy, and some life cycles involve an alternation of morphologically different generations with the same ploidy level (Steenstrup alternation of phases). For instance, in rhodophytes (red algae), the diploid carposporohyte produces mitotically diploid spores, which develop into diploid tetrasporophytes, producing haploid spores through meiosis (Bell 1994).

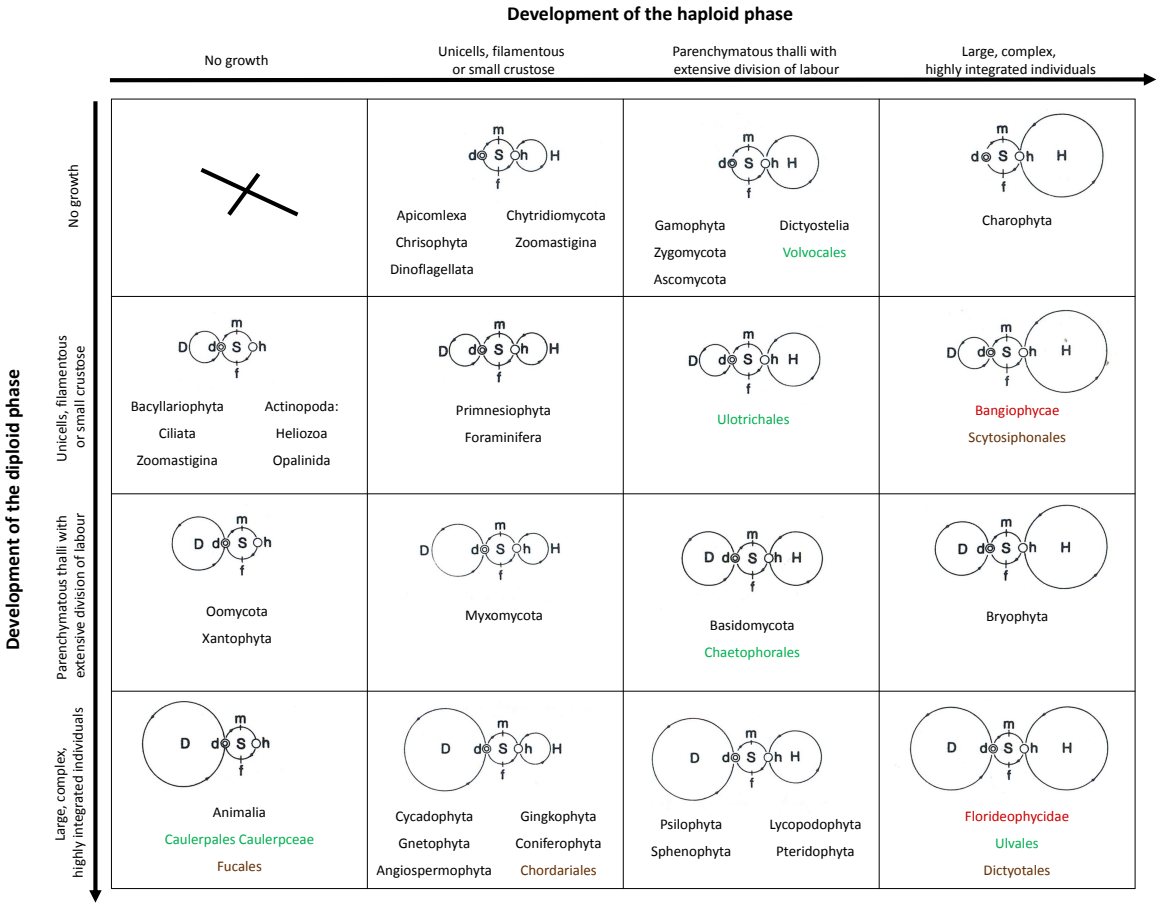

Figure 2: *Modified from Bell (1994).* Diversity of life cycles among eukaryotes. Life cycles are classified according to the size and complexity of the haploid and the diploid phases. Phaeophytes, rhodophytes and chlorophytes exhibit a large range of life cycles and examples are given in brown, red and green respectively. Abbreviations: H, haploid vegetative growth; D, diploid vegetative growth; S, sexual cycle; f, gamete fusion; m, meiosis; h, haplospores; d, diplospores.

1.2 Diversity of life cycles

The relative development of the haploid and the diploid phases is extremely variable between species (Figure 2). When both phases present a vegetative development, the life cycle may be isomorphic, with morphologically similar haploid and diploid generations, or heteromorphic. For instance, in *Dictyotales*, the gametophyte and the sporophyte have morphologically identical thalli, in *Porphyra*, the gametophyte is macroscopic while the sporophyte is microscopic, and in *Laminaria*, the gametophyte is microscopic. Some large phyla are fixed for haploid or diploid life cycles (charophytes and zygomycetes are purely haplontic, while animals and ciliates are strictly diplontic), suggesting that life cycle evolution may be narrowly constrained in those phyla. However, in phyla where both haploid and diploid individuals undergo mitotic development, their is a great deal of variation in the relative degree of development of both phases. For instance, chlorophyta (green algae), phaeophyta (brown algae) and rhodophytoa

(red algae) display a large range of intermediate life cycles (Figure 2). In particular, in phaeophyta, isomorphic haploid diploid life cycle (some ectocarpales) to purely diploid life cycles (fucales) are present. When the different types of brown algal life cycles are mapped onto a phylogenetic tree (Figure 3), the distribution pattern suggests that there has been considerable switching between different life cycle strategies during the evolution of this group (Cock et al. 2014). Embryophyte (land plants) is another group presenting a wild diversity of life cycles, ranging from dominantly haploid (mosses, liveworts) to dominantly diploid life cycles (ferns, seed plants). Phylogenetical studies reveal that land plants life cycle evolve from a dominantly haploid life cycles (Niklas and Kutschera 2010, Qiu et al. 2012).

At a lower taxonomic level, some groups also present a large diversity (Bell 1994): in the Ulvales, *Ulvaria* present an isomorphic alternation between generations, while in *Monostroma*, the sporophyte is extremely reduced, contrary to *Kornmannia*, whose gametophyte is a tiny disc. Finally, in animals, several groups have independently evolved variation in ploidy as a sex determination system: in arrhenotokous species, females are diploid and males develop from an unfertilized egg and are haploid. This sex determination system is systematic in hymenoptera and monogonont rotifers, and can be found sporadically in hemiptera, in coleoptera (bark beetles) and in one mite specie (see references in Mable and Otto 1998, Immler and Otto 2014). Note that even in classic XY or WZ sex determination systems, the heterogametic sexual chromosome is always under haploid selection.

2 Evolution of life cycles

2.1 Physiological differences between phases

Changes in ploidy can have immediate phenotypic effects, which may favor one or the other phase. First, diploids, with 2 copies of each gene, have the material to produce twice as many proteins and should be able to grow faster than haploids in rich environments (Bell 1989). On the other hand, replication of a large genome requires more nutrients, which should favor haploid cells when the environment is limitant. Then, genome and cell size tend to be positively correlated (Cavalier Smith 1985). For example, many large unicellular eukaryotes are diploid (ciliates, foraminiferans) or polyploid (euglenids, amoebas) (Bell 1994). Conditions favoring large cells should therefore favor diploidy (and reciprocally, conditions favoring small cells should favor haploidy). Small cells have a higher surface to volume ratio, which confers an advantage to haploidy when nutrients are limitant (Lewis 1985), but should favor diploids in toxic environments.

However, Bessho et al. (2015), using a mathematical model, found that the advantage of diploids in rich medium depends not only of their size, but also on their energy conversion

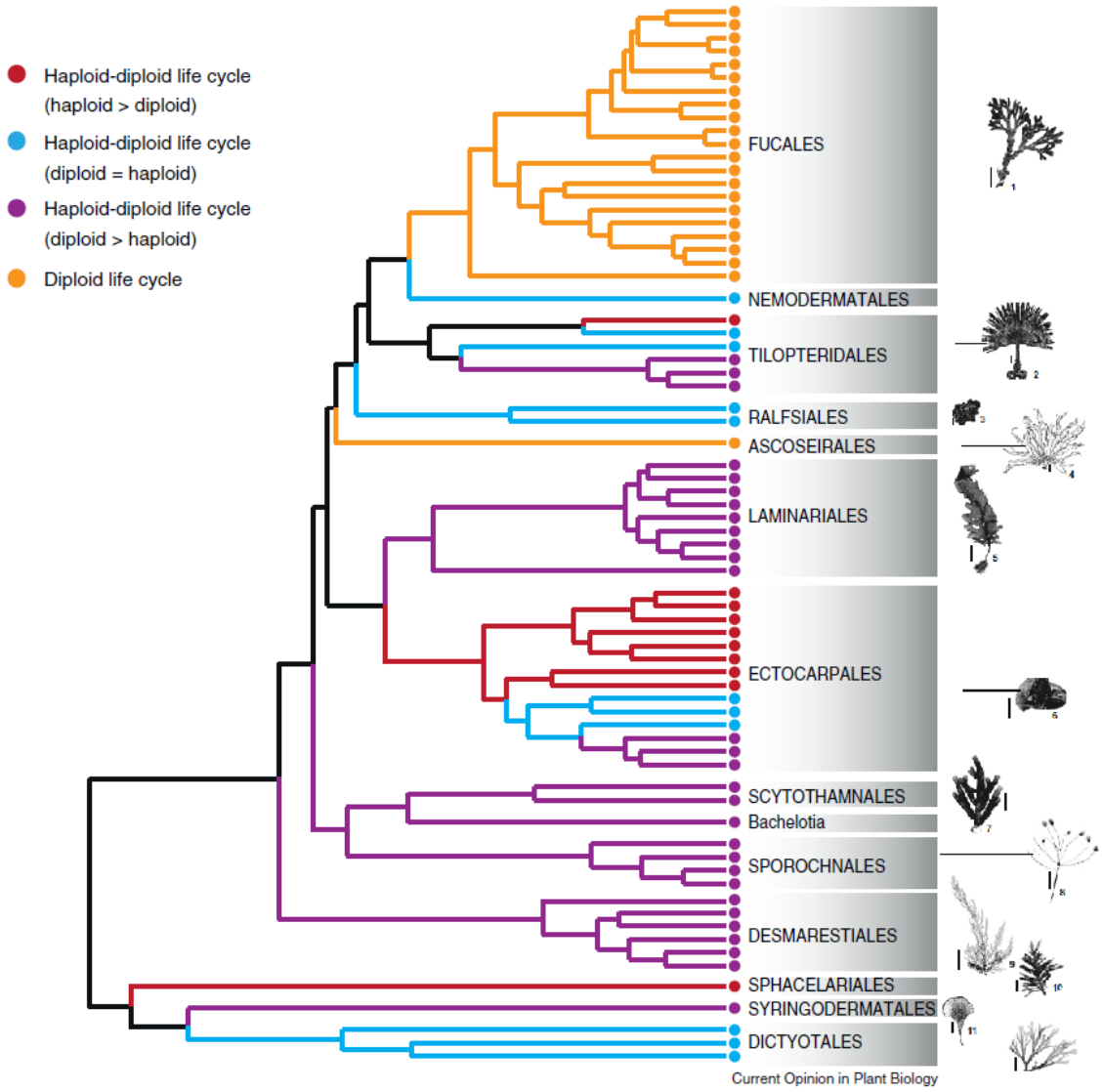

Figure 3: *Modified from Cock et al. (2014).*The brown algae exhibit a broad range of life cycle types. Life cycle type has been mapped onto the phylogenetic tree of the brown algae (*Phaeophyceae*)

efficiency and on the relation between cell size and mortality. Finally, in multicellular species, cell geometry is unlikely to influence the evolution of ploidy, as having more somatic cell may compensate for their smaller size. Finally, there is little experimental evidence in support of the theory assuming that life cycles evolve in response to cell size difference between ploidy phases (Zeyl 2004).

2.2 Genetic models

In addition to the physiological consequences discussed above, a change in ploidy level is a major genomic change. Life cycles are often characterized is term of size, duration and complexity of the haploid and the diploid phases (see Figure 2), but to explain the effect of genetic factors on the evolution of life cycles, models need to recast the relative importance of each phase in term of opportunity for selection within each phase. In fact, short life length and small size may strongly limit the opportunity for selection in one phase, in particular when its growth depends on the other phase. For instance, in many female animals, there is virtually no haploid phase (since the last meiotic division of the egg only takes place at fertilization). By contrast, much selection can occur among male gametes, even though they have very limited development (Joseph and Kirkpatrick 2004). In this case, although no mitotic division occurs during the haploid phase, it may have a strong evolutionary importance.

In distinct populations of haploids and diploids, theoretical models (Haldane 1937) predict that a deleterious allele generated by recurrent deleterious mutation generates a load that is twice as big in diploid populations than in haploid ones ($L_h = u$ while $L_d = 2u$, where u is the mutation rate from the wild type to the deleterious allele), as long as the dominance coefficient of the mutation (h) is significantly greater than zero. Indeed, although recessive mutations have smaller fitness effects ($h\ s$) in diploids, they rise to higher equilibrium frequencies in diploid than haploid populations, and a diploid may carry twice as many mutations as a haploid. Under synergetic epistasis (i.e, when several deleterious mutations have a greater effect than if they acted independently), the mutation load is reduced, and this reduction is more important in diploids than in haploids. The diploid mutation load can even become lower than the haploid load with truncation selection (Kondrashov and Crow 1991).

Modifier models

The relative mutation load of haploids and diploids cannot be used to make predictions about the evolution of life cycles, because in a biphasic population, ploidy phases are linked by sexual reproduction. To consider the dynamic of a haploid-diploid population subject to mutations, theoretical studies use modifier models: they analyze the evolution of a modifier

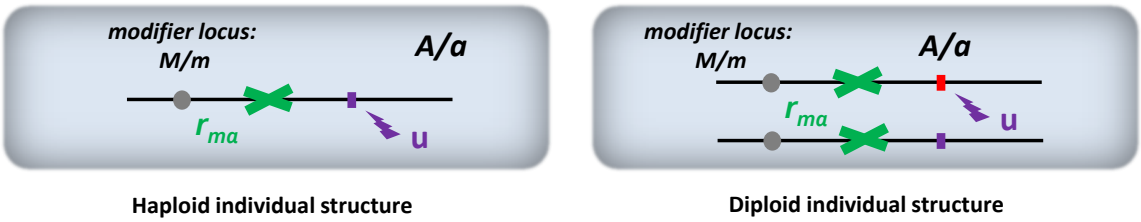

Figure 4: Individuals structure in the modifier model. Abbreviations: r_{ma}, recombination rate between alleles m and a; u, mutation rate from A to a.

locus affecting either the probability to undergo diploid rather than haploid selection (e.g. Otto and Goldstein 1992) or the relative length d of the diploid phase (e.g. Jenkins 1993, Jenkins and Kirkpatrick 1995).

Two-locus models consider a modifier locus with allele M (resident) or m (mutant) linked with recombination rate r_{ma} to a locus at which two alleles (A and a) directly affect fitness. This generates 4 possible haplotypes: MA, Ma, mA and ma. The association between alleles m and a at both loci can be measured by the linkage disequilibrium $D_{ma} = p_{ma} - p_m p_a$.

The model dynamics can be analyzed using two different methods: linear stability analysis and Quasi Linkage Equilibrium (QLE) approximation. Linear stability analysis is used to determine when a mutant allele can increase in frequency from rarity. At the equilibrium where the modifier allele M is fixed ($p_m = 0$), one considers the introduction of a small number of individual with a new modifier allele m ($p_m = p_{mA} + p_{ma}$). Since the perturbation is small, quadratic terms in p_{ma} and p_{mA} can be ignored. Linear recursions for p_{mA} and p_{ma} are produced and the eigenvalue λ is determined. If $\lambda > 1$, the equilibrium with M fixed is unstable and an increased (if $d_{Mm} > d_{MM}$) or decreased ($d_{Mm} < d_{MM}$) level of diploidy is favored. Equilibrium strategies M^* are resistant to the invasion of any other modifier allele. The equilibrium is convergence stable if $\partial\lambda/\partial d \left|_{d=d_{M^*M^*}} < 0\right.$. Otherwise, M^* is a repelling point. The equilibrium is then evolutionary stable if $\partial\lambda^2/\partial d^2 \left|_{d=d_{M^*M^*}} < 0\right.$, otherwise, strategy M^* is a branching point: the population reaches M^*, then splits into several strategies.

In the QLE analysis, selection is assumed to be weak relative to recombination ($s << r_{ma}$) so that linkage disequilibrium equilibrates fast relative to alleles frequencies. D_{ma} can therefore be expressed as a function of p_a and p_m. The change in frequency of a new modifier allele m, Δp_m is expressed as a function of alleles frequencies p_a, p_m and of the linkage disequilibrium D_{ma} at QLE. This allows one to obtain an approximate expression for the change in frequency of a modifier allele from any frequency.

Indirect selection at a ploidy modifier locus generated by a single deleterious allele at mutation-

selection balance should generally be small (of the order of the mutation rate towards the deleterious allele). However, the overall effect of deleterious alleles occurring at many loci can become more important. Assuming that the effects of different viability alleles are multiplicative, two-locus results can be extrapolated to obtain an expression for the change in frequency of the modifier when deleterious mutations occur at a large number of loci. Neglecting linkage disequilibria between loci at which deleterious alleles are segregating, the overall effect of all selected loci on the change in frequency of the modifier can be approximated by the sum of the individual effects of these loci (e.g. Jenkins and Kirkpatrick 1995, Otto and Bourguet 1999, Hough et al. 2013).

Masking and purging deleterious mutations

As long as inbreeding is not too high and mutation is weak relative to selection ($u << s$), mutations are mostly present in the heterozygous state and masked in diploids (Crow and Kimura 1965, Perrot et al. 1991). Since deleterious mutations are often partially recessive ($h < 1/2$), this confers a direct advantage to diploidy: a zygote carrying a deleterious mutation in the heterozygous state will have a higher average fitness if it develops as a diploid, than if it undergoes meiosis and produce a haploid individual with a probability $1/2$ to carry the deleterious allele (assuming that the deleterious allele has the same fitness effect in haploids and in homozygous diploids).

Previous modifier models (Otto and Goldstein 1992, Jenkins 1993, Jenkins and Kirkpatrick 1995) have explored the effect of dominance and recombination on the evolution life cycles. They showed that linkage disequilibrium tends to build up between alleles coding for a longer diploid phase and deleterious alleles at other loci, due to the fact that mutations are generally partially masked ($h < 1$) and natural selection is less efficient among diploids. In particular, Otto and Goldstein (1992) showed that diploidy is favored when:

$$(1-2h) - \frac{s\,(1-h)\,(d\,h + (1-d)(1-r_{ma}))}{r_{ma} + (1-r_{ma})\,s\,(d\,h + (1-d)\,(1-r_{ma}))} > 0.$$

The first term corresponds to the masking advantage of diploids: a modifier locus increasing the probability to undergo selection as a diploid is favored when dominance $h < 1/2$. The second term describes the effect of linkage disequilibrium, which tends to build between deleterious alleles and modifier alleles increasing the level of diploidy when $h < 1$. If linkage is sufficiently tight (r_{ma} close to 0), this effect may overcome the masking advantage of diploids, and favor modifier alleles increasing the haploid phase, while diploidy is favored under looser linkage. As a consequence, reproductive systems that reduce the effective recombination rate (inbreeding, partial asexuality) tend to increase selection for haploidy (Otto and Marks 1996).

Evolution of life cycles in adapting populations

The evolution of life cycles in sexual organisms appears to be similarly influenced by beneficial mutations. Otto (1994) and Orr and Otto (1994) showed that diploidy is favored during sweeps of beneficial mutations that are partially dominant. An increase in the length of the diploid phase of the life cycle leads to an increase in the amount of selection experienced by heterozygotes, which have higher fitness than the average fitness of the two component haploids when $h > 1/2$. In addition, because deleterious mutations are masked in diploids, diploid populations retains a higher genetic diversity than haploid populations. This variability may be beneficial in the face of environmental changes (Bell 1982). Furthermore, in haploids, gene duplication must precede the evolution of new function, while in diploids, it may be faster because new alleles can appear in an existing locus: at the heterozygous state, the ancestral allele preserve the first function (Lewis 1979).

Other genetics factors affecting the evolution of life cycles

Diploids may benefit from more efficient repair of DNA damages, due to the presence of a homologous chromosome that may serve as a template (Michod and Wojciechowski 1994). In host-parasite interactions, diploidy should be favored in the host (due to the benefit of a higher number of recognition alleles) and haploidy in the parasite (Nuismer and Otto 2004).

Finally, a direct advantage of haploid diploid life cycles emerges from the cost of sex: with equal generation (sporophytic and gametophytic) length, haploid-diploid populations complete a sexual cycle half as often as either haplonts or diplonts. This intrinsic advantage favors biphasic life cycles when the cost of sex is high (Richerd et al. 1993). However, the evolution of asexual reproduction seems more likely in response to such a cost (Mable and Otto 1998).

Polymorphic versus alternation models

Two types of models for the evolution life cycles have been developed. In the first type (Figure 5A), mitotic development occurs either in the haploid or the diploid phase: individuals undergo selection in the diploid phase with probability d, and the models analyse the evolution of this probability (Perrot et al. 1991, Otto and Goldstein 1992, Otto and Marks 1996, Hall 2000). In the second type (Figure 5 B), mitotic growth occurs in both phases, and the model explores the evolution of the relative length of the haploid $(1 - d)$ and diploid (d) phases (Jenkins 1993, Jenkins and Kirkpatrick 1995).

Different timing of mutations have also been used. Most of the polymorphic models considered

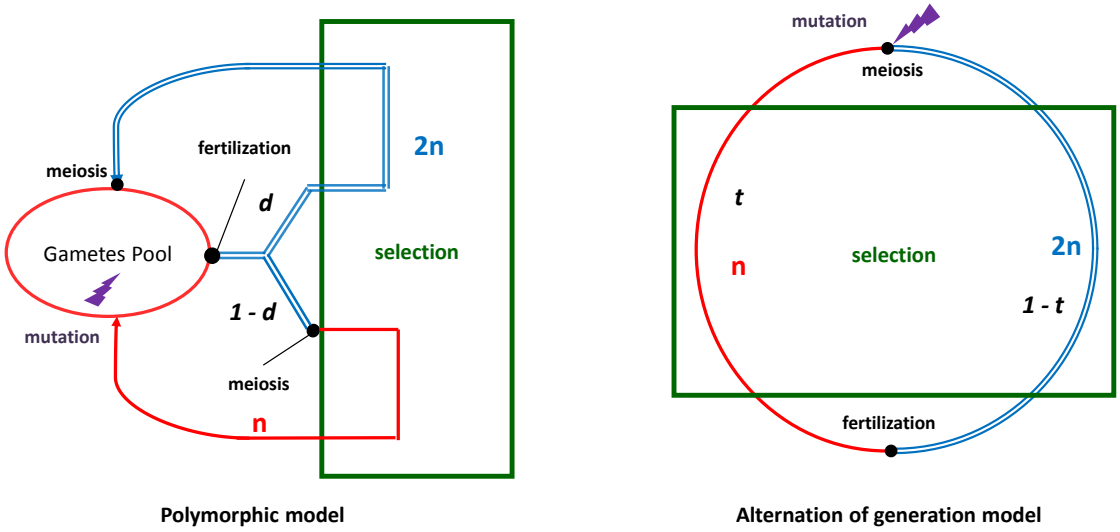

Figure 5: Polymorphic and alternation of generation models

mutations occurring at fertilization, while most of alternation of generation models placed mutations at meiosis. This last case is more consistent with mutation rates measured in mice: Russell and Russell (1996) found that at some loci, up to 50% of mutations occur between the last premeiotic and the first postmeiotic mitosis. In a panmictic population, the time at which mutation occurs does not affect much the predictions of the model, although using a polymorphic model similar to Otto and Goldstein's (1992) model in which mutations occur at meiosis, Hall (2000) found that biphasic life cycles could be maintained under restricted conditions.

2.3 Ecological models

Depending on the balance between masking and purging, genetic models predict an evolution toward diploidy or haploidy, but in most cases, haploid-diploid life cycles are not stable (but see Hall 2000). Therefore, additional arguments must be provided to explain the wide diversity of haploid-diploid life cycles.

Biphasic (haploid-diploid) populations should exploit a broader range of environmental niches than haploid or diploid populations (Willson 1981). Indeed, in many biphasic species, the gametophyte and the sporophyte differ in morphology, and presumably also in physiology and ecology. For instance, in algal species, one phase is often large and well adapted to exploit favorable conditions, while the other phase is small and resistant to environmental stress such as grazing (Klinger 1993). Hughes and Otto (1999) showed in a density dependent model that niche differentiation between the haploid and the diploid phases favors biphasic life cy-

cles over purely haploid or diploid cycles, since a mutant using a relatively empty niche tends to increase in frequency. Even if no differentiation seems apparent in isomorphic species, cryptic differences can occur among phases, such as in the brown algae *Ectocarpus siliculosus*, where gametophytes and sporophytes grow on different substrata (rock/shells versus other algal species, Couceiro et al. 2015). Temporal variations of the relative abundance (Bolton and Joska 1993, Otaiza et al. 2001, Dyck and De Wreede 2006) or fecundity (Santos and Duarte 1996) of haploids and diploids have also been reported, suggesting that environmental fluctuations may favor alternatively diploid sporophytes or haploid gametophytes. Furthermore, even if adult haploids and diploids are ecologically identical, their unicellular propagules serve different functions (spore dispersal vs. gamete fusion, Bell 1997) and may be different (Clayton 1992, Destombe et al. 1992). In their ecological model, Hughes and Otto (1999) showed that ecological differentiation of the juvenile phase is sufficient to maintain haploid diploid life cycles.

Models assumptions and future extensions

Most genetic models are based on several simplifying assumptions. First, they considered that mutations have the same fitness effect in haploids and in homozygous diploids. However, many mutations may have different effects on both phases: transcriptomic studies on haploid-diploid species show that a fraction of genes is expressed in one phase only (Coelho et al. 2007, Von Dassow et al. 2009, Rokitta et al. 2011), and mutations in these genes should thus have no effect on fitness in the other phase. More generally, selective pressures on different genes may not be equally strong in both phases, leading to different selection coefficients of deleterious alleles. One can also imagine that selection at some loci may favor different alleles in haploids and diploids: such ploidally antagonistic selection can maintain stable polymorphism (Ewing 1977, Immler et al. 2012), which may have important effects on the evolution of the relative duration of both phases. Relaxing this hypothesis can have important consequences on models predictions. In particular, antagonistic sexual selection between phases (in one sex, one allele is deleterious in haploid and beneficial in diploid, and inversely in the other sex) can explain the evolution of stable biphasic life cycle with different ploidy levels in male and female, such as arrhenotoky (Immler and Otto 2014).

Then, genetic models also assume that the baseline fitness of haploids and diploids is the same, although physiological differences between phases have been measured in some biphasic species: for example, differences in growth rate and survival are observed between haploid and diploid phases of the isomorphic red algae *Gracilaria verrucosa* and *Chondracanthus squarrulosus* in some laboratory conditions (Destombe et al. 1993, Pacheco-Ruíz et al. 2011). Intrinsic fitness advantage of diploids due to physiological differences between phases has been

proposed to explain the evolution of ploidy cycles.

Finally, genetic models always assume that haploids and diploids compete directly against each other. However, in heteromorphic species, ecological niches of haploids and diploids are often differentiated. Ecological models revealed that niche differentiation between phases generates a direct advantage for biphasic life cycles (Hughes and Otto 1999). In addition, reducing direct competition between haploids and diploids should decrease the masking advantage of diploids (or more generally, selection stemming from differences in mean fitness between haploids and diploids).

During this thesis, I relaxed these assumptions and extended previous genetic models to more realistic ecological conditions. I used QLE and stability analyses on modifier models including differential, even antagonistic effect of mutations between haploids and diploids. In Chapter 1, the effect on niche differentiation and intrinsic fitness difference between phases was explored in an analytical model considering infinite populations, I checked the analytical predictions by multilocus simulations. I additionally analyzed the consequences of a distribution of selection coefficients and of epistasis across loci on the evolution of life cycles. In Chapter 2, I use a more explicit density-dependent context to explore further consequences of interactions between genetic and ecological factor on the evolution of life cycles.

3 *S. cerevisiae*, a model organism to study ploidy evolution

3.1 The budding yeast *Saccharomyces cerevisiae*

The unicellular fungus *S. cerevisiae* is an ascomycota with a predominantly diploid biphasic life cycle. Haploids and diploids are heterotrophic cells of 6 to 12 µm which can grow both aerobically or anaerobically. *S. cerevisiae* has a haploid sex determination system with two mating types: a and α, which spontaneously mate to form diploid cells. Diploids undergo meiosis under stressful conditions.

S. cerevisiae has emerged as a model system to study ploidy evolution since the seventies (Adams and Hansche 1974). The budding yeast was already a well developed eukaryotic model for genetic research. It offers great advantages for experimental evolution: it has a short generation time (doubling time: 2 hours at 30 °C) and can be easily cultured both in liquid and solid (agar) medium. Axenic cultures can be maintained for thousands of generations and periodically frozen, and sexual reproduction is easily induced.

The yeast genome was the first eukaryotic genome to be sequenced (1986, see Cherry et al. 2012) and this model benefits from many genetic tools. Fluorescent strains derived from an ancestral genotype can be used as common competitors to measure the relative competitive

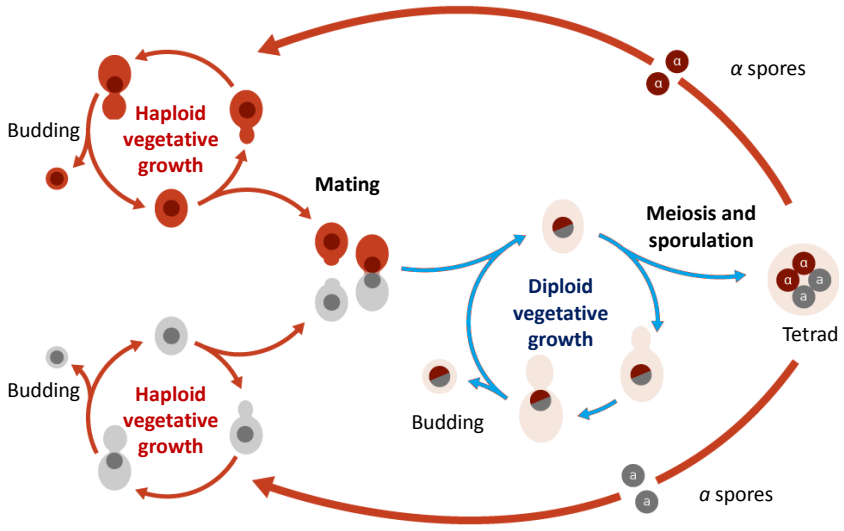

Figure 6: The haploid diploid life cycle of *Sacharomyces cerevisiae*

fitness of evolved strains. In the wild, haploids are heterothallic and undergo rapidly fusion, therefore the haploid phase is extremely reduced. Deletion of the HO gene prevents mating type switching and offers the possibility to control the ploidy level. Mating type switching can then be restored, allowing the construction of isogenic haploid and diploid cells to compare the fitness effect of mutations in haploids and diploids.

3.2 Testing the physiological effect of ploidy level

Genetic models on the evolution of lifes cycle assume that in the absence of genetic variations, haploids and diploids have the same intrinsic fitness, although physiological differences between phases may generate a direct advantage of one phase over the other. In *S. cerevisiae*, haploids are generally smaller, with a larger surface to volume ratio than diploids (Weiss et al. 1975). However, Weiss et al. (1975) showed that cell geometry in yeast depends more on environmental conditions than on ploidy: in poor medium, the volume of diploid cells is the same as the volume of haploid cells. In addition, diploid cells are more eccentric, which increases their surface to volume ratio.

In the lab, isogenic strains of *S. cerevisiae* do not clearly show intrinsic differences in fitness between ploidy levels. Adams and Hansche (1974) showed that in chemostat cultures of the yeast *Saccharomyces cerevisiae*, haploid and diploid growth rates are equal in rich medium, but haploids grow faster when phosphate is limiting. A more recent experimental test of the nutrient scarcity hypothesis did not find any correlation between ploidy level and medium richness (Mable 2001). Moreover, the comparison of isogenic haploid and diploid strains of *S.*

cerevisiae and its relative *S. paradoxus* in different environments revealed the same ploidy-environment interactions in both species and found no consistent effect of ploidy on either toxin tolerance or nutrient utilization (Zörgö et al. 2013). Finally, there is little experimental evidence in support of the theory that in yeast, life cycles evolve in response to difference in cell size between ploidy phases (Zeyl 2004), and genetic arguments may have played a more important role.

3.3 Measuring mutational effects in haploids and diploids

Previous models on life cycle evolution (e.g. Otto and Goldstein 1992, Jenkins 1993, Otto 1994) assumed that mutations have the same effect in haploids and homozygous diploids and showed that diploids benefit from a masking advantage when deleterious mutations are recessive ($h < 1/2$). However, this threshold does not hold if mutations have different effects in haploids and diploids. Recently, Gerstein (2013) showed that in isogenic haploid and diploid strains of yeast, beneficial mutations often have different fitness effects in haploids and homozygous diploids.

3.4 Testing the masking advantage of diploids

EMS (a mutagenic agent) treatment induces a more important decrease in haploid than in diploid growth rates, indicating that induced mutations have a stronger deleterious effect in haploids than in heterozygous diploids (Mable and Otto 2001), which is consistent with the masking theory. More precisely, Korona (1999*b*) measured the dominance coefficient of spontaneous deleterious mutations. For this purpose, he constructed isogenic homozygous and heterozygous diploids from haploid mutation accumulation lines and found a negative correlation between the magnitude and the dominance of deleterious mutations, and estimated the average dominance coefficient $h \approx 0.08$. More recent growth rates measurements in more than 4500 heterozygotes and homozygote *S. cerevisiae* strains carrying one single deletion were also consistent with the recesivity of deleterious mutations and estimated $h \approx 0.2$ (Agrawal and Whitlock 2010, Manna et al. 2012).

3.5 Rate of adaptation in haploids and diploids

Because diploids have twice as many genes as haploids, beneficial mutations have twice the opportunity to appear in a diploid than in a haploid population. However, beneficial mutations may be partially masked in diploids ($h < 1$) and have a greater chance to be lost. The first experiment testing for the relative rates of adaptation of haploids and diploids revealed a diploid advantage in chemostat (Paquin and Adams 1983), suggesting that a higher mutation

production in diploids confers a larger advantage than the greater probability of fixation in haploids. However, Zeyl (2004) showed that the adaptation rate of Paquin and Adam's diploid populations were incompatible with yeast mutations rate and fitness effects. He concluded that these populations were initially polymorphic and that adaptation relied on standing variation. A more recent experiment has shown that in small populations ($N_e < 100$), the rate of mutation appearance is limiting, and diploids adapt faster than haploids. However, in population large enough to produce at least one beneficial mutation per generation ($N_eU > 1$), haploids adapt faster than diploids due to the fact that the probability of fixation of new beneficial mutations is higher (and the fixation time lower) in haploid than in diploid populations (Zeyl et al. 2003).

4 Some evolutionary consequences of ploidy level

While mutations are fully exposed in haploids, in diploids, they may be present at the heterozygous state. Their effect therefore depends on the dominance coefficient. In a diploid population, the relative proportion of alleles in the homozygous and the heterozygous state depends on the population size and on the reproductive system: while diploid individuals from asexual populations should be heterozygous at all loci (neglecting loss of heterozygosity due, for instance, to mitotic recombination, see Gerstein et al. 2014), sexual reproduction changes intralocus alleles combinations through segregation. During sexual reproduction, recombination and segregation affect both the mean (short-term effect) and the variance (long-term effect) of fitness in the population. In haploids, recombination is the only sexual process that affects fitness, while in diploids, segregation is an additional intervening factor. The degree of ploidy of individuals will thus have important consequences on the evolution of sex and mating systems.

4.1 Ploidy and the dominance of fixed beneficial mutations

In populations adapting to a new environment, the fixation probability of an adaptative mutation increases with its fitness effect. In diploid panmictic populations, adaptative mutations are initially mainly heterozygous, and dominant mutations have a lower risk to be lost by drift (Haldane's sieve, 1937). In asexual populations, mutations remain heterozgous (in the absence of mitotic recombination), and dominant to overdominant ($h > 1$) mutations are expected to fix. On the contrary, in haploid adapting populations, dominance does not affect the fixation probability of a mutation and we do not expect a particular pattern of dominance for these mutations (Gerstein et al. 2014).

Theoretical models based on the Fisher's geometric model (FGM, see Figure 7) predict that

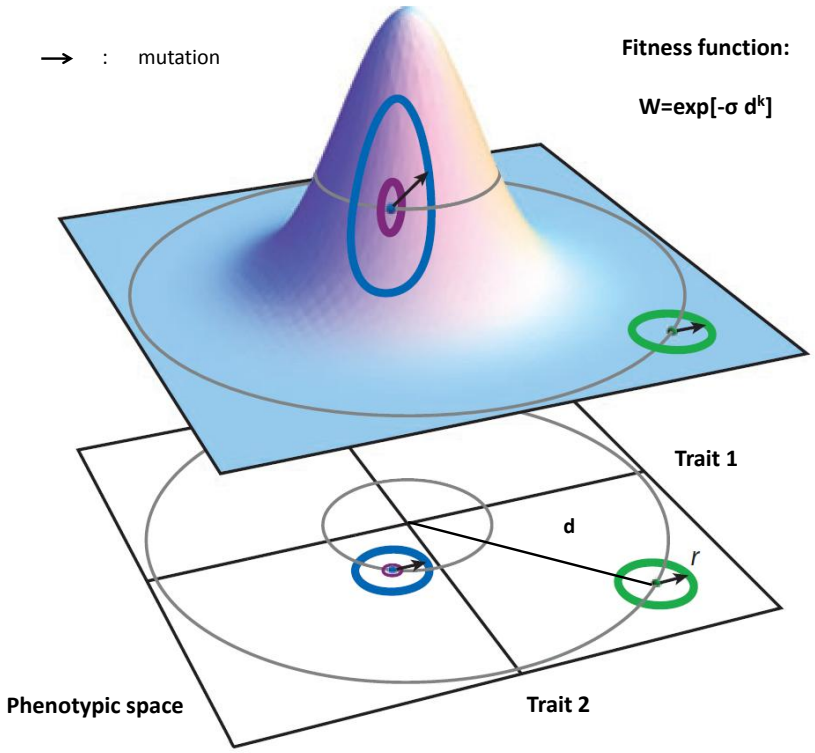

Figure 7: *Modified from Tenaillon (2014).* Fitness landscape of Fishers geometric model, with fitness on the vertical axis, and the phenotypic space (here, in two dimensions) below. Fitness W is a decreasing function of the distance to the optimal phenotype (d). Mutation of phenotypic size r moves the phenotype in the phenotypic space

the fitness effect of mutations fixed during an adaptative walk should follow an exponential distribution (Orr 1998). In an ongoing project, I explored the distribution of dominance coefficients using numerical simulations based on FGM. I also generated predictions on fitness distributions of haploid, diploid heterozygous and diploid homozygous offspring produced from haploid and diploid adapted populations. I intended to test these predictions using lines from Zeyl et al. (2003) experiment on S *cerevisiae* (evolved in minimal medium for 2000 generations), and already performed a number of crosses on these lines. Unfortunately, this project is not advanced enough to constitute a chapter of this thesis manuscript, and I will only discuss our preliminary results in the section devoted to perspectives.

4.2 Mating system and inbreeding depression

In diploid populations, offspring from mating between relatives often have a reduced fitness compared to the offspring produced under random mating. This phenomenon, called inbreeding depression (Charlesworth and Charlesworth 1987; 1999), is mostly due to the presence

of deleterious recessive alleles maintained at non negligible frequencies ($\approx u/hs$ in a infinite, panmictic population). While haploids do not suffer from such a cost, in diploids, the genetic load with inbreeding is determined by the dominance coefficients (Whitlock 2002).

Diploid and haploid selfing increase the number of potential partners for any gamete, which may provide an important advantage compared with outcrossing. In diploids, inbreeding depression is the major factor for the evolution of reproductive system preventing selfing, such as heterothallism (see Billiard et al. 2012), mating types or self incompatibility systems. This factor should not play a role in haploids and we may thus expect selfing rates to be higher in haploids than in diploids.

4.3 Outbreeding depression and heterosis

Deleterious mutations are also responsible for heterosis, i.e. the higher fitness of diploid hybrids compared to offspring of within-populations crosses. However, crosses between populations may also reveal genetic incompatibilities when populations are sufficiently divergent.

Indeed, in diploid and haploid allopatric populations, mutations appear and may reach fixation depending on their fitness effect in the genetic background of the population in which they occurred. Hybridization between two isolated populations brings within the same individual several mutations that have never been tested together before. Some of these new combinations may be deleterious, leading to a reduced fitness of hybrids (outbreeding depression, Lynch 2012).

While reproductive incompatibilities are fully express in the first generation (F1) of haploid hybrids, diploid F1 hybrids are heterozygous at all loci. In finite populations, deleterious mutations of small effect may reach fixation by drift. In this case, diploid F1 hybrids may benefit from heterosis (Lynch and Walsh 1998). Incompatibilities involving additive by additive epistatic effects may overcome the heterosis effect (Burton 1990, Lynch 2012, Edmans 1999), but outbreeding depression is more likely to be observed in the F2 generation, when recombination exposes recessive alleles involved in incompatibilities (Lynch 2012, Edmans 1999; 2007).

While diploids F1 hybrids are partially protected against recessive incompatibilities, in species with chromosomic sex determination systems, the heterogametic sex (XY or ZW) suffers from fully exposed incompatibilities involving alleles on the X or the W chromosome, leading to the preferential sterility and inviability of hybrids of the heterogametic (XY or ZW) sex (Haldane's rule, Haldane 1922).

In biphasic species with haploid sex determination (e.g. *S. cerevisiae*), this asymmetry between sexes in hybrid fitness is not expected, but we predict a stronger decline of hybrid

fitness with the divergence time in haploids than in diploids. In Chapter 3, I compare the dynamics of reproductive isolation in haploid and diploid small populations of the yeast *Saccharomyces cerevisiae*. In finite populations, deleterious mutations can reach fixation by drift, while compensatory mutations are selected for, leading to a dynamic mutation selection drift equilibrium (Poon and Otto 2000, Sella and Hirsh 2005, Silander et al. 2007, Sella 2009). Under this scenario, models based on Fisher's geometric model (see Figure 7) predict a linear decline of hybrid fitness (Chevin et al. 2014), and a more important oubreeding depression in haploids than in diploids (Barton 2001). We used mutation accumulation lines of yeast strains with high mutation rates in order to reach this equilibrium, and analyze the dynamics of the fitness decline in haploid F1, and diploid F1 and F2 hybrids.

Chapter 1

Introducing ecological components into genetic models for the evolution of life cycles

Within this first chapter, I present a polymorphic model exploring the effect of more realistic ecological and genetic components on the evolution of life cycles. This work combines a preliminary version of the published paper presented in Chapter 2, and several results from a paper (submitted to Evolution and presented in Appendix D) written in collaboration with Michael Scott, in which I performed the multilocus simulations.

Contents

1 Introduction

A prominent theory for the evolution of either haploid or diploid life cycles involves the direct consequences of ploidy level on the expression of deleterious mutations. In diploids, the fitness effect of a deleterious mutation can be partially hidden by the homologous gene copy, which is advantageous if a heterozygous diploid has a higher fitness than the average fitnesses of the two possible haploids. In the short term, diploidy is favoured when deleterious mutations are partially recessive and haploidy is favoured when deleterious mutations are partially dominant (Perrot et al. 1991, Otto and Goldstein 1992, Jenkins and Kirkpatrick 1994; 1995). As a consequence of mutations being partially concealed, an expanded diploid phase allows mutations to reach a higher frequency (Crow and Kimura 1965, Kondrashov and Crow 1991). Modifier models, in which the extent of haploid and diploid phases is determined by a second,'modifier' locus, have found that low recombination rates favour haploidy because modifier alleles that expand the haploid phase remain longer associated with purged genetic backgrounds (Otto and Goldstein 1992). As a consequence, reproductive systems that reduce the effective recombination rate (inbreeding, partial asexuality) tend to increase selection for haploidy (Otto and Marks 1996).

In general, genetic models predict evolution towards either haploidy or diploidy (depending for example on the degree of dominance of deleterious mutations or on the mating system), but cannot explain the evolutionary stability of haploid-diploid life cycles unless considering additional mechanisms directly favoring biphasic cycles. By contrast, models incorporating an ecological differentiation between phases can explain the maintenance of biphasic cycles. If the haploid and the diploid ecological niches are sufficiently separated, a mutant using a relatively empty niche can be favored, and depending on the degree of niche overlap, the evolutionarily stable strategy may consist in producing both haploid and diploid individuals (Hughes and Otto 1999). Even in isomorphic species, slight metabolic differences may have important ecological effects, leading to some degree of differentiation between the haploid and diploid niches (Couceiro et al. 2015, Bell 1997)

In this chapter, we relax several assumptions of previous genetic models exploring the effect of deleterious mutations on the evolution of ploidy levels (Otto and Goldstein 1992, Otto and Marks 1996). Previous models assumed that deleterious alleles have the same fitness effect in haploids and in homozygous diploids. However, many mutations may have differential, or even antagonistic effects on both phases (Gerstein 2013). Moreover, genetic models suppose that haploids and diploids have the same baseline fitness, which may not necessarily be the case. Following a series of substitution events, the overall (intrinsic) fitness of a haploid and a diploid should not be equal. For instance, in *Saccharomyces* yeast, differences between haploid and diploid growth rates measured by Zörgö et al. (2013) range from being negligible

to substantial (one phase can have growth rates up to 1.75 times higher) in different environments. Similar differences in growth rate and survival are observed between haploid and diploid phases of the red algae *Gracilaria verrucosa* and *Chondracanthus squarrulosus* in some laboratory conditions (Destombe et al. 1993, Pacheco-Ruíz et al. 2011). More generally, the phase with higher fitness and the magnitude of fitness differences varies widely and is heavily dependent on the environmental context (Mable and Otto 1998, Thornber 2006, Zörgö et al. 2013).

In a first modifier model (of the 'polymorphic' type, see Introduction), we assume that haploids and diploids may occupy different ecological niches (which affects the degree of competition between individuals having different ploidy level). We use two- and three-locus models to explore the effects of differences in strength of selection or epistasis between phases. We will see that selection on the modifier can be decomposed into a term stemming from differences in mean fitness between ploidy levels ('short-term' effect, which is proportional to the degree of niche overlap between haploids and diploids), and terms involving linkage disequilibria between the modifier and selected loci, favoring the phase in which natural selection is the most efficient.

In a second model (of the 'alternation of generations' type, see Introduction), we consider the effect of deleterious mutations on the life cycle when haploids and diploid have different intrinsic fitnesses. We performed an linear stability analysis to determine the stable life cycle strategies under different conditions of dominance and intrinsic fitness differences between phases.

Analytical predictions from both models are checked by multilocus, individual-based simulations representing mutations occurring at a large number of loci, located at different genetic distances from the modifier locus. As we will see, analytical models and simulations also allow us to explore effects of variances and covariance across loci of the distribution of selection coefficients in haploids and diploids.

2 Models

Previously, two general life cycles have been used to examine the evolution of haploid versus diploid phases (Figure 1.1). In the polymorphic models considered by Perrot et al. (1991), Otto and Goldstein (1992), Otto and Marks (1996) and Hall (2000), selection occurs once per generation and a modifier locus affects the probability ot undergo selection during the haploid or diploid phase (polymorphic models), while in the alternation of generations models, (Jenkins and Kirkpatrick 1994; 1995, Otto 1994, see Figure 1.1a), selection occurs continuously throughout the life cycle, first in the haploid, then in the diploid phase (alternation of

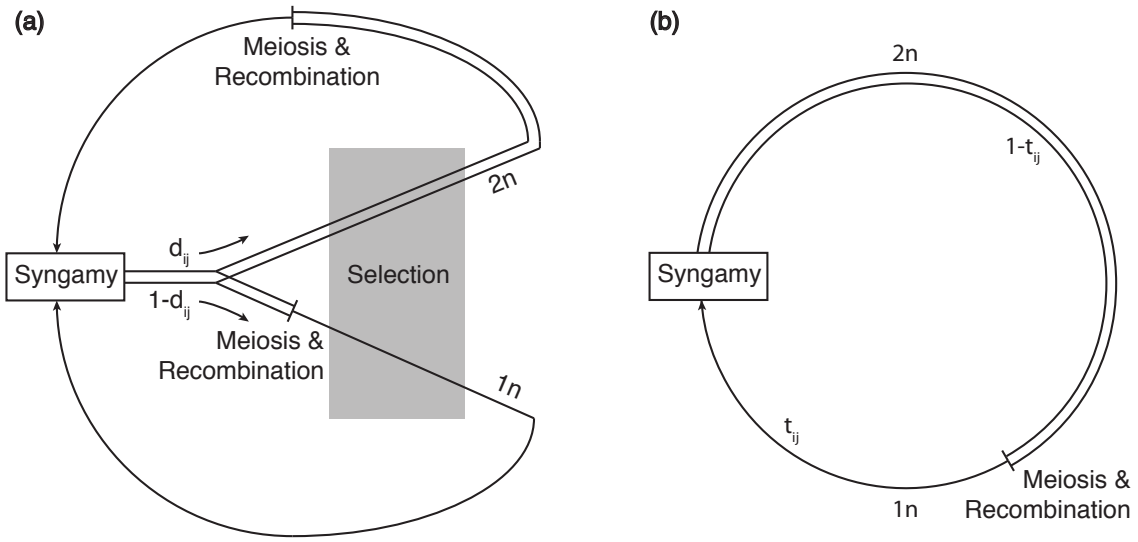

Figure 1.1: Model (a) polymorphic model and (b) alternation of generation model. Single lines represent haploid phases and doubled lines indicate diploid phases. In (a), modified from Perrot et al. (1991) and Otto and Goldstein (1992), zygotes with the modifier genotype ij undergo selection as diploids with probability d_{ij} or undergo meiosis and recombination before experiencing selection as haploids with probability $(1 - d_{ij})$. In (b), after Jenkins and Kirkpatrick (1994; 1995) and Otto (1994), all zygotes with genotype ij experience viability selection as a diploid for a proportion $(1 - t_{ij})$ of their life cycle before undergoing meiosis and recombination and then experiencing viability selection as a haploid for the remainder of the life cycle.

generation models, Figure 1.1b).

Both types of model will be considered in this chapter. In both cases, we will assume that mutations may have differential effect in a haploid and a homozygous diploid and analyze the effect of a distribution of selection coefficients across loci. In the polymorphic model, I explored the effect of niche differentiation between phases, and epistasis, while the effect of intrinsic fitness differences between haploids and diploids was analyzed in both types of models (see Appendix D), but only results from the alternation of generation model are detailed here.

2.1 Polymorphic model

Analytical model

This model represents an infinite, panmictic population with non-overlapping generations. The first event of each generation corresponds to gamete union, assumed to be random. As in Otto and Goldstein (1992), we assume that diploid zygotes can either enter meiosis immediately to form new haploid individuals, which then undergo selection, or remain diploid and undergo selection before meiosis (see Figure 1.1). The probability to undergo selection as a diploid or a haploid is controlled by a modifier locus with two alleles M and m: zygotes MM, Mm and mm develop as diploids with probabilities d, $d + h_m\delta$ and $d + \delta$, respectively (h_m thus measures the level of dominance of allele m).

During selection, the fitness (i.e. survival or fecundity) of individuals depends on their genotype at a second locus, with two alleles A and a (the recombination rate between both loci is denoted r_{ma}). We assume that allele a is produced by mutation from A at rate u per generation, and has a deleterious effect in both phases. Fitnesses of AA, Aa and aa diploid individuals are given by 1, $1 - hs$ and $1 - s$, while fitnesses of A and a haploids are given by 1 and $1 - \alpha s$ (Table 1.2). The parameter α thus modulates the fitness of a individuals compared with homozygote diploids aa.

Finally, a parameter γ determines to what extent diploids and haploids compete against each other: we assume that before selection, a proportion $1 - \gamma$ of diploids enter an ecological niche where they do not compete against haploids (the proportion of individuals entering this niche is thus $(1 - \gamma)\,\bar{d}$, where $\bar{d}$ is the average propensity to remain diploid); similarly, a proportion $1 - \gamma$ of haploids occupy a niche where they do not compete against diploids (proportion $(1 - \gamma)\left(1 - \bar{d}\right)$ of the total population). Finally a proportion γ of individuals enter a 'common' niche where haploids and diploids compete against each other. When $\gamma = 1$, competition occurs at the scale of the whole population (i.e., haploids and diploids have the same ecological niche) as in Otto and Goldstein (1992), while $\gamma = 0$ corresponds to the case where haploids and diploids occupy different niches and thus do not compete against each

Table 1.1: Fitnesses of different genotypes in the three-locus polymorphic model

	AA	Aa	aa
BB	1	$1-hs$	$1-s$
Bb	$1-hs$	$(1-hs)^2(1+e_{a\times a})$	$(1-hs)^2(1+e_{a\times a})^2(1+e_{a\times d})$
bb	$1-s$	$(1-hs)^2(1+e_{a\times a})^2(1+e_{a\times d})$	$(1-hs)^2(1+e_{a\times a})^4(1+e_{a\times d})^4(1+e_{d\times d})$

other. Importantly, we assume that population regulation occurs within each niche ('soft' selection, e.g. Wallace 1975) so that the output of each niche is proportional to the proportion of individuals entering the niche (independently of the mean fitness of these individuals). The case of 'hard' selection (output proportional to mean fitness) would be formally equivalent to $\gamma = 1$.

We explored the dynamics of this model using two approaches: a local stability analysis, providing conditions under which allele m increases in frequency when rare, and a quasi-linkage equilibrium (QLE) analysis, which yields a simple expression for the change in frequency of m when recombination is strong relative to selection, so that linkage disequilibrium equilibrates fast relatively to allele frequencies. Variables p_a and p_m denote the frequencies of alleles a and m in the whole population, at the time of zygote formation, while D_{ma} is the linkage disequilibrium (LD) between the two loci ($D_{ma} = p_{ma} - p_m p_a$, where p_{ma} is the frequency of ma gametes). Finally, $q_a = 1 - p_a$ and $q_m = 1 - p_m$ are the frequencies of alleles A and M. As we will see, two-locus results can be extrapolated to obtain an expression for the change in frequency of the modifier when deleterious mutations occur at a large number of loci, assuming that the effects of different deleterious alleles on fitness are multiplicative. To explore the effect of epistasis among selected loci we also analyzed a three-locus polymorphic model, including a second selected locus with two alleles b and B. We assume that loci are in order m-a-b along the chromosome and that the selected loci recombine at rate r_{ab}. The fitnesses of the different haploid and diploid genotypes are given in Table 1.1: a parameter e measures epistasis between selected loci in haploids, while $e_{a\times a}$, $e_{a\times d}$ and $e_{d\times d}$ measure additive-by-additive, additive-by-dominance and dominance-by-dominance epistasis within diploids (see Table 1.1). Because deleterious alleles are mostly present in the heterozygous state in diploids (under our assumption of random mating), we will see that $e_{a\times d}$ and $e_{d\times d}$ have negligible effects on the evolution of the modifier.

Multilocus simulations

We used individual-based simulations to test extrapolations from our analytical model when mutations segregate at a large number of loci. Our C++ program is based on multilocus simulation programs used in previous papers (Roze 2009, Roze and Michod 2010, Roze and

Otto 2012). At the start of every generation, the population is made of N individuals, which may either be haploid or diploid. Each individual carries either one or two copies of a chromosome (depending on its ploidy level) along which mutations occur at a rate U per generation: more precisely, the number of new mutations per chromosome is sampled from a Poisson distribution with parameter U, while the position of each new mutation along the chromosome is random (the number of possible sites being quasi-infinite). Each deleterious mutation is characterized by its position and its fitness effects in haploids and diploids, which may either be fixed for all mutations or sampled in a probability distribution. In the absence of epistasis, the fitness of an individual is given by:

$$W_{haplo} = \prod_i (1 - s_{h,i}) \qquad \text{and} \qquad W_{diplo} = \prod_i (1 - s_{d,i}) \prod_j (1 - h s_{d,i})$$

.

In the case of haploids, the product is over all mutations i present in the individual (where $s_{h,i}$ is the deleterious effect of mutation i in a haploid), while in the case of diploids the first product is over all mutations in the homozygous state, and the second over all mutations in the heterozygous state. Note that in the case where selection varies across loci we assume that dominance coefficients (h) are the same at all loci; although this assumption seems obviously not realistic, deleterious alleles should almost always be present in the heterozygous state under our assumption of random mating (as long as $s_{d,i} >> u, 1/N$) in which case the strength of diploid selection only depends on the $hs_{d,i}$ product (which is distributed when $s_{d,i}$ is distributed). Data available on the distribution of fitness effects of mutations indicate that gamma or log-normal distributions provide reasonable fits (Loewe and Charlesworth 2006). Here we use a log-normal distribution: for each mutation i, the log of $s_{h,i}$ and $s_{d,i}$ are sampled in a bivariate normal distribution with means (m_h, m_d), variances (σ_h, σ_d) and correlation ρ.

We also used the program to explore effects of epistasis among deleterious alleles, but only under fixed selection and epistatic coefficients; in this case, fitnesses are given by:

$$W_{haplo} = (1 - \alpha s)^n (1 + e)^{\frac{n(n-1)}{2}}$$
$$W_{diplo} = (1 - s)^{n_{Ho}} (1 - hs)^{n_{He}} (1 + e_{a\times a})^{n_1 + 2n_2 + 4n_3} (1 + e_{a\times d})^{n_2 + 4n_3} (1 + e_{d\times d})^{n_3}$$

where n is the number of mutations in a given haploid, while n_{Ho}, n_{He} are the number of homozygous and heterozygous mutations in a given diploid, $n_1 = n_{He}(n_{He} - 1)/2$, $n_2 = n_{He} n_{Ho}$ and $n_3 = n_{Ho}(n_{Ho} - 1)/2$ the numbers of interactions between two heterozygous mutations, one heterozygous and one homozygous mutation, and two homozygous mutations, respectively (Roze 2009). We only considered negative epistasis, since combinations of mutations quickly

become advantageous when epistasis is positive and U is not small, in which case mutations accumulate rapidly over the course of the simulation.

At the start of each generation, diploid zygotes undergo meiosis or not with probabilities depending on their alleles at the modifier locus. To eliminate direct selection at the modifier locus, we assume that each diploid produces two haploid individuals (when meiosis occurs before selection). Then, individuals enter the common ecological niche with probability γ or their ploidy-specific niche (haploid or diploid) with probability $1 - \gamma$. The $2N$ gametes contributing to the next generation are then produced as follows. For each of these gametes, a parent is sampled randomly among all individuals. If its relative fitness w_i/w_{max} is higher than a random number sampled in a uniform distribution in $[0, 1]$, the individual is selected; otherwise, we sample another individual from the same niche and repeat the process. If the selected individual is diploid, a recombinant haplotype is generated.

During a given number of preliminary generations (generally 5000), the rate of diploidy is fixed so that the population reaches mutation-selection balance; we then introduce mutations at the modifier locus at a rate μ_M per generation. When a mutation occurs, the rate of diploidy coded by the mutant allele is sampled from a uniform distribution between $d_{old} - 0.1$ and $d_{old} + 0.1$, where d_{old} is the value of the parent allele; if the new value is negative or higher than 1, it is set to zero or 1, respectively. Simulations generally lasted 20000 generations, which was sufficient in most cases for the rate of diploidy to reach equilibrium.

2.2 Alternation of generation model

The alternation of generation model (see Figure 1.1b) was used to assess the effect of intrinsic fitness difference between phases and considers that haploids and diploids compete within the same ecological niche (which corresponds to $\gamma = 1$ in the previous model).

In this model, zygotes are formed during synchronous random mating and the diploid genotype (ij) at the modifier locus determines the timing of meiosis and hence the proportion of time each individual spends as a diploid (d_{ij}) and as a haploid $(1 - d_{ij})$. Here, S_h and S_d represent selection acting across the genome due to intrinsic fitness differences between haploids and diploids. The analytical model focuses on the selection experienced at each of L selected loci, and we defined $\varphi_h = S_h/L$ and $\varphi_d = S_d/L$ as the intrinsic fitnesses per viability locus. When $\varphi_h > \varphi_d$, haploids have higher fitness than diploids and the fitness of diploids is higher when $\varphi_d > \varphi_h$.

Mutations and recombination occur as described in the previous model. The deleterious allele a reduce haploid fitness by αs, homozygous diploid fitness by s and heterozygous diploid fitness by hs, but here, haploids and diploids fitnesses, presented in Table 1.2, depend on the time

Table 1.2: Fitnesses of different genotypes in the alternation model (with intrinsic fitness differences between phases, φ_h and φ_d, and in the polymorphic model

Genotype	Polymorphic model	Alternation model
A	1	$w_A(1-d_{ij}) = \exp[(1-d_{ij})\varphi_h]$
a	$1-\alpha s$	$w_a(1-d_{ij}) = \exp[(1-d_{ij})(\varphi_h - \alpha s)]$
AA	1	$w_{AA}(d_{ij}) = \exp[d_{ij}(\varphi_d)]$
Aa	$1-hs$	$w_{Aa}(d_{ij}) = \exp[d_{ij}(\varphi_d - hs)]$
aa	$1-s$	$w_{aa}(d_{ij}) = \exp[d_{ij}(\varphi_d - s)]$

spent on each phase(d_{ij} and $1-d_{ij}$), and on their intrinsic fitness (φ_h and φ_d).

In the multilocus simulations, the fitness of a haploid carrying n deleterious alleles is now given by $W_h = \exp[S_h + s_h n]$, while the fitness of a diploid carrying n_{he} deleterious alleles in the heterozygous state, and n_{ho} in the homozygous state is given by $W_d = \exp[S_d + n_{he} h s_d + n_{ho} s_d]$. At the start of each generation, all N individuals are diploid. To produce the $2N$ gametes that will form the diploids of the next generation, a diploid individual is sampled randomly among all diploids of the previous generation, and undergoes meiosis to produce a haploid; the number and positions of cross-overs are determined in the same way as in the polymorphic model. If a random number sampled from a uniform distribution between 0 and 1 is lower than ${w_d}^{1-t}{w_h}^t$ (where w_d and w_h are the fitnesses of the diploid parent and haploid offspring), divided by its maximal possible value, then the haploid is retained; otherwise another diploid parent is sampled, until the condition is fulfilled.

At the beginning of the simulations, the modifier locus is fixed for an allele coding for diploidy level (probability to undergo selection as a diploid, or length of the diploid phase) d_{init} and all selected loci are fixed for allele 0. Then, deleterious mutations are introduced at rate U per chromosome (the length of the diploid phase being still fixed to d_{init}) until the population reaches mutation-selection equilibrium (after generally 2000 generations). After that, mutations at the modifier locus are introduced at a rate μ_M per generation with the same distribution of effects as in the polymorphic model. The mutation rate was of order 0.1 and simulations generally lasted 100 000 generations.

3 Results

In the following, the QLE analysis is detailed for the polymorphic model with niche differentiation between phases. The linear stability analysis for the alternation of generation model with intrinsic fitness differences between phases is presented in Appendix D.

3.1 Equilibrium frequency of deleterious mutations

In the polymorphic model, when allele M is fixed at the modifier locus, the change in frequency of allele a due to selection is given by (see Appendix A):

$$\Delta p_a = p'_a - p_a = -(1-\gamma)s\left[dh + (1-d)\alpha\right]p_a q_a - \gamma s \frac{d\overline{W_d}h + (1-d)\overline{W_h}\alpha}{d\overline{W_d} + (1-d)\overline{W_h}} p_a q_a \tag{1.1}$$

where p'_a is the frequency of allele a after selection, and $\overline{W_d}$, $\overline{W_h}$ are the mean fitnesses of diploids and haploids, respectively. If these mean fitnesses are close to 1, we have to the first order in s:

$$\Delta p_a = -s\left[d\left(h + (1-2h)p_a\right) + (1-d)\alpha\right] \tag{1.2}$$

which is independent of the niche overlap parameter (γ). After mutation, the frequency of a is changed to $u + (1-u)p_a$. From this, the the mutation-selection balance is approximately (assuming $s\left[dh + (1-d)\alpha\right] >> u$):

$$\tilde{p}_a = \frac{u}{s\left[dh + (1-d)\alpha\right]} \tag{1.3}$$

In the alternation of generation model, selection in the haploid and diploid phases multiply, giving a slightly different result:

$$\tilde{p}_a = \frac{u\,\mathrm{e}^{(1-d)s\alpha}}{1 - \mathrm{e}^{s[dh+(1-d)\alpha]}} \tag{1.4}$$

Extrapolating equations 1.3 and 1.4 to the case of many loci subject to deleterious mutations (and assuming that selection parameters are the same at all loci and that epistasis is absent), the average number of mutations per haploid genome is given by:

$$\tilde{N}_{mut} = \frac{U}{s\left[dh + (1-d)\alpha\right]}\ , \qquad \tilde{N}_{mut} = \frac{U\,\mathrm{e}^{(1-d)s\alpha}}{1 - \mathrm{e}^{s[dh+(1-d)\alpha]}} \tag{1.5}$$

in the polymorphic and in the alternation of generation model, respectively, with U the sum of mutation rates over all loci. Figure 1.2 shows that this prediction is confirmed by multilocus simulations, except when U is relatively large and $\gamma > 0$. Indeed when U is not small, the terms $\overline{W_d}$ and $\overline{W_h}$ that appear in the second term of equation 1.1 cannot be neglected. Under the assumption of negligible LD among selected loci and large number of loci, numbers of mutations per genome should be Poisson-distributed, which yields $\overline{W_h} = \exp(-s\alpha N_{mut})$,

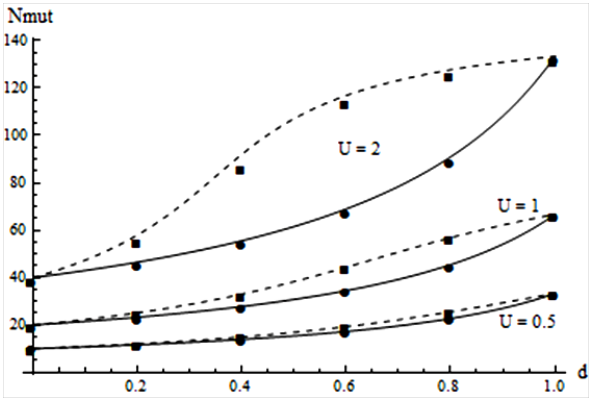

Figure 1.2: Average number of mutations per chromosome as a function of the fraction of the population undergoing selection in the diploid phase (d). Solid curves: analytical solution (6) for $\gamma = 0$. Dashed curves: numerical solution of equation 1.6 for $\gamma = 1$. Dots: multilocus simulations for $\gamma = 1$ (squares) and $\gamma = 0$ (circles). Other parameter values: $s = 0.05$, $h = 0.3$, $\alpha = 1$

$\overline{W_d} = \exp(-2shN_{mut})$. Plugging these expressions into equation 1.1 (and neglecting terms in p_a^2) gives an equation which can be solved numerically to obtain N_{mut} at equilibrium:

$$U = s\tilde{N}_{mut}\left\{(1-\gamma)\left[dh + (1-d)\alpha\right] + \gamma\frac{dh\,\mathrm{e}^{-2sh\tilde{N}_{mut}} + (1-d)\alpha\,\mathrm{e}^{-s\alpha\tilde{N}_{mut}}}{d\,\mathrm{e}^{-2sh\tilde{N}_{mut}} + (1-d)\,\mathrm{e}^{-s\alpha\tilde{N}_{mut}}}\right\} \tag{1.6}$$

Figure 1.2 shows that the numerical solution matches well the simulation results for $\gamma = 1$. These results show that a biphasic population ($d < 0 < 1$) has a smaller average number of mutations at equilibrium when haploids and diploids do not compete against each other. This can be understood from the fact that the efficiency of selection in decreasing depends on the covariance between the number of mutations within an individual and its fitness (e.g., Price 1970). When haploids and diploids occupy different niches, selection takes place independently in these two niches and depends on the average covariance between number of mutations and fitness within each niche. In contrast, when haploids and diploids compete against each other, the covariance between number of mutations and fitness (and thus the efficiency of selection) is reduced due to the fact that diploids typically carry larger numbers of mutations than haploids, but may have similar or even higher fitnesses when mutations are partially recessive.

3.2 Niche differentiation between phases

The effect of niche differentiation between phases was studied using the polymorphic model (Figure 1.1a). In the following we assume that selection s, mutation rate u and the modifier effect δ are weak (of order ϵ), and express the change in frequency of the modifier Δp_m to the leading order in ϵ as a function of allele frequencies and the linkage disequilibrium D_{ma}. We

then use a QLE approximation to express D_{ma} in terms of allele frequencies, and obtain an expression for Δp_m in terms of the model parameters (see Appendix A for derivations):

$$\begin{aligned}\Delta p_m =& \gamma \frac{\overline{W_d} - \overline{W_h}}{d\overline{W_d} + (1-d)\bar{W}_h} d_m p_m q_m \\ &- \gamma s \frac{d\overline{W_d} h + (1-d)(1-r_{ma})\overline{W_h}\alpha}{d\overline{W_d} + (1-d)\overline{W_h}} D_{ma} \\ &- (1-\gamma) s \left[dh + (1-d)(1-r_{ma})\alpha\right] D_{ma}\end{aligned} \tag{1.7}$$

where $d_m = [h_m(1-p_m) + (1-h_m)p_m]$ (simplifying to $\delta/2$ when $h_m = 1/2$). The first term of equation 1.7 is proportional to the degree of niche overlap γ and favors diploidy when diploids have a higher mean fitness than haploids, and haploidy otherwise. To the first order in s, it simplifies to $s(\alpha - 2h)\gamma$ and corresponds to the second term of the expression 1.11 in the alternation of generation model.

The last two terms are proportional to D_{ma} and represent indirect selection on m through its association with allele a within the common niche (second term) and the separate niches (third term). If $\overline{W_d}$ and $\overline{W_h}$ are close to 1, the second term simplifies and the effect of indirect selection becomes independent of γ (to leading order in s), yielding:

$$\Delta p_m = \gamma \left(\overline{W_d} - \overline{W_h}\right) - s \left[dh + (1-d)(1-r_{ma})\alpha\right] D_{ma} \tag{1.8}$$

Linkage disequilibrium at QLE is given by (see Appendix A for derivation):

$$\tilde{D}_{ma} = \gamma s \frac{\left(\overline{W_h}\alpha - \overline{W_d}h\right)}{d\overline{W_d} + (1-d)\overline{W_h}} \frac{p_m q_m p_a q_a d_m}{r_{ma}} + (1-\gamma) s \frac{d_m}{r_{ma}} (\alpha - h) p_m q_m p_a q_a \tag{1.9}$$

When $1 - \overline{W_d}$ and $1 - \overline{W_h}$ are of order ϵ, the terms in γ and $1-\gamma$ become equivalent to leading order, which gives:

$$\tilde{D}_{ma} = s \frac{d_m}{r_{ma}} (\alpha - h) p_m q_m p_a q_a \tag{1.10}$$

Equation 1.10 indicates that D_{ma} is generated by the difference in selection between the phases: when allele a is deleterious in both phases, it tends to be associated with the modifier allele increasing diploidy if selection is more efficient in haploids than in diploids, i.e. if $h < \alpha$. The second and third term of equation 1.7 thus represent a 'long-term effect' and favors the ploidy level which maximizes purging. It is the only contributing term when the haploid and diploid niches are fully separated.

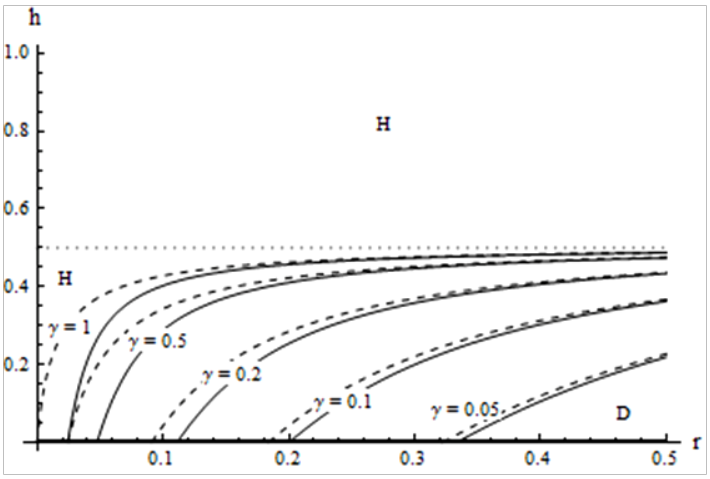

Figure 1.3: Value of the dominance coefficient (h) above which increased haploidy is favored (and below which increased diploidy is favored) as a function of the recombination rate r_{ma}: solid and dashed curves show predictions from the QLE and local stability analysis, respectively. Selection coefficients are the same for haploids and homozygote diploids: $\alpha = 1$, $s = 0.05$, $d = 0.5$. Niche differentiation (smaller γ) tends to favor haploids, and diploidy never invades when $\gamma = 0$.

Extrapolating to many, freely recombining deleterious alleles, the overall strength of selection on the modifier can be obtained by replacing p_a by the mean number of mutations per haploid genome N_{mut} (given by equation 1.6) and r_{ma} by 1/2 in equations 1.8 and 1.10. Depending on the values of hs(diploid selection) and $s\alpha$ (haploid selection), the population evolves either towards haploidy ($d = 0$) or diploidy ($d = 1$): as in previous genetical models (Otto and Goldstein 1992, Otto and Marks 1996), the maintenance of a biphasic cycle is never predicted; this last result was confirmed by multilocus simulations with various map lengths.

Results from the local stability analysis are given in Appendix C. The leading eigenvalue of the linearized system (at the equilibrium corresponding to fixation of allele M) gives similar results to the one obtained from the QLE analysis, except that it does not diverge when r_{ma} tends to zero. Figure 1.3 shows parameter regions in which modifier alleles increasing haploidy or diploidy are favored, according to the predictions obtained from both methods and assuming that the strength of selection is the same in both phases ($\alpha = 1$). When niches become more differentiated (γ decreases), the relative importance of purging increases (see Figure 1.3) and the parameter region where diploidy is favored tends to shrink.

3.3 Effect of epistasis

Results from the three-locus polymorphic model (including epistatic effects, see Appendix B) are shown in Figure 1.4. Overall, the results are similar to the two-locus case: the change in frequency of the modifier decomposes into a term proportional to the degree of niche overlap γ and the difference between the mean fitness of diploids and haploids, and terms involving

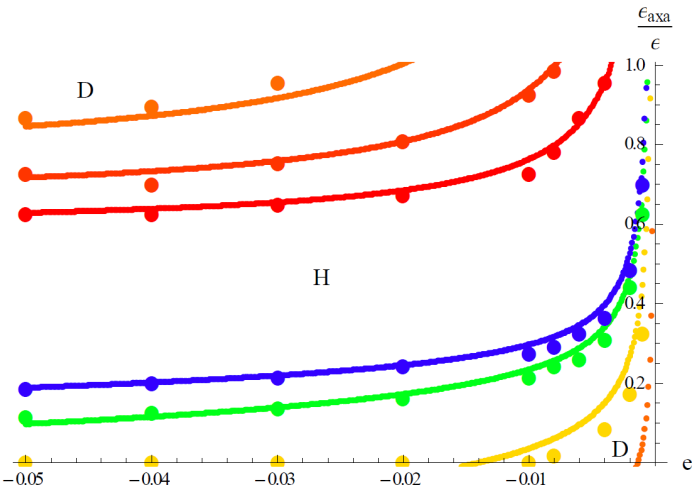

Figure 1.4: Evolution of ploidy as a function of (negative) epistasis e, and of the relative strength of epistasis in heterozygous diploids compared with haploids, for different levels of niche differentiation. Curves : Analytical results. Dots : Multilocus simulations. From blue to red: $\gamma = \{1, 0.5, 0.2, 0.1, 0.05, 0\}$. When haploids and diploids compete for more than 15% in a common niche (blue, green and yellow curves), weak epistasis advantages diploids through a direct competition advantage (below the curves). When niche differentiation is high (orange and red curves), the purge effect is dominant and diploidy will be favored if epistasis is higher than in haploids (above the curves). Other parameters: $s = 0.05$, $U = 1$, $h = 0.3$, $\alpha = 1$, $N = 20000$.

linkage disequilibria, representing indirect selection for modifier alleles allowing a better elimination of unfavorable alleles and genotypes. Appendix 2 provides analytical and simulation results assuming that epistasis is negative, and is the same for all pairs of mutations (indeed under positive epistasis, combinations of mutations quickly start to become advantageous and spread in the population). When the mutation rate U is sufficiently high, both the direct and indirect terms are mainly determined by the values of epistasis in haploids and in heterozygous diploids (e and $e_{a\times a}$, respectively), independently of s, h and α. In that case, direct competition favors diploidy when $e < 4e_{a\times a}$, and haploidy otherwise (see Appendix B for more details). Two types of indirect terms arise: a first term favoring the ploidy level that maximizes elimination of deleterious alleles (haploidy if $e < e_{a\times a}/2 < 0$, and diploidy otherwise), and a second term favoring the modifier allele that is more associated to the best two-locus combinations (Ab, aB, which under negative epistasis have higher mean fitness than coupling haplotypes AB, ab): haploidy if $e < e_{axa}$ and diploidy otherwise.

3.4 Intrinsic fitness differences between phases

The effect of intrinsic fitness differences between phases was studied using the alternation of generation model (see Appendix D for more details). Life cycle evolution is considered by introducing an allele (m) at the modifier locus that controls the timing of meiosis and evaluating whether its frequency increases when rare.

Mutants are able to invade when the leading eigenvalue of the system described by equations D.12c and D.12d in Appendix D, λ_l, is greater than one. When the per locus fitness difference between haploids and diploids ($|\varphi_d - \varphi_h|$) is of similar magnitude to the per locus mutation rate, $O(\epsilon^2)$, selection (s) is of a larger order of magnitude, $O(\epsilon)$, and linkage is loose (r of $O(1)$):

$$\lambda_l \approx 1 + h_m \delta \left(\varphi_d - \varphi_h\right) + s \left(2h - \alpha\right) \hat{p}_a + O(\epsilon^3) \tag{1.11}$$

The first term of equation 1.11 corresponds to the intrinsic fitness advantage of one phase over the other, and would favor modifiers alleles increasing the length of the haploid phase when $\varphi_h > \varphi_d$, and the diploid phase otherwise. The second term corresponds to the direct effect of the deleterious allele a: assuming panmixia, allele a is mainly present in heterozygous Aa zygotes. If an Aa individual remains diploid, its fitness is $1 - hs$, while if it enters meiosis its haploid offspring will have an average fitness of $1 - s\alpha/2$. Therefore, diploidy has an advantage in terms of mean fitness of offspring when $h < \alpha/2$ (this corresponds to the benefit of masking partially recessive mutations usually described for $\alpha = 1$).

Replacing p_a by 1.4 in equation 1.11, we found that equilibrium length of the diploid $\tilde{d}$ phase should satisfy:

$$\varphi_h - \varphi_d = s \left(2h - \alpha\right) \frac{u \, \mathrm{e}^{(1-\tilde{d})s\alpha}}{1 - \mathrm{e}^{s[\tilde{d}h + (1-\tilde{d})\alpha]}}$$

In the absence of intrinsic fitness differences between phase, only diploidy ($\tilde{d} = 1$) or haploidy ($\tilde{d} = 0$) evolve (depending on $(2h - \alpha)$). While Hall (2000) showed that that biphasic haploid-diploid life cycles could evolve in a polymorphic model (figure D.1a) when mutations occurred at meiosis, it is not the case in the alternation of generation model (figure D.1b). Intermediate equilibria ($0 < \tilde{d} < 0$) do exist when diploids have higher intrinsic fitness ($\varphi_d > \varphi_h$) and deleterious mutations are effectively partially dominant ($2h > \alpha$) or when haploids are favoured ($\varphi_h > \varphi_d$) and deleterious mutations are effectively partially recessive ($2h < \alpha$).

However, life cycles will only converge upon this strategy if $\varphi_d > \varphi_h$ and $2h > \alpha$. Otherwise, the singular strategy is a repelling point, see figure 1.5. Indeed, when diploids have higher intrinsic fitness ($\varphi_d > \varphi_h$) and deleterious mutations are effectively partially dominant ($2h > \alpha$), deleterious mutations (second term of equation 1.11) favor haploidy and their effect is stronger when the diploid phase is longer, because mutations reach higher frequencies in diploids: term 1 and 2 of equation 1.11 equilibrate. On the contrary, when haploids have higher intrinsic fitness ($\varphi_h > \varphi_d$) and deleterious mutations are effectively partially dominant ($2h < \alpha$), the diploid advantage of masking increases with the length of the diploid phase, and

decreases when the haploid phase is longer, because the equilibrium frequency of deleterious mutations is lower.

After convergence on a haploid-diploid strategy, we can then ask whether this singular strategy is evolutionarily stable. Using the same weak selection approximations as above, evolutionary stability is given by:

$$\frac{\delta^2 \lambda_l}{\delta d^2}\Big|_{d=\tilde{d}} = \frac{(\varphi_d - \varphi_h) s (2h - \alpha)(1 - r_{ma}) w_a[\tilde{d}] w_{Aa}[\tilde{d}]}{w_A[\tilde{d}] w_{AA}[\tilde{d}] - (1 - r_{ma}) w_a[\tilde{d}] w_{Aa}[\tilde{d}]} \tag{1.12}$$

When convergence is stable (requiring that $\varphi_d > \varphi_h$ and $2h > \alpha$), the singular strategy is evolutionarily unstable (1.12 is positive). Thus we expect weak disruptive selection after this singular point is reached. Indeed, our multilocus simulations sometimes displayed branching after 100,000 generations, such that there was a proportion $\tilde{d}$ of haploid alleles ($d = 1$), and a proportion $(1 - \tilde{d})$ of diploid alleles ($d = 0$). Increasing the number of generations always lead to branching when it was not observed by this time.

We extrapolated our two-locus result to consider deleterious mutations across L viability loci by assuming that these loci are loosely linked, autosomal and nonepistatic. In this case, the overall strength of selection on the modifier can be obtained by replacing p_a by the mean number of mutations per haploid genome N_{mut} (given by equation 1.5). Analytical predictions are checked with multilocus simulations in figure 1.5.

3.5 Distribution of mutational effects

The previous results assumed identical effects of mutations at all loci. However, these analysis can be extended to incorporate distributions of haploid and diploid selection coefficients across loci. In the following we denote $\overline{s_h}$ and $\overline{s_d}$ the average selection coefficients against deleterious alleles in haploids and in heterozygous diploids, respectively ($s\alpha$ and hs in the previous sections). Assuming that the variance of haploid and diploid selection coefficients s_h and s_d is weak, at a given locus we have $s_h = \overline{s_h} + \epsilon_h$, $s_d = \overline{s_d} + \epsilon_d$ (where ϵ_h and ϵ_d vary across loci). Replacing $s\alpha$ by s_h and hs by s_d in equation 1.8 (after expressing allele frequency p_a in terms of selection coefficients) and taking a Taylor series to the second order in ϵ_h, ϵ_d, one obtains an expression for the change in frequency of the modifier in terms of $\overline{s_h}$, $\overline{s_d}$, $\mathrm{E}\left[\epsilon_h^2\right] = \mathrm{Var}[s_h] = \sigma_h^2$, $\mathrm{E}\left[\epsilon_d^2\right] = \mathrm{Var}[s_d] = \sigma_d^2$ and $\mathrm{E}\left[\epsilon_h \epsilon_d\right] = \mathrm{Cov}[s_h, s_d] = \rho_{h_d} \sigma_h \sigma_d$. The term involving σ_h and σ_d writes:

$$\frac{d_m p q_m U}{\bar{s}^3} \begin{bmatrix} \gamma(2-d)\left[d\overline{s_h}\sigma_d(\sigma_d - \rho_{hd}\sigma_h) - (1-d)\overline{s_d}\sigma_h(\sigma_h - \rho_{hd}\sigma_d)\right] \\ +d(1-d)\left[\left(\overline{s_h}\sigma_d - \overline{s_d}\sigma_h\right)^2 + 2\left(1-\rho_{hd}\right)\overline{s_h}\sigma_d\overline{s_d}\sigma_h\right] \end{bmatrix} \tag{1.13}$$

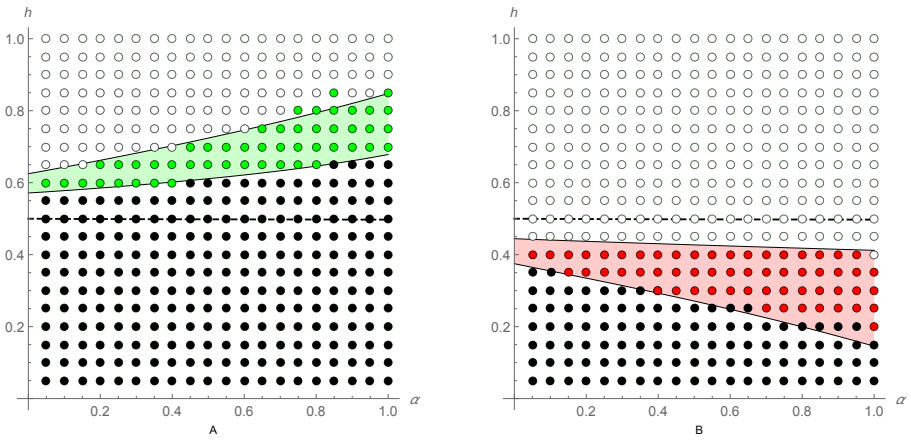

Figure 1.5: Parameter space where haploidy, diploidy and haploid-diploid life cycles are favoured where the relative strength of selection in haploids compared to homozygous diploids(α) and dominance h vary. Background colors: prediction from the two-locus stability analysis extrapolated to multiple loci. Circles: multilocus simulation results starting from three different initial haploidy rates (d_{init} = 0.01, 0.5, or 0.99), with population size 20,000. White: evolution toward haplonty. Green: convergence stable haploid-diploid life cycles. Red: either haploidy or diploidy is favoured, with a repelling state in between. Black and gray: evolution toward diplonty. (A): Diploids have higher intrinsic fitness ($S_h = 0$, $S_d = 0.025$) and (B): haploids have higher intrinsic fitness ($S_h = 0.025$, $S_d = 0$) Map length: $R = 100$. The dashed lines show where haploidy (above dashed lines) and diploidy (below dashed lines) are favoured when there is no difference in intrinsic fitness ($S_h = S_d = 0$). Mutants change the life cycle by a small amount ($|\delta| = 0.002$) and the genome-wide haploid mutation rate, $U = 0.1$.

with $\bar{s} = d\overline{s_d} + (1-d)\overline{s_h}$.

The term on the first line of equation 1.13 corresponds to the effect of σ_h and σ_d on selection on the modifier through the direct competition effect, showing that the variance in diploid selection coefficients favors diploidy, while the variance in haploid selection coefficients favors haploidy. Indeed, it can be shown that increasing the variance in selection coefficients in a given phase gives rise to several mutations with small effects, which increases the mean number of mutations per chromosome, but also increases the mean fitness of individuals in this phase, while decreasing the mean fitness of individuals in the other phase (as long as $0 < d < 1$). This effect is reduced when the correlation between haploid and diploid fitness effects of mutations (ρ_{hd}) increases. Qualitatively similar results were found using the alternation of generation model (see Appendix D).

The term on the second line of equation 1.13 corresponds to the effect of σ_h and σ_d on indirect selection. Due to non-linearities in s_d and s_h of the term representing indirect selection (involving linkage disequilibrium), one can show that both σ_d and σ_h tend to favor diploidy during selection in diploid individuals, and favor haploidy during selection in haploids. However, the second effect is weaker because recombination occurs before haploid selection, thereby reducing the amount of linkage disequilibrium and the importance of indirect selection. As a consequence, variances in mutational effects tend to favor diploidy through the indirect term.

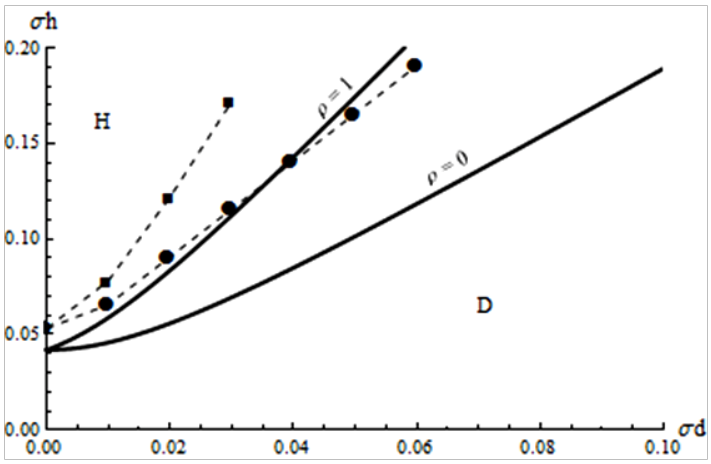

Figure 1.6: Effect of variances and covariances of selection coefficients across loci on the evolution of ploidy. Solid curves: analytical predictions from 1.8 and 1.13; dashed curves: multilocus simulation results with a genetic map length R of 10 Morgans, for $\rho = 0$ (dots) and $\rho = 1$ (squares). Parameter values: $s = 0.05$, $U = 1$, $h = 0.3$, $\alpha = 1$, $N = 20000$.

Multilocus simulations (incorporating distributions of mutational effects) confirm these predictions when $\gamma = 1$ (see Figure 1.6) note that when $\gamma = 1$ the indirect term has a negligible effect on the dynamics.

4 Discussion

In this chapter, we explored the effect of several genetic and ecological differences between the haploid and the diploid phases on the evolution of life cycles. Previous models on the effect of deleterious mutations on the evolution of ploidy cycles always assumed that mutations have the same fitness effect in haploids and homozygous diploids. However, one may expect that many mutations may have different effects on haploid and diploid individuals, as illustrated by a recent experiment on the yeast *Saccharomyces cerevisiae* (Gerstein 2013). Here, we showed that the 'masking' effect favors diploidy whenever ($2hs_d < s_h$), and haploidy otherwise. Conversely, indirect selection favors haploidy whenever ($s_h < s_d$), and diploidy otherwise. While there is no obvious reason why on average deleterious alleles should have higher or lower effects in haploids or in homozygous diploids (at least in isomorphic species), it would be particularly interesting to obtain more estimates on properties of spontaneous deleterious mutations in haploid-diploid species, and compare average effects in both phases. In heteromorphic species, a higher number of genes may be expressed in one phase than in the other, in which case many mutations may be deleterious in the phase where more genes are expressed, but neutral in the other.

We additionally explored the effects of variances and covariances of s_d and s_h across loci, and showed that increasing the variance of mutational effects in a given phase tends to favor

that phase through the direct competition term. As shown by our analytical model, what matters (in the absence of inbreeding) is the variance of hs_d in diploids, versus the variance of s_h in haploids. Interestingly, yeast deletion data indicate that the heterozygous effects of deleterious mutations may be much less variable than their homozygous effects, due to a negative correlation between h and s (Phadnis and Fry 2005, Agrawal and Whitlock 2011, Manna et al. 2011). Even if s_d and s_h are on average the same, it may thus be that the variance of hs_d is much lower than the variance of s_h, which would tend to decrease selection for diploidy (through direct competition). By contrast, variances in selection coefficients tend to increase selection for diploidy through indirect effects (involving linkage disequilibria).

Previous genetic models also assumed that haploid and diploid baseline fitness were equal, which is not generally the case (e.g Thornber 2006, Zörgö et al. 2013). Here, we showed that when diploids have a higher intrinsic fitness than haploids and when deleterious mutations are effectively partially dominant (haploid advantage), stable haploid-diploid life cycles can be maintained. In Appendix D, we show that these conditions become less restrictive when recombination is lower. Previous models predicting the evolution of biphasic haploid-diploid life cycles have posited indirect benefits from decreasing senescence by reducing phase-specific generation time (Jenkins 1993), reducing the frequency of sexual reproduction (Richerd et al. 1993), or exploiting different ecological niches (Bell 1997, Hughes and Otto 1999). However, haploid-diploid life cycles are not a unique way of accessing to these benefits. For example, diploid or haploid species can reduce generation times or the frequency of sexual reproduction without evolving haploid-diploid life cycles. Similarly, differentiated life cycle stages (Steenstrup alternations), phenotypic plasticity or genetic polymorphism can allow diploid or haploid species to exploit multiple ecological niches without tying growth form to the sexual cycle. Here, we used a population genetic model to show that haploid-diploid life cycles can evolve as a direct consequence of ploidy if the intrinsic fitnesses of haploids and diploids are not equal. When the balance between intrinsic fitness differences and the effect of mutations favours convergence on haploid-diploid strategies, disruptive selection then arises such that polymorphisms can evolve with alternative alleles coding for longer haploid and longer diploid phases (i.e., a polymorphic strategy of specialists).

Introducing a parameter representing the degree of niche differentiation among haploid and diploid individuals (γ) illustrates the fact that 'short-term effects' (that stem from differences in mean fitness between phases) can affect the evolution of life cycles only if there is direct competition among haploid and diploid individuals. If these individuals occupy different niches, and if population regulation occurs within each niche, short-term effects vanish. By contrast, our model shows that indirect selection ('purging' effect) is not affected much by niche differentiation among phases. Contrarily to the model developed by Hughes and Otto (1999), here niche differentiation does not lead to stable biphasic life-cycles. However, Hughes

and Otto's model included a demographic component and density-dependence terms, so that an individual entering a more vacant niche has more chances to survive and contribute to the next generation. This ecological component is absent from our model, which assumes that an infinite number of individuals enter each niche at every generation. Whatever the parameters, our model predicts evolution towards either a diploid or a haploid life cycle, and individual-based simulations confirm this result (in particular, having many selected loci located at different genetic distances from the modifier does not lead to stable haploid-diploid cycles).

In the absence of intrinsic fitness differences between phases, deleterious mutations tend to destabilize biphasic life cycle, while ecological differentiation with density dependence should maintain them. In order to quantify the relative effect of both factors in the evolution of life cycle, we need to introduce mutations in an ecological scenario including density-dependence. Furthermore, haploids and diploids exploiting different ecological niches should differ in other ecological parameters, wich can affect first their baseline fitness, and the strength of selection. We have already seen that intrinsic fitness differences between phase can lead to stable biphasic life cycles. In Chapter 2, we will analyzed a polymorphic model in a density-dependent scenario to explore the interplay between deleterious mutations and the ecological advantage of biphasic life cycles when the haploid and the diploid niches are differentiated.

In this chapter, we only considered deleterious mutations. The alternation model is extended to beneficial mutations occurring during adaptation in Appendix D, and shows that beneficial mutations favor haploidy, diploidy or intermediate life cycles following the same mechanisms as deleterious mutations. In particular, biphasic life cycles can be maintained when diploids intrinsic fitness is higher than haploids fitness, and when beneficial alleles are recessive. One can also imagine that selection at some loci may favor different alleles in haploids and diploids: such ploidally antagonistic selection can maintain stable polymorphism (Ewing 1977, Immler et al. 2012), which may have important effects on the evolution of the relative duration of both phases. The model presented in Chapter2 explores the effect of selection coefficients of ploidally antagonistic selection.

Finally, in the alternation of generation model with intrinsic fitness difference between phases, we predicted evolutionary branching when intermediate level of diploidy are favored. In their ecological model, Hughes and Otto (1999) found that when haploid-diploid life cycles were maintained, many alleles at the modifier locus could coexist as long as the average proportion of haploids and diploids was optimal to exploit both ecological niches. In Chapter 2, we assess the occurrence of evolutionary branching when mutations occur in a density-dependent model. In addition to the spatial differentiation of the haploid and the diploid phases, temporal variations of the relative abundance (Bolton and Joska 1993, Otaiza et al. 2001, Dyck and De Wreede 2006) or fecundity (Santos and Duarte 1996) of haploids and diploids have been

reported, suggesting that environmental fluctuations may favor alternatively diploid sporophytes or haploid gametophytes. We analyze the effect of temporal fluctuations of the haploid and the diploid niche size in the next chapter.

Chapter 2

Interactions between genetic and ecological effects on the evolution of life cycles

Contents

Interactions between Genetic and Ecological Effects on the Evolution of Life Cycles

Marie Rescan,[1,*] Thomas Lenormand,[2] and Denis Roze[1]

1. CNRS, Unité Mixte Internationale 3614, Evolutionary Biology and Ecology of Algae, Roscoff, France; and Sorbonne Universités, Université Pierre et Marie Curie, University of Paris 6, Roscoff, France; 2. Centre d'Ecologie Fonctionnelle et Evolutive Unité Mixte de Recherche 5175, CNRS, Université de Montpellier, Université Paul-Valéry Montpellier, École Pratique des Hautes Études, 1919 Route de Mende, F-34293 Montpellier, Cedex 5, France

Submitted February 25, 2015; Accepted July 20, 2015; Electronically published November 25, 2015

Online enhancement: appendix. *Dryad data:* http://dx.doi.org/10.5061/dryad.40qp5.

abstract: Sexual reproduction leads to an alternation between haploid and diploid phases, whose relative length varies widely across taxa. Previous genetical models showed that diploid or haploid life cycles may be favored, depending on dominance interactions and on effective recombination rates. By contrast, niche differentiation between haploids and diploids may favor biphasic life cycles, in which development occurs in both phases. In this article, we explore the interplay between genetical and ecological factors, assuming that deleterious mutations affect the competitivity of individuals within their ecological niche and allowing different effects of mutations in haploids and diploids (including antagonistic selection). We show that selection on a modifier gene affecting the relative length of both phases can be decomposed into a direct selection term favoring the phase with the highest mean fitness (due to either ecological differences or differential effects of mutations) and an indirect selection term favoring the phase in which selection is more efficient. When deleterious alleles occur at many loci and in the presence of ecological differentiation between haploids and diploids, evolutionary branching often occurs and leads to the stable coexistence of alleles coding for haploid and diploid cycles, while temporal variations in niche sizes may stabilize biphasic cycles.

Keywords: evolution of life cycles, density dependence, deleterious mutations, multilocus model, evolutionary branching.

Introduction

Alternation of meiosis and syngamy in sexual eukaryotes results in the alternation of haploid and diploid generations, whose relative duration and degree of development vary largely among taxa. Most animals and some protists (e.g., diatoms, oomycetes) have diploid life cycles: the haploid phase is reduced to a single cell, the gamete. Other organisms (e.g., ascomycetes, charophytes, dinoflagellates) have haploid cycles, where the diploid phase is reduced to the zygote, which undergoes meiosis before any mitotic development. Finally, many species present haploid-diploid life cycles, where somatic development occurs in both haploid and diploid phases. While the relative development of the haploid (gametophytic) generation is rather limited in seed plants (spermatophytes)—with only a few cell divisions to form the pollen grain and the embryo sac—it is much more important in many fungi, mosses, and macroalgae. In particular, the life cycle of many red algae involves an alternation between haploid and diploid individuals (which may have very different morphologies), while many different life cycles are observed among brown algae, from the isomorphic, haploid-diploid cycle of Dictyotales to the diploid cycle of Fucales.

The limited development and/or short duration of one phase may strongly limit the opportunity for selection. For instance, in many female animals, there is virtually no haploid phase (since the last meiotic division of the egg takes place only at fertilization). By contrast, much selection can occur among male gametes, even though they have very limited development (Joseph and Kirkpatrick 2004). As a consequence, the problem of the evolution of life cycles (i.e., the relative degrees of development of the haploid and diploid phases) is often recast in terms of the opportunity of selection within each phase. From the early nineties, different theoretical studies have explored how genetical or ecological factors may affect the evolution of life cycles (for reviews, see, e.g., Valero et al. 1992; Mable and Otto 1998; Otto and Gerstein 2008). From a genetical perspective, diploids may benefit from more efficient repair of DNA damage because of the presence of a homologous chromosome that may serve as a template (Michod and Wojciechowski 1994). In addition, diploids may benefit from an increased fitness as a result of the masking of deleterious mutations: for this, the fitness effect of mutations in the heterozygous state must be sufficiently low to compensate for the fact that a diploid

* Corresponding author; e-mail: marie.rescan@sb-roscoff.fr.

Am. Nat. 2016. Vol. 187, pp. 19–34.
DOI: 10.1086/684167

tends to carry twice as many mutations as a haploid. Under random mating, and assuming that mutations have the same effect in haploids and homozygous diploids, this requires only that deleterious alleles are partially recessive on average (Perrot et al. 1991), which seems to be the case (Halligan and Keightley 2009; Manna et al. 2011). However, mutations increasing the relative length of the diploid phase may not necessarily be favored in this situation. In particular, Otto and Goldstein (1992) showed that modifier alleles coding for a longer diploid phase tend to be associated with more heavily loaded genomes because selection is less efficient among diploids. If linkage is sufficiently tight, this effect may favor modifier alleles increasing the haploid phase, while diploidy is favored under looser linkage because of the masking effect (as long as deleterious alleles are partially recessive). As a consequence, reproductive systems that reduce the effective recombination rate (inbreeding, partial asexuality) tend to increase selection for haploidy (Otto and Marks 1996). Similar results are obtained when considering the spread of beneficial alleles within a population: selection should generally be more efficient in haploids (which was confirmed by evolution experiments on *Saccharomyces cerevisiae*; e.g., Zeyl et al. 2003; Gerstein et al. 2010), but partially dominant beneficial alleles may favor diploidy in sexual populations (Orr and Otto 1994). Finally, Nuismer and Otto (2004) proposed that host-parasite interactions should favor diploidy in the host (because of the benefit of a higher number of recognition alleles) and haploidy in the parasite (for the opposite reason). In general, these genetic models predict evolution toward either haploidy or diploidy (depending, for example, on the degree of dominance of deleterious mutations or on the mating system) but cannot explain the evolutionary stability of haploid-diploid life cycles unless considering additional mechanisms directly favoring biphasic cycles.

Importantly, most of these previous models assume that deleterious alleles have the same fitness effect in haploids and in homozygous diploids. However, haploids and diploids often differ in terms of physiology, morphology, or ecology (Thornber 2006). Transcriptomic studies on haploid-diploid species show that a fraction of genes is expressed in one phase only (Coelho et al. 2007; Von Dassow et al. 2009; Rokitta et al. 2011), and mutations in these genes should thus have no effect on fitness in the other phase. More generally, selective pressures on different genes may differ quantitatively in both phases, leading to different selection coefficients of mutations in both phases. The fact that mutations may have different fitness effects in haploids and diploids is illustrated by several experimental studies on yeast: in particular, Szafraniec et al. (2003) found that ethyl methanesulfonate–induced spontaneous mutations were more deleterious in haploids than in homozygous diploids, while Gerstein (2013) showed that mutations conferring tolerance to nystatin often have larger fitness effects in haploids than in homozygous diploids. Furthermore, Thompson et al. (2006) observed that beneficial mutations fixed during the adaptation of haploid and diploid mutator populations have different properties, with the mutations fixed in diploids being more generalist. One can also imagine that selection at some loci may favor different alleles in haploids and diploids; such ploidally antagonistic selection can maintain stable polymorphism (Ewing 1977; Immler et al. 2012) and was recently shown to have the potential to drive (in combination with sexually antagonistic selection) the evolution of ploidy differences between sexes (Immler and Otto 2014). However, the overall impact of quantitative and qualitative variations of mutational effects across phases on the evolution of the relative duration of these phases has received very limited attention.

One reason why selection may differ among phases is that haploids and diploids may not be ecologically equivalent. These differences may be cryptic and occur even among morphologically similar haploids and diploids, such as in the isomorphic red alga *Gracilaria gracilis* (Destombe et al. 1993; Hughes and Otto 1999) or the nearly isomorphic brown alga *Ectocarpus crouaniorum*, where sporophytes and gametophytes are typically found on different substrata (rock/shells versus other algal species; Couceiro et al. 2015). In addition, temporal variations of the relative abundance (Bolton and Joska 1993; Otaiza et al. 2001; Dyck and De Wreede 2006) or fecundity (Santos and Duarte 1996) of haploids and diploids have been reported, suggesting that environmental fluctuations may favor alternatively diploid sporophytes or haploid gametophytes. As shown by Hughes and Otto (1999) using a model incorporating density dependence effects, differentiation between the haploid and the diploid ecological niches may favor biphasic life cycles over purely haploid or diploid cycles (since a mutant using a relatively empty niche tends to increase in frequency). This model thus provides a plausible mechanism for the maintenance of biphasic cycles, given that many haploid-diploid species (such as algae or mosses) are often found in dense populations, in which individuals may be strongly affected by intraspecific competition (e.g., Reed 1990; Paalme et al. 2013). However, in the presence of genetic variability for fitness, ecological differentiation between phases may also affect the relative importance of purging and masking effects, depending in particular on the level of competition within and between phases. For example, the masking advantage associated with diploidy (when deleterious alleles are partially recessive) should vanish when the haploid and diploid niches are fully separated (i.e., when haploids do not compete against diploids) and when selection is soft within each niche, so that the total reproductive output from a niche is not affected by selection (e.g., Wallace 1975; Agrawal 2010). Furthermore, different experimental studies (to which we will return in the discussion) suggest that the overall strength of selection against deleterious

alleles may increase with density, in which case ecological differences between haploids and diploids may affect the relative efficiency of selection in the two phases.

In this article, we explore the interplay between ecological and genetic effects on the evolution of life cycles. As in Hughes and Otto (1999), our ecological model is based on a logistic model of population growth, with a variable degree of overlap between the haploid and diploid niches. We assume that deleterious alleles affect the efficiency with which individuals compete for resources within each niche and allow them to have different effects in haploids and homozygous diploids (including ploidally antagonistic selection). In the following, we first use a two-locus model to derive analytical results on the strength of selection on a modifier gene affecting the probability of developing as a diploid or as a haploid (using the approach of Otto and Goldstein 1992), and we explore the effects of niche differentiation and differences in mutational effects between phases. In general, we will see that selection on the modifier can be decomposed into a term stemming from differences in mean fitness between ploidy levels (short-term effect, due to differences in niche availability and in the fitness effect of mutations between phases) and terms involving linkage disequilibria between the modifier and selected locus (indirect selection), favoring the phase in which natural selection is most efficient. We will then extrapolate from our two-locus model to consider deleterious alleles occurring at a large number of loci and check our analytical predictions using multilocus, individual-based simulations. These simulations show that in the absence of deleterious mutation, many different strategies may coexist in the population, provided that the haploid and diploid niches are sufficiently differentiated (in agreement with predictions from Hughes and Otto 1999). With deleterious mutations, however, only extreme strategies (corresponding to purely haploid or purely diploid cycles) are maintained, either alone or coexisting. Finally, in both cases (with or without deleterious mutation), adding temporal fluctuations in the relative sizes of ecological niches may lead to the fixation of biphasic life cycles, in agreement with bet-hedging theory (e.g., Philippi and Seger 1989).

Model

Analytical Model

Our model represents a panmictic population undergoing a biphasic life cycle with nonoverlapping generations. The first event of each generation corresponds to gamete union, assumed to be random. As in the study by Otto and Goldstein (1992), diploid zygotes can either enter meiosis immediately to form new haploid individuals or develop as diploids. The probability to develop as a diploid or as a haploid is controlled by a modifier locus with two alleles, M and m: zygotes MM, Mm, and mm develop as diploids with probabilities d, $d + h_m\delta$, and $d + \delta$, respectively; h_m thus measures dominance of allele m (the different parameters and variables of the model are summarized in table 1). The population follows a discrete-time logistic model of population growth. The fitness of the different genotypes in the different phases depends on several parameters. We first consider the fitness of haploids (W_h) and diploids (W_d), defined as the number of offspring—or half the number of gametes—that will participate to the next generation, in the absence of deleterious mutation. It is given by

$$W_h = 1 + r_h\left(1 - \frac{N_h + \gamma_{hd}N_d}{K_h}\right),$$
$$W_d = 1 + r_d\left(1 - \frac{N_d + \gamma_{dh}N_h}{K_d}\right), \quad (1)$$

where N_h and N_d are the numbers of parental haploids and diploids (haploid individuals produce gametes by mitosis, while diploids produce gametes by meiosis). As can be seen from equation (1), K_h and K_d correspond to the carrying capacities of a purely haploid ($d = 0$) and purely diploid ($d = 1$) population, which depends on resource abundance within the haploid and diploid niches and on the efficiency with which individuals use those resources. Coefficient γ_{dh} (respectively, γ_{hd}) measures the efficacies by which haploids (diploids) compete for the diploid (haploid) resources (Hughes and Otto 1999). If $\gamma_{dh} = \gamma_{hd} = 1$, both phases use the same resources and thus compete directly against each other, while $\gamma_{dh} = \gamma_{hd} = 0$ corresponds to the case where haploids and diploids occupy different ecological niches. As shown by Hughes and Otto (1999), haploid-diploid life cycles (i.e., intermediate values of d) may be favored when $\gamma_{dh}\gamma_{hd} < 1$. Finally, note that the baseline reproductive factors $1 + r_h$ and $1 + r_d$ (corresponding to the fitness of haploids and diploids in the absence of intraspecific competition) incorporate all effects of the environment on fertility and mortality (e.g., gamete or zygote mortality) that do not depend on density.

We then introduce a second locus affecting the sensitivity of individuals to competition. Following Christiansen and Loeschcke (1980), we consider two different forms of selection. In the first scenario, genotypes differ in their efficiency in using limited resources, so that individuals carrying more deleterious alleles need more resources to produce offspring. This may be represented by multiplying the coefficients K_d and K_h in equation (1) by factors that depend on the genotype of the individual. Note that selection affects demography under this scenario because it changes carrying capacities. The second scenario corresponds to a situation where genotypes differ in their competitivity for resources: individuals suffer more from competition with genotypes carrying fewer deleterious alleles and less from

Table 1: Parameters and variables used in the model

Symbol	Definition
N	Population size
d	Probability that a zygote with genotype *MM* at the modifier locus develops as a diploid
δ	Change in the probability to develop as a diploid caused by the modifier allele m in the homozygous state
h_m	Dominance coefficient of allele m
W_{h}^i, $\overline{W_{\mathrm{h}}}$	Fitness (number of offspring at the next generation) of a haploid individual with genotype i and average fitness of haploids
W_{d}^i, $\overline{W_{\mathrm{d}}}$	Fitness of a diploid with genotype i and average fitness of diploids
r_{h}, r_{d}	Baseline growth rate of haploids and diploids ($W_{\mathrm{h}} = 1 + r_{\mathrm{h}}$ and $W_{\mathrm{d}} = 1 + r_{\mathrm{d}}$ in the absence of competition)
K_{h}, K_{d}	Haploid and diploid carrying capacities
γ_{hd}, γ_{dh}	Competitive effect of diploids on haploids (γ_{hd}) and of haploids on diploids (γ_{dh})
C_{h}, C_{d}	Strength of density-dependent competition acting on haploids and diploids (see eq. [7])
λ, τ	Amplitude (λ) and period (τ) of the temporal fluctuations of K_{h} and K_{d}
ω_{h}^i, $\overline{\omega_{\mathrm{h}}}$	Sensitivity to competition of a haploid with genotype i and average sensitivity to competition of haploids
ω_{d}^i, $\overline{\omega_{\mathrm{d}}}$	Sensitivity to competition of a diploid with genotype i and average sensitivity to competition of diploids
β	Degree of softness of selection on sensitivity to competition
α	Effect of allele a on the sensitivity to competition of homozygous diploids
h	Dominance coefficient of allele a
$\rho\alpha$	Effect of allele a on the sensitivity to competition of haploids (ρ is thus the ratio between the effect of allele a in haploids and diploids)
s_{h}, s_{d}	Strength of selection against allele a in haploids and diploids (see eq. [10])
u	Mutation rate from allele A to allele a
r_{ma}	Recombination rate between the (M, m) and (A, a) locus
p_m, p_a, D_{ma}	Frequencies of alleles m and a and linkage disequilibrium between these alleles
U	Genomic deleterious mutation rate (multilocus model)
R	Genome map length (multilocus model)

competition with genotypes carrying more deleterious alleles. Under this scenario, selection may not affect demography (soft selection): indeed, the fact that some individuals leave fewer offspring (because they are poor competitors) may be compensated by the fact that better competitors leave more offspring. If we consider for a moment that haploids and diploids do not compete against each other ($\gamma_{\mathrm{dh}} = \gamma_{\mathrm{hd}} = 0$), the fitnesses of haploids with genotype i for the two scenarios just mentioned may be written as

$$W_{\mathrm{h}}^i = 1 + r_{\mathrm{h}}\left(1 - \frac{N_{\mathrm{h}}}{K_{\mathrm{h}}}\frac{\omega_{\mathrm{h}}^i}{1 - \beta + \beta\overline{\omega_{\mathrm{h}}}}\right). \quad (2)$$

The same equation holds for diploids, with all h subscripts replaced by d subscripts. In this equation, ω_{h}^i represents the sensitivity to competition of haploid genotype i, and $\overline{\omega_{\mathrm{h}}}$ is the average sensitivity to competition of the different haploid genotypes. When $\beta = 1$, the genotype with the lowest sensitivity is favored by selection. However, the overall demography is not affected by variation in sensitivities, since the decrease in net growth rate of high-sensitivity genotypes is exactly compensated by the opposite increase in growth rate of low-sensitivity genotypes (the division by $\overline{\omega_{\mathrm{h}}}$ ensures this behavior). This situation thus represents the soft selection regime mentioned above (selection through difference in competitivity). When $\beta = 0$, the genotype with the lowest sensitivity is still favored by selection, but selection now affects the overall demography, as the carrying capacity of genotype i becomes $K_{\mathrm{h}}/\omega_{\mathrm{h}}^i$. This situation represents variation among genotypes in their efficiency in using resources, as mentioned above. In fact, changing the parameter β allows one to tune the softness of selection and consider selective scenarios that are intermediate between the two extreme scenarios just described (selection through differences in competitivity for $\beta = 1$ or differences in efficiency of using resources for $\beta = 0$).

Now, genetic variation in sensitivity to competition can be combined with partial competition between haploids and diploids introduced in equation (1), yielding the full fitness functions

$$W_{\mathrm{h}}^i = 1 + r_{\mathrm{h}}\left[1 - \frac{\omega_{\mathrm{h}}^i}{K_{\mathrm{h}}}\left(\frac{N_{\mathrm{h}}}{1 - \beta + \beta\overline{\omega_{\mathrm{h}}}} + \gamma_{\mathrm{hd}}\frac{N_{\mathrm{d}}}{1 - \beta + \beta\overline{\omega_{\mathrm{d}}}}\right)\right],$$
$$W_{\mathrm{d}}^i = 1 + r_{\mathrm{d}}\left[1 - \frac{\omega_{\mathrm{d}}^i}{K_{\mathrm{d}}}\left(\gamma_{\mathrm{dh}}\frac{N_{\mathrm{h}}}{1 - \beta + \beta\overline{\omega_{\mathrm{h}}}} + \frac{N_{\mathrm{d}}}{1 - \beta + \beta\overline{\omega_{\mathrm{d}}}}\right)\right], \quad (3)$$

where W_{h}^i and W_{d}^i are the fitnesses of haploids and diploids with genotype i. Overall, fitness is density dependent as in a classical logistic model but also depends on (1) the degree of competition between haploid and diploids (γ_{dh}, γ_{hd} mea-

suring the degree of niche overlap); (2) individual sensitivity to competition, which is genotype dependent (ω_h^i, ω_d^i); and (3) the degree of softness of selection against individuals with higher sensitivities (β). In our two-locus model, sensitivities depend on genotype at a single locus with two alleles A and a (located at recombination distance r_{ma} from the modifier locus) and are written as

$$\begin{aligned} \omega_d^{AA} &= 1, \\ \omega_d^{Aa} &= 1 + h\alpha, \\ \omega_d^{aa} &= 1 + \alpha, \\ \omega_h^{A} &= 1, \\ \omega_h^{a} &= 1 + \rho\alpha. \end{aligned} \tag{4}$$

The parameter α thus measures the effect of allele a in homozygous diploids, h is the dominance coefficient of a, and ρ measures the effect of a in haploids relative to homozygous diploids. We will treat separately situations where $\rho > 0$ (selection has the same direction in both phases) and $\rho < 0$ (ploidally antagonistic selection).

In the two-locus analysis, we assume that the modifier effect and the strength of selection against the deleterious allele are weak (α, δ small), so that selection acting at both loci has a weak effect on population size at equilibrium. Assuming that both population size and the frequency of allele a have reached equilibrium, we express the change in frequency of the modifier Δp_m as a function of the frequencies of alleles a and m (p_a and p_m) and of the linkage disequilibrium D_{ma} ($D_{ma} = p_m - p_a p_m$, where p_{ma} is the frequency of ma haplotypes). In a second step, we use a quasi-linkage equilibrium (QLE) approximation to express the linkage disequilibrium in terms of allele frequencies; this approximation assumes that selection is weak relative to recombination, so that D_{ma} equilibrates quickly relative to the rate of change of allele frequencies.

As we will see, these two-locus results can be extrapolated to obtain an expression for the change in frequency of the modifier when deleterious mutations ($\alpha > 0$, $\rho > 0$) occur at a large number of loci, assuming that the effects of different deleterious alleles on sensitivity to competition are multiplicative (see "Multilocus Simulations"). When selection affects the efficiency of resource use ($\beta = 0$), mutations may have a strong effect on population size, which must be taken into account. As shown in the appendix, available online, the equilibrium population size and mean number of deleterious alleles per genome can be obtained by solving numerically a system of equations.

Multilocus Simulations

We used individual-based simulations (C++ program available in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.40qp5; Rescan et al. 2015) to test predictions from our analytical model when deleterious mutations ($\alpha > 0$, $\rho > 0$) segregate at a large number L of loci. Each individual carries either one or two copies of a chromosome (depending on its ploidy level) represented by a modifier locus (located at the midpoint of the chromosome) and a sequence of L bits (0 or 1) corresponding to the different loci. Mutations occur at a rate U per generation: the number of new mutations per chromosome is sampled from a Poisson distribution with parameter U and are distributed randomly; alleles at mutant loci are switched from 0 to 1 or from 1 to 0. Mutation and back mutation thus occur at the same rate, but back mutations should generally have negligible effects under the parameter values that we use (as deleterious alleles remain at low frequencies). We assume that all deleterious alleles have the same effects on sensitivity to competition (α, h, ρ) and that these effects multiply across loci: the sensitivity of a haploid carrying n deleterious alleles is given by $\omega_h = (1 + \rho\alpha)^n$, while the sensitivity of a diploid carrying n_{he} deleterious alleles in the heterozygous state and n_{ho} in the homozygous state is given by $\omega_d = (1 + h\alpha)^{n_{he}}(1 + \alpha)^{n_{ho}}$.

At the start of each generation, diploid zygotes undergo meiosis or not with probabilities depending on their alleles at the modifier locus. We assume additivity among modifier alleles (which can take any value between 0 and 1): a zygote with alleles coding for rates of diploidy d_1 and d_2 develops as a diploid with probability $(d_1 + d_2)/2$. If the individual develops as a haploid, meiosis occurs to produce a recombinant haplotype: the number of crossovers is sampled from a Poisson distribution with parameter R, while the position of each crossover is sampled from a uniform distribution. The next generation of zygotes is then generated as follows: the number of successful gametes (gametes that will participate to the next generation) produced by each individual is sampled from a Poisson distribution with parameter set to twice the fitness of the individual (calculated as explained above; see also appendix). If the individual is diploid, a recombinant haplotype is generated for each of these gametes, while gametes produced by haploid parents carry the same genotype as the parent. Finally, gametes fuse randomly to form the next generation of zygotes.

During the first few preliminary generations (generally 200), the modifier locus is fixed for an allele coding for an initial diploidy rate d_{init}, while all selected loci are fixed for allele 0, so that the population can reach its ecological equilibrium in the absence of mutation. Then, deleterious mutations are introduced at rate U per chromosome (the diploidy rate being still fixed to d_{init}), so that the population reaches mutation-selection equilibrium (after generally 2,000 generations). After that, mutations at the modifier locus are introduced at a rate μ_M per generation. When a mutation occurs, the rate of diploidy coded by the mu-

tant allele is sampled from a uniform distribution between $d_{old} - 0.1$ and $d_{old} + 0.1$, where d_{old} is the value of the parent allele; if the new value is negative or higher than 1, it is set to 0 or 1, respectively. Simulations generally lasted 20,000 generations, which was sufficient in most cases for the rate of diploidy to reach equilibrium.

Finally, in additional simulations, we explored the effect of temporal fluctuations in the relative niche sizes of haploids and diploids. For this, we set $K_d = \overline{K_d}[1 + \lambda\sin(\pi t/\tau)]$ and $K_h = \overline{K_h}[1 - \lambda\sin(\pi t/\tau)]$ (where t is time in generations), so that both K_d and K_h undergo oscillations with amplitude λ ($\in [0, 1]$) and period τ, while the ratio K_d/K_h also fluctuates over time.

By default, growth rates r_h and r_d were set to 1.8 so that population size reaches a stable equilibrium and the population remains viable for values of the deleterious mutation rate U up to 1 or 2. Carrying capacities K_h and K_d were set to 15,000, so that population size reaches values that are compatible to what may be observed in real populations (in particular, large enough so that drift is not too strong). The degree of softness of selection was set to either 0 or 1, in order to contrast the two scenarios mentioned above concerning the effect of selection on demography. Finally, the default values of α and h (0.05 and 0.3) generate selection and dominance coefficients of deleterious alleles that are in the range of estimated values from mutation accumulation studies (e.g., Halligan and Keightley 2009; Manna et al. 2011).

Results

The change in frequency of a modifier allele affecting the probability to undergo selection as a diploid depends on population size, the frequency of allele a (that may be deleterious in both phases or under antagonistic selection), and linkage disequilibrium between the two loci. We first compute the equilibrium population size and equilibrium frequency of allele a and then use a QLE approximation to express linkage disequilibrium D_{ma}. Details of the mathematical derivations are given in the appendix. $\overline{W_d}$ and $\overline{W_h}$ stand for the mean fitnesses of diploids and haploids (averages of W_d^i and W_h^i given by eq. [3] over all diploids and haploids). All results are derived under the assumption that the modifier effect (δ) and selection acting on allele a (through its effect α on sensitivity to competition) are weak.

Equilibrium Population Size

Neglecting the effect of the modifier, population size at the next generation is given by $N' = N[d\overline{W_d} + (1 - d)\overline{W_h}]$. At equilibrium, we thus have

$$d\overline{W_d} + (1 - d)\overline{W_h} = 1. \quad (5)$$

Neglecting the effect of selection acting on allele a and the modifier effect, we have from equation (3)

$$\overline{W_d} = 1 + r_d - NC_d,$$
$$\overline{W_h} = 1 + r_h - NC_h, \quad (6)$$

where C_d and C_h measure the strength of competition acting on diploids and haploids:

$$C_d = \frac{r_d}{K_d}\left[d + \left(1 - d\right)\gamma_{dh}\right],$$
$$C_h = \frac{r_h}{K_h}\left[\left(1 - d\right) + d\gamma_{hd}\right]. \quad (7)$$

In absence of genetic variation for the rate of diploidy and the sensitivity to competition ($\delta = \alpha = 0$), population size reaches an equilibrium value $\hat{N}_0$:

$$\hat{N}_0 = \frac{d\,r_d + (1 - d)r_h}{d\,C_d + (1 - d)C_h}. \quad (8)$$

Population size increases with the diploid and haploid carrying capacities K_d and K_h and is maximized for intermediate levels of diploidy d when the haploid and diploid niches are at least partly separated ($\gamma_{dh}\gamma_{hd} < 1$).

Equilibrium Frequency of Allele a

As discussed in "Model," we consider two different situations: (1) allele a is deleterious in both phases ($\alpha > 0$, $\rho > 0$) and (2) allele a is beneficial in one phase but deleterious in the other (ploidally antagonistic selection; $\rho < 0$). In the first case, we assume that allele a is generated by mutation from allele A at a rate u per generation (and neglect back mutation), while in the second case, we focus on situations where polymorphism is maintained by selection only, and we neglect mutation. In both situations, we assume that population size is sufficiently large so that genetic drift can be ignored. As shown in the appendix, a first-order approximation for the change in frequency of a over one generation is given by

$$\Delta p_a \approx \{d\,s_d[h + (1 - 2h)p_a] + (1 - d)s_h\}p_a q_a + uq_a, \quad (9)$$

where s_d and s_h measure the strength of selection against allele a in haploids and diploids, respectively, and $q_a = 1 - p_a$. To the first order in α, s_d and s_h are given by

$$s_h = \rho\alpha NC_h,$$
$$s_d = \alpha NC_d. \quad (10)$$

As expected, density-dependent selection becomes more intense when competition is harsher (as measured by NC_h, NC_d); however, for a given value of N, s_h and s_d are not affected by the degree of softness of selection (β).

From equation (9), one obtains three possible equilibria for p_a. The first, $p_a = 1$, is trivial and corresponds to the fixation of allele a. The second equilibrium corresponds to mutation-selection balance when a is deleterious in both phases. The third corresponds to polymorphism maintained by antagonistic selection ($\rho < 0$).

Mutation-Selection Balance. Assuming that p_a is small at equilibrium, we can neglect terms in p_a^2 in equation (9). Furthermore, we can replace N by $\hat{N}_0$ in equation (10) to express Δp_a to the first order in α at the demographic equilibrium, which yields

$$p_a^{\text{del}} = \frac{u}{d\,h\,s_d + (1-d)s_h} = \frac{u}{\hat{N}_0\alpha[d\,h\,C_d + (1-d)\rho C_h]}. \quad (11)$$

This equilibrium frequency takes the same form as in Otto and Goldstein's (1992) model and decreases as the intensity of competition among individuals increases. In a purely haploid population ($d = 0$), we have $p_a^{\text{del}} = u/s_h$ with $s_h = \rho\alpha r_h$, while in a purely diploid population ($d = 1$), $p_a^{\text{del}} = u/h s_d$, with $s_d = \alpha r_d$ (note that baseline fecundities r_h and r_d determine the intensity of competition among haploid or diploid offspring in populations at demographic equilibrium).

The previous expressions neglect the effect of the deleterious allele on population size at equilibrium. While this is legitimate in the case of a single mutation, deleterious alleles occurring at many loci are more likely to affect population size (in particular when selection acts through differences in the efficiency of resource use, i.e., small β). In the appendix, we consider a situation where all deleterious alleles have the same effect (α, ρ, h) and where epistasis is absent. Neglecting linkage disequilibria among loci, the equilibrium values of population size and of the mean number of mutations per haploid genome can be obtained by solving numerically a system of two equations. As expected, genetic variation in competitivity ($\beta = 1$) has virtually no effect on population size. On the contrary, variation in the efficiency of resource use ($\beta = 0$) has much stronger effects, which are reasonably well captured by our analytical results (fig. A1, available online).

Antagonistic Selection. Ploidally antagonistic selection may maintain polymorphism in the absence of recurrent mutation. From equation (9), we have at equilibrium

$$p_a^{\text{anta}} = \frac{d\,h\,C_d + (1-d)\rho C_h}{(2h-1)d\,C_d}. \quad (12)$$

This equilibrium is biologically relevant (i.e., between 0 and 1) if

$$-\frac{d\,C_d}{(1-d)C_h}\max(1-h,h) < \rho < -\frac{d\,C_d}{(1-d)C_h}\min(1-h,h).$$

It is stable if the allele that is disfavored in diploids (a if $\alpha > 0$, A if $\alpha < 0$) is partially recessive. If allele a is totally recessive (dominant if $\alpha < 0$) and if the population is mainly diploid, the parameter range allowing a stable polymorphism is wide, while it tends to shrink when the proportion of haploids increases (small d), when competition in haploids is stronger ($C_h > C_d$), or when mutations become more additive (h close to 1/2), which corroborates previous results by Ewing (1977) and Immler et al. (2012).

Evolution of the Ploidy Level

Change in Frequency at the Modifier Locus. To leading order in the effects of alleles m and a (δ and α), the change in frequency of the modifier allele m can be decomposed into two terms (for derivation, see appendix):

$$\Delta p_m = \delta h_m^*(\overline{W_d} - \overline{W_h})p_m q_m + [d\,s_d h_a^* + (1-d)(1-r_{ma})s_h]D_{ma}, \quad (13)$$

where $q_m = 1 - p_m$, $h_m^* = h_m(1-p_m) + (1-h_m)p_m$ (simplifying to 1/2 when $h_m = 1/2$), and $h_a^* = h(1-p_a) + (1-h)p_a$ (which is approximately h when p_a is small). The first term of equation (13) is proportional to the difference in mean fitness between haploids and diploids and favors diploidy when $\overline{W_d} > \overline{W_h}$. The second term is proportional to D_{ma} and represents indirect selection on m through its association with allele a; when the direction of selection on allele a is the same in both phases, this term disfavors the modifier allele that tends to be associated with the deleterious allele. Following previous models (e.g., Cailleau et al. 2010), we will denote the first term direct selection (selection to produce more of the phase that has the highest mean fitness), while the second term will be denoted indirect selection (effect of the linkage disequilibrium).

Direct Selection. Direct selection in turn decomposes into two terms: a term of order δ representing differences in mean fitness between haploids and diploids due to different ecological parameters, and a term of order $\delta\alpha p_a q_a$ representing additional differences in mean fitness due to the selection acting at the locus affecting the sensitivity to competition, for example, increased mean fitness of diploids due to the fact that recessive deleterious alleles are partially masked (Otto and Goldstein 1992; Cailleau et al. 2010). This second term is generally negligible relative to the first when only a single

locus segregates for a deleterious allele (as shown by eq. [11], $\alpha p_a q_a$ is of order u at mutation-selection balance); however, we will see that the overall effect of deleterious alleles on mean fitnesses may become more important when they segregate at many loci.

Effect of ecological differences between haploids and diploids. In the absence of selection at the second locus ($\alpha = 0$ or $p_a q_a = 0$), our model corresponds to a simplified version of Hughes and Otto's (1999) model and confirms that ecological differences between haploids and diploids generate direct selection on the modifier locus. From equations (6) and (7), one obtains

$$\overline{W_{\mathrm{d}}} - \overline{W_{\mathrm{h}}} = \frac{C_{\mathrm{d}} C_{\mathrm{h}}}{d C_{\mathrm{d}} + (1-d) C_{\mathrm{h}}} \left(\frac{r_{\mathrm{h}}}{C_{\mathrm{h}}} - \frac{r_{\mathrm{d}}}{C_{\mathrm{d}}} \right) = S_{\mathrm{eco}}, \quad (14)$$

while the change in frequency of the modifier simplifies to $\Delta p_m = \delta h_m S_{\mathrm{eco}} p_m q_m$. The term S_{eco} thus represents the strength of selection on the modifier generated by ecological differences between phases (favoring the phase where fecundity is highest and competition lowest). This term cancels when $r_{\mathrm{d}}/C_{\mathrm{d}} = r_{\mathrm{h}}/C_{\mathrm{h}}$, which occurs when the rate of diploidy d equals

$$d_0 = \frac{K_{\mathrm{d}} - K_{\mathrm{h}} \gamma_{\mathrm{dh}}}{K_{\mathrm{d}}(1-\gamma_{\mathrm{hd}}) + K_{\mathrm{h}}(1-\gamma_{\mathrm{dh}})}. \quad (15)$$

In addition,

$$\left. \frac{d(S_{\mathrm{eco}})}{d(d)} \right|_{d_0} < 0,$$

and d_0 thus represents the evolutionarily stable rate of diploidy (evolutionarily stable strategy [ESS]; Maynard Smith 1982) when it is comprised between 0 and 1. We can note that the ratios $r_{\mathrm{d}}/C_{\mathrm{d}}$ and $r_{\mathrm{h}}/C_{\mathrm{h}}$ are independent of r_{d} and r_h (see eq. [7]); therefore, differences in the baseline fecundity of haploids and diploids do not affect the sign of S_{eco} or the ESS. Furthermore, carrying capacities K_{d} and K_{h} affect only d_0 through the ratio $K_{\mathrm{d}}/K_{\mathrm{h}}$. With weak or no ecological differentiation (γ_{dh} and γ_{hd} close to 1), the population should evolve toward the phase with the highest carrying capacity (i.e., most efficient resource usage). Confirming Hughes and Otto's (1999) results, we predict that intermediate rates of diploidy may be maintained when the ecological niches of haploids and diploids are sufficiently differentiated (see fig. 1*A*). In this situation, Hughes and Otto (1999) showed that the population may consist of a mixture of genotypes coding for different rates of diploidy (as long as this combination of genotypes fully exploits the available resources). Indeed, figure 1*B* shows that when $\gamma_{\mathrm{dh}} = \gamma_{\mathrm{hd}} = 0.5$ and $K_{\mathrm{d}} = K_{\mathrm{h}}$ (so that the predicted ESS rate of diploidy is $d_0 = 0.5$), a high level of polymorphism is maintained at the modifier locus: although the average value of d is 0.5, many alleles coding for values between 0 and 1 coexist in the population. In a stable environment, niche differentiation between haploids and diploids therefore does not lead to the fixation of a strategy corresponding to a biphasic life cycle (intermediate value of d). However, results from bet-hedging theory (e.g., Philippi and Seger 1989) suggest that genotypes coding for intermediate values of d may be favored when the relative abundance of resources used by haploids and diploids fluctuate over time (indeed, producing both haploid and diploid offspring may be seen as a bet-hedging strategy). As shown in figure 1*C* and 1*D*, this prediction is confirmed by introducing fluctuations in the relative sizes of the haploid and diploid niches ($K_{\mathrm{d}}/K_{\mathrm{h}}$) over time: as the amplitude of these fluctuations increases, the distribution of values of d in the population narrows around $d = 0.5$ (decreasing the period of oscillations also reduces the variance in d; simulations performed with $\tau \in \{1, 2, 4, 8, 16, 64\}$ and $\lambda = 1/15$, not shown).

Effect of genetic variation in the sensitivity to competition. Genetic variation for fitness (e.g., due to recurrent deleterious mutations) may not have the same quantitative effect on the mean fitness of haploid and diploid individuals, generating direct selection on ploidy even in the absence of ecological difference between haploids and diploids (Otto and Goldstein 1992). This introduces a new component to the direct selection term, which becomes $dh_m^*(S_{\mathrm{eco}} + S_{a,\mathrm{direct}}) p_m q_m$, where $S_{a,\mathrm{direct}}$ is the effect of allele a on $\overline{W_{\mathrm{d}}} - \overline{W_{\mathrm{h}}}$. At mutation-selection balance, $S_{a,\mathrm{direct}}$ is given by

$$S_{a,\mathrm{direct}} \approx \frac{\alpha(\rho - 2h) C_{\mathrm{d}} C_{\mathrm{h}}}{d C_{\mathrm{d}} + (1-d) C_{\mathrm{h}}} \times \left\{ (1-\beta) + \beta \frac{d^2 \gamma_{\mathrm{hd}} + (1-d)^2 \gamma_{\mathrm{dh}} + 2d(1-d)\gamma_{\mathrm{dh}}\gamma_{\mathrm{hd}}}{(1 - d + d\gamma_{\mathrm{hd}})[d + (1-d)\gamma_{\mathrm{dh}}]} \right\} \hat{N}_0 \hat{p}_a^{\mathrm{del}} \quad (16)$$

(see appendix).

Interestingly, the sign of $S_{a,\mathrm{direct}}$ is entirely determined by the sign of $\alpha\ (\rho - 2h)$ (because $d C_{\mathrm{d}} + (1-d) C_{\mathrm{h}}$ and the term within curly brackets are always positive), so that ecological parameters (γ_i, β, r_i, and K_i) modulate only the strength of $S_{a,\mathrm{direct}}$. Equation (16) bears some similarity with previous results on the benefits of diploidy stemming from masking of deleterious alleles (Otto and Goldstein 1992); in particular, when $\rho = 1$ (so that allele a has the same effect on the sensitivity to competition of haploids and homozygous diploids), diploidy is favored when the deleterious allele is partially recessive ($h < 1/2$). However, in Otto and Goldstein's model, the sign of direct selection is deter-

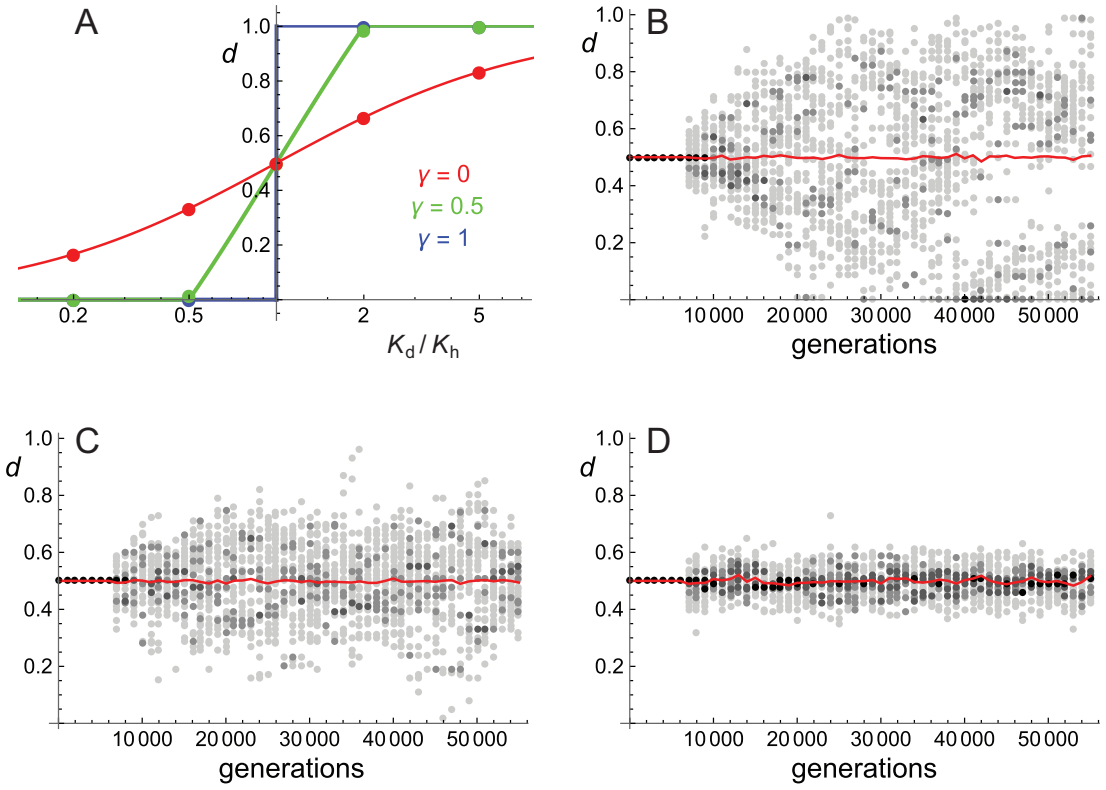

Figure 1: *A*, Evolutionarily stable strategy rate of diploidy (d_{eq}) as a function of the ratio of carrying capacities K_d/K_h and for different levels of differentiation between phases ($\gamma_{dh} = \gamma_{hd} = \gamma$). Lines, analytical prediction in the absence of deleterious mutations (eq. [15]). Circles, multilocus simulations in the absence of deleterious mutations. Blue, $\gamma = 1$; green, $\gamma = 0.5$; red, $\gamma = 0$. *B–D*, Distribution of alleles present at the modifier locus throughout the simulation, in the absence of deleterious mutation, with $K_d = K_h = 15{,}000$, $r_d = r_h = 1.8$, and $\gamma_{hd} = \gamma_{dh} = 0.5$. Shading corresponds to the frequencies f of different ranges of values of d in the population, from light gray to black: $f \in [1\%, 5\%]$, $f \in [5\%, 10\%]$, $f \in [10\%, 20\%]$, and $f \geq 20\%$. *B*, When the environment stays constant, a large range of values of d is maintained in the population. *C*, Fluctuations of the relative sizes of the haploid and the diploid niches of low amplitude ($\lambda = 1/15$, $\tau = 4$) narrow this range. *D*, Increasing the amplitude of fluctuations ($\lambda = 1/2$, $\tau = 4$) further narrows the range around the ESS value of d. In all cases, the average value of d in the population stays near 0.5 (red lines).

mined by the fitness effect of the deleterious allele in haploids and heterozygous diploids: $s_h - 2hs_d$, which should thus depend on ecological parameters (given that $s_h = \rho\alpha NC_h$ and $s_d = \alpha NC_d$). By contrast, in our model, a deleterious allele may generate direct selection for a given ploidy phase even if it has a stronger fitness effect in that phase (provided that $s_h - 2hs_d$ and $\rho - 2h$ have opposite signs). This seemingly surprising result comes from the fact that in our model of density-dependent selection, there is not a simple correspondence between the fitness effect of adding a mutation in a given haploid or diploid (s_h or s_d, to the first order in α) and the mean fitness of haploids and diploids at equilibrium.

When polymorphism at the selected locus is maintained by ploidally antagonistic selection, terms in p_a^2 cannot be neglected, and the term $\rho - 2h$ in equation (16) is replaced by $\rho - 2h_a^* = \rho - h(1 - p_a) - (1 - h)p_a$, which is always negative (recall that $\rho < 0$ under antagonistic selection). Therefore, $S_{a,\text{direct}}$ has the sign of $-\alpha$: direct selection favors the phase in which ploidally antagonistic selection increases mean fitness (haploidy if $\alpha > 0$ and diploidy if $\alpha < 0$).

Finally, as expected, $S_{a,\text{direct}}$ vanishes when the ecological niches of haploids and diploids are fully disjoint ($\gamma_{dh} = \gamma_{hd} = 0$) and selection is soft (selection through differences in competitivity: $\beta = 1$; see eq. [16]). When deleterious mutations impact resource use ($\beta = 0$), however, $S_{a,\text{direct}}$ does not depend on the degree of overlap between the haploid and diploid niches: mutations generate a direct selection component favoring the phase whose carrying capacity is least impacted by the presence of deleterious mutations. Between these two extreme situations, equation (16) indicates

that direct selection on the modifier decreases linearly with β (as long as $\gamma_{hd}\gamma_{dh} \leq 1$).

Indirect Selection. Finally, genetic variation for fitness generates indirect selection (through the linkage disequilibrium between the modifier and the selected locus) favoring the ploidy phase in which selection is more efficient. From equation (13), the change in frequency of the modifier is affected by linkage disequilibrium (D_{ma}) through the term

$$S_{a,\text{indirect}} = [d s_d h_a^* + (1-d)(1-r_{ma})s_h]D_{ma}. \quad (17)$$

At QLE, and to the first order in α and δ, D_{ma} is given by

$$D_{ma} \approx \frac{\delta h_m^*}{r_{ma}} \left(\frac{s_h}{\overline{W_h}} - \frac{h_a^* s_d}{\overline{W_d}} \right) \overline{W_h}\,\overline{W_d} p q_{ma} \quad (18)$$

(for derivation, see appendix), where $pq_{ma} = p_m q_m p_a q_a$, r_{ma} is the recombination rate, and $\overline{W_h}$ and $\overline{W_d}$ are given by equation (6) (replacing N by $\hat{N}_0$). When allele a is deleterious in both phases, it thus tends to be associated with the modifier allele increasing the phase in which selection is less efficient (diploidy if $hs_d/\overline{W_d} < s_h/\overline{W_h}$, haploidy otherwise). When the rate of diploidy d is such that the direct selection term cancels (i.e., when $\overline{W_d} = \overline{W_h}$), D_{ma} has the sign of $\delta(s_h - hs_d)$. Importantly, and in contrast to the direct selection term $S_{a,\text{direct}}$ discussed in "Direct Selection," the sign of indirect selection depends on ecological parameters r_i, γ_i, and K_i (through s_d and s_h): it is the effect of the deleterious allele on fitness (and not on sensitivity to competition) that matters.

Under ploidally antagonistic selection ($\alpha < 0$), inserting the expression of the equilibrium frequency of a given by equation (12) into equation (18) yields

$$D_{ma} \approx \frac{\delta h_m^*}{d\, r_{ma}} s_h p q_{ma}, \quad (19)$$

indicating that the modifier allele increasing a given ploidy level tends to be associated with the allele that is favored in this ploidy level; that is, $D_{ma} > 0$ if m increases diploidy and a is advantageous in diploids ($s_h < 0$, $s_d > 0$) or if m increases haploidy and a is advantageous in haploids ($s_h > 0$, $s_d < 0$). From equation (17), the indirect selection term is then

$$S_{a,\text{indirect}} \approx (1-d) r_{ma} s_h D_{ma}, \quad (20)$$

which has the sign of d_m: therefore, indirect selection generated by ploidally antagonistic polymorphisms always favor modifier alleles that increase diploidy. This result can be understood as follows: during selection, each modifier allele benefits from the increase in frequency of the selected allele with which it is associated (within the ploidy phase where this allele is favored); however, this hitchhiking effect is weaker for the allele that increases haploidy, because recombination occurs before haploid selection (reducing the linkage disequilibrium before selection).

As can be seen from the equations above, indirect selection is not affected by the parameter β (indicating the degree of softness of selection). Under complete differentiation of the haploid and diploid niches ($\gamma_{dh} = \gamma_{hd} = 0$) and soft selection ($\beta = 1$), the selected locus therefore affects the evolution of the modifier through indirect selection only, since the direct selection term $S_{a,\text{direct}}$ vanishes in this situation (eq. [16]). In other situations, the effect of indirect selection (of order $\delta\alpha^2 p_a q_a$) is expected to be weak relative to direct selection ($S_{a,\text{direct}}$, of order $\delta\alpha p_a q_a$). However, stronger effects of indirect selection may arise at mutation-selection balance either when linkage is tight (as shown previously in Otto and Goldstein 1992) or when ecological parameters of haploids and diploids (in particular, their baseline fecundities r_h and r_d) are very different, so that density-dependent selection is much stronger in one phase than in the other (since the strength of indirect selection is proportional to $\rho/C_d - h/C_h$).

Extrapolation to Many Loci at Mutation-Selection Balance. As discussed previously, the effect of a single deleterious allele should generally be negligible compared with direct selection stemming from ecological differences between haploids and diploids (unless these ecological differences are very slight). However, the overall effect of deleterious alleles occurring at many loci could be more important. To explore that, we extrapolated results from our two-locus model to an arbitrary number of selected loci, assuming a deleterious mutation rate U per haploid genome per generation. Neglecting linkage disequilibria between loci at which deleterious alleles are segregating, the overall effect of all selected loci on the change in frequency of the modifier can be approximated by the sum of the individual effects of these loci (provided by the two-locus analysis above). When the average number of deleterious alleles per genome (n) is large (so that $s\,n$ may be of order 1), more accurate results can be obtained numerically as explained in the appendix.

Figure 2 compares the ESS rate of diploidy predicted by our analytical model with multilocus simulation results (with deleterious alleles occurring at 5,000 loci and a map length of 10 Morgans). Note that analytical extrapolations assume unlinked loci rather than a continuous genetic map, since our QLE results (eq. [18]) diverge for tightly linked loci. However, simulations indicate that linkage does not have much effect on the ESS rate of diploidy for the parameter values used in the figures (simulating free recombination yields very similar results; not shown).

In the absence of genetic variation for sensitivity to competition ($U = 0$), the evolution of ploidy is only driven by ecological differences between haploids and diploids (S_{eco}). Because $K_d = K_h$ while $\gamma_{dh} = \gamma_{hd} = \gamma$ for all graphs of

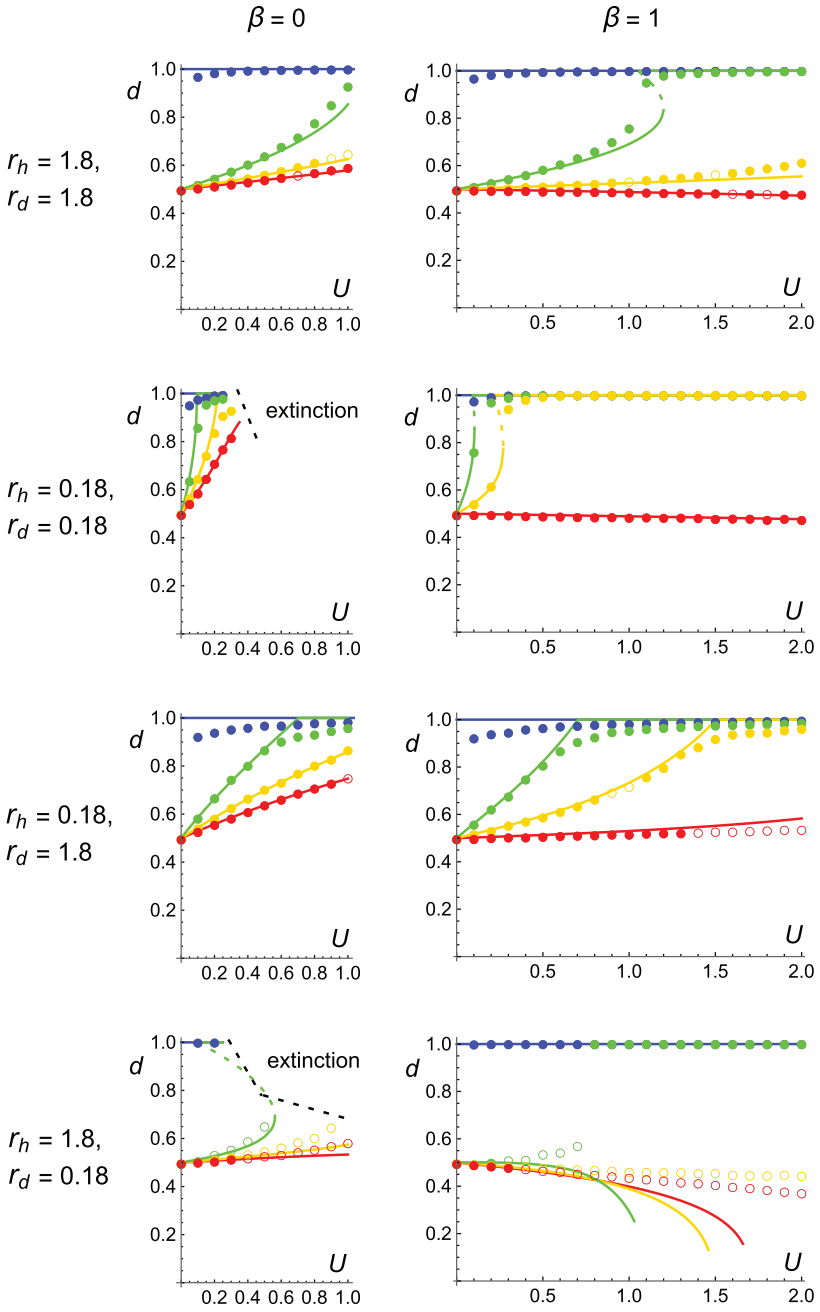

Figure 2: Evolutionarily stable strategy ploidy level as a function of the mutation rate U when mutations affect the efficiency of resource use ($\beta = 0$) or the competitivity ($\beta = 1$). Lines, analytical results from equations (13), (A5), (A6), and (A7) (eqq. [A5], [A6], and [A7] in the appendix, available online); dashed lines correspond to unstable equilibria predicted by the model, while solid lines correspond to stable equilibria. Circles, multilocus simulations with 5,000 loci under selection and a map length of 10 Morgans. Simulations are run for 20,000 generations, with a mutation rate of 10^{-2} at the modifier locus. Open circles correspond to cases where evolutionary branching occurred during the simulation. Blue, $\gamma = 1$; green, $\gamma = 0.5$; yellow, $\gamma = 0.2$; red, $\gamma = 0$. Other parameters values: $K_d = K_h = 15{,}000$, $\rho = 1$, $h = 0.3$, $\alpha = 0.05$. Error bars (computed using the spAMM1.2.4 R package [Rousset and Ferdy 2014] to deal with autocorrelated data) were smaller than the size of symbols in most cases and are thus not shown.

figure 2, the evolutionarily stable rate of diploidy is $d = 0.5$ if $\gamma < 1$, while selection on ploidy vanishes when $\gamma = 1$. Indeed, we can see on figure 2 that d always converges to 0.5 as U tends to 0, as long as $\gamma < 1$.

When $U > 0$, we have seen that deleterious alleles generate direct selection for diploidy (term $S_{a,\text{direct}}$) whenever $\rho > 2h$ (and selection for haploidy otherwise), which vanishes when selection is soft ($\beta = 1$) and haploids and diploids occupy totally separated niches ($\gamma = 0$). In figure 2, deleterious alleles have the same effect on sensitivity to competition in haploids and homozygous diploids ($\rho = 1$) and are partially recessive ($h = 0.3$), generating direct selection for diploidy. Increasing the degree of niche overlap (γ) or the mutation rate U generally increases the relative effect of deleterious alleles ($S_{a,\text{direct}}$) over ecological differentiation (S_{eco}) and displaces the ESS rate of diploidy d from 0.5 toward 1. Interestingly, lower baseline fecundities (r_d, $r_h = 0.18$) also increase the relative effect of $S_{a,\text{direct}}$, leading to higher rates of diploidy. Indeed, decreasing baseline fecundities decreases the strength of the ecological component of selection S_{eco} by reducing competition between individuals (see eq. [14]), while the direct selection advantage generated by deleterious alleles ($S_{a,\text{direct}}$) is less affected by r_d and r_h because of the fact that deleterious alleles are more frequent when fecundities are lower (eq. [11]).

Finally, indirect selection generated by deleterious mutations ($S_{a,\text{indirect}}$) is generally negligible but causes slight displacements from $d = 0.5$ when $S_{a,\text{direct}}$ vanishes; that is, selection is soft ($\beta = 1$), and haploids and diploids occupy totally separated niches ($\gamma = 0$; red lines in fig. 2). In this case, deleterious alleles tend to favor haploidy, unless selection is stronger in diploids ($r_d \gg r_h$; see eqq. [7], [10]). Although our analytical model predicts stronger deviations toward haploidy when $r_d \ll r_h$ and U is large, d remains close to 0.5 in the simulations (this discrepancy may be due to the fact that our analytical model neglects higher-order associations between selected loci).

Evolutionary Branching. As we have seen before, high polymorphism may be maintained at equilibrium at the modifier locus under ecological selection alone ($U = 0$; fig. 1*B*). A different pattern can be observed in the presence of deleterious alleles ($U > 0$), however. An example is shown in figure 3*A*: in a stable environment (K_h, K_d constant), evolutionary branching occurs once the population has reached the ESS rate of diploidy d (here close to 0.5), and a proportion d of the modifier alleles evolves toward 1 (diploidy) while a proportion $1 - d$ evolves toward 0 (haploidy); note that the average proportion of diploids in the population remains unchanged. As illustrated in figure 4, this evolutionary branching occurs more easily when selection against deleterious alleles is stronger in one phase than in the other ($s_h - hs_d$ large in absolute value; upper left/lower right in fig. 4*A*). This effect may be understood as follows. A modifier allele coding for an intermediate value of d undergoes selection alternatively in the diploid and the haploid phases and therefore experiences periods of higher fitness (in the phase where mutations have a weaker effect) followed by periods of lower fitness (in the phase where mutations have a stronger effect). These temporal fluctuations reduce geometric mean fitness and favor modifiers that tend to stay in the same phase (so that selection is always weak or always strong). However, part of the benefit of this specialization is lost under random mating (since heterozygous genotypes at the modifier locus are produced each generation), suggesting that assortative mating should be favored under this sce-

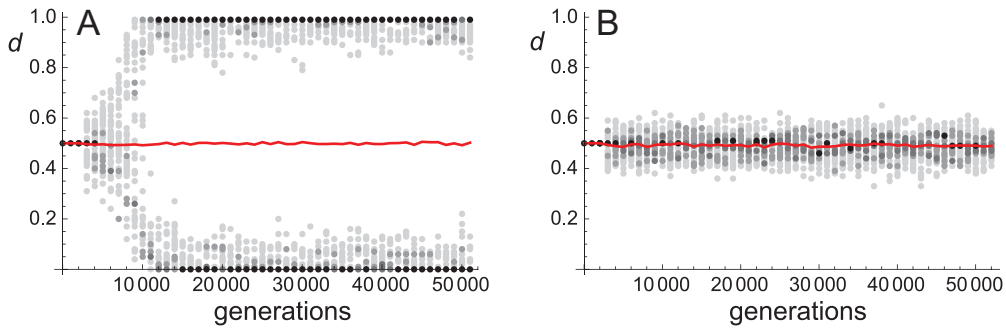

Figure 3: Distribution of effects of alleles at the modifier locus through time (50,000 generations) when deleterious mutations occur at rate $U = 1$ in a stable (*A*) or fluctuating (*B*) environment. In this example, deleterious mutations affect competitivity ($\beta = 1$), and there is limited niche overlap ($\gamma_{dh} = \gamma_{hd} = 0.05$). In both cases, the average rate of diploidy at equilibrium is close to 0.5 (red lines). *A*, When the environment stays constant ($\lambda = 0$), evolutionary branching occurs, and at equilibrium, only alleles coding for values of d close to either 0 or 1 are maintained in the population. *B*, Fluctuations of the relative sizes of the haploid and diploid niches ($\lambda = 1/2$, $\tau = 4$) prevent this branching by favoring genotypes coding for intermediate rates of diploidy. Parameters values: $U = 1$, $K_d = K_h = 15{,}000$, $r_d = 1.8$, $r_h = 1.8$, $\rho = 1$, $h = 0.3$, $\alpha = 0.05$, mutation rate at the modifier locus: 10^{-2}.

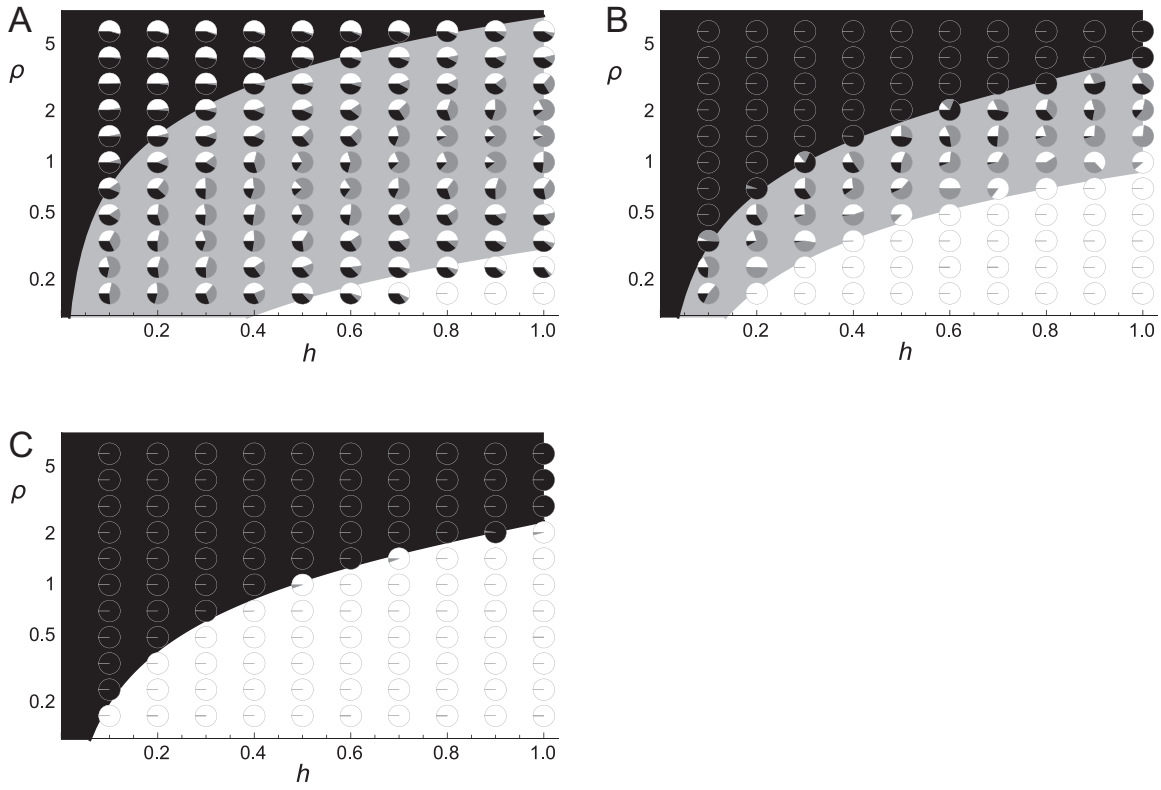

Figure 4: Evolution of the ploidy level when mutations have differential effects between phases and with different degrees of niche overlap under soft selection ($\beta = 1$). *A*, $\gamma_{dh} = \gamma_{hd} = 0.05$. *B*, $\gamma_{dh} = \gamma_{hd} = 0.5$. *C*, $\gamma_{dh} = \gamma_{hd} = 1$. On the *X*-axes, dominance (h) of deleterious mutations varies between 0 (mutations are fully recessive) and 1 (mutations are fully dominant). On *Y*-axes, the relative effect of mutations in haploids compared with homozygote diploids (ρ) varies between 1/6 and 6 (shown on a log scale). Background color corresponds to analytical predictions: evolution toward diploidy (black), haploidy (white), or maintenance of an intermediate rate of diploidy (gray). Pie charts represent the distribution of modifier alleles in simulated populations after 20,000 generations of evolution, with a mutation rate of 10^{-3} at the modifier locus (black, $d < 0.1$; gray, $d \in [0.1, 0.9]$; white, $d > 0.9$). Other parameters values: $U = 1$, $K_d = K_h = 15{,}000$, $r_d = r_h = 1.8$, $\alpha = 0.05$.

nario. For the parameter values used in figure 4, branching was not observed when the strength of selection against deleterious alleles was similar in haploids and diploids; however, in several cases, we ran simulations over larger numbers of generations and observed branching. Similarly, increasing the mutation rate at the modifier locus (so that evolution occurs more rapidly) greatly increases the number of cases where branching occurs, and this suggests that it may eventually occur in most cases where $U > 0$ and where an intermediate value of d is predicted at the ESS. This is confirmed in figure 2 using a mutation rate of 0.01 at the modifier locus: evolutionary branching occurs in many cases when the average value of d in the population at equilibrium is intermediate, in particular when $s_h - hs_d$ is higher in absolute value ($r_d = 1.8$, $r_h = 0.18$); it is possible that branching would occur in more cases if the simulations could run over larger numbers of generations. Finally, as we have seen in the absence of deleterious mutation (fig. 1), temporal fluctuations in the relative sizes of the haploid and diploid niches (K_h, K_d) may inhibit this branching phenomenon (as illustrated in fig. 3*B*) by favoring intermediate values of d (which may be seen as a bet-hedging strategy).

Discussion

Different types of selective forces may affect the evolution of the relative degree of development of the haploid and diploid phases of the eukaryotic sexual life cycle. In the presence of genetic variation in fitness, changing the ploidy level of individuals may affect their mean fitness as a result of dominance interactions among selected alleles (in particular, diploidy benefits from an immediate advantage in the presence of partially recessive deleterious alleles). However, ploidy also affects the genetic variance among individuals

and thus the efficiency of natural selection: in general, selection is expected to be more efficient among haploids, in which all mutations are expressed. As shown by Otto and Goldstein (1992) and Otto and Marks (1996), either haploid or diploid life cycles are expected to be favored under the combined action of these two forces, depending in particular on the importance of recombination within genomes and on the mating system. However, an important assumption of these genetic models is that haploids and diploids have the same ecological properties, being fully equivalent in terms of survival, fecundity, and competition exerted on other individuals in the absence of genetic variation in fitness. As shown by Hughes and Otto (1999), relaxing this hypothesis introduces direct selection on life cycle variants and can explain the evolutionary stability of biphasic life cycles when haploids and diploids have different ecological niches: indeed, under density-dependent competition, an individual entering a more vacant niche has greater chances to survive and contribute to the next generation. Furthermore, ecological niche differences between haploids and diploids may affect relative fitness effect of deleterious mutations in the two phases, in turn affecting the predictions from the genetic models cited above (in which mutations are assumed to have the same fitness effect in haploids and homozygote diploids).

In this article, we used a simple demographic model to explore the interaction between the effects of ecological differentiation and of genetic variation in fitness on the evolution of ploidy cycles. We focused on a scenario where selection is density dependent, with deleterious alleles reducing the success of individuals when competing with conspecifics. This particular form of selection is certainly restricted to a subset of all possible deleterious alleles, with the effect of most mutations causing developmental abnormality or lethality being probably little affected by population density. Nevertheless, it appears likely that the deleterious effect of a substantial proportion of mutations may increase with the strength of intraspecific competition: in particular, several experimental studies (e.g., table 2 in Agrawal and Whitlock 2010) reported stronger average effects of deleterious alleles at higher density, while other studies measured stronger inbreeding depression at higher densities, which may possibly be due to stronger effects of deleterious alleles (e.g., Cheptou et al. 2000; Meagher et al. 2000). Exploring the effect of density-dependent selection appears particularly interesting in the context of models combining ecological and genetical effects, since the interplay between these different components is not intuitively obvious.

In our model, the evolution of a modifier gene affecting the ploidy of individuals is controlled by three different effects: (1) differences in mean fitness between ploidy levels generated by intrinsic ecological differences between haploid and diploid individuals (independently of their genotype); (2) additional differences in mean fitness caused by deleterious alleles, which may have different effects on the sensitivity to competition in haploids and homozygous diploids; and (3) differences among phases in the efficiency of selection against these deleterious alleles. As we have seen, the results show that these effects scale differently with the strength of competition within and between ploidy phases. In particular, at demographic equilibrium, the strength of ecological selection depends on differences in the carrying capacities of haploids and diploids (reflecting differences in the availability of resources and/or in the efficiency with which individuals use these resources), on the degree of overlap between their ecological niches, and on the average baseline fecundities of individuals (r_d and r_h, controlling the intensity of competition among offspring). By contrast, changing baseline fecundities has little effect on differences in mean fitness caused by deleterious alleles. As a consequence, the relative importance of effects 1 and 2 depends on these baseline fecundities, with effect 1 being relatively stronger when r_d and r_h are higher and relatively weaker when r_d and r_h are lower. Therefore, we predict that biphasic life cycles should be more easily maintained (by ecological differentiation among phases) when the intensity of competition among offspring is strong, while deleterious alleles affecting the sensitivity to competition of individuals may have stronger destabilizing effects (favoring either haploid or diploid life cycles) when lower fecundities reduce the intensity of competition (fig. 2). It would thus be interesting to test whether biphasic (haploid-diploid) life cycles tend to be associated with higher intensities of competition among offspring in phylogenetic groups where different types of life cycles coexist (such as algae).

Another result is that ecological differences between haploids and diploids may affect the relative importance of the two components of selection on ploidy generated by deleterious alleles (effects 2 and 3). In our model, the effect of deleterious alleles on the sensitivity of individuals to competition determines whether they generate a direct selection advantage for haploidy or diploidy (effect 2). By contrast, whether indirect selection (effect 3) favors haploidy or diploidy depends on the effect of deleterious alleles on the overall fitness of individuals (number of offspring produced), which under density-dependent selection is affected by the ecological parameters of haploids and diploids. As a consequence, a deleterious allele having the same effect on the competitivity of haploids and homozygous diploids ($\rho = 1$) would favor diploidy through a masking advantage if it is partially recessive ($h < 1/2$) and could also favor diploidy through a purging advantage if competition is stronger among diploids, leading to stronger density-dependent selection ($r_d h > r_h$). More generally, these results outline the fact that selection may not necessarily be more efficient among haploids when mutations have the same phenotypic effects in haploids and homozygous diploids, provided that

selection is affected by density and that haploids and diploids experience different degrees of density-dependent competition. This may explain the observation that purifying selection does not seem to be more efficient in haploid-expressed genes than in diploid-expressed genes of *Arabidopsis thaliana* and of the moss *Funaria hygrometrica* (Szoevenyi et al. 2013), and it outlines the need for more systematic measures of fitness effects of deleterious mutations at different population densities.

We also explored the case of antagonistic selection, where different alleles at the same locus are favored in haploids and in diploids. As with deleterious alleles, the direct competition term favors the phase with the highest mean fitness—a result also described in a recent article by Immler and Otto (2014). In addition, we found that independently of the ecological parameters, indirect selection always favors diploidy when the locus under antagonistic selection is at its polymorphic equilibrium, because of the fact that a modifier allele increasing diploidy tends to benefit more from hitchhiking than an allele increasing haploidy, whose association with the haploid-beneficial allele is broken by recombination before selection. In the case of heteromorphic species where haploid and diploid individuals occupy different niches, many loci may possibly be under ploidally antagonistic selection (because of morphological and physiological differences between ploidy phases), and this should thus tend to favor modifiers increasing the relative importance of the diploid phase.

Our simulations showed that in the absence of deleterious mutations, ecological differentiation between phases may lead to the stable coexistence of a high diversity of life cycles (from fully haploid to fully diploid cycles), in agreement with predictions by Hughes and Otto (1999). However, when deleterious alleles are introduced and when the ecological component of selection is sufficiently strong to favor biphasic life cycles (despite the destabilizing effect of deleterious alleles), eventually the population evolves to a state where alleles coding for fully haploid and fully diploid life cycles stably coexist. In this case, we expect that assortative mating should be favored, ultimately leading to the coexistence of a haploid and a diploid species. Therefore, ecological differentiation between phases does not seem sufficient to explain the stable maintenance of truly biphasic life cycles (involving an obligatory alternation between haploid and diploid individuals), unless additional factors favor such an alternation, for example, constraints due to the different biology of spores and gametes (e.g., Stebbins and Hill 1980; Bell 1997) or temporal variability of the environment, as explored here. Finally, we assumed a preexisting ecological differentiation between haploid and diploid individuals. It may be of interest to extend our model in order to explore the ecological and genetic factors that may drive this differentiation.

Acknowledgments

We thank A. Agrawal, Y. Michalakis, S. Otto, M. Valero, and two anonymous reviewers for helpful discussions and comments. M.R. and D.R. were funded by the Agence Nationale de la Recherche (project BiCycle ANR10-BLAN-1727).

Literature Cited

Agrawal, A. F. 2010. Ecological determinants of mutation load and inbreeding depression in subdivided populations. American Naturalist 176:111–122.

Agrawal, A. F., and M. C. Whitlock. 2010. Environmental duress and epistasis: how does stress affect the strength of selection on new mutations? Trends in Ecology and Evolution 25:450–458.

Bell, G. 1997. The evolution of the life cycle of brown seaweeds. Biological Journal of the Linnean Society 60:21–38.

Bolton, J. J., and M. A. P. Joska. 1993. Population studies on a South African carrageenophyte: *Iridaea capensis* (*Gigartinaceae*, *Rhodophyta*). Hydrobiologia 261:191–195.

Cailleau, A., P. O. Cheptou, and T. Lenormand. 2010. Ploidy and the evolution of endosperm of flowering plants. Genetics 184:439–453.

Cheptou, P. O., E. Imbert, J. Lepart, and J. Escarre. 2000. Effects of competition on lifetime estimates of inbreeding depression in the outcrossing plant *Crepis sancta* (Asteraceae). Journal of Evolutionary Biology 13:522–531.

Christiansen, F. B., and V. Loeschcke. 1980. Evolution and intraspecific exploitative competition. 1. One-locus theory for small additive gene effects. Theoretical Population Biology 18:297–313.

Coelho, S. M., A. F. Peters, B. Charrier, D. Roze, C. Destombe, M. Valero, and J. M. Cock. 2007. Complex life cycles of multicellular eukaryotes: new approaches based on the use of model organisms. Gene 406:152–170.

Couceiro, L., M. Le Gac, H. Hunsperger, S. Mauger, C. Destombe, M. Cock, S. Ahmed, et al. 2015. Evolution and maintenance of haploid-diploid life cycles in natural populations: the case of the marine brown alga *Ectocarpus*. Evolution 69:1808–1822.

Destombe, C., J. Godin, M. Nocher, S. Richerd, and M. Valero. 1993. Differences in response between haploid and diploid isomorphic phases of *Gracilaria verrucosa* (*Rhodophyta*, *Gigartinales*) exposed to artificial environmental conditions. Hydrobiologia 261:131–137.

Dyck, L. J., and R. E. De Wreede. 2006. Seasonal and spatial patterns of population density in the marine macroalga *Mazzaella splendens* (*Gigartinales*, *Rhodophyta*). Phycological Research 54:21–31.

Ewing, E. P. 1977. Selection at haploid and diploid phases: cyclical variation. Genetics 87:195–207.

Gerstein, A. C. 2013. Mutational effects depend on ploidy level: all else is not equal. Biology Letters 9:20120614.

Gerstein, A. C., L. A. Cleathero, M. A. Mandegar, and S. P. Otto. 2010. Haploids adapt faster than diploids across a range of environments. Journal of Evolutionary Biology 24:531–540.

Halligan, D. L., and P. D. Keightley. 2009. Spontaneous mutation accumulation studies in evolutionary genetics. Annual Review of Ecology, Evolution, and Systematics 40:151–172.

Hughes, J. S., and S. P. Otto. 1999. Ecology and the evolution of biphasic life cycles. American Naturalist 154:306–320.

Immler, S., G. Arnqvist, and S. P. Otto. 2012. Ploidally antagonistic selection maintains stable genetic polymorphism. Evolution 66:55–65.

Immler, S., and S. P. Otto. 2014. Driven apart: the evolution of ploidy differences between the sexes under antagonistic selection. American Naturalist 183:96–107.

Joseph, S. B., and M. Kirkpatrick. 2004. Haploid selection in animals. Trends in Ecology and Evolution 19:592–597.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

Mable, B. K., and S. P. Otto. 1998. The evolution of life cycles with haploid and diploid phases. Bioessays 20:453–462.

Manna, F., G. Martin, and T. Lenormand. 2011. Fitness landscape: an alternative theory for the dominance of mutations. Genetics 189: 923–937.

Maynard Smith, J. 1982. The century since Darwin. Nature 296:599–601.

Meagher, S., D. J. Penn, and W. K. Potts. 2000. Male-male competition magnifies inbreeding depression in wild house mice. Proceedings of the National Academy of Sciences of the USA 97:3324–3329.

Michod, R. E., and M. F. Wojciechowski. 1994. DNA repair and the evolution of transformation. 4. DNA damage increases transformation. Journal of Evolutionary Biology 7:147–175.

Nuismer, S. L., and S. P. Otto. 2004. Host-parasite interactions and the evolution of ploidy. Proceedings of the National Academy of Sciences of the USA 101:11036–11039.

Orr, H. A., and S. P. Otto. 1994. Does diploidy Increase the rate of adaptation? Genetics 136:1475–1480.

Otaiza, R. D., S. R. Abades, and A. J. Brante. 2001. Seasonal changes in abundance and shifts in dominance of life history stages of the carrageenophyte *Sarcothalia crispata* (*Rhodophyta*, *Gigartinales*) in south central Chile. Journal of Applied Phycology 13:161–171.

Otto, S. P., and A. C. Gerstein. 2008. The evolution of haploidy and diploidy. Current Biology 18:R1121–R1124.

Otto, S. P., and D. B. Goldstein. 1992. Recombination and the evolution of diploidy. Genetics 131:745–751.

Otto, S. P., and J. C. Marks. 1996. Mating systems and the evolutionary transition between haploidy and diploidy. Biological Journal of the Linnean Society 57:197–218.

Paalme, T., J. Kotta, and P. Kersen. 2013. Does the growth rate of drifting *Furcellaria lumbricalis* and *Coccotylus truncatus* depend on their proportion and density? Proceedings of the Estonian Academy of Sciences 62:141–147.

Perrot, V., S. Richerd, and M. Valero. 1991. Transition from haploidy to diploidy. Nature 351:315–317.

Philippi, T., and J. Seger. 1989. Hedging ones evolutionary bets, revisited. Trends in Ecology and Evolution 4:41–44.

Reed, D. C. 1990. An experimental evaluation of density dependence in a subtidal algal population. Ecology 71:2286–2296.

Rescan, M., T. Lenormand, and D. Roze. 2015. Data from: Interactions between genetic and ecological effects on the evolution of life cycles. American Naturalist, Dryad Digital Repository, http://dx.doi.org/10.5061/dryad.40qp5.

Rokitta, S. D., L. J. de Nooijer, S. Trimborn, C. de Vargas, B. Rost, and U. John. 2011. Transcriptome analyses reveal differential gene expression patterns between the life-cycle stages of *Emiliania huxleyi* (Haptophyta) and reflect specialization to different ecological niches. Journal of Phycology 47:829–838.

Rousset, F., and J. B. Ferdy. 2014. Testing environmental and genetic effects in the presence of spatial autocorrelation. Ecography 37: 781–790.

Santos, R., and P. Duarte. 1996. Fecundity, spore recruitment and size in *Gelidium sesquipedale* (Gelidiales, Rhodophyta). Hydrobiologia 327:223–228.

Stebbins, G. L., and G. J. C. Hill. 1980. Did multicellular plants invade the land? American Naturalist 115:342–353.

Szafraniec, K., D. M. Wloch, P. Sliwa, R. H. Borts, and R. Korona. 2003. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. Genetical Research 82:19–31.

Szoevenyi, P., M. Ricca, Z. Hock, J. A. Shaw, K. K. Shimizu, and A. Wagner. 2013. Selection is no more efficient in haploid than in diploid life stages of an angiosperm and a moss. Molecular Biology and Evolution 30:1929–1939.

Thompson, D. A., M. M. Desai, and A. W. Murray. 2006. Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. Current Biology 16:1581–1590.

Thornber, C. S. 2006. Functional properties of the isomorphic biphasic algal life cycle. Integrative and Comparative Biology 46: 605–614.

Valero, M., S. Richerd, V. Perrot, and C. Destombe. 1992. Evolution of alternation of haploid and diploid phases in life cycles. Trends in Ecology and Evolution 7:25–29.

Von Dassow, P., H. Ogata, I. Probert, P. Wincker, C. Da Silva, S. Audic, J. M. Claverie, et al. 2009. Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. Genome Biology 10:R114.

Wallace, B. 1975. Hard and soft selection revisited. Evolution 29: 465–473.

Zeyl, C., T. Vanderford, and M. Carter. 2003. An evolutionary advantage of haploidy in large yeast populations. Science 299:555–558.

Associate Editor: Jeff Leips
Editor: Susan Kalisz

Chapter 3

Evolution of post zygotic reproductive isolation between small populations of *Saccharomyces cerevisiae* at mutation selection drift equilibrium

Marie Rescan[1,2], Denis Roze[1,2] and Jean-Nicolas Jasmin[3]

[1] CNRS, Unité Mixte Internationale 3614, Evolutionary Biology and Ecology of Algae, Roscoff, France

[2] Sorbonne Universités, Université Pierre et Marie Curie, University of Paris 6, Roscoff, France

[3] INRA, Unité Mixte de Recherche 1083 Sciences Pour l'Oenologie, Montpellier, France

Manuscript in preparation

Contents

1 Introduction

1.1 Post zygotic reproductive isolation

Understanding the mechanisms and dynamics of reproductive isolation between incipient species remains a key issue in evolutionary biology. Post-zygotic reproductive isolation is generally considered as a by-product of evolutionary processes occurring in different subpopulations. The simplest scenario corresponds to genetic divergence between two fully isolated populations (allopatry), although an important number of models showed how reproductive isolation can emerge in the presence of gene flow. Without gene flow, each population follows an independent evolutionary trajectory: mutations appear and may reach fixation depending on their fitness effect in the genetic background of the population in which they occurred. During a secondary contact between two such populations, hybrids that combine derived mutations from both parental populations may be formed. Some of these new combinations may be deleterious, leading to a reduced fitness of hybrids and further genetic

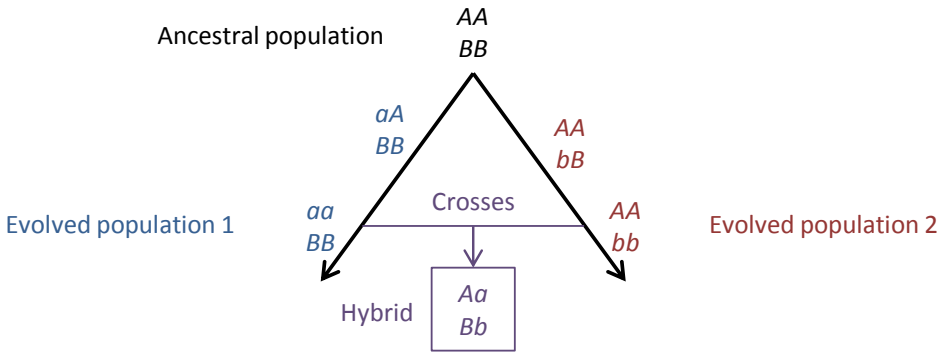

Figure 3.1: The Dobhzansky and Muller diploid model. An incompatibility (a and b) can evolve without the need for parental populations to pass through a fitness valley.

divergence. This phenomenon should increase with the divergence time of parental populations, finally leading to complete reproductive isolation and to the formation of new species (or conversely, to admixture and the decay of genetic differentiation, Barton 2013).

Reproductive isolation is often classified as extrinsic or intrinsic. When parental populations have adapted to different ecological niches, hybrids may present an intermediate phenotype. These hybrids may thus be fitter than parental populations in an intermediate environment (e.g. Wand et al. 1997). However, such habitats are generally not available, leading to low hybrid fitness. This corresponds to extrinsic, or environment-dependent, post zygotic isolation. Intrinsic postzygotic isolation relates to physiological or developmental problems reducing hybrid fitness in all environments (Coyne and Orr 1998). Extrinsic post zygotic isolation can emerge even under strictly additive gene action, because epistasis is not required to produce intermediate phenotypes. On the other hand, intrinsic reproductive isolation often involves epistatic interactions. Considering a single locus, if the cross between AA and aa individuals produces unfit Aa hybrids, it means either that multiple substitutions have occurred at this locus (Bordenstein and Drapeau 2001), or that one of the populations has passed through the unfit heterozygote genotype. Bateson (1909), Dobzhansky (1936) and Muller (1939; 1940; 1942) showed how intrinsic postzygotic incompatibilities can evolve without the need to pass through a fitness valley (Figure 3.1). In the simplest case, an allele a reaches fixation in one population, while at another locus, an allele b fixes in a second population. Alleles a and b have never been tested together by selection and their epistatic effect may be deleterious: in this case, the resulting (aB/Ab) hybrid has a reduced fitness.

The effect of such incompatibilities can be observed in introgression experiments, where a small portion of the genome of one population is introduced into the genome of another population (Coyne and Orr (2004), Table S1 Fraïsse et al. (2016)). These experiments show that incompatibilities can be quite strong (and thus probably stronger than the fitness effect of

the causal mutations in their original background). Predicting the dynamics of reproductive isolation over time requires additional hypotheses about how genes interact. For example, (Orr 1995*a*) developed a model of the evolution for reproductive isolation through the accumulation of pairwise incompatibilities, assuming that each new substitution has a given probability of being incompatible with previous substitutions that occurred in the other population. In that case, any new substitution has more chances to cause incompatibilities than the previous one: the n^{th} substitution may cause up to $n-1$ incompatibilities, and the expected number of incompatibilities is proportional to n^2 when n is large. This generates a snowball effect, reproductive isolation increasing faster than linearly with the divergence between parental populations. If incompatibilities arise between more than two loci, the number of incompatibilities rises even faster (Orr 1995*a*, Turelli et al. 2001). To date, evidence for the snowball effect in nature is still mixed (e.g., Table S1 in Fraïsse et al. 2016). Finally, one can note that in these models, the evolution of reproductive isolation does not depend on the evolutionary forces that have shaped the evolution of the parental populations: the mutations involved in incompatibilities may have been fixed either by selection or by drift in the parental populations.

1.2 Using Fisher's geometric model to study reproductive isolation

Other authors have used quantitative trait models to generate predictions on the dynamics of reproductive isolation (Barton 2001, Chevin et al. 2014, Fraïsse et al. 2016). These are based on Fisher's geometric model (FGM), representing evolution in a multidimensional phenotypic space, where the effects of mutations at different loci on phenotypic traits are supposed additive, and where fitness declines monotonously with the Euclidean distance from the phenotypic optimum. Interestingly, FGM provides a natural way of introducing distributions of epistatic interactions among mutations, and several predictions based on this model show a remarkable fit with experimental data (e.g. Martin and Lenormand 2006). Speciation models based on FGM generally consider two haploid isolated populations that may be confronted to the same change in environment represented by a change in the phenotypic optimum, or to divergent environmental changes. In some cases, the environment stays constant but the populations constantly move around the optimum due to genetic drift. As shown by Barton (2001) and Chevin et al. (2014), the average fitness of hybrids between these parental populations can be decomposed into two effects: (1) their intermediate position between parental populations in phenotypic space and (2) their increased phenotypic variance. The first factor depends on the environment. The fact that the hybrids mean phenotype is centered on the mean parental phenotype tends to increase the fitness of hybrids relative to parents in a fitness landscape with a single optimum (Barton 2001), but naturally generates extrinsic reproductive isolation when parental populations are adapted to different optima, with a fitness valley in-between.

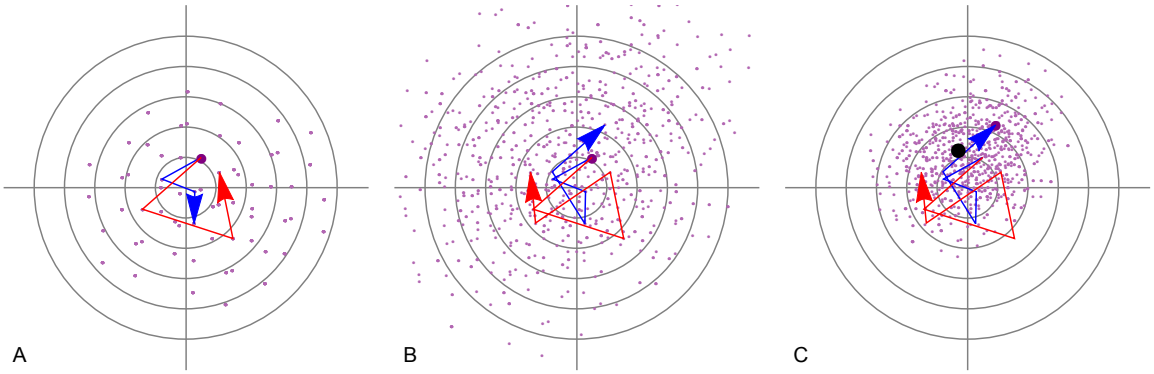

Figure 3.2: Parental trajectories (blue and red arrows) and phenotype distribution in hybrids (purple dots) in a 2 dimensional version of Fisher's geometric model. Blue and Red arrow tips correspond to parental positions in the phenotypic space at the time of hybridization. Axes cross at the fitness optimum and circles represent fitness isoclines. A. Parental trajectories and phenotypes of 500 random haploid hybrids after 3 substitutions. B. Parental trajectories and phenotype of 500 random haploid hybrids after 5 substitutions. C. Same as B in a diploid population (assuming that mutations have the same fitness effect in haploids and homozygous diploids): phenotype of the diploid F1 hybrids (black dot) and of 500 F2 hybrids.

The second factor reduces fitness in all environments (due to unfit combinations of mutations from both populations) and causes intrinsic reproductive isolation: hybridization increases the phenotypic variance (segregation variance, Wright 1968, Slatkin and Lande 1994), which reduces fitness as soon as there is stabilizing selection around a fitness optimum. This second effect increases with the divergence between parental populations; this increase is linear when the fitness function is Gaussian or quadratic (Chevin et al. 2014), but may be faster than linear when the fitness function is more 'plateau-like' (Fraïsse et al. 2016). When populations evolve by drift (fixing deleterious alleles) and compensatory mutations in a constant environment (mutation-selection-drift equilibrium, MSDE), expressions for the rate of decline of hybrid fitness over time can be obtained assuming that parental populations are monomorphic most of the time (based on the results of Sella and Hirsh 2005, Sella 2009). Figures 3.2A. and B. show the distribution of hybrids (in a two-dimensional phenotypic space) when parental populations have experienced 3 and 5 substitutions (respectively) in a constant environment, showing that the segregation variance increases with the number of substitutions.

1.3 Hybrid breakdown in haploids and diploids

The models by Chevin et al. (2014) and Fraïsse et al. (2016) considered haploid populations, in which the segregation variance is expressed and reduces hybrid fitness from the first generation (F1). By contrast, in diploids, the first hybrid generation is monomorphic (assuming that parents from the same source population all share the same genotype). Moreover, because the phenotypic effects of mutations are assumed to be additive, it is precisely at the mean of parental phenotypes (Figure 3.2C, the black dot is equidistant to blue and red arrow tips).

The mean fitness of the F1 is therefore only affected by this intermediate phenotype, (effect (1) defined above), which may confer a fitness advantage to hybrids in FGM geometry (see Figure 3.2C, black dot). In particular, when two finite diploid populations move around a fitness optimum due to mutation, selection and drift, FGM predicts a hybrid advantage over the mean parental fitness (heterosis), whose strength fluctuates around an equilibrium value. In diploids, the segregation variance is expressed in the second hybrid generation (F2). Heterosis followed by hybrid breakdown in later generations is often reported in intra or interspecific crosses (Burton 1990, Edmans 1999, Johansen-Morris and Latta 2006, Rogers and Bernatchez 2006). If we assume that the effects of mutations in haploids are equivalent to their effects in homozygous diploids (and halved in heterozygotes due to additivity), the segregation variance should be halved in a diploid F2 compared with a haploid F1, for the same level of divergence between parental populations (Barton 2001). This is illustrated in Figure 3.2B and C. As in haploids, at MSDE, hybrid fitness is supposed to decrease linearly with the divergence between parents when the fitness function is Gaussian or quadratic, but the hybrid breakdown is lower.

1.4 Experimental evolution of RI in haploids and diploids using *Saccharomyces cerevisiae*

The above predictions about the dynamics of reproductive isolation have received few experimental tests. Analyzing F1 sterility and viability versus the time of divergence between species across various taxa (bacillus, drosophila, starfish, shrimps), Gourbiere and Mallet (2010) did not find a general accelerating decrease of hybrid fitness traits. Presgraves (2010) argued that the snowball prediction from Dobzhansky Muller incompatibilities involves the number and not the cumulative fitness effect of incompatibilities. In particular, large effect incompatibilities cannot cause an accelerating fitness decrease. Evidence for a snowball effect has therefore to be sought in genomic data. Indeed, genetic mapping of different species of drosophila (Matute et al. 2010) and tomatoes (Moyle and Nakazato 2010) found that the number of incompatibilities increases faster than linearly with the time of divergence between sister species. Kondrashov et al. (2002) studied incompatibilities in protein evolution. They showed that when sequence divergence between species corresponds to the fixation of slightly deleterious and compensatory mutations (which corresponds to the MSDE described above), the number of incompatibilities does not snowball.

Here we test several predictions on reproductive isolation between isolated populations at MSDE made by FGM using mutation accumulation (MA) lines of mutator strains of the yeast *Saccharomyces cerevisiae*. *S. cerevisiae* has sbeen used previously to demonstrate the emergence of reproductive isolation between diploid populations adapting to different envi-

ronments (Dettman et al. 2007). It can develop and reproduce in the haploid or diploid phase, enabling us to explore the effect of ploidy on reproductive isolation. The present study was based on experimental yeast populations undergoing strong drift (due to repeated bottlenecks), in order to generate rapid genetic divergence due to the accumulation of deleterious and compensatory mutations, and explore patterns of reproductive isolation between haploids and diploids at MSDE. Mutator strains were used to further accelerate the rate of genetic divergence among replicate lines. We initially planned to study how selection and drift influence the evolution of RI by maintaining parental populations at MSDE at different average distances to the fitness optimum. For this, we used two different levels of bottleneck: 48 lines were carried with regular bottlenecks of one colony (formed by a single cell), and 24 lines with bottlenecks of three colonies, during 45 and 35 transfers (respectively). The fitness of MA lines with bottlenecks of 3 colonies stabilized after 15 transfers, while the fitness of the MA lines with bottlenecks of one colony was still declining at transfer 35 (no equilibrium). Therefore, the reproductive isolation analysis was carried out only on MA lines with bottlenecks of 3 colonies. In addition to the fitness decline, the capacity to undergo sexual reproduction was affected in both treatments: mating success, sporulation success and spore survival all decreased during the asexual MA lines, especially in the lines with bottlenecks of one colony. However, five pairs of haploid mutator lines (with bottlenecks of 3 colonies) could mate and produce enough segregants at transfer 15, 35 and 35 for the RI analysis. As explained below, reproductive isolation in diploids was analyzed using the haploid lines, by producing isogenic homozygous diploid parents from the haploids. Mating between haploid parents provided a diploid F1 generation, while a diploid F2 generation was produced by mating among F1 haploid segregants. These different crosses allowed us to compare the effect of genetic incompatibilities on the fitness of haploid and diploid hybrids. From transfer 15, we found that the fitness of haploid hybrids was reduced compared to parental mean fitness, although the average hybrid fitness was overestimated in our analysis: lethality of haploid genotype could not be distinguished from lethality due to a loss of spore viability and was neglected. Unfortunately, the number of successful crosses at different time-points during mutation accumulation were insufficient for testing the linearity of the fitness decline in haploid hybrids between populations at MSDE (mainly because most lines accumulated too many deleterious mutations affecting sexual traits). In diploids, the F1 generation displays a strong heterosis effect. The F2 generation average fitness was intermediate between parents and F1, and much less variable than haploid F1 fitnesses.

2 Material and Methods

2.1 Mutator strains

We constructed six haploid mutator strains carrying one of two mutator alleles ($msh2\Delta$ or $pms1\Delta$) in different combinations of MAT and $ADE2$ genotypes ($ADE2$ or $ade2\Delta$) of S. cerevisiae North American oak strain YPS670. This strain was chosen because of its fast growth rate in YPD medium. One-step gene deletions were used to replace HO with $HPMX$ and $ADE2$ with $KANMX$ in diploid YPS670. Deletions were then segregated and PCR-verified. Deletions of DNA repair genes PMS1 and MSH2 were performed in YPS670 MATa $ho :: HPHMX$ and transferred to other backgrounds by mating and tetrad dissections. Six haploid genotypes *ho::HPHMX* were isolated to initiate the MA lines: $pms1\delta$ in all four combinations of $ade2\Delta$ vs. $ADE2$ and MATα vs. MATα, and $msh2\Delta$ with MATa $ade2\Delta$ and MATα $ADE2$. Strains were frozen at $-80\,^{\circ}\mathrm{C}$ in 25% glycerol and used to initiate replicate mutation accumulation (MA) lines.

2.2 Mutation accumulation lines

MA lines were propagated on YPG + adenine agar plates ($20\,\mathrm{g\,L^{-1}}$ glycerol, $20\,\mathrm{g\,L^{-1}}$ bacto-peptone, $10\,\mathrm{g\,L^{-1}}$ yeast extract, $20\,\mathrm{g\,L^{-1}}$ agar, $60\,\mathrm{mg\,L^{-1}}$ adenine). Using glycerol (instead of the usual glucose) as a carbon substrate prevents the fixation of spontaneous respiratory-deficient mutations, which cannot metabolize glycerol (Sherman and Ephrussi 1962, Ogur and StJohn 1956, Zeyl and DeVisser 2001). Adenine was added to limit the fitness effect of the $ade2$ deletion ($ade2$ cells form red colonies smaller than wild-type on non-supplemented plates). Each plate was shared by six replicate lines and, to detect cross-contamination, neighboring lines had different genotypes ($ADE2$ vs $ade2\Delta$, MATa vs MATα). No cross-contamination events were detected.

Lines were transferred once every three days under two different bottleneck levels. In treatment MA1, a single colony was randomly chosen and streaked onto a fresh plate. In treatment MA3, three colonies were transferred. Treatment MA1 was applied to 12 replicates of genotype $MATa$ $msh2\Delta$ $ade2\Delta$, 12 replicates of $MAT\alpha$ $msh2\Delta$ $ADE2$, and six replicates of each of the four $pms1\Delta$ genotypes. Half as many replicate lines of each genotype underwent treatment MA3. The MA1 and MA3 treatments thus comprised 48 and 24 MA lines, respectively.

The MA3 lines were transferred 35 and MA1 lines 45 times. After 5, 15, 25 and 35 transfers, lines were frozen in 25% glycerol for later analyses. During the course of the experiment, the red color of $ade2\Delta$ lines was sometimes altered or lost, especially in the MA1 treatment, how-

ever these changes did not coincide with reversion to adenine prototrophy. We also observed several changes to auxotrophy in the prototrophic lines. However spontaneous auxotrophic lines were unable to grow on DM + adenine medium ($20\,\mathrm{g\,L^{-1}}$ dextose, $5\,\mathrm{g\,L^{-1}}$ Yeast Nitrogen Base, $1.7\,\mathrm{g\,L^{-1}}$ Ammonium Sulfate, $60\,\mathrm{mg\,L^{-1}}$ adenine), indicating new auxotrophies rather than cross-contamination by adenine auxotroph lines.

2.3 Crosses among lines and mutation segregation

Three MA1 lines went extinct during the course of the experiment, and 18 out of 24 MA1 pairs of lines loss their mating or sporulation ability after 15 transfers. More details about the evolution of viability in MA1 lines are provided in the Results section, but most of our analyses focused on MA3 lines.

In the MA3 treatment, 12 pairs of lines with the same mutator genotype but of opposite sex and ADE2 genotype ($MATa$ $msh2\Delta$ $ade2\Delta$ lines were mated with $MAT\alpha$ $msh2\Delta$ $ADE2$ lines, and analogously for $pms1\Delta$ lines - mating were carried with one single clone of each line, although 3 clones were used for transfers) were mated after 15, 25 and 35 transfers, giving 36 diploid hybrids (3 time-points $\times$ 12 pairs = 36 crosses). The same 12 pairs of lines were crossed at each time-point. Sporulation-competent (diploid) heterozygous clones were frozen at $-80\,^{\circ}\mathrm{C}$ in 25% glycerol for later assays. Diploid homozygotes were also obtained from the same 72 haploids parents (3 time-points 12 pairs $\times$ 2 parents / pair = 72 crosses). To do this, haploid parents were transformed with plasmid pSH3 carrying HO and $NATMX$. Clones resistant to cloNAT were isolated and tested for sporulation following a single-cell bottleneck on YPD solid medium to ensure isogeny. Most of the homozygous strains had lost their sporulation ability, and consequently, their ploidy level was verified by flow cytometry on four transformants. The clones were grown overnight in YPD liquid, and cultures were diluted 100-fold in YPD and grown for 5 hours with agitation to obtain exponentially growing populations. Cultures were centrifuged and fixed in 1mL EtOH 70% for 24 hours. Cells were then washed with 50 mM citrate sodium, sonicated, and incubated at $37\,^{\circ}\mathrm{C}$ in citrate sodium (50 mM) with $200\,\mu\mathrm{g\,mL^{-1}}$ RNAase A and $200\,\mu\mathrm{g\,mL^{-1}}$ Proteinase K. After 10 hours, samples were centrifuged and cell pellets were resuspended in citrate sodium solution with $50\,\mu\mathrm{mol}$ of Sytox Green, incubated for 1 hour at ambient temperature, and diluted 50-fold in PBS solution. $FL1-A$ distribution was measured by flow cytometry. The two modes of the distribution of fluorescence (height of FL1 values) were used to determine ploidy (cells in $G1$ and $G2$ each form a mode) (Delobel and Tesniere 2014).

Diploid hybrids were sporulated and haploid segregants were isolated by tetrad dissection (16 to 18 tetrads for each of the hybrids). Tetrads were dissected on YPD plates following 40 to 60 days on sporulation plates (preliminary experiments on ancestral genotypes showed that

segregants viability increased over time: 52% versus 61% of the segregants formed visible colonies 9 versus 54 days following sporulation induction). Segregants forming visible colonies after incubating at 30 °C for 7 days and at ambient temperature for an additional 14 days (to allow slow growing genotypes to form colonies) were frozen for fitness assays. Fitness differences between parents and F1 hybrids were analyzed when crosses produced at least 15 segregants at transfer 15, and 15 segregants at transfer 25 or 35, which was the case for 2 of 6 pairs of $msh2\Delta$ and 3 of 6 pairs of $pms1\Delta$ MA3 haploid parents (see Table 3.1). Finally, F2 diploids were produced at transfer 15 for 3 pair of lines, and at transfer 35 for one of them (see Table 3.1). For these four pair of lines, 40 mating among F1 haploid segregants were conducted. To obtain a F2 generation approaching Hardy Weinberg equilibrium, we followed these rules: (1) Equal representation of each tetrad and, as possible, all segregants should be used, and represented the same number of times. However, in order to mate every segregants, some individuals with a rare mating type × auxotrophy combinations have been used more than others; (2) no intra-tetrad mating, because haplotypes resulting from the same meiosis are not independent. Mating success was checked by ploidy measurements (see above), and F2 diploids were frozen for fitness assays.

2.4 Fitness assays and estimation of competitive fitness

Competitive fitness of the haploid parents, autodiploids parents, heterozygous F1, F1 haploid segregants, and F2 diploids were measured against a YFP-fluorescent competitor. The common competitor was a direct derivative of the ancestor expressing the fluorescent protein mCitrine under the strong promoter of the actin gene ACT1. It was obtained by PCR amplifying $ura3D :: PACT1-ymCitrine-tADH1-URA3$ from strain yJHK226 of (Koschwanez et al. 2011).The PCR product was transformed into YPS670 $ho :: HPHMX$ using $URA3$ as a selectable marker and a standard lithium acetate/single stranded carrier DNA/polyethylene glycol protocol (Gietz and Woods 1991). Haploid MATa and MATα lines were competed against the haploid competitor of the same mating type and diploid strains were competed against the isogenic diploid competitor. Haploid fitness assays were performed in YPG liquid because MA lines were propagated on that medium. However, diploid growth was very weak in YPG because growth on non-fermentable carbon substrates such as glycerol induces sporulation, therefore diploid fitness was measured in YPD liquid. This precluded a direct comparison between haploid and diploid fitness. Cultures were agitated on an orbital shaker at 250 rpm during competitions.

For each competitive assay, 5 µL of each frozen strain was inoculated into 1 ml YPD. After 48h of incubation at 30 °C without agitation, 10 to 200 µL (the volume was increased for cultures with low cell densities) of each culture were transferred to 1mL YPG liquid, covered with a

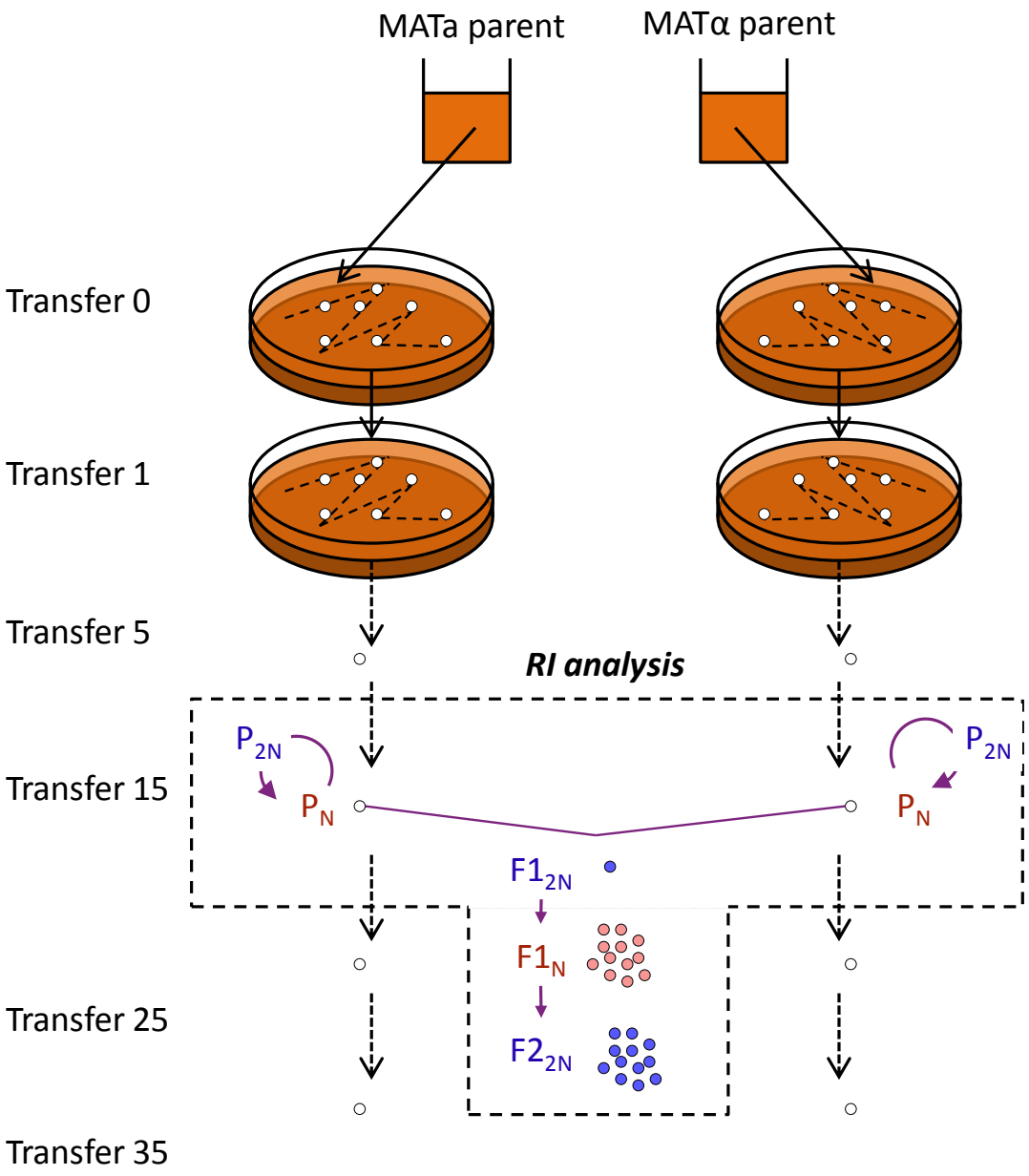

Figure 3.3: Experimental plan for a pair of MA lines. At transfer 15, 25 and 35, crosses were made between parental $MAT\alpha$ and $MAT\alpha$ lines (P_N). Autodiploidization of the haploids parents gives homozygote diploids parents (P_{2N}). Diploid F1 ($F1_{2N}$) were sporulated to obtain haploid F1 segregants ($F1_N$), segregants from the same diploid F1 were mated to obtain diploid F2 ($F2_{2N}$). Fitness of all genotypes (parents and hybrids) was measured. The dashed box show the protocol for reproductive isolation analysis at transfer 15. The same protocol was used at transfer 25 and 35.

Breath Easy®film to favor gas exchanges while limiting evaporation, and grown for 48 hours at 30 °C and 250rpm. A similar number of cells of the focal genotypes and the competitor were then mixed by adjusting the volume of focal genotype transferred (20, 50, 200, 400, or 600 µm) and completing the mixture to 1.1mL with YPG liquid. Mixed cultures were incubated on 24 wells plates at 30 °C, shaken at 250 rpm and covered with Breath Easy®filters. After 24h of incubation, 10 µL of the competition cultures was transferred to 1mL YPG and incubated for another 24h under the same conditions. One hour after the transfer, we sampled 30 µL of the cultures, diluted these samples 10 fold in PBS and processed them on a BD Accuri cytometer (20 000 cells were measured from each culture). Diluted cultures were kept at 4 °C during cytometry analyzes to prevent additional cells divisions before measurements. The same measurements were made after 24h of competition.

Data were analyzed using the BDSampler software. Small debris were excluded with an initial gate, and gates were drawn around the two clusters formed by non-fluorescent and fluorescent cells on plots of $FL1-A$, $FL2-A$ and $FSC-A$ to estimate the number of fluorescent vs non-fluorescent cells. The competitive logfitness W_m was determined for each assay as:

$$W_m = \ln\left[\frac{p_2\,(1-p_1)}{p_1\,(1-p_2)}\right] / \ln\left[\frac{1000\,(1-p_2)\,d_2}{10\,(1-p_1)\,d_1}\right]$$

Where p_i and d_i are the frequencies of non-fluorescent cells and the total number of events measured in the competition mix after i days of competition. The relative selection coefficient of the focal strain is given by $\log 2\left[p_2\,(1-p_1)\,/\,(p_1\,(1-p_2))\right]/gen$ where gen is the number of generations (Chevin 2011), and $(10\,(1-p_1)\,d_1)\,/\,(1000\,(1-p_2)\,d_2)$ gives the ratio between the number of competitor cells in the mix before and after the 24h of competition. The number of doublings of the competitor is then given by $\log 2\left[1000\,(1-p_2)\,d_2/\,(10\,(1-p_1)\,d_1)\right]$.

Four blocks of fitness assays were performed and analyzed independently. In the first block, parental haploid fitness was measured for treatments MA1 and MA3 (ten blocks each, in each block, fitness assays were carried within the same 2 weeks). In the second experiment, fitness of the haploid segregants and their corresponding parents was measured (4 blocks). In the last experiment, fitness of the diploid F2, F1 and parents was measured (2 blocks). Measurements were duplicated for all blocks. For each experiment, we used a linear model to correct for the effect of the block: $W_m = W_g + X_{block} + \epsilon$, where W_m is the competitive logfitness measured, W_g is the estimation of the logfitness for each genotype, X_{block} the effect of the block (variation in relative fitness due to small changes in environmental conditions, cell densities, etc), and ϵ the standard measurement error. We performed an ANOVA to obtain W_g.

Table 3.1: Experimental design for fitness assays. Ancestral genotypes of the five pairs of line selected for reproductive isolation analysis: M1, M2: *msh*2Δ, MAT*a* *ade*2Δ × MATα *ADE*2; P1, P2: *pms*1Δ, MAT*a* *ade*2 × MATα *ADE*2; P3: *pms*1Δ, MAT*a* *ADE*2 × MATα *ade*2Δ

Exp.	Treatment	Block	Genotypes measured (twice)	Transfer
MA	MA1 lines	A	All haploid parents	0 and 5
		B		0, 5 and 15
		C		0, 15 and 25
		D		0, 25 and 35
		E		0 and 35
	MA3 lines	F		0 and 5
		G		0, 5 and 15
		H		0, 15 and 25
		I		0, 25 and 35
		J		0 and 35
N	MA3 lines Haploid F1 and parents	A	Lines M1, M2, P1, P2, P3	0 and 15
		B	Lines M1, M2, P1, P3	0 and 25
		C	Lines M1, M2, P1, P2	0 and 35
		D	Missing Data - F1 and parents	0, 15 and 25
2N	MA3 lines Diploid F1, F2 and parents	A	Lines M1, M2, P1, P2, P3	0 and 15
		A	Lines M2	0 and 35

Table 3.2: ANOVA of the effect of bottleneck treatment (MA1 / MA3), mutator genotype ($msh2\Delta$ / $pms1\Delta$), mating-type ($MAT\mathbf{a}$ / $MAT\alpha$) and adenine auxotrophy ($ADE2$ / $ade2\Delta$) on the fitness of the haploid MA lines at transfer 35

	DF	SS	MS	F-Statistic	P-Value
Bottleneck	1	1.86	1.86	47.1	< 0.0001
Mutator	1	0.0394	0.0394	0.100	0.322
Mating-type	1	0.0766	0.0766	1.94	0.169
ADE2	1	0.00843	0.00843	0.213	0.646
Error	67	2.65	0.0396	Null	Null
Total	71	4.64	Null	Null	Null

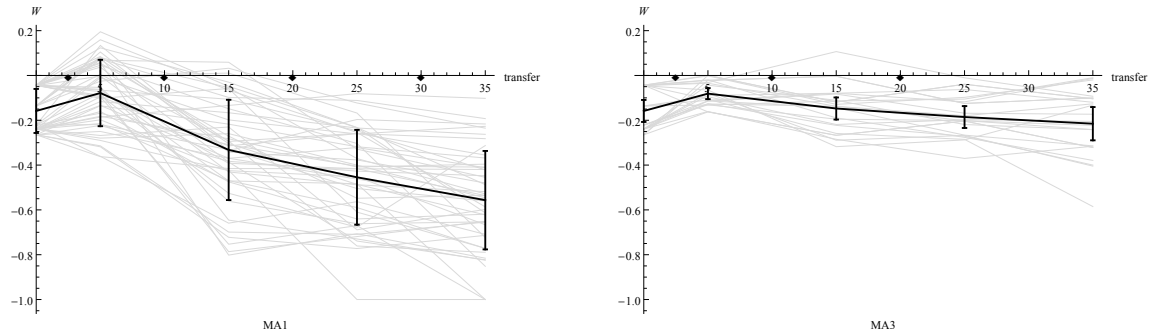

Figure 3.4: Competitive fitness trajectories of the haploid parental lines under treatment MA1 and MA3 treatment. The grey curves show individuals trajectories. The thick curves indicate the average fitness of the lines. Errors bars show standard deviation.

3 Results

3.1 Mutation selection drift equilibrium

Fitness declined in 59 out of 72 MA lines (45/48 for MA1 and 14/24 for MA3) and bottleneck size was the only factor explaining the final fitness of the MA lines (ANOVA analysis, see Table 2). The mating-type, the ADE2 genotype and even the mutators genotype ($pms1\Delta$ or $msh2\Delta$) did not affect final fitness (Table 3.2). Analysis of the MA3 and MA1 treatments were therefore carried out after pooling lines from different ancestral genotypes.

Competitive fitness trajectories of the MA lines under both MA1 and MA3 treatments display strikingly different patterns (Figure 3.4). After a significant increase in fitness at transfer 5 (student t-test, $p < 10^{-3}$), fitness declined in both treatments. In the MA1 treatment, the average fitness at $t = 35$ was reduced by $s = 0.40$ compared with the initial fitness. This is almost 7 times greater than the fitness reduction of $s = 0.06$ observed in the lines undergoing the MA3 treatment ($p < 0.0001$, see Table 3.2).

Table 3.3: AIC and criterion for the linear and the exponential model of the fitness trajectory of the haploid MA lines under treatment MA1 and MA3

	MA1	MA3
Linear	-56.99	-155.9
Exponentiel	-62.63	-156.1

To test whether the fitness trajectories of the lines undergoing the MA1 or the MA3 treatment declined linearly or approached an equilibrium, we compared the fit of (1) a linear and (2) an exponential model: (1) $W = a + bT$ (2) $W = a + e^{bT}$

For both treatments, the exponential model is better (AIC criterion), indicating a slowdown in the fitness decrease of the lines ($b < 0$). Indeed, in the MA1 and MA3 treatments, respectively 53 and 50% of the fitness decrease occurs between transfers 5 and 15. Linear regression on the fitness trajectories between transfer 15 and transfer 35 showed that MA1, but not MA3 lines were still significantly declining after T15 (MA1:$DF = 46$, $p = 2.0310^{-6}$, MA3: $DF = 22$, $p = 0.057$). The extinction of three mutator lines $msh2\Delta$ between transfer 15 and 35, and of two $msh2\Delta$ and two $pms1\Delta$ additional lines before transfer 45 (not shown) suggests mutational meltdown in the MA1 lines.

3.2 Loss of sex during mutation accumulation

Table 3.4 summarizes the ability to carry out each step of the sexual cycle for the 24 MA1 and 12 MA3 pairs of haploid parental lines. Each step of the sexual cycle declined rapidly, especially for MA1 lines, most of which had lost the ability to sporulate by $t = 25$ and to mate by $t = 35$. Under the MA3 treatment, 75% of the pairs of lines could still mate at transfer 35, and 5 out of 8 were still able to sporulate. However, spore survival declined between transfer 15 and transfer 25. One pair of lines mated at transfer 25, but not at transfer 15, possibly because the clone tested at $t = 25$ was not a direct ancestor of the clone tested at $t = 15$ (this is possible in MA3 lines).

The loss of mating and sporulation abilities in our crosses can result from the accumulation of mutations directly impacting these traits, or from incompatibilities between parental genotypes. In order to distinguish between these effects, we estimated the mating success of haploid parents from the same line, and the sporulation success of the resulting autodiploids. In MA1 lines, selfing success was reduced compared with mating between lines (9 success out of 48 attempts, in comparison to 11/24 for hybrids), and only 2 out of 9 autodiploids sporulated. In MA3 lines, mating success was nearly equal for selfed versus hybrid lines (20 successes out of 24 attempts), but sporulation rates were greatly reduced: only one MA3

Table 3.4: Reproductive success of the crosses between pairs of haploid MA lines under treatment MA3 and MA1 at transfer 15, 25 and 35. (Number of pair of lines succeeding in $_{/\text{number of strains tested}}$). Dashes indicate the absence of data

	MA1			MA3		
Transfer	15	25	35	15	25	35
Mating	$11_{/24}$	$11_{/23}$	$7_{/21}$	$11_{/12}$	$9_{/12}$	$8_{/12}$
Sporulation	$7_{/11}$	$0_{/4}$	-	$8_{/11}$	$9_{/9}$	$5_{/8}$
Spores survival $> 20\%$	$4_{/7}$	-	-	$6_{/8}$	$4_{/9}$	$3_{/5}$

homozygote sporulated. At transfer 25, mating success had decreased to 25%. These results indicate that the loss of mating and sporulation is at least partly due to mutations directly affecting these traits, that accumulated within the MA lines (as those mutations were not selected against).

3.3 Evolution of post zygotic incompatibilities in haploids

Because MA1 lines did not reach a fitness plateau and yield very few segregants, only haploid lines from the MA3 treatment were used for the analysis of reproductive isolation between lines. Five pairs of lines, described in table 1, produced more than 15 segregants at least at two transfers (15, 25 and/or 35). Competition experiments reveal that for 12 out of 13 crosses, F1 hybrids are on average less fit than their parents. This difference is significant in 9 crosses (Kruskal-Wallis non parametric test). Note that the distribution of hybrid fitness is large and that despite the reduction of the hybrid mean fitness some individuals have very high fitness value. The fitness of some segregants was so low that they were not detected in the competition mix in the first measurement of the fitness assay ($p_1 = 0$). We arbitrarily attribute a relative logfitness of $W_{F1} - W_P = -1$ to these segregants.

In 4 out of 5 pairs of lines (M1, M2, P2 and P3), the fitness of F1 hybrids relative to their parents ($W_{F1}W_P$) decreased with the number of transfers. This change was significant in lines M1 and P2 (linear regression, M1: $DF = 82$, $p = 0.014$; P2: $DF = 68$, $p = 0.022$). In pair M2, a surprisingly high fitness of hybrids at transfer 35 lead to a non significant fitness change between transfer 15 and 35. In pair P1, the hybrid fitness relative to the parents increased significantly with the number of transfers since divergence (linear regression, $DF = 103$, $p = 0.028$). All genotypes taken together, we found a small but significant negative effect of the number of transfers since divergence on the hybrid fitness relative to the parents (linear regression, $DF = 410$, $p = 0.025$).

Spore survival decreased with the number of transfers in the 5 pairs analyzed (Figure 3.5, lower line). Assuming that the ratio death due to haploid null fitness/ death due to a loss of

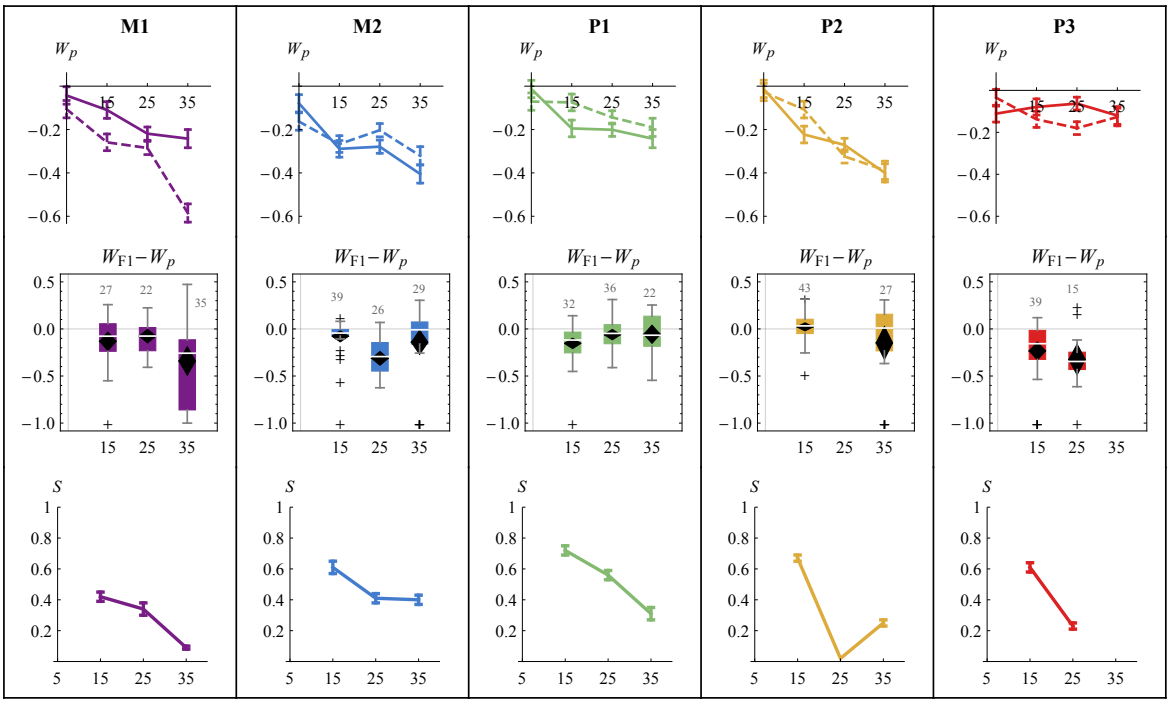

Figure 3.5: Reproductive isolation analysis at transfer 15, 25 and 35 (x-axis) in five MA3 line pairs (columns). Two pairs of mutator lines (M1 and M2) are $msh2\Delta$ and 3 pairs of mutator lines (P1, P2 and P3) are $pms1\Delta$. The upper panels display parental competitive logfitness trajectories, where the solid curve represents the MATa parent and the dashed curve the MATα parent. Error bars show the standard error of the fitness estimations. The middle panels represent the relative competitive logfitness of the haploid F1 segregants, shown in Tukey Boxplot (parents have $W = 0$). Diamonds represent F1 hybrids mean relative fitness and its interval confidence (95%). The number of hybrids analysed is displayed above the boxes. The lower panels show the mean spore survival (S) after crosses between haploid parents, and error bars represent the standard error.

sex is constant, the number of lethal haploid genotype ($W_h - W_p = -1$ in Figure 3.5) increased between transfer 15 and 35. We therefore neglected an increased number of hybrids with a null fitness with the number of transfers. This can explain the fitness increase in crosses from pair P1, but not the high fitness of hybrids from pair M2 at transfer 35, given that spore mortality is the same at transfer 25 and 35.

3.4 Heterosis and incompatibilities in diploid hybrids

The five pairs of haploid lines analyzed for haploid reproductive isolation were also used to compare the fitness of homozygous autodiploid parents with F1 hybrids (crosses between lines - five pairs of lines at $t = 15$, and one pair at $t = 35$), and F2 offspring (crosses among segregants 3 pairs of line at $t = 15$, one at $t = 35$) (Figure 3.6).

Fitness measurements revealed that F1 hybrids are fitter than both homozygous parental ancestors even at $t = 0$ (Figure 3.6). This can be partially explained by the fact that in all crosses, one of the parents was auxotroph *ade*2Δ. Indeed, ANOVA analysis revealed that auxotrophy but not mutator genotype affected the diploid ancestor fitness and that *ade*2Δ ancestors were significantly less fit than the isogenic *ADE*2 diploids ($\overline{W_{ADE2}} - \overline{W_{ade2\Delta}} = -0.32$, ANOVA analysis, $DF = 3$,$p < 5.10^{-3}$). F1 hybrids, with one functional copy at the ade2 locus, do not suffer from a fitness reduction, indicating that *ade*2Δ deletion is recessive. In addition, F1 hybrids are on average fitter than their prototrophic parent by $s = 0.10$ (ANOVA analysis, $DF = 3$, $p = 0.01$).

At transfer 15, all lines except P3 displayed an increased heterosis effect (Figure 3.6). Although all auxotroph autodiploids parents increased in fitness between transfer 0 and 15 (mean change: +0.11, see Figure 3.6A), the fitness decrease of the prototroph lines was more important and the mean fitness of autodiploids parents tended to decrease ($s = -0.04$, paired t-test, $DF = 9$,$p = 0.43$) from transfer 0 to transfer 15. In contrast, the fitness of F1 hybrids increase slightly ($s = +0.05$, t-test, $DF = 4$, $p < 0.01$) (Figure 3.6A), showing that the deleterious mutations fixed in the autodiploids parents are recessive. Incompatibilities may also be present, but there are at the heterozygous state and their deleterious effect is negligible compared with the heterosis effect. In lines M2 (blue line), heterosis in F1 hybrids at transfer 35 was lower than at transfer 15, although the mean fitness of autodiploid parents did not change significantly between transfer 15 and 35. Fitness analysis of the haploid F1 segregants from the pair of lines M2 reveals the presence of post zygotic incompatibilities between parental lines (see Figure 3.5, $W_{F1}W_P$ was significantly lower than 0), however, the overall effect of these incompatibilities did not change between transfer 15 and 35.

In all crosses, F2 hybrid fitness was intermediate between the parental and F1 fitnesses. The ratios $W_{F2}/(W_{F1} + W_P)$ was significantly higher from 0.5 in line M2 at transfer 15 (Figure

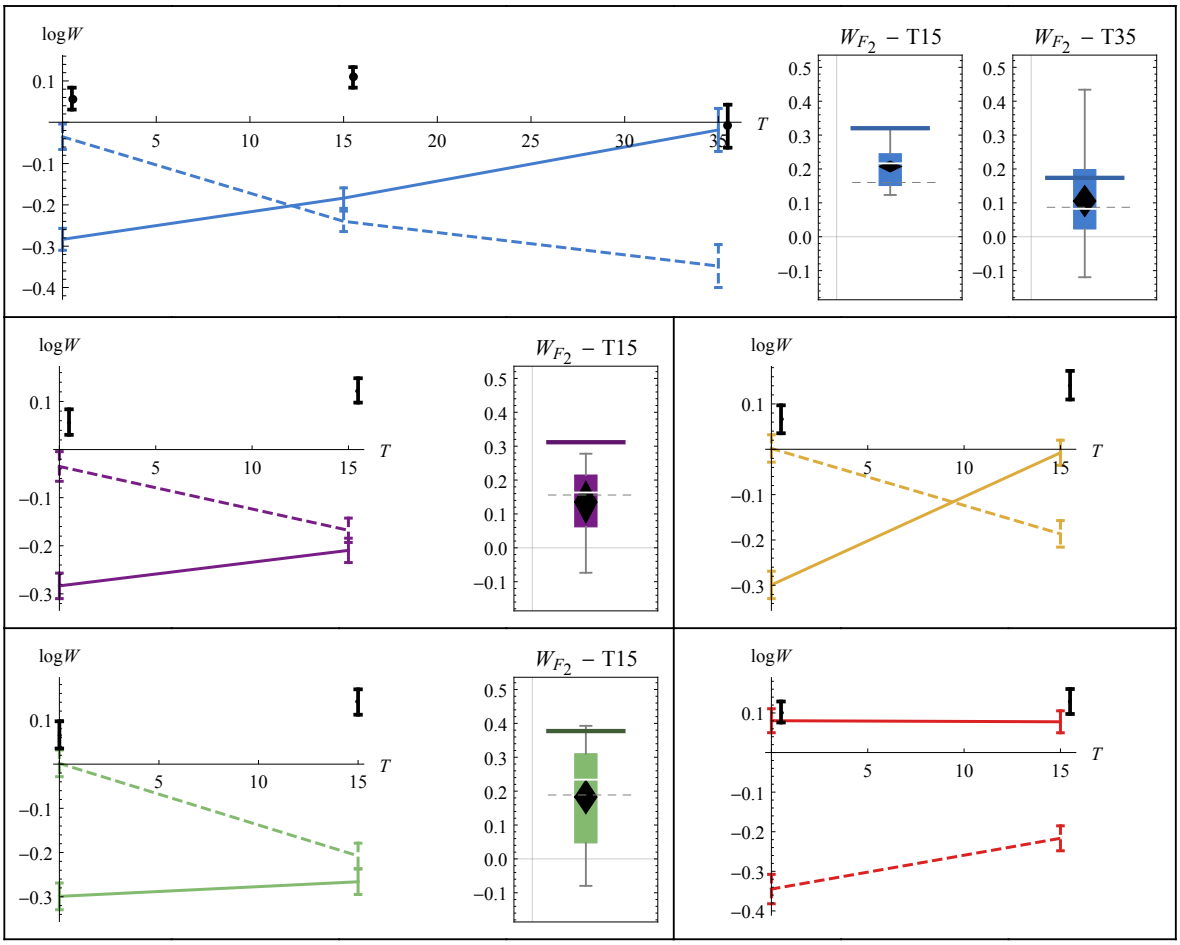

Figure 3.6: Fitness analysis of autodiploids parents, F1 and F2 diploid hybrids. Colors are the same as in Figure 3.5. Graphs represent the logfitness trajectories of homozygous diploid parents (Solid lines, autodiploids from haploid MATa, Dashed lines, autodiploids from haploid MATα) and the logfitness of F1 hybrids (black dots). Error bars represent the standard error. Note that in all lines but P3 (red curve), the MATa parent is also *ade*2Δ. Box plots present the logfitness distribution in F2 hybrids relatively to mean parental logfitness. Diamonds represent F2 hybrids mean relative fitness and its interval confidence (95%). F1 hybrid fitness (darker bars), and the average between F1 and parents mean fitness (grey dashed bar) are plotted for comparison.

3.6, blue box plots) (Kruskal Wallis test, $DF = 34$, $p = 0.01$). In lines M1 and P1 at transfer 15 (Figure 3.6, purple and green box plots), F2 hybrids tend to be less fit than the average between F1 hybrids and parents, possibly due to recessive incompatibilities masked in the F1 hybrids and exposed in the F2. In line M2 at transfer 15 and 35, F2 hybrids tend to be fitter than the average between F1 hybrids and parents.

4 Discussion

The goal of this study was to explore the dynamics of reproductive isolation between populations diverging rapidly due to the fixation of deleterious and compensatory mutations. For this, we used mutator yeast MA lines with two different levels of bottlenecks. With bottlenecks of three colonies (MA3 treatment), the mean fitness of lines equilibrated after 15 transfers. Auxotroph *ade*2Δ lines showed a reduced fitness at the beginning of the lines, but no significant differences between initial genotypes could be detected after 35 transfers, which confirms that the dynamic fitness equilibrium does not depend on the initial fitness (Silander et al. 2007). Unlike prototroph line, the fitness of *ade*2Δ lines increased between transfer 0 and 5 under both bottleneck treatments. With bottlenecks of one colony (MA1 treatment), fitness was still decreasing between transfer 25 and 35, and the average fitness after 35 transfers was about 30% lower than in the lines with bottlenecks of 3 colonies (0.57 vs 0.81). Final fitness was also much variable in this treatment. In particular, before 45 transfers, 7 lines over 48 got extinct, which was significantly higher than in the MA lines with a bottleneck of 3 colonies (p ¡ 10-6, likelihood ratio test). These independent extinctions suggest that, as observed in a previous experiment mutator *msh*2Δ MA lines (Zeyl et al. 2001), mutational meltdown occurs during these MA lines.

Mutations affecting the capacity to undergo sexual reproduction accumulated during the experiment (due to the fact that these mutations were not counter-selected within MA lines), causing a steady decline in sexual capacity. Remarkably, the different sexual traits (mating, sporulation and spore survival) were more severely impaired in the lines bottlenecked at one than at three colonies: this may be due to pleiotropic effects of mutations affecting survival and asexual reproduction on traits associated with sex. Indeed, Enyenihi and Saunders (2003) found in a library of single-gene deletion mutants of S. cerevisiae that only 17% of genes necessary for full sporulation were sporulation specific, the remainder being also involved in asexual propagation. In a previous experiment, Hill and Otto (2007) showed that sporulation success increased in large asexual populations of S. cerevisiae, suggesting that positive pleiotropy may help to maintain sex. Here, we observed the same positive correlation between asexual fitness and sporulation in diploids, and other sexual traits that were not measured by Hill and Otto (2007): mating success and spore survival. By contrast, Zeyl et al. (2005) found antago-

nistic pleiotropy between sexual and asexual fitness in adapting populations of S. cerevisiae, suggesting that the form of pleiotropy may strongly depend on the mutations.

The fixation of mutations directly affecting sporulation and spore viability prevented us to estimate the additional contribution of epistatic interactions between mutations present in different lines on spore lethality, which was high in the experiment (up to 98% for the M1 lines at transfer 35). To disentangle both effects, it would have been interesting to analyze spores survival after sporulation of the autodiploidized strains (diploid parents), however these were in most cases unable to sporulate or spore viability extremely low, so we could not perform this analysis. Under both treatments, crosses between pairs of lines were more successful than crosses within lines (reproductive success of autodiploidized parents), which indicates that these mutations are at least partially recessive, and that heterozygous diploids benefit from a masking advantage compared with autodiploids. However, without knowing quantitatively the dominance coefficient of these mutations, we cannot exclude that genetic incompatibilities also reduce the success of crosses between lines. We eventually had to neglect lethal genotypes in our hybrid fitness analysis, although some of them may correspond to incompatible combinations of mutations causing lethality. The mean fitness of hybrids was therefore overestimated in our analysis, and hybrid breakdown probably stronger than our estimates (again, due to the fact that we did not take lethals into account).

Due to the important decrease of the sexual success mentioned above, only 5 pairs of lines under the MA3 treatment could produce enough segregants for the hybrid fitness analysis in haploids. F1 hybrids were found less fit (on average) than their parents in 12 out of 13 crosses (the fitness was significantly lower in 9 crosses), indicating negative epistasis between fixed mutations. Furthermore, segregants from the five pairs of lines analyzed showed a slight but significant decrease in fitness (relative the fitness of parents) with the number of transfers. In four out of five pairs of lines, hybrid fitness decreased with the number of transfer (significantly for two pairs), while in one pair, hybrid relative fitness increased significantly. Chevin et al. (2014) showed that in Fisher's geometric model (with a Gaussian fitness function), segregation variance increases linearly with genetic divergence, causing hybrid mean fitness to decrease linearly (Figure 3.2 A and B). Due to our limited number of data points, we were not able to test the overall shape of the decline in relative fitness of hybrids over time. In the first line analyzed (M1), hybrid fitness fell at the last transfer. In the second line (P1), hybrid fitness increased linearly with the time of divergence, while spore survival decreased over time. In this case, spore mortality can mask the presence of lethal haploid genotypes: a stronger purge during the sexual treatment can explain this dynamic. In the third lines (M2), hybrid relative fitness decreased, then increased strongly. Spore lethality was the same between the last time points and cannot explain this dynamics. Two other arguments may be advanced to explain this unexpected change. First, in the MA3 lines, transfers imply

3 colonies, while one single clone is used for crosses: individuals used at transfer 25 may not be an ancestor of the individual crossed at transfer 35. Then, in FGM, hybrid mean position follows a constant distribution and does not change, in average, hybrids fitness over time. However, when parents are far apart in the fitness landscape, hybrids mean position can have a strong impact. Here, the fitness of parental lines under MA3 treatment at MSDE was, in average, reduced by 59% compared to the highest fitness measured in the haploids (hybrid genotype). Populations were therefore far from the optimum and the mean hybrid position in the fitness landscape may add an important noise in addition to the segregation variance effect. Finally, the variance in fitness of segregants was generally important, some F1 hybrids having higher fitness than their parents: these correspond to recombinant genotypes carrying fewer mutations, or combinations of compensatory mutations placing them closer to the fitness optimum than their parents. The possibility to produce autodiploids from haploid parents enabled us to explore the effect of hybridization between homozygous diploid lineages carrying different sets of mutations. This may thus be considered as mimicking diploid populations evolving under strong drift (in the presence of sexual segregation leading to homozygosity), with the caveat that mutations fixed in our haploid lines may have different effects in haploids and in homozygous diploids. Previous work on S. cerevisiae showed that both deleterious (Szafraniec et al. 2003) and beneficial (Gerstein 2013) mutations can have different effects in haploids and homozygous diploids, but they these were either EMS induced mutations, or mutations conferring resistance to a specific antibiotic. By contrast, Korona (1999*a*) found a strong correlation between the fitness of genetically loaded haploids and the derived homozygous autodiploids.

We observed a strong heterosis effect in diploid F1 hybrids, relative to their homozygous parents. Hybrid fitness was higher than parental fitness from transfer 0, indicating that the *ade*2Δ deletetion was recessive, but also that ancestral diploid strains were not totally isogenic, because diploids F1 were fitter than their prototrophic parents. Deleterious mutations, induced during the haploid ancestral strains construction or during the autodiploidisation, were probably present in all ancestral strains, and masked in the F1 hybrids. At transfer 15, 4 out of 5 lines displayed an increased heterosis effect. In the lines analyzed at transfer 35, heterosis was weakened compared with transfer 15.At MSDE, parental positions fluctuate around the optimum, and FGM predicts that diploid F1, located at the mean between parental positions, also follow a constant distribution. Although our data are insufficient to test the constant distribution of the diploid F1 fitness, the increased F1 fitness observed between transfer 0 and transfer seems consistent with this prediction.

F2 hybrids were produced for three pair of lines at transfer 15 and one pair of line at transfer 35. Their mean fitness was always reduced compared with F1 fitness, while all mutations are heterozygous in diploids F1, under random mating, the population reaches Hardy Weinberg

equilibrium after a single round of sexual reproduction. We mimic this panmictic sexual cycle to obtain a F2 generation as close as possible to Hardy Weinberg equilibrium. In the absence of epistasis, the mean F2 hybrid fitness at Hardy Weinberg equilibrium should be equal to the average between the F1 and mean parental fitness (Figure 6, dashed grey lines). Except from the M2 lines at transfer 15, the F2 hybrid mean fitness was never significantly different from this value, indicating no significant effect of the incompatibilities detected in the haploid phase on the diploid F2 generation. In the M2 lines, the F2 generation was even significantly fitter than the average between parents and F1 diploids, which would indicate positive epistasis. However, it is important to note that our protocol introduces a bias against observing low-fitness F2 genotypes, due to selection occurring among the haploid segregants produced by the F1 generation: most segregants that failed to develop or grew too slowly could not be used to generate F2 diploids. In addition, only a fraction of all segregants could mate to produce F2 diploids, and segregants carrying higher numbers of mutations may have failed to mate. Overall, there was no correlation between the fitness of diploid F2 individuals and the fitness of their haploid F1 parents (Figure 3.7), although a significant positive correlation was found between autodiploids and haploids parents fitnesses (linear regression, $DF = 16$, $p = 0.047$). It probably indicates a negative correlation between the strength and the dominance of the incompatibilities, which would not be surprising, since a negative correlation between strength and dominance of deleterious mutations have already been revealed (Korona 1999*a*, Phadnis and Fry 2005, Agrawal and Whitlock 2011, Manna et al. 2011).

Overall, our experiment outlines the possible benefits of hybridization between heavily loaded populations, through the production of some fit recombinants (in haploids) and through masking of deleterious alleles (heterosis in diploids). In the long run, however, one expects that such populations evolving under strong drift would become incompatible due to the presence of different epistatic combinations of deleterious and compensatory mutations (the proportion of recombinants fitter than the parents becoming smaller and smaller as divergence between the parents increase), but our experiment did not allow us to explore the dynamics of reproductive isolation over a sufficiently long time. It would be interesting to modify this experimental setting by regularly selecting for the capacity to undergo sex (in order to avoid fixing mutations causing the loss of sex), and perhaps also to explore different regimes leading to a lower fitness decline of the parental lines (higher bottleneck size, or strains with a lower mutation rate).

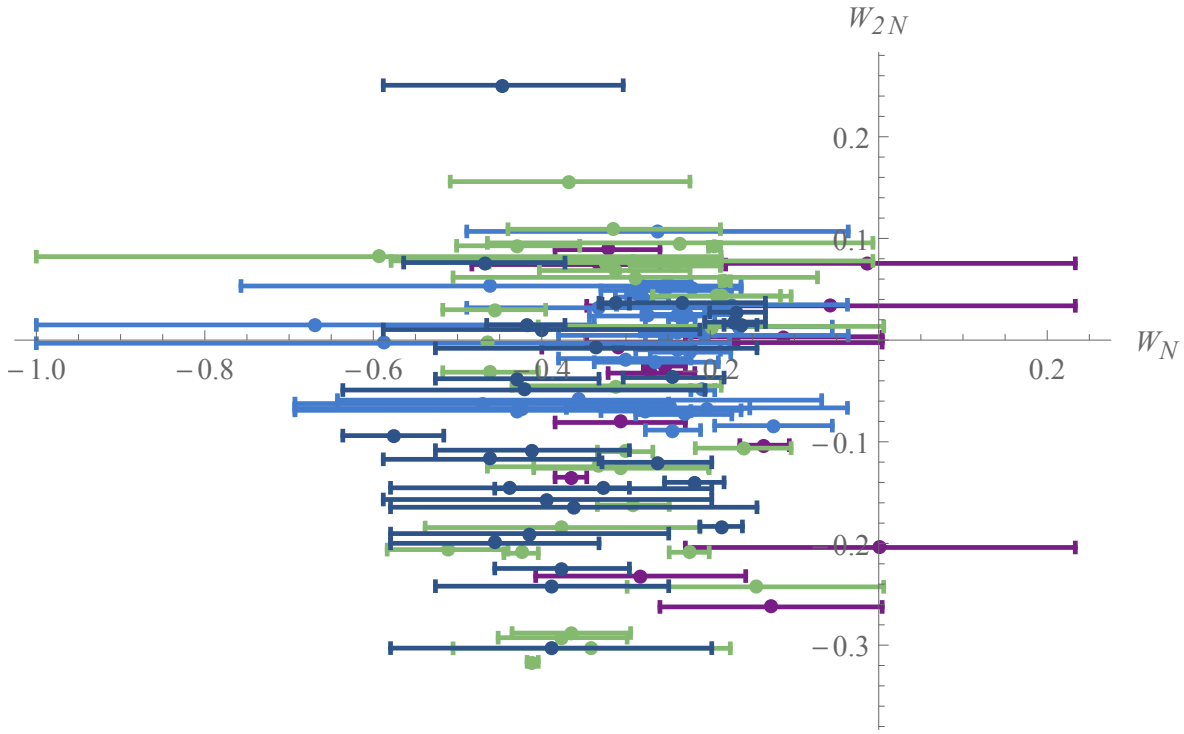

Figure 3.7: Relation between diploid F2 individuals logfitness and the mean logfitness of their haploid F1 parents. Colors are the same as in Figure 3.6, except for lines M6 at transfer 35 which is represented in dark blue. Extremities of the error bars indicate both haploid parents fitnesses.

Conclusions and perspectives

The important diversity of life cycles in the wild can be explained by both genetic and ecological factors. Previous genetic models have shown that mutations influence the evolution of life cycles through a direct fitness advantage of one phase over the other, corresponding to the diploid masking advantage, and an indirect effect due to genetic linkage between modifier alleles affecting the life cycle and deleterious alleles. Depending on the relative importance of these factors, genetic models predict an evolution toward haploidy or diploidy, but do not predict the maintenance of biphasic life cycles (Otto and Goldstein 1992, Otto 1994, Jenkins and Kirkpatrick 1995). During this PhD, I included more realism into genetic models and considered the effect of (1) differential or antagonistic selection between phases, (2) distribution of selection coefficients across loci, (3) epistatic interactions between deleterious mutations and (4) multiple loci distributed at different genetic distances from the modifier locus. In all cases, haploid diploid cycles were not stable, confirming that genetic factors may favor haploidy or diploidy, but cannot explain the wide diversity of biphasic life cycles.

On the contrary, ecological effects may stabilize haploid-diploid life cycles. Hughes and Otto (1999) showed that biphasic life cycles are favored when haploids and diploids compete within even slightly different niches with density dependent selection. In Chapter 2, we confirmed this result but showed that the occurrence of deleterious mutations across the genome narrows the level of niche differentiation allowing the maintenance of biphasic life cycles. We additionally showed that under certain conditions (depending on the dominance of deleterious alleles), the interaction between mutations and intrinsic fitness difference between phases may favor such life cycles. Intrinsic fitness differences and niche differentiation between phases require physiological or morphological differences between haploids and diploids. However, the evolution of different morphs adapted to different resources may also be achieved without any difference in ploidy between morphs, as in life cycles involving an alternation of generations with the same ploidy level (organisms with a larval phase, cnidarians...). In a sense, theories based on niche differentiation do not really explain the maintenance of biphasic life cycles, unless additionally assuming that the specialization of haploids and diploids on different resources is easier to evolve than the specialization of different morphs within the same ploidy level. As

in Hughes and Otto (1999), the model presented in Chapter 2 assumes a pre-existing niche differentiation between haploid and diploid individuals; it would be interesting to explore how such a differentiation would evolve, or more generally study the evolution of adaptation to different resources in species with haploid diploid life cycles.

In their ecological model, Hughes and Otto (1999) found that the modifier locus could be highly polymorphic at equilibrium, provided that the number of haploids and diploids coincides to an optimal exploitation of the resources available. As we saw in Chapter 2, adding mutations into this ecological scenario leads to evolutionary branching and to the coexistence of modifier alleles coding for haploid and diploid life cycles. This effect is due to the fact that the alternation between a ploidy phase in which selection is weaker (and mutations tend to increase in frequency), and a ploidy phase in which deleterious alleles have stronger effects reduces the geometric mean fitness. Although we did not check this, one can expect that introducing alleles under antagonistic selection between phases increases this cost and accelerates branching.

Assuming additive effects of alleles at the modifier locus (no dominance), heterozygous zygotes carrying one allele coding for haploidy and one allele coding for diploidy suffer from the same alternation cost. We therefore expect the evolution of assortative mating (Otto et al. 2008), and eventually the evolution of complete isolation between a haploid and a diploid new species from an ancestral biphasic species. Exploring the evolution of assortative mating when haploid diploid life cycles are maintained would be a possible next step to this modeling work. In collaboration with Michael Scott, I am currently modifying the simulation programs used to study the evolution of life cycles in order to include an assortment modifier locus. We are using a 'group based model' modified from Otto et al. (2008): depending on its genotype at the modifier locus, each gamete enters a haploid, a diploid or a common group, and mating occurs within each group. In the absence of assortative mating, all individuals enter the common group, while with complete assortment, mating occurs only within the diploid and the haploid group.

To my knowledge, no haploid and diploid sister species with an ancestral biphasic ancestor have been identified. For instance, although the Phaeophyceae present a high diversity of life cycles, the relation between life cycle and phylogeny has been clarified (Cock et al. 2014) and does not provide support for this prediction. However, if future models exploring the evolution of assortative mating in biphasic life cycles predict complete reproductive isolation between phases under a range of realistic parameters, it would be interesting to test this prediction by comparative studies in other groups presenting a wide range of life cycles, such as Rhodophytes or Chlorophytes. Until now, the lack of empirical support for the branching theory may suggest that other factors prevent branching.

One such factor could be the genetic architecture of life cycle variation: if the probability d to undergo selection as a diploid (or the timing of meiosis) was a polygenic trait coded by many loci with small effects, sexual recombination could maintain a distribution of values of d close to the ESS value and thus prevent branching. Furthermore, in Chapter 2, we showed that alleles coding for biphasic life cycles benefit from a bet-hedging advantage when the environment is alternatively favorable to diploids and haploids. While in the polymorphic model, such modifier alleles would suffer from the production of unfit offspring, in the alternation of generation model, one can expect that modifier alleles coding for a timing of meiosis synchronized to the environmental change would benefit from a direct fitness advantage and should fix more easily in the population. For example, in the Coccolithophore *Emiliania huxleyi*, it has been suggested that diploid individuals may undergo meiosis to escape from viral infections, as the virus does not attack haploid cells ('Cheshire cat strategy' Frada et al. 2008).

Another possible ecological factor leading to the maintenance of evolutionary stable haploid diploid life cycles relies on the different function of gametes and haploid spores. Haploids produce gametes that are adapted to fusion, which implies a close proximity, while diploid produce spores, that are adapted for dispersal. The maintenance of two ploidy phases with different morphologies could then evolve in response to these different functions of propagules (Bell 1997). This factor may easily explain the maintenance of biphasic life cycles with a diploid macrothallus and haploid microthallus, but cannot explain the maintenance of isomorphic haploid-diploid cycles (where the sporophyte and gametophyte have the same morphology), as in *Ulvaria* or *Gracilaria*, or in life cycles with a larger haploid phase. Nevertheless, introducing such biological constraints into the models may prevent branching and allow the stable maintenance of biphasic cycles. In addition, as spores and gametes dispersal would affect the level of competition and relatedness between individuals within each niche, it would be interesting to study the effect of population structure in the on the evolution of haploid diploid life cycles.

In addition to providing new insights on the evolution of life cycles, this modeling work allowed me to compare two types of models: the polymorphic model (hereafter denoted P.) considers the evolution of the probability to undergo selection as a diploid or an haploid, while the alternation of generation model (hereafter denoted A.) considers the evolution of the relative length of each phase. From a practical point of view, P. generates simpler equations and it is easier to include ecological components (niche differentiation between phases, density dependence...) into this framework. However, it seems strange to many scientists (during conferences, I was often asked if my results would hold in A.), and except in red algae, I know no organisms that can skip development in one phase. The red algal life cycle consists of the alternation of two diploid phase (one macroscopic and one much smaller, attached to the haploid) and one haploid phase. In *Gracilaria gracilis* (Destombe et al. 1989) and

some other species (Hughes and Otto 1999), some individuals can effectively allow diploids to skip the haploid phase of their life cycle to some degree: haploid spores are retained on the diploid thallus and undergo reduced vegetative stage before reproducing. On the other hand, A. looks more realistic, but it is difficult to tune the level of competition between phases in this type of model, which always considers hard selection. Fortunately, the same results are generally obtained from both models (see Appendix D). However, Hall (2000) found using P. that biphasic life cycles could be maintained when mutations occurs at meiosis, while this result is not obtained when using A. It could be interesting to see if branching occur in this purely genetic model, or if branching requires ecological components. In addition, while mating systems reducing the efficiency of recombination favor haploidy in the P. model (Otto and Marks 1996), Michael Scott (Appendix D) showed that selfing does not provide the same result in A., because in this case, although selfing decreases the efficiency of recombination, the effect of linkage is reduced because selfing generates homozygotes and increases purging. Using P., we found that linkage should favor diploidy when when selection is antagonistic between phases, because in haploids, recombination reduces linkage disequilibrium before selection. It would be interesting to test if the same results holds in A.

More generally, the modeling part of this PhD underlines the possibly important effects of introducing ecological factors into genetic models. In a density dependent model, selection coefficients scale with the degree of competition within niches, and selection may not necessarily be more efficient among haploids. In addition, Chapter 2 showed the importance of niche differentiation between phases on the relative effect of masking and purging. Finally, we found that masking depends on the dominance of mutations affecting competitivity, therefore, on mutational effects on a phenotypic trait, while purging depends on the selection coefficient of mutations, i.e, on their fitness effect. This suggests that it may be interesting to use the approach of Chapter 2 to explore the evolution of other traits, such as reproductive system or dispersion, in which mutations may have both direct and indirect effects.

Exploring ploidy effects with model organisms

Experimental evolution enables to test the assumptions and predictions of models for the evolutionary consequences of ploidy and the evolution of life cycles. The haploid-diploid yeast *Saccharomyces cerevisiae* has already been used to test a number of theoretical results and measure some important quantities: for instance, the differential effects of beneficial mutations in haploids vs. homozygous diploids (e.g. Gerstein 2013), the dominance of deleterious mutations (Agrawal and Whitlock 2011, Manna et al. 2012), the masking diploid advantage (Mable and Otto 2001) or the rate of adaptation in haploids vs. diploids (Zeyl et al. 2003). However, many questions remain unanswered, and during this PhD, I have tried to explore

some of them.

Drift and reproductive isolation in *Saccharomyces cerevisiae*

In Chapter 3, we explored the effect of ploidy on the evolution of reproductive isolation in haploid and diploid yeasts. In allopatric populations, mutations appear and may reach fixation depending on their fitness effect in the genetic background of the population in which they occurred. Hybridization between two isolated populations brings within the same individual several mutations that have never been tested together before and some of these new combinations may be deleterious, leading to a reduced fitness of hybrids (outbreeding depression, Lynch 2012). Despite an important loss of sex in our mutation accumulation lines (performed on mutator strains) and strong selection during the sexual phase, in haploids, we could measure an hybrid depression from less than 300 generations of divergence, and see that hybrid fitness (relative to their parents) tends to decrease as divergence time increases. In diploids, F1 hybrids benefit from strong heterosis, and the F2 generation was still fitter than the homozygous diploids parent. The high fitness of diploid hybrids compared with haploid hybrids may slow down, in the short term, reproductive isolation between divergent populations subject to strong drift. However in a longer term, a diploid sexual hybrid population would be penalized: diploid hybrids carry recessive, partially masked incompatibilities that would be revealed during following generations (Barton 2001). On the contrary, in haploids, all alleles are expressed, and in species with high rate of selfing such as *S. cerevisiae*, a hybrid population derived from a particularly fit haploid F1 hybrid would settle easily. Overall, our experiment outlines the possible benefits of hybridization between heavily loaded populations, through the production of some fit recombinants (in haploids) and through masking of deleterious alleles (heterosis in diploids). However, due to the quick loss of sexual capacity, we could not achieve the initial goal of the experiment, which was to explore the dynamics of reproductive isolation over time in haploids and diploids drifting around a fitness optimum.

Although carrying mutation accumulation lines is much easier in asexual than sexual strains of yeasts, it would thus be interesting to modify this experimental setting by regularly selecting for the capacity to undergo sex. In addition, many reproductive incompatibilities involve sexual traits, and few studies have used experimental evolution to compare the effect of reproductive isolation on sexual and asexual fitness (Dettman et al. 2007, Catillo et al. 2015). Moreover, although our experiment used two bottlenecks regimes, reproductive isolation could be studied in one treatment only, partly because strains under the narrower bottleneck (bottlenecks of 1 colony) treatment did not reach a fitness equilibrium. Using strains with lower mutation rate and larger bottlenecks would allow one to test theoretical predictions concerning the effects of these parameters (e.g. Chevin et al. 2014).

Dominance of mutations fixed during adaptation in haploids and diploids

Finally, in a ongoing project, we have started to measure the dominance and fitness effect (in haploids and diploids) of beneficial alleles that have fixed in experimental haploid and diploid populations (Zeyl et al. 2003). In populations adapting to a new environment, the fixation probability of an adaptative mutation increases with its fitness effect. In diploid asexual populations, mutations remain heterozygous (in the absence of loss-of-heterozygosity, see Gerstein et al. 2014) and selection should favor the fixation of dominant, even overdominant adaptative mutations (Haldane's sieve, 1937). On the contrary, in haploid evolving populations, dominance does not affect the fixation probability of a mutation and we do not expect a particular pattern of dominance for these mutations. We thus expect to find higher dominance in adaptative mutations fixed in diploid than in haploid populations.

We are using yeast strains from a previous evolution experiment of adaptation of 5 haploid and 5 diploid asexual populations to minimal medium (Zeyl et al. 2003). In haploids, 3 to 5 mutations were fixed after 2000 generations (Zeyl 2005). We first planned to isolate each adaptative mutation by performing backcrosses with the ancestor, and measure the fitness effect of each mutation in haploids, homozygous and heterozygous diploids. However, measuring the effect of each single mutation in the ancestral background may not reflect its effect in the genetic background where the mutation appeared (and given the current composition of the population). Epistatic interactions between mutations can render the estimation of the effect of a mutation occurring in a particular background difficult to estimate: for this, one would need to know the fixation order of all mutations.

Instead of trying to measure individual effects of mutations, we thus decided to obtain a general measure of the effect of mutations fixed in diploid (heterozygous) populations on haploids and on homozygous diploids (and also of the effect of mutations fixed in haploid populations on both types of diploids) by performing different types of crosses (to produce haploid, homozygous and heterozygous diploid offspring from evolved haploids and diploids). I also performed simulations of adaptive walks using Fisher's geometric model to obtain numerical predictions for fitness distributions in these different types of offspring. For this, I simulated adaptation in haploids and diploids, and performed crosses at different distance to the optimum to produce haploid and diploid offspring. Preliminary simulation results showed that the dominance of adaptative mutations fixed during diploid adaptation increased with the distance to the optimum, with more overdominant haplotypes being present as the population approaches the optimum.

Model organisms remain constrained by their evolutionary history

In addition to the experiments described above, we tried unsuccessfully to explore other consequences of ploidy, using first the recent model for brown algae, *Ectocarpus siliculosus*, then *Saccharomyces cerevisiae*.

Although multicellular organisms present a large variety of life cycles, they have never been used for evolution experiments on the effects of ploidy. In a first project, we planned to use the brown algal model, *Ectocarpus siliculosus*, to test the masking and purging theory: we wanted to follow the fitness of isolated haploid and diploid large populations, and perform competition between haploids and diploids faced to different mutation rates, induced by U.V. light. However, there is no standard fitness assay for *Ectocarpus*, and more importantly, the maintenance of a large population in the lab for a long time generated endodiploidization of the haploids and aneuploidy, and we thus renounced to this model organism and turned to yeast.

One of our initial goals with yeast was also to test theoretical predictions on the equilibrium mutation load in haploid and diploid populations, using mutator strains (obtained by deleting two DNA repair genes). For this, we maintained large haploid and diploid populations during a large number of generations,but realized that although mating type switching had been prevented in our strains, all haploid populations had become diploid before the 200th generation of adaptation. In addition, evolved diploid mutator strains could not produce viable spores. We therefore decided to abandon this project.

Finally, the many problems encountered during the experimental part of this PhD show that although external conditions are narrowly controlled in experimental evolution, model organisms have specificities proper to their clade. *Saccharomyces cerevisiae* is a powerful system to explore the effects of ploidy, but wild type strains are predominantly diploid. Convergence toward diploidy from haploid and tetraploid asexual strains have been observed in previous experiments (Gerstein et al. 2006) and in the mutator strains mentioned above, revealing constraints acting on the ploidy level, selection favouring the ploidy level typical of yeast's evolutionary past. Contrary to *S. cerevisiae*, other yeast species, such as *Saccharomyces rouxii* or *Hansenula wingei* display equivalent vegetative growth and rapid alternation of haploid and diploid phases (Mable and Otto 1998). Until now, evolution experiments on yeast have explored the consequences of ploidy level, but could not let the life cycle itself (proportion of time spent in the haploid vs. diploid phase) evolve. In yeast species with more symmetric life cycles, it is possible that the life cycle would be less constrained, and that such species could be used to test theoretical scenarios on the evolution of the sexual cycle.

References

Adams, J., and P. Hansche. 1974. Population studies in microorganisms. 1. evolution of diploidy in *Saccharomyces cerevisiae*. Genetics 76:327338.

Agrawal, A. F., and M. C. Whitlock. 2010. Environmental duress and epistasis: how does stress affect the strength of selection on new mutations? Trends in Ecology & Evolution 25:450–458.

———. 2011. Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 187:553–U274.

Barton, N. H. 2001. The role of hybridization in evolution. Molecular ecology 10:551–568.

———. 2013. Does hybridization influence speciation? Journal of evolutionary biology 26:267–269.

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Journal of evolutionary biology 127:229–255.

Bateson, N. 1909. Heredity and variation in modern lights. Cambridge University Press, Cambridge, UK.

Bell, G. 1982. The masterpiece of nature: The evolution of genetics of sexuality. University of Califonia Press, California.

———. 1989. Sex and death in Protozoa. Cambridge University Press, Cambridge.

———. 1994. The comparative biology of the alternation of generations. Lectures on Mathematics in the Life Sciences 25:1–26.

———. 1997. The evolution of the life cycle of brown seaweeds. Biological Journal of the Linnean Society 60:21–38.

Bessho, K., Y. Iwasa, and T. Day. 2015. The evolutionary advantage of haploid versus diploid microbes in nutrient-poor environments. Journal of Theoretical Biology 383:116–129.

Bhattacharya, D. 1985. The demography of fronds of *Chondrus crispus* Stackhouse. Journal of experimental marine biology and ecology 91:217–231.

Billiard, S., M. López-Villavicencio, M. Hood, and T. Giraud. 2012. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. Journal of Evolutionary Biology 25:10201038.

Bolton, J. J., and M. A. P. Joska. 1993. Populations studies on a south-african carrageenophyte - *Iridaea capensis* (gigartinaceae, rhodophyta). Hydrobiologia 261:191–195.

Bordenstein, S. R., and M. D. Drapeau. 2001. Genotype-environment interaction and the dobzhansky-muller model of postzygotic isolation. Journal of evolutionary biology 11:490–501.

Burton, R. S. 1990. Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. Evolution 44:18141822.

Catillo, D., M. Burger, C. Lively, and L. Delph. 2015. Experimental evolution: Assortative mating and sexual selectino, independent of local adaptation, lead to reproductive isolatin in the nematode *Caenorhabditis remanei*. Evolution 69:3141–3155.

Cavalier Smith, T. 1985. The evolution of genome size. Wiley, Chichester, New Hampshire.

Charlesworth, A., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Annual Review of Ecology, Evolution, and Systematics 18:237–268.

Charlesworth, B., and D. Charlesworth. 1999. The genetic basis of inbreeding depression. Genetic Research 74:329–340.

Cherry, J., E. Hong, C. Amundsen, R. Balakrishnan, G. Binkley, E. Chan, K. Christie, M. Costanzo, S. Dwight, S. Engel, D. Fisk, J. Hirschman, B. Hitz, K. Karra, C. Krieger, S. Miyasato, R. Nash, J. Park, M. Skrzypek, M. Simison, S. Weng, and E. Wong. 2012. *Saccharomyces* genome database: the genomics resource of budding yeast. Nucleic acid research 40:700–705.

Chevin, L. M. 2011. On measuring selection in experimental evolution. Biology Letters 7:210213.

Chevin, L. M., G. Decorzent, and T. Lenormand. 2014. Niche dimensionality and the genetics of ecological speciation. Evolution 68:1244–1256.

Clayton, M. N. 1992. Propagules of marine macroalgae : structure and development. British Phycological Journal 27:219–232.

Cock, J. M., O. Godfroy, N. Macaisne, A. F. Peters, and S. M. Coelho. 2014. Complex life cycles of multicellular eukaryotes: New approaches based on the use of model organisms. Current Opinion in Plant Biology 17:1–6.

Coelho, S. M., A. F. Peters, B. Charrier, D. Roze, C. Destombe, M. Valero, and J. M. Cock. 2007. Complex life cycles of multicellular eukaryotes: New approaches based on the use of model organisms. Gene 406:152–170.

Couceiro, L., M. Le Gac, H. Hunsperger, S. Mauger, C. Destombe, M. Cock, S. Ahmed, S. Coelho, M. Valero, and A. F. Peters. 2015. Evolution and maintenance of haploid-diploid life cycles in natural populations: The case of the marine brown alga *Ectocarpus*. Evolution 69:1808–1822.

Coyne, J., and A. Orr. 1998. The evolutionary genetics of speciation. Philosophical Transactions of the Royal Society B 253:287–305.

———. 2004. Speciation. Sinauer Associates Incorporation, Sunderland, MA, USA.

Crow, J. F., and M. Kimura. 1965. Evolution in sexual and asexual populations - a reply. American Naturalist 103:89–&.

Delobel, P., and C. Tesniere. 2014. A simple fcm method to avoid misinterpretation in *Saccharomyces cerevisiae* cell cycle assessment between g0 and sub-g1. Evolution 9:e84645.

Destombe, C., J. Godin, C. Lefebvre, O. Dehorter, and P. Vernet. 1992. Differences in dispersal abilities of haploid and diploid spores of *Gracilaria verrucosa* (gracilariales, rhodophyta). Botanica Marina 35:93–98.

Destombe, C., J. Godin, M. Nocher, S. Richerd, and M. Valero. 1993. Differences in response between haploid and diploid isomorphic phases of *Gracilaria verrucosa* (rhodophyta, gigartinales) exposed to artificial environmental conditions. Hydrobiologia 261:131–137.

Destombe, C., M. Valero, P. Vernet, and D. Couvet. 1989. What controls the haploid-diploid ratio in the red alga, *Gracilaria verrucosa*? Journal of Evolutionary Biology 2:317–338.

Dettman, J., C. Sirjusingh, L. Kohn, and J. Anderson. 2007. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. Nature 447:585588.

Dobzhansky, T. 1936. Studies on hybrid sterility. ii. localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics 21:113–135.

Dyck, L. J., and R. E. De Wreede. 2006. Seasonal and spatial patterns of population density in the marine macroalga *Mazzaella splendens* (gigartinales, rhodophyta). Phycological Research 54:21–31.

Edmans, S. 1999. Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. Evolution 53:17571768.

———. 2007. Between a rock and a hard place: evaluating the relative risk of inbreeding and outbreeding for conservaiton and management. Molecular Ecology 16:463–475.

Engel, C., P. Åberg, O. E. Gaggiotti, and C. Destombe. 2001. Population dynamics and stage structure in a haploid-diploid red seaweed, *Gracilaria gracilis*. Journal of Ecology 89:436–450.

Enyenihi, A. H., and W. S. Saunders. 2003. Large-scale functional genomic analysis of sporulation and meiosis in *Saccharomyces cerevisiae*. Genetics 163:4754.

Ewing, E. P. 1977. Selection at haploid and diploid phases - cyclical variation. Genetics 87:195–207.

Frada, M., I. Probert, M. Allen, W. Wilson, and C. de Vargas. 2008. The 'cheshire cat' escape strategy of the coccolithophore *Emiliania huxleyi* in response to viral infection. Proceedings of the National Academy of Sciences 105:1944–1949.

Fraïsse, C., P. Gunnarsson, D. Roze, N. Bierne, and J. Welch. 2016. The genetics of speciation: Insights from fisher's geometric model. Evolution .

Gerstein, A. C. 2013. Mutational effects depend on ploidy level: all else is not equal. Biology Letters 9.

Gerstein, A. C., H. J. E. Chun, A. Grant, and S. P. Otto. 2006. Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. Plos Genetics 2:1396–1401.

Gerstein, A. C., A. Kuzmin, and S. P. Otto. 2014. Loss-of-heterozygosity facilitates passage through haldane's sieve for *Saccharomyces cerevisiae* undergoing adaptation. Nature communications 5:3819.

Gourbiere, S., and J. Mallet. 2010. Are species real? the shape of the species boundary with exponential failure, reinforcement, and the 'missing snowball. Evolution 64:1–24.

Haig, D., and A. Wilczek. 2006. Sexual conflict and the alternation of haploid and diploid generations. Philosophical Transactions of the Royal Society B: Biological Sciences 361:335–343.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in animal hybrids. Journal of Genetics 12:101109.

———. 1937. The effect of variation on fitness. American Naturalist 71:337–349.

Hall, D. W. 2000. The evolution of haploid, diploid and polymorphic haploid-diploid life cycles: The role of meiotic mutation. Genetics 156:893–898.

Halligan, D. L., and P. D. Keightley. 2009. Spontaneous mutation accumulation studies in evolutionary genetics, vol. 40. Annual Reviews, Palo Alto.

Hill, J. A., and S. P. Otto. 2007. The role of pleiotropy in the maintenance of sex in yeast. Genetics 175:1419–1427.

Hofmeister, W. 1851. Vergleichende untersuchungen der keimung, entfaltung und fruchtbildung höherer kryptogamen: (Moose, Farrn, Equisetaceen, Rhizocarpeen und Lycopodiaceen) und der samenbildung der coniferen. Leipzig.

Hough, J., S. Immler, S. C. H. Barrett, and S. P. Otto. 2013. Evolutionarily stable sex ratios and mutation load. Evolution 67:1915–1925.

Hughes, J. S., and S. P. Otto. 1999. Ecology and the evolution of biphasic life cycles. American Naturalist 154:306–320.

Immler, S., G. Arnqvist, and S. P. Otto. 2012. Ploidally antagonistic selection maintains stable genetic polymorphism. Evolution 66:55–65.

Immler, S., and S. P. Otto. 2014. Driven apart: the evolution of ploidy differences between the sexes under antagonistic selection. American Naturalist 183:96–107.

Jenkins, C. D. 1993. Selection and the evolution of genetic life-cycles. Genetics 133:401–410.

Jenkins, C. D., and M. Kirkpatrick. 1994. Deleterious mutation and ecological selection in the evolution of life cycles. Lect Math Life Sci 25:53–68.

———. 1995. Deleterious mutation and the evolution of genetic life-cycles. Evolution 49:512–520.

Johansen-Morris, A. D., and R. G. Latta. 2006. Fitness consequences of hybridization between ecotypes of *Avena barbata*: hybrid breakdown, hybrid vigor, and transgressive segregation. Evolution 60:15851595.

Joseph, S. B., and M. Kirkpatrick. 2004. Haploid selection in animals. Trends in Ecology & Evolution 19:592–597.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

Klinger, T. 1993. The persistence of haplodiploidy in algae. Trends in Ecology & Evolution 8:256258.

Kondrashov, A., S. Sunyaev, and F. Kondrashov. 2002. Dobzhansky-muller incompatibilities in protein evolution. Proceedings of the National Academy of Sciences of the United States of America 99:1487814883.

Kondrashov, A. S., and J. F. Crow. 1991. Haploidy or diploidy - which is better. Nature 351:314–315.

Korona, R. 1999*a*. Genetic load of the yeast *Saccharomyces cerevisiae* under diverse environmental conditions. Evolution 53:1966–1971.

———. 1999*b*. Unpredictable fitness transitions between haploid and diploid strains of the genetically loaded yeast *Saccharomyces cerevisiae*. Genetics 151:77–85.

Koschwanez, J. H., K. R. Foster, and A. W. Murray. 2011. Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. Plos Biology 9:e1001122.

Lewis, W. M. 1979. Diploidy, evolution and sex. Journal of theoretical biology 78:425–438.

———. 1985. Nutrient scarcity as an evolutionary cause of haploidy. American Naturalist 125:692–701.

Loewe, L., and B. Charlesworth. 2006. Inferring the distribution of mutational effects on fitness in drosophila. Biology Letters 2:426–430.

Lynch, M. 2012. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution 45:622–629.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. 1st ed. Sinauer Associates.

Mable, B. K. 2001. Ploidy evolution in the yeast *Saccharomyces cerevisiae*: a test of the nutrient limitation hypothesis. Journal of Evolutionary Biology 14:157–170.

Mable, B. K., and S. P. Otto. 1998. The evolution of life cycles with haploid and diploid phases. Bioessays 20:453–462.

———. 2001. Masking and purging mutations following ems treatment in haploid, diploid and tetraploid yeast (*Saccharomyces cerevisiae*). Genetical Research 77:9–26.

Manna, F., R. GALLET, G. Martin, and T. Lenormand. 2012. The high-throughput yeast deletion fitness data and the theories of dominance. Journal of Evolutionnary Biology 25:892903.

Manna, F., G. Martin, and T. Lenormand. 2011. Fitness landscapes: An alternative theory for the dominance of mutation. Genetics 189:923–U303.

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

Matute, D. R., I. A. Butler, D. A. Turissini, and J. Coyne. 2010. A test of the snowball theory for the rate of evolution of hybrid incompatibilities. Science 329:15181521.

Michod, R. E., and M. F. Wojciechowski. 1994. Dna-repair and the evolution of transformation. 4. dna-damage increases transformation. Journal of Evolutionary Biology 7:147–175.

Moyle, L., and T. Nakazato. 2010. Hybrid incompatibility 'snowballs between *Solanum* species. Science 329:15211523.

Muller, H. 1939. Reversibility in evolution considered from the standpoint of genetics. Biological Reviews of the Cambridge Philosophical Society 14:261–280.

———. 1940. Bearing of the *Drosophila* work on systematics. Clarendon Press.

———. 1942. Isolating mechanisms, evolution and temperature. Biology Symposium 6:71–125.

Niklas, K. J., and U. Kutschera. 2010. The evolution of the land plant life cycle. The New Phytologist 185:27–41.

Nuismer, S. L., and S. P. Otto. 2004. Host-parasite interactions and the evolution of ploidy. Proceedings of the National Academy of Sciences of the United States of America 101:11036–11039.

Ogur, M., and R. StJohn. 1956. A differential and diagnostic plating method for population studies of respiration deficiency in yeast. Journal of bacteriology 72:500–504.

Orr, H. A. 1995*a*. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:18051813.

———. 1995*b*. Somatic mutation favor the evolution of diploidy. Genetics 139:1441–1447.

———. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. Evolution 52:935–949.

Orr, H. A., and S. P. Otto. 1994. Does diploidy increase the rate of adaptation. Genetics 136:1475–1480.

Otaiza, R. D., S. R. Abades, and A. J. Brante. 2001. Seasonal changes in abundance and shifts in dominance of life history stages of the carrageenophyte *Sarcothalia crispata* (rhodophyta, gigartinales) in south central chile. Journal of Applied Phycology 13:161–171.

Otto, S. P. 1994. The role of deleterious and beneficial mutaitons in the evolution of ploidy levels. Lectures on Mathematics in the Life Sciences 25:69–98.

———. 2007. The evolutionary consequences of polyploidy. Cell 131:452–462.

Otto, S. P., and D. Bourguet. 1999. Balanced polymorphisms and the evolution of dominance. The American Naturalist 153:561574.

Otto, S. P., and A. C. Gerstein. 2008. The evolution of haploidy and diploidy. Current Biology 18:R1121–R1124.

Otto, S. P., and D. B. Goldstein. 1992. Recombination and the evolution of diploidy. Genetics 131:745–751.

Otto, S. P., and J. C. Marks. 1996. Mating systems and the evolutionary transition between haploidy and diploidy. Biological Journal of the Linnean Society 57:197–218.

Otto, S. P., M. R. Servedio, and S. L. Nuismer. 2008. Frequency-dependent selection and the evolution of assortative mating. Genetics 179:2091–2112.

Pacheco-Ruíz, I., A. Cabello-Pasini, J. Zertuche-González, S. Murray, J. Espinoza-Avalos, , and M. Dreyfus-Leon. 2011. Carpospore and tetraspore release and survival in *Chondracanthus squarrulosus* (rhodophyta: Gigartinaceae) from the gulf of california. Botanica Marina 54:127–134.

Paquin, C., and J. Adams. 1983. Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. Nature 302:495–500.

Perrot, V., S. Richerd, and M. Valero. 1991. Transition from haploidy to diploidy. Nature 351:315–317.

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

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

Presgraves, D. C. 2010. Speciation genetics: Search for themissing snowball. Current biology 20:R1073R1074.

Price, G. R. 1970. Selection and covariance. Nature 227:520–&.

Qiu, Y. L., A. B. Taylor, and M. H. A. 2012. Evolution of the life cycle in land plants. Journal of Systematics and Evolution 50:171194.

Rescan, M., T. Lenormand, and D. Roze. 2016. Interaction between genetic and ecological effects on the evolution of life cycles. American Naturalist 187:19–34.

Richerd, S., D. Couvet, and M. Valero. 1993. Evolution of the alternation of haploid and diploid phases in life-cycles .2. maintenance of the haplo-diplontic cycle. Journal of Evolutionary Biology 6:263–280.

Rogers, S. M., and L. Bernatchez. 2006. The genetic basis of intrinsic and extrinsic post-zygotic reproductive isolation jointly promoting speciation in the lake whitefish species complex (*Coregonus clupeaformis*). Journal of evolutionary biology 19:19791994.

Rokitta, S. D., L. J. de Nooijer, S. Trimborn, C. de Vargas, B. Rost, and U. John. 2011. Transcriptome analyses reveal differential gene expression patterns between the life-cycle stages of *Emiliania huxleyi* (haptophyta) and reflect specialization to different ecological niches. Journal of Phycology 47:829–838.

Roze, D. 2009. Diploidy, population structure, and the evolution of recombination. American Naturalist 174:S79–S94.

Roze, D., and R. E. Michod. 2010. Deleterious mutations and selection for sex in finite diploid populations. Genetics 184:1095–U396.

Roze, D., and S. P. Otto. 2012. Differential selection between the sexes and selection for sex. Evolution 66:558–574.

Russell, L. B., and W. L. Russell. 1996. Spontaneous mutations recovered as mosaics in the mouse specific-locus test. Proceedings of the National Academy of Sciences 93:1307213077.

Santos, R., and P. Duarte. 1996. Fecundity, spore recruitment and size in *Gelidium sesquipedale* (gelidiales, rhodophyta). Hydrobiologia 327:223–228.

Schoen, D. J. 2005. Deleterious mutation in related species of the plant genus *Amsinckia* with contrasting mating systems. Evolution 59:2370–2377.

Sella, G. 2009. An exact steady state solution of fisher's geometric model and other models. Theoretical Population Biology 75:30–34.

Sella, G., and E. Hirsh. 2005. The application of statistical physics to evolutionary biology. Proceedings of the National Academy of Sciences of the USA 102:95419546.

Sherman, F., and B. Ephrussi. 1962. Relationship between respiratory deficiency and suppressiveness in yeast as determined with segregational mutants. Genetics 47:695–&.

Silander, O. K., O. Tenaillon, and L. Chao. 2007. Understanding the evolutionary fate of finite populations: The dynamics of mutational effects. Plos Biology 5:922–931.

Simmons, M. J., and J. F. Crow. 1977. Mutations affecting fitness in *Drosophila* populations. Annual Review of Genetics 11:49–78.

Slatkin, M., and R. Lande. 1994. Segregation variance after hybridization of isolated populations. Genetics Research 64:5156.

Strasburger, E. 1894. The periodic reduction of the number of chromosomes in the life history of living organisms. Annals of Botany 8:281–316.

Szafraniec, K., D. M. Wloch, P. Sliwa, R. H. Borts, and R. Korona. 2003. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. Genetical Research 82:19–31.

Tenaillon, O. 2014. The utility of fishers geometric model in evolutionary genetics. Annual Review of Ecology, Evolution, and Systematics 45:179–201.

Thornber, C. S. 2006. Functional properties of the isomorphic biphasic algal life cycle. Integrative and Comparative Biology 46:605–614.

Thornber, C. S., and S. D. Gaines. 2004. Population demographics in species with biphasic life cycles. Ecology 85:1661–1674.

Turelli, M., N. Barton, and J. Coyne. 2001. Theory and speciation. Trends in Ecology & Evolution 16:330–343.

Valero, M., S. Richerd, V. Perrot, and C. Destombe. 1992. Evolution of alternation of haploid and diploid phases in life-cycles. Trends in Ecology & Evolution 7:25–29.

Von Dassow, P., H. Ogata, I. Probert, P. Wincker, C. Da Silva, S. Audic, J. M. Claverie, and C. de Vargas. 2009. Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. Genome Biology 10.

Wallace, B. 1975. Hard and soft selection revisited. Evolution 29:465–473.

Weiss, R., J. Kukora, and J. Adams. 1975. The relationship between enzyme activity, cell geometry, and fitness in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America 72:794798.

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

Willson, M. F. 1981. On the evolution of complex life-cycles in plants - a review and an ecological perspective. Annals of the Missouri Botanical Garden 68:275–300.

Wolfram Research Inc. 2010. Mathematica. Version 8.0 ed. Wolfram Research, Inc., Champaign, Illinois.

Wright, S. 1968. Evolution and the genetics of populations, volume 1: genetic and biometric foundations. University of Chicago Press, Chicago.

Zeyl, C. 2004. Experimental studies of ploidy evolution in yeast. Fems Microbiology Letters 233:187–192.

———. 2005. The number of mutations selected during adaptation in a laboratory population of *Saccharomyces cerevisiae*. Genetics 169:1825–1831.

Zeyl, C., C. Curtin, K. Karnap, and E. Beauchamp. 2005. Antagonism between sexual and natural selection in experimental populations of *Saccharomyces cerevisiae*. Evolution 59:2109–2115.

Zeyl, C., and J. DeVisser. 2001. Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*e. Genetics 157:53–61.

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

Zeyl, C., T. Vanderford, and M. Carter. 2003. An evolutionary advantage of haploidy in large yeast populations. Science 299:555–558.

Zörgö, E., K. Chwialkowska, A. B. Gjuvsland, E. Garré, P. Sunnerhagen, G. Liti, A. Blomberg, S. W. Omholt, and J. Warringer. 2013. Ancient evolutionary trade-offs between yeast ploidy states. PLOS Genetics 9:e1003388.

Appendix A

QLE analysis of the two-locus model

To compute recursions on allele frequency and linkage disequilibrium, we will use the following variables: X_a is an indicator variable that equals 1 if allele a is present in a given haploid at the selected locus, and 0 otherwise, while $X_{a,1}$ and $X_{a,2}$ are similar indicator variables for the two genes at the selected locus in a given diploid. Similarly, variables $X_{m,1}$ and $X_{m,2}$ equal 1 if allele m is present at the modifier locus in a haploid, or in each haplotype of a diploid. Genetic associations between genes present on different haplotypes of a diploid zygote can be defined as follows: defining $\zeta_{a,1} = X_{a,1} - p_a$ and $\zeta_{a,2} = X_{a,2} - p_a$, and similar variables $\zeta_{m,1}$, $\zeta_{m,2}$ for the modifier locus, the genetic association between genes in sets U and V, present respectively in the first and second haplotype of a diploid individual is defined as (e.g. Kirkpatrick et al. 2002):

$$D_{U,V} = E\left[\prod_{i\in U}\zeta_{i,1}\prod_{i\in V}\zeta_{i,2}\right]$$

where E stands for the average over all zygotes (since our model does not include any asymmetry between sexes, we have $D_{U,V} = D_{V,U}$ for all U, V). Because we assume random mating, associations $D_{m,m}$, $D_{a,a}$ and $D_{ma,a}$ equal zero, while $D_{ma,ma} = D_{ma}^2$, which will be negligible under the QLE assumption.

Different superscripts will be used to denote variables measured in haploids and diploids, at different stages of the life cycle: p_a^d, p_m^d and $D_{U,V}^d$ denote the frequencies of a, m and the genetic association $D_{U,V}$ before selection, within the subset of individuals that will undergo selection as diploids, while $p_a^{d'}$, $p_m^{d'}$ and $D_{U,V}^{d'}$ denote the same variables measured after selection. Similarly, p_a^h, p_m^h and $D_{U,V}^h$ represent allele frequencies and LD among haploids, before selection, while $p_a^{h'}$, $p_m^{h'}$ and $D_{U,V}^{h'}$ represent the same variables measured after selection.

Change in frequency at the selected locus

We first derive expressions for equilibrium allele frequencies at the selected loci under the different scenarios considered (deleterious allele, antagonistic selection), and when allele M is fixed at the modifier locus. In this case, the frequency of allele a is the same among haploids and diploids before selection: $p_a^d = p_a^h = p_a$. To compute changes in frequencies within the haploid and diploid sub-populations ($\Delta p_a^d = p_a^{d'} - p_a$, $\Delta p_a^h = p_a^{h'} - p_a$), it is useful to express fitnesses of individuals in terms of their ζ variables: $\zeta_a = X_a - p_a$ for haploids, $\zeta_a = X_a - p_a$ and $\zeta_{a,2} = X_{a,2} - p_a$ for diploids (e.g. Kirkpatrick et al. 2002):

$$\begin{aligned} W_h &= \overline{W_h}\left[1 + b_a^h \zeta_a\right] \\ W_d &= \overline{W_d}\left[1 + b_a^d(\zeta_{a,1} + \zeta_{a,2}) + b_{a,a}^d(\zeta_{a,1}\zeta_{a,2})\right] \end{aligned} \tag{A.1}$$

where $\overline{W_h}$, $\overline{W_d}$ represent the mean fitnesses of haploids and diploids, b_a^h and b_a^d measure the effect of selection for (or against) allele a in haploids and diploids (respectively) while $b_{a,a}^d$ is a measure of dominance. To the first order in s, these coefficients are given by:

$$b_a^h = -s\alpha \ , \qquad b_a^d = -s\left[h + (1-2h)p_a\right] \ , \qquad b_{a,a}^d = -s(1-2h)$$

Note that when allele a is at low frequency, $b_a^d \approx -sh$. The effect of selection on allele frequencies within diploids and haploids is then given by:

$$\Delta p_a^d = \mathrm{E}_d\left[\overline{W_d}\bar{W}_d \frac{\zeta_{a,1} + \zeta_{a,2}}{2}\right] \ , \qquad \Delta p_a^h = \mathrm{E}_d\left[\frac{W_h}{\overline{W_h}}\zeta_a\right] \tag{A.2}$$

where E_d and E_h stand for the average over all diploid and haploid individuals. Using the fact that $E\left[\zeta_{a,1}^2\right] = E\left[\zeta_{a,2}^2\right] = E\left[\zeta_a^2\right] = p_a q_a$ (where $q_a = 1 - p_a$) and that $E\left[\zeta_{a,1}\zeta_{a,2}\right] = D_{a,a} = 0$, one obtains:

$$\Delta p_a^d = b_a^d p_a q_a \ , \qquad \Delta p_a^h = b_a^h p_a$$

Note that the same result holds when selection acts at multiple loci, as long as selection is multiplicative among loci and that we can neglect linkage disequilibria. Finally, the frequency of a in the whole population after selection is given by:

$$p_a^{'} = (1-\gamma)\left[d p_a^{d'} + (1-d)p_a^{h'} + \gamma \frac{d\overline{W_d}p_a^{d'} + (1-d)\overline{W_h}p_a^{h'}}{d\overline{W_d} + (1-d)\overline{W_h}}\right] \tag{A.3}$$

which yields:

$$\Delta p_a = p'_a - p_a = (1-\gamma)\left[d b_a^d + (1-d) b_a^h\right] p_a q_a + \gamma \frac{d\overline{W_d} b_a^d + (1-d)\overline{W_h} b_a^h}{d\overline{W_d} + (1-d)\overline{W_h}} p_a q_a \qquad \text{(A.4)}$$

which simplify to 1.1 in the main text when p_a is deleterious in both phases so that terms in p_a^2 can be neglected.

Change in frequency at the modifier locus

To derive recursions in terms of genetic associations, it is helpful to express the rate of diploidy of a focal individual (denoted hereafter d^*) in terms of its $\zeta_{m,1}$, $\zeta_{m,2}$ variables:

$$d^* = \bar{d} + d_m\left(\zeta_{m,1} + \zeta_{m,2}\right) + d_{m,m}\zeta_{m,1}\zeta_{m,2}$$

where $\bar{d}$ is the average rate of diploidy in the population, $d_m = [h_m(1-p_m) + (1-h_m)p_m]$ and $d_{m,m} = \delta(1-2h)$. Before selection, the frequencies of m within the subpopulation that remains diploid and the subpopulation that has switched to haploidy are given by:

$$p_m^d = p_m + \mathrm{E}\left[\frac{d^*}{\bar{d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right], \qquad p_m^h = p_m + \mathrm{E}\left[\frac{1-d^*}{1-\bar{d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right]$$

where E stands for the average over all zygotes (at the start of the generation). After simplification (and using the fact that $D_{m,m} = 0$), we have:

$$p_m^d = p_m + \frac{d_m}{\bar{d}} p_m q_m\ , \qquad p_m^h = p_m - \frac{d_m}{1-\bar{d}} p_m q_m$$

Using a similar argument, one obtains:

$$p_a^d = p_a + \frac{d_m}{\bar{d}} D_{ma}\ , \qquad p_a^h = p_m - \frac{d_m}{1-\bar{d}} D_{ma} \qquad \text{(A.5)}$$

However, because D_{ma} is of order δ (since it should be zero when m is neutral), at QLE $d_m D_{ma}$ is of order δ^2 and can be neglected; therefore we have $p_a^d \approx p_a^h \approx p_a$. Similarly, one can show that $D_{ma}^d \approx D_{ma}$, $D_{ma}^h \approx (1-r_{ma})D_{ma}$ (the neglected terms being of order δ^2, see below). The changes in frequency of m within the diploid and haploid sub-populations are given by equation A.2, replacing 'a' by 'm' subscripts; after simplification, this gives:

$$\Delta p_m^d = b_a^d D_{ma}\ , \qquad \Delta p_m^h = b_a^d (1-r_{ma}) D_{ma}$$

Finally, the frequency of m in the whole population after selection is given again by equation A.3, replacing 'a' by 'm' subscripts, which yields equation 1.7 in Chapter 1.

Linkage disequilibrium at QLE

In the following we compute D_{ma} at quasi-linkage equilibrium, to the first order in s and to the first order in δ (one can show easily that $D_{ma} = 0$ at QLE if either $s = 0$ or $\delta = 0$). The linkage disequilibrium within individuals that remain diploid, before selection is given by:

$$D^d_{ma} = \mathrm{E}\left[\frac{d^*}{\bar{d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right] - (p^d_m - p_m)(p^d_a - p_a) \tag{A.6}$$

(where E means the average over all zygotes). The term $(p^d_m - p_m)(p^d_a - p_a)$ comes from the fact that allele frequencies may change between the population of zygotes and the population of individuals that remain diploid; however, because $p^d_a - p_a$ is of order δ (see equation A.5), this term can be neglected under our weak modifier assumption. Using the fact that under random mating $D_{ma,m}$ equals zero, equation A.6 yields:

$$D^d_{ma} = D_{ma} + \frac{d_m}{\bar{d}}(1-2p_m)D_{ma}$$

However, since D_{ma} and d_m are both of order δ, we have (to the first order in δ) $D^d_{ma} \approx D_{ma}$, while $D^h_{ma} \approx (1-r_{ma})D_{ma}$. After selection, we have:

$$\begin{aligned} D^{d'}_{ma} &= \mathrm{E}\left[\frac{W_d}{\overline{W_d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right] - (p^{d'}_m - p^d_m)(p^{d'}_a - p^d_a) \\ &= D^d_{ma} + b^d_a(1-2p_a)D_{ma} - (b^d_a D_{ma})(b^d_a p_a q_a) \end{aligned}$$

However, because both D_{ma} and b^d_a are of order s, we have to the first order in s $D^{d'}_{ma} \approx D_{ma}$. Similarly, one obtains $D^{h'}_{ma} \approx (1-r_{ma})D_{ma}$. Finally, after recombination the linkage disequilibrium within gametes produced by individuals that underwent selection as diploids is $D^{d''}_{ma} \approx (1-r_{ma})D_{ma}$. Therefore, to the first order in δ and s the linkage disequilibrium within gametes produced by diploids ($D^{d''}_{ma}$) and within gametes produced by haploids ($D^{h''}_{ma}$) both equal $(1-r_{ma})D_{ma}$. However, these linkage disequilibria are defined as:

$$\begin{aligned} D^{h'}_{ma} &= \mathrm{E}_h\left[\left(X_m - p^{h'}_m\right)\left(X_a - p^{h'}_a\right)\right] \\ D^{d'}_{ma} &= \mathrm{E}_d\left[\left(X_m - p^{d'}_m\right)\left(X_a - p^{d'}_a\right)\right] \end{aligned}$$

(where E_h and E_d are averages over all haploids after selection, and over all gametes produced

by diploids), while the linkage disequilibrium among all gametes is given by:

$$D_{ma}^{'} = (1-\gamma)\left[\bar{d}D_{ma}^{h*} + (1-\bar{d})D_{ma}^{d*} + \gamma\frac{\overline{dW_d}D_{ma}^{h*} + (1-\bar{d})\overline{W_d}D_{ma}^{d*}}{\overline{dW_d} + (1-\bar{d})\overline{W_d}}\right]$$

with

$$\begin{aligned} D_{ma}^{h*} =& D_{ma}^{d*} = E_h\left[\left(X_m - p_m^{'}\right)\left(X_a - p_a^{'}\right)\right] \\ D_{ma}^{d*} =& D_{ma}^{d*} = E_d\left[\left(X_m - p_m^{'}\right)\left(X_a - p_a^{'}\right)\right] \end{aligned}$$

D_{ma}^{h*} can be expressed in terms of $D_{ma}^{'}$ as follows:

$$\begin{aligned} D_{ma}^{h*} =& \mathrm{E}_h\left[\left(X_m - p_m^{h'} - (p_m^{'} - p_m^{h'})\right)\left(X_a - p_a^{h'} - (p_a^{'} - p_a^{h'})\right)\right] \\ =& D_{ma}^{'} + \left(p_m^{'} - p_m^{h'}\right)\left(p_a^{'} - p_a^{h'}\right) \end{aligned}$$

(since $\mathrm{E}_h\left[X_m - p_m^{h'}\right] = 0$, $\mathrm{E}_h\left[X_a - p_a^{h'}\right] = 0$). Similarly, $D_{ma}^{d*} = D_{ma}^{'} + \left(p_m^{'} - p_m^{d'}\right)\left(p_a^{'} - p_a^{d'}\right)$ Finally, because we want to compute D_{ma} to the first order in δ and to the first order in s, it is sufficient to express $p_m^{'} - p_m^{h'}$, $p_m^{'} - p_m^{d'}$ to the first order in δ, when $s = 0$, and $p_a^{'} - p_a^{h'}$, $p_a^{'} - p_a^{d'}$ to the first order in s, when $\delta = 0$. At the equilibrium for p_a, this gives:

$$\begin{aligned} p_m^{'} - p_m^{d'} &= -\frac{d_m}{\bar{d}}p_m q_m, \ , \quad p_a^{'} - p_a^{d'} = -b_a^d p_a q_a \\ p_m^{'} - p_m^{h'} &= \frac{d_m}{1-\bar{d}}p_m q_m, \ , \quad p_a^{'} - p_a^{h'} = -b_a^h p_a q_a \end{aligned}$$

Putting everything together finally yields:

$$D_{ma}^{'} = (1-r_{ma})D_{ma} + d_m\left[(1-\gamma)\left(b_a^d - b_a^h\right) + \gamma\frac{\overline{W_d}b_a^d - \overline{W_h}b_a^h}{d\overline{W_d} + (1-d)\overline{W_h}}\right]$$

Appendix B

QLE analysis of the three-locus model

The analyses of Appendix A can be extended to the three-locus model (with two selected loci). For this, the fitness of a haploid individual can be written as (Barton and Turelli 1991):

$$W_h = \overline{W_h}\left[1 + b_a^h \zeta_a + b_b^h \zeta_b + b_{ab}^h(\zeta_{ab} - D_{ab})\right]$$

with $\zeta_{ab} = \zeta_a \zeta_b$, and where (to the first order in s and e)

$$b_a^h \approx -s\alpha + e p_b \ , \quad b_b^h \approx -s\alpha + e p_a \ , \quad b_{ab}^h \approx e \tag{B.1}$$

Extrapolating to many loci under recurrent deleterious mutation, the selection coefficient b_i^h against the deleterious allele at locus i is given by $b_i^h = -s\alpha + eN_{mut}$, where $N_{mut} = \sum_j p_j$ is the average number of deleterious mutations per haplotype.

The fitness of a diploid is written as (e.g. Kirkpatrick et al. 2002)

$$W_d = \overline{W_d} \sum_{U,V} b_{U,V}^d \left(\zeta_{U,1}\zeta_{V,2} - D_{U,V}\right)$$

where the sum is over all possible sets of loci $\emptyset, a, b$ and $b_{U,V}$ represents the effect of selection on the set U and V of loci present on the first and second haplotypes of a diploid individual, respectively. Under our random mating assumption, we will see that only the coefficients b_a^d, b_b^d and b_{ab}^d will contribute. To the first order in s and e, and neglecting terms in p_a^2, p_b^2, these coefficients are given by:

$$b_a^d \approx -sh + 2e_{a\times a}p_b \ , \quad b_b^h \approx -sh + 2e_{a\times a}p_a \ , \quad b_{ab}^d \approx e_{axa} \tag{B.2}$$

Extrapolating to many loci, the diploid selection coefficient b_i^d against the deleterious allele at locus i is given by $b_i^d = -sh + 2e_{a\times a}N_{mut}$. Changes in allele frequencies and linkage disequilibria can be obtained using the same methods than in Appendix A; in the following we only provide the final results.

Frequency of deleterious alleles

To leading order in s, the change in frequency of allele a is still given by equation A.4 in A, where b_a^h and b_a^d are now given by the expressions above. In the case where $\overline{W_h} \approx \overline{W_d} \approx 1$(or if $\gamma = 0$), we can use the simpler equation 1.1 in Chapter 1, which yields:

$$\Delta N_{mut} \approx \left[d\left(-sh + 2e_{a\times a}N_{mut}\right) + (1-d)\left(-s\alpha + eN_{mut}\right)\right] + U \tag{B.3}$$

(where ΔN_{mut} is the change over selection and mutation). At equilibrium, one obtains:

$$\tilde{N}_{mut} = \frac{2U}{s\left[dh + (1-d)\alpha\right] + \sqrt{s^2\left[dh + (1-d)\alpha\right]^2 - 4U\left((1-d)e + 2de_{a\times a}\right)}} \tag{B.4}$$

In other cases, more precise solutions can be obtained numerically be replacing $\overline{W_h}$ and $\overline{W_d}$ in equation 1.1 by:

$$\overline{W_h} = \mathrm{e}^{-s\alpha N_{mut} + e\frac{N_{mut}^2}{2}} \ , \qquad \overline{W_d} = \mathrm{e}^{-2shN_{mut} + 2e_{a\times a}N_{mut}^2} \tag{B.5}$$

and solving numerically for N_{mut}.

Change in frequency of the modifier

The change in frequency of the modifier is obtained as in Appendix A, with:

$$\begin{aligned}
\Delta p_m^d &= b_a^d D_{ma} + b_b^d D_{mb} + b_{ab}^d D_{mab} \\
\Delta p_m^h &= b_a^h(1-r_{ma})D_{ma} + b_b^h(1-r_{mb})D_{mb} + b_{ab}^h(1-r_{ma})(1-r_{ab})D_{mab}
\end{aligned}$$

When $\overline{W_h} \approx \overline{W_d} \approx 1$, this yields:

$$\begin{aligned}\Delta p_m =& \gamma\left(\overline{W_d}-\overline{W_h}\right)p_m q_m \\ &+\left[d\ b_a^d+(1-d)(1-r_{ma})b_a^h\right]D_{ma}+\left[d\ b_b^d+(1-d)(1-r_{mb})b_b^h\right]D_{mb} \\ &+\left[d\ b_{ab}^d+(1-d)(1-r_{ma})(1-r_{ab})b_{ab}^h\right]D_{mab}\end{aligned} \tag{B.6}$$

A more precise expression (without assumption on $\overline{W_h}$, $\overline{W_d}$) can be obtained as in Appendix A.

Linkage disequilibria

To leading order, D_{ma} is given by the same expression as in Appendix A in terms of coefficients b_a^d and b_a^d (which now depend on epistasis as shown above); D_{mb} is given by the same expression, exchanging 'a' and 'b' subscripts. Finally, as for D_{ma}, D_{mb}, one obtains that the three-locus disequilibrium D_{mab} within gametes produced by diploids after selection and within gametes produced by haploids, after selection equal $D_{mab}^{d'}=D_{mab}^{h'}=(1-r_{ma})(1-r_{ab})D_{mab}$ to leading order. The overall value of D_{mab} within all gametes is given by:

$$D_{mab}^{'}=(1-\gamma)\left[\bar{d}D_{mab}^{d*}+(1-\bar{d})D_{mab}^{h*}\right]+\gamma\frac{d\overline{W_d}D_{mab}^{d*}+(1-\bar{d})\overline{W_h}D_{mab}^{h*}}{d\overline{W_d}+(1-\bar{d})\overline{W_h}}$$

where D_{mab}^{d*} and D_{mab}^{h*} are linkage disequilibria measured in gametes produced by diploids and haploids, but using as 'reference values' alleles frequencies in the whole population after selection; these are given by:

$$\begin{aligned}D_{mab}^{d*}&=(1-r_{ma})(1-r_{ab})D_{mab}+\frac{d_m}{\bar{d}}\left[(1-r_{ab})D_{ab}+b_{ab}^d pq_{ab}\right]\\ D_{mab}^{h*}&=(1-r_{ma})(1-r_{ab})D_{mab}-\frac{d_m}{1-\bar{d}}\left[(1-r_{ab})D_{ab}+b_{ab}^h pq_{ab}\right]\end{aligned}$$

When $\overline{W_h}\approx\overline{W_d}\approx 1$, one obtains at QLE: $D_{mab}=\frac{d_m}{r_{mab}}\left(b_{ab}^d-b_{ab}^h\right)pq_{mab}$

with $r_{ab}=1-(1-r_{ma})(1-r_{ab})$ and $pq_{mab}=p_m q_m p_a q_a p_b q_b$.

Extrapolation to many loci and multilocus simulations

We consider that epistasis is negative and that mutations are deleterious in both phases, so that mutation-selection equilibrium B.4 holds. Extrapolating equation B.6 to the whole genome and assuming free recombination among loci, the change in frequency of the modifier becomes:

$$\Delta p_m = d_m p_m q_m (S_m^{competition} + S_m^{D_{ma}} + S_m^{D_{mab}}) \tag{B.7}$$

with:

$$S_m^{competition} = \gamma \left[s\left(\alpha - 2h\right) N_{mut} - \left(e - 4e_{a\times a}\right) \frac{N_{mut}^2}{2} \right]$$

$$S_m^{D_{ma}} = -\frac{N_{mut}}{2} \left[s\left(\alpha - h\right) - \left(e - 2e_{a\times a}\right) N_{mut} \right] \left[d\left(sh - e_{a\times a} N_{mut}\right) + \frac{\left(1-d\right)\left(s\alpha - eN_{mut}\right)}{2} \right]$$

$$S_m^{D_{mab}} = -\frac{N_{mut}^2}{2} \left(e - e_{a\times a}\right) \left(d e_{a\times a} + (1-d)(1-r_{ma})e\right)$$

While the number of mutations per chromosome N_{mut} is large, interactions terms (in e and $e_{a\times a}$) become predominant and overcome the terms in s (unless e and $e_{a\times a}$ are very small). Figure 1.4 shows that even when selection coefficients lead to a diploid advantage, the evolution of ploidy is mainly determined by the relation between $e_{a\times a}$ and e when epistasis is sufficiently strong. When haploids and diploids compete mainly in a common ecological niche and when recombination is high compared with selection and epistasis, $S_m^{D_{ma}}$ and $S_m^{D_{mab}}$ are negligible. In this case, direct competition favors diploidy when $e < 4e_{a\times a} < 0$, as illustrated for $\gamma > 0.5$ and $e < -0.01$ (see Figure 1.4, blue and green curves). Indeed, diploids have twice more chances than haploids of carrying a deleterious allele at each locus, and thus four times more chances of carrying two deleterious alleles at a given pair of locus; therefore, the mean fitness of diploids is higher than that of haploids only if epistasis is four times weaker in diploids.

A high ecological differentiation ($\gamma < 0.1$) reveals the effect of linkage disequilibria. Here $S_m^{D_{ma}}$ still favors the modifier allele that extends the phase in which purging is more efficient. When selection against deleterious alleles is mainly driven by epistatic interactions (high N_{mut}), selection is stronger in diploids if $2e_{a\times a} < e < 0$ (because each mutation interacts with twice as many mutations as in haploids). Finally, epistasis also affects the evolution of the modifier locus through the term $S_m^{D_{mab}}$, which tends to favor the modifier allele that is more associated to the best two-locus combinations (i.e. AB and ab if epistasis is positive, Ab and aB otherwise). As shown by the expressions above, the sign of this term depends on the difference $e_{a\times a} - e$: indeed, when $e_{a\times a} < e < 0$, diploid selection is more efficient in generating Ab, aB combinations (which have higher fitness on average than ab, AB combinations), while haploid selection is more efficient when $e < e_{a\times a} < 0$.

Appendix C

Stability analysis of the two locus model

In the following we denote d_{11}, d_{12} and d_{22} the probability to undergo selection before meiosis for MM, Mm and mm zygotes, respectively: $d_{11} = d$, $d_{12} = d + h_m\delta$, $d_{22} = d + \delta$. In the two-locus model, we denote genotype frequencies among gametes as $x_1 = p_{AM}$, $x_2 = p_{aM}$, $x_3 = p_{Am}$ and $x_4 = p_{am}$. The general recursions for these frequencies are:

$$
\begin{aligned}
x_1^{'} =& (1-u)\left[(1-\gamma)\left(\frac{x_{1,D}^{sel}}{\overline{W_D}} + \frac{x_{1,H}^{sel}}{\overline{W_H}}\right) + \gamma\frac{x_{1,D}^{sel} + x_{1,H}^{sel}}{\overline{W_D} + \overline{W_H}}\right] \\
x_1^{'} =& (1-\gamma)\left(\frac{x_{2,D}^{sel}}{\overline{W_D}} + \frac{x_{2,H}^{sel}}{\overline{W_H}}\right) + \gamma\frac{x_{2,D}^{sel} + x_{2,H}^{sel}}{\overline{W_D} + \overline{W_H}} + u\left[(1-\gamma)\left(\frac{x_{1,D}^{sel}}{\overline{W_D}} + \frac{x_{1,H}^{sel}}{\overline{W_H}}\right) + \gamma\frac{x_{1,D}^{sel} + x_{1,H}^{sel}}{\overline{W_D} + \overline{W_H}}\right] \\
x_3^{'} =& (1-u)\left[(1-\gamma)\left(\frac{x_{3,D}^{sel}}{\overline{W_D}} + \frac{x_{3,H}^{sel}}{\overline{W_H}}\right) + \gamma\frac{x_{3,D}^{sel} + x_{3,H}^{sel}}{\overline{W_D} + \overline{W_H}}\right] \\
x_4^{'} =& (1-\gamma)\left(\frac{x_{4,D}^{sel}}{\overline{W_D}} + \frac{x_{4,H}^{sel}}{\overline{W_H}}\right) + \gamma\frac{x_{4,D}^{sel} + x_{4,H}^{sel}}{\overline{W_D} + \overline{W_H}} + u\left[(1-\gamma)\left(\frac{x_{3,D}^{sel}}{\overline{W_D}} + \frac{x_{3,H}^{sel}}{\overline{W_H}}\right) + \gamma\frac{x_{3,D}^{sel} + x_{3,H}^{sel}}{\overline{W_D} + \overline{W_H}}\right]
\end{aligned}
$$

with

$$
\begin{aligned}
x_{1,D}^{sel} &= x_1 \left[d_{11} \left(x_1 + x_2(1-hs) \right) + d_{12} \left(x_3 + x_4(1-r)(1-hs) \right) \right] + r d_{12} x_2 x_3 (1-hs) \\
x_{2,D}^{sel} &= x_2 \left[d_{11} \left(x_1(1-hs) + x_2(1-s) \right) + d_{12} \left(x_3(1-r)(1-hs) + x_4(1-s) \right) \right] + r d_{12} x_1 x_4 (1-hs)) \\
x_{3,D}^{sel} &= x_3 \left[d_{12} \left(x_1 + x_2(1-hs)(1-r) \right) + d_{22} \left(x_3 + x_4(1-hs) \right) \right] + r d_{12} x_1 x_4 (1-hs)) \\
x_{4,D}^{sel} &= x_4 \left[d_{12} \left(x_1(1-hs)(1-r) + x_2(1-s) \right) + d_{22} \left(x_3(1-hs) + x_4(1-s) \right) \right] + r d_{12} x_2 x_3 (1-hs)) \\
x_{1,H}^{sel} &= x_1 \left[d_{11} \left(x_1 + x_2 \right) + d_{12} \left(x_3 + x_4(1-r) \right) \right] + r d_{12} x_2 x_3 \\
x_{2,H}^{sel} &= x_2(1-s\alpha) \left[d_{11} \left(x_1 + x_2 \right) + d_{12} \left(x_3(1-r) + x_4 \right) \right] + (1-s\alpha) r d_{12} x_1 x_4 \\
x_{3,H}^{sel} &= x_3 \left[d_{12} \left(x_1 + x_2(1-r) \right) + d_{22} \left(x_3 + x_4 \right) \right] + r d_{12} x_1 x_4 \\
x_{2,H}^{sel} &= x_4(1-s\alpha) \left[d_{12} \left(x_1(1-r) + x_2 \right) + d_{22} \left(x_3 + x_4 \right) \right] + (1-s\alpha) r d_{12} x_2 x_3 \\
\overline{W_D} &= \sum_{i=1}^{4} x_{i,D}^{sel} \ , \qquad \overline{W_H} = \sum_{i=1}^{4} x_{i,H}^{sel}
\end{aligned}
$$

At the equilibrium where allele M is fixed and when allele a is maintained at mutation-selection balance, assuming s and δ of order ϵ, at the third order in ϵ the leading eigenvalue the linearized system equals:

$$
\lambda_{del} = 1 + p_a \left[\gamma(\alpha - 2h) + s \frac{(h-\alpha) \left[dh + (1-d)(1-r)\alpha \right]}{r + (1-r)s \left[dh + (1-d)(1-r)\alpha \right]} \right] \tag{C.1}
$$

In this case, $\lambda_{del} - 1$ is equivalent to the selection coefficient for allele m (e.g. Otto and Bourguet 1999) and can be compared with the QLE results.

Appendix D

Evolution of haplont, diplont or haploid-diploid life cycles when haploid and diploid fitnesses are not equal

Michael F Scott[1], Marie Rescan[2,3]

[1] Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC, Canada V6T 1Z4

[2] CNRS, Unité Mixte Internationale 3614, Evolutionary Biology and Ecology of Algae, Roscoff, France

[3] Sorbonne Universités, Université Pierre et Marie Curie, University of Paris 6, Roscoff, France

email: mfscott@biodiversity.ubc.ca.

Keywords: alternation of generations, life cycle evolution, diplohaplontic, modifier model, multilocus simulations

Running Title: Haploid-Diploid Evolution

Abstract

Many organisms spend a significant portion of their life cycle as haploids and as diploids (a haploid-diploid life cycle). However, the evolutionary processes that could maintain this sort of life cycle are unclear. Most previous models of ploidy evolution have assumed that the fitness effects of new mutations are equal in haploids and homozygous diploids, however, this equivalency is not supported by empirical data. With different mutational effects, the overall (intrinsic) fitness of a haploid would not be equal to that of a diploid after a series of substitution events. Intrinsic fitness differences between haploids and diploids can also arise directly, e.g., because diploids tend to have larger cell sizes than haploids. Here, we include intrinsic fitness differences into genetic models for the evolution of time spent in the haploid versus diploid phases, in which ploidy affects whether new mutations are masked. Life cycle evolution can be predominantly determined by intrinsic fitness differences between phases, genetic effects, or a combination of both. We find parameter ranges where these two selective forces act and show that the balance between them can favour convergence on a haploid-diploid life cycle, which is not observed in the absence of intrinsic fitness differences.

Introduction

Sexual reproduction in eukaryotes requires an alternation of haploid and diploid phases in the life cycle. Across taxa, there is a great deal of variation in the amount of growth (and time spent) in each of the haploid and diploid phases (see Valero et al. 1992, Klinger 1993, Richerd et al. 1993, Bell 1994; 1997, Mable and Otto 1998, Coelho et al. 2007). Some organisms, including almost all animals, are diplontic (somatic development occurs only in the diploid phase) and others, including ascomycete fungi, dictyostelid slime moulds, and some green algae (e.g., *Chara*), are haplontic (somatic development occurs only in the haploid phase). However, a large and phylogenetically diverse group of eukaryotes, including most land plants, basidiomycete fungi, most brown algae, red algae and some green algae, undergo some mitotic growth in both the haploid and diploid phases, which is referred to as a haploid-diploid life cycle here (sometimes called diplohaplontic or haplodiplontic) to avoid confusion with arrhenotoky ('haplodiploid' sex determination). While several theoretical studies have explored the conditions that should favour expansion of the haploid or diploid phases, there are still relatively few studies that show how a haploid-diploid life cycle could be maintained by selection.

A prominent theory for the evolution of either haplont or diplont life cycles involves the direct consequences of ploidy level on the expression of deleterious mutations. The fitness effects of a deleterious mutation can be partially hidden by the homologous gene copy in diploids, which is favourable if a heterozygote has a higher fitness than the average fitness of the

two component haploids. Thus modifier models, in which the extent of haploid and diploid phases is determined by a second locus, have found that diplonty is favoured when deleterious mutations are partially recessive and haplonty is favoured when deleterious mutations are partially dominant (Perrot et al. 1991, Otto and Goldstein 1992, Jenkins and Kirkpatrick 1994; 1995). As a consequence of mutations being partially concealed, an expanded diploid phase allows mutations to reach a higher frequency and thus increases mutation load (Crow and Kimura 1965, Kondrashov and Crow 1991). Therefore, reduced recombination favours haplonty because the association between the modifiers that expand the haploid phase and a higher quality, purged genetic background is retained for longer (Otto and Goldstein 1992).

The evolution of life cycles in sexual organisms appears to be similarly influenced by beneficial mutations. Using a numerical simulation approach, Otto (1994) and Orr and Otto (1994) show that diplonty is favoured during sweeps of beneficial mutations that are partially dominant. Increasing the length of the diploid phase of the life cycle increases the amount of selection experienced by heterozygotes and, with partial dominance, heterozygotes have higher fitness than the average fitness of the two component haploids. Conversely, haplonty is favoured when beneficial mutations are partially recessive. Again, lower recombination rates between the life cycle modifier and beneficial mutations broaden the parameter range over which haplonty is favoured because of associations between the modifiers expanding the haploid phase and higher quality genetic backgrounds that evolve when beneficial mutations are not masked.

These models typically assume that the overall fitness of haploids or diploids is the same. However, even with identical genomes, haploid and diploid cells typically differ in size and often in shape (e.g., Mable 2001), and growth and survival often differs between haploid and diploid phases. The phase with higher fitness and the magnitude of fitness differences varies widely and is heavily dependent on environmental context (Mable and Otto 1998, Thornber 2006). In *Saccharomyces* yeast, differences between haploid and diploid growth rates measured by Zörgö et al. (2013) range from being negligible to substantial (one phase can have growth rates up to 1.75 times higher) in different environments. Similar differences in growth rate and survival are observed between haploid and diploid phases of the red algae *Gracilaria verrucosa* and *Chondracanthus squarrulosus* in some laboratory conditions (Destombe et al. 1993, Pacheco-Ruíz et al. 2011). In addition, the fitness effect of new mutations may be unequal when present in haploids or in homozygous diploids, as reported by Gerstein (2013) and Zörgö et al. (2013). Therefore, following a series of substitution events, the overall (intrinsic) fitness of a haploid and a diploid should not be equal, as explored here.

The models discussed above assume that selection is independent of the densities of haploid and diploid individuals. These models also predict that either haplonty or diplonty evolves but not biphasic, haploid-diploid life cycles. Hughes and Otto (1999) and Rescan et al. (2016) consider density-dependent selection in which haploids and diploids occupy different ecological

niches and show that haploid-diploid life cycles can evolve in order to exploit both the haploid and diploid ecological niches. In this study, we complement these studies by considering only density independent selection in order to focus on intrinsic fitness differences between haploids and diploids.

The effect of intrinsic fitness differences on the evolution of the life cycle may seem obvious - selection should favour expansion of whichever phase (haploid or diploid) has higher fitness, as found by Jenkins and Kirkpatrick (1994; 1995). However, Jenkins and Kirkpatrick (1994; 1995) only considered the case where the differences in intrinsic fitness is either much larger or much smaller than the genome-wide deleterious mutation rate. Here, we consider the case where the two forces are of similar strength and quantify the parameters (e.g., mutation rate) for which this is true. In addition, we consider the effect of beneficial mutations on life cycle evolution when there are intrinsic fitness differences between haploids and diploids. We show that haploid-diploid life cycle can evolve even in the absence of density dependent selection due to a balance between intrinsic fitness differences between phases and the genetic effects of masking/revealing mutations. We also determine whether haploid-diploid life cycles favour intermediate amounts of time in each phase or mixtures of haploid and diploid specialists by considering the branching conditions (see also Rescan et al. 2016).

Model

We consider life cycle evolution using a modifier model in which the proportion of time spent in the haploid and diploid phases depends on the genotype at a modifier locus. Selection on the modifier results from viability selection on a set of L other loci. We first present a two-locus model, in which there is one viability locus and one modifier locus (as in the previous models by Perrot et al. 1991, Otto 1994, Jenkins and Kirkpatrick 1994; 1995, Hall 2000). We then extrapolate our results to the evolution of a modifier locus linked to many loci under selection; selection on a modifier caused by many loci is well approximated by the sum of the selective effect of each pairwise interaction considered separately (e.g., Jenkins and Kirkpatrick 1995, Otto and Bourguet 1999, Hough et al. 2013), assuming that the viability loci are loosely linked, autosomal and nonepistatic and the modifier has a small effect. We then test this approach by comparing our results to an explicit multi-locus simulation. Finally, we show that beneficial mutations can generate selection on the life cycle similar to that caused by deleterious mutations.

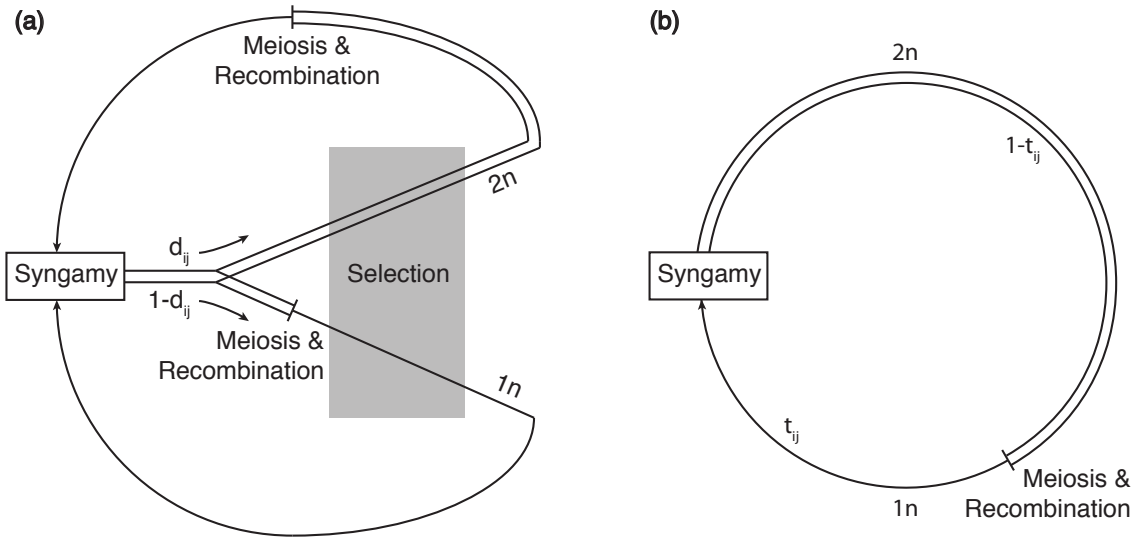

Figure D.1: Model (a) discrete selection and (b) continuous selection haploid-diploid life cycles. Single lines represent haploid phases and doubled lines indicate diploid phases. In (a), modified from Perrot et al. (1991) and Otto and Goldstein (1992), zygotes with the modifier genotype ij undergo selection as diploids with probability d_{ij} or undergo meiosis and recombination before experiencing selection as haploids with probability $(1 - d_{ij})$. In (b), after Jenkins and Kirkpatrick (1994; 1995) and Otto (1994), all zygotes with genotype ij experience viability selection as a diploid for a proportion $(1 - t_{ij})$ of their life cycle before undergoing meiosis and recombination and then experiencing viability selection as a haploid for the remainder of the life cycle.

Analytical Model

Previously, two general life cycles have been used to examine the evolution of haploid versus diploid phases. In the model considered by Perrot et al. (1991), Otto and Goldstein (1992), Otto and Marks (1996), Hall (2000) and Rescan et al. (2016), selection occurs once per generation and modifiers affect whether selection occurs during the haploid or diploid phase, figure D.1a. Jenkins and Kirkpatrick (1994; 1995) and Otto (1994) assume that selection occurs continuously throughout the life cycle, figure D.1b. In addition, some models have assumed that mutations occur upon gamete production (e.g., Otto and Goldstein 1992), and others assume that mutations occur at meiosis (e.g., Hall 2000). This leads to four possible life cycles, which can be found in the appendix.

In the main text, we primarily present results from the life cycle in which selection occurs continuously throughout the life cycle, see figure D.1b, and mutations occur at meiosis. We obtain generally similar results using the alternative models (e.g., discrete selection and mutations at gamete production), these analyses can be found in the supplementary *Mathematica* file (Wolfram Research Inc. 2010). However, discrete and continuous selection models can differ in whether/when convergence upon a haploid-diploid life cycle occurs. In particular, using

Table D.1: Fitnesses of different genotypes.

Genotype	Fitness
A	$w_A(t_{ij}) = \exp[t_{ij}\sigma_h]$
a	$w_a(t_{ij}) = \exp[t_{ij}(\sigma_h + s_h)]$
AA	$w_{AA}(t_{ij}) = \exp[(1 - t_{ij})(\sigma_d)]$
Aa	$w_{Aa}(t_{ij}) = \exp[(1 - t_{ij})(\sigma_d + hs_d)]$
aa	$w_{aa}(t_{ij}) = \exp[(1 - t_{ij})(\sigma_d + s_d)]$

the discrete selection model (figure D.1a), Hall (2000) showed that 'polymorphic' haploid-diploid life cycles can evolve if mutations occur at meiosis. However, as shown below, meiotic mutation does not favour haploid-diploid life cycles in the continuous selection model (figure D.1b). In addition, the convergence properties of discrete and continuous selection life cycles can differ when selfing occurs, see appendix.

In the continuous selection model (figure D.1b), zygotes are formed during synchronous random mating. The diploid genotype (ij) at the modifier locus (MM, Mm, or mm) determines the timing of meiosis and hence the proportion of time each individual spends as a diploid ($1 - t_{ij}$) and as a haploid (t_{ij}). Here, S_h and S_d represent selection acting across the genome due to intrinsic fitness differences between haploids and diploids. As our initial focus will be on the selection experienced at each of L selected loci, we also define $\sigma_h = S_h/L$ and $\sigma_d = S_d/L$ as the intrinsic fitnesses per viability locus. When $\sigma_h > \sigma_d$, haploids have higher fitness than diploids and the fitness of diploids is higher when $\sigma_d > \sigma_h$. At each viability locus, we consider a wild type and mutant allele (alleles A and a). The mutant allele at each viability locus, a, can have a different effect on fitness when present in a haploid (s_h) or in a homozygous diploid (s_d). The fitness of heterozygous diploids depends on the dominance of these mutations, given by h. When considering deleterious mutations, s_h and s_d are both negative, and when considering beneficial mutations, s_h and s_d are both positive. The fitnesses of the various genotypes are given in table D.1, where fitness terms are in the exponent. Recombination between the modifier and viability locus (at rate r) and mutation (from A to a, at rate μ per viability locus) occur at meiosis followed by haploid selection and then gamete production. The frequencies of genotypes MA, Ma, mA and ma are censused in the gametes (given by x_1, x_2, x_3 and x_4 respectively). The recursion equations given this life cycle are provided in the appendix.

Multilocus Simulations

We used individual-based simulations (C++ program available in the Dryad Digital Repository) to test predictions from our analytical model when deleterious mutations segregate at

L loci. Each individual carries either one or two copies of a chromosome (depending on its ploidy level) represented by a modifier locus (located at the midpoint of the chromosome) and a sequence of L bits (0 or 1) corresponding to the different loci.

Mutations occur at a rate U per generation: the number of new mutations per chromosome is sampled from a Poisson distribution with parameter U and distributed randomly across the genome; alleles at mutant loci are switched from 0 to 1 or from 1 to 0. Mutation and back mutation thus occur at the same rate, but back mutations should generally have negligible effects under the parameter values that we use, as deleterious alleles remain at low frequencies. We assume that all deleterious alleles have the same effects on fitness (s_d, s_h, and h are constant) and that these effects multiply across loci: the fitness of a haploid carrying n deleterious alleles is given by $w_h = \exp[S_h + s_h n]$, while the fitness of a diploid carrying n_{he} deleterious alleles in the heterozygous state, and n_{ho} in the homozygous state is given by $w_d = \exp[S_d + n_{he} h s_d + n_{ho} s_d]$.

At the start of each generation, all N individuals are diploid. To produce the 2N gametes that will form the diploids of the next generation, a diploid individual is sampled randomly among all diploids of the previous generation, and undergoes meiosis to produce a haploid; the number of cross-overs is sampled from a Poisson distribution with parameter R, while the position of each cross-over is sampled from a uniform distribution. If a random number sampled from a uniform distribution between 0 and 1 is lower than ${w_d}^{1-t} {w_h}^t$ (where w_d and w_h are the fitnesses of the diploid parent and haploid offspring), divided by its maximal possible value, then the haploid is retained; otherwise another diploid parent is sampled, until the condition is fulfilled.

At the beginning of the simulation, the modifier locus is fixed for an allele coding for an initial length of the haploid phase t_{init} (all simulations were performed for t_{init} values of 0.1, 0.5 and 0.9) and all selected loci are fixed for allele 0. Then, deleterious mutations are introduced at rate U per chromosome (the length of the haploid phase being still fixed to t_{init}) until the population reaches mutation-selection equilibrium (after generally 2,000 generations). After that, mutations at the modifier locus are introduced at a rate m_M per generation. When a mutation occurs, the length of the haploid phase coded by the mutant allele is sampled from a uniform distribution between $t_{old} - 0.1$ and $t_{old} + 0.1$, where t_{old} is the value of the parent allele; if the new value is negative or higher than 1, it is set to 0 or 1, respectively. We assume additivity among modifier alleles such that a zygote with alleles t_1 and t_2 will have a haploid phase of length $t = (t_1 + t_2)/2$. Simulations initially lasted 100,000 generations, which was sufficient in most cases for the average rate of diploidy to reach steady state, $\bar{t}$. We categorized the life cycle that evolved at the end of the simulation as haplont ($\bar{t} > 0.95$, white circles in figures 2 and 3b), diplont ($\bar{t} < 0.05$, black circles), or haploid-diploid ($0.05 < \bar{t} < 0.95$, green circles). In some cases, there was a repelling state such that the population evolved to

haplonty or diplonty depending on t_{init} (red circles).

Results

Deleterious Mutations

We first find the frequency of deleterious mutations at mutation-selection balance ($\hat{q}_a$) when the modifier locus is fixed for a particular resident allele (MM fixed, so that the length of the haploid phase is t_{MM}). Assuming that the per locus mutation rate (μ) is small, terms of the order of the square of the per locus mutation rate can be ignored, yielding

$$\hat{q}_a = \frac{\mu \exp[t_{MM} s_h]}{1 - \exp[t_{MM} s_h + (1 - t_{MM}) h s_d]}, \tag{D.1}$$

assuming there is some haploid or diploidy heterozygous expression so the denominator isn't near zero. When deleterious mutations are partially masked by the homologous gene copy in diploids ($hs_d/s_h < 1$), the frequency of deleterious mutations ($\hat{q}_a$) is higher when the diploid phase is longer (lower t_{MM}).

Life cycle evolution is considered by introducing an allele (m) at the modifier locus that controls the timing of meiosis and evaluating whether its frequency increases when rare. Mutants are able to invade when the leading eigenvalue of the system described by equations D.12c and D.12d, λ_l, is greater than one. Jenkins and Kirkpatrick (1994) derive a version of λ_l when $s_d = s_h$, however, they only discuss per locus intrinsic fitness differences that are of a much greater magnitude than the mutation load ($|\sigma_d - \sigma_h| \gg \mu$). To investigate the interaction between these selective forces we first present an approximation of λ_l in which the per locus fitness difference between haploids and diploids ($|\sigma_d - \sigma_h|$) is of similar magnitude to the per locus mutation rate, $O(\epsilon^2)$, the selective disadvantage of mutants (s_d and s_h) is of a larger order of magnitude, $O(\epsilon)$, and linkage is loose (r of $O(1)$) yielding

$$\lambda_l \approx 1 + (t_{Mm} - t_{MM}) \left(\sigma_h - \sigma_d + 2(-s_h)\hat{q}_a \left(\frac{h s_d}{s_h} - \frac{1}{2} \right) \right) + O(\epsilon^3). \tag{D.2}$$

Because mutation rates are small, deleterious mutations are found at low frequencies, therefore life cycle evolution depends only on the fitness of heterozygous mutants and not homozygous mutants (i.e., s_d is always found with the dominance coefficient, h). Consequently, life cycle evolution depends only on the 'effective dominance', $h_e = hs_d/s_h$, rather than dominance per se.

Using this approximation, haploid-diploid life cycles are evolutionarily singular strategies when $\sigma_h - \sigma_d = 2(s_h)\hat{q}_a(h_e - 1/2)$. Without intrinsic fitness differences, there is no intermediate

value of t_{MM} that solves this condition, hence either haplont or diplont life cycles are favoured. Thus, whereas Hall (2000) shows that biphasic haploid-diploid life cycles can evolve if selection occurs once per generation (figure D.1a) and mutations occur at meiosis (as considered here), haploid-diploid life cycles in the continuous selection model (figure D.1b) do not evolve in the absence of intrinsic fitness differences.

Intermediate singular strategies do exist when diploids have higher intrinsic fitness ($\sigma_d > \sigma_h$) and deleterious mutations are effectively partially dominant ($hs_d/s_h > 1/2$) or when haploids are favoured ($\sigma_h > \sigma_d$) and deleterious mutations are effectively partially recessive ($hs_d/s_h < 1/2$). However, life cycles will only converge upon this strategy if $\sigma_d > \sigma_h$ and $hs_d/s_h > 1/2$. Otherwise, the singular strategy is a repelling point and organisms should evolve either haplonty or diplonty, see figures D.2c, D.2d, and D.3a.

After convergence on a haploid-diploid strategy, we can then ask whether this singular strategy is evolutionarily stable. Using the same weak selection approximations as above, evolutionary stability is given by:

$$\frac{\delta^2 \lambda_l}{\delta {t_{Mm}}^2}\Big|_{t_{Mm}=t^*} = \frac{2(-s_h)(\sigma_d - \sigma_h)(hs_d/s_h - 1)(1-r)w_a[t^*]w_{Aa}[t^*]}{w_A[t^*]w_{AA}[t^*] - (1-r)w_a[t^*]w_{Aa}[t^*]}, \tag{D.3}$$

where t^* indicates the singular strategy for t, the length of the haploid phase. When convergence is stable (requiring that $\sigma_d > \sigma_h$ and $hs_d/s_h < 1$, see below), the singular strategy is evolutionarily unstable (D.3 is positive). Thus we expect weak disruptive selection after this singular point is reached. Indeed, our multilocus simulations sometimes displayed branching after 100,000 generations, such that there was a proportion t^* of haploid alleles ($t_1 = 1$), and a proportion $(1 - t^*)$ of diploid alleles ($t_2 = 0$). Increasing the number of generations always lead to branching when it was not observed by this time.

The weak selection approximation above demonstrates that intermediate haploid-diploid strategies can evolve when diploids have higher intrinsic fitness and when the difference in intrinsic fitness is of the same order of magnitude as the per locus mutation rate (although these strategies may be evolutionarily unstable). However, this analysis assumes that the recombination rate is large relative to selection. To consider tighter linkage and/or stronger selection we can use the more accurate expression of λ_l

$$\lambda_l = \exp[(t_{Mm} - t_{MM})(\sigma_h - \sigma_d)]\left(1 + \frac{\mu K_1}{K_2 K_3}\right), \tag{D.4}$$

where

$$
\begin{aligned}
K_1 = 1 &- (1-r)\exp[-(t_{Mm}-t_{MM})hs_d] \\
&- r\exp[(t_{Mm}-t_{MM})(s_h-hs_d)] \\
&+ (1-2r)\{\exp[(1-t_{Mm}-(t_{Mm}-t_{MM}))hs_d + t_{Mm}s_h] \\
&\qquad - \exp[(1-t_{Mm})hs_d + t_{Mm}s_h]\} \\
K_2 = 1 &- \exp[-(1-t_{MM})hs_d - t_{MM}s_h] \\
K_3 = 1 &- (1-r)\exp[(1-t_{Mm})hs_d + t_{Mm}s_h],
\end{aligned}
$$

in which the per locus mutation rate (μ) is assumed to be small, so that terms on the order of the square of the mutation rate can be ignored.

Equation (D.4) shows that singular strategies can exist without intrinsic fitness differences when recombination rates are low, $r < 1/2$, see figures D.2b and D.2d). As above, these singular strategies are always repelling points when $\sigma_d = \sigma_h$ (see supplementary *Mathematica* file) such that differences in intrinsic fitness are required for haploid-diploid life cycles to evolve. Convergence upon a haploid-diploid life cycle still requires that diploids have higher intrinsic fitness ($\sigma_d > \sigma_h$, see supplementary *Mathematica* file). However, as selection becomes less weak relative to recombination rates (such that the approximation in D.2 is not appropriate), haploid-diploid life cycles can evolve when $hs_d/s_h < 1/2$, see figure D.2b. In addition, convergence stability requires $hs_d/s_h < 1$, such that the frequency of deleterious mutations ($\hat{q}_a$) increases with the length of the diploid phase, see figure D.3a.

We next extend our two-locus result to consider deleterious mutations across L viability loci by assuming that these loci are loosely linked, autosomal and nonepistatic. With these assumptions (e.g., Jenkins and Kirkpatrick 1995, Otto and Bourguet 1999, Hough et al. 2013, Rescan et al. 2016), invasion of a modifier of weak effect is given by

$$\lambda_{net} = 1 + \sum_{l=1}^{L}(\lambda_l - 1). \tag{D.5}$$

In figures D.2 and D.3a we plot where this approximation predicts haplont, diplont or haploid-diploid life cycles to evolve for comparison to the explict multi-locus simulation (described above). In figures D.2 and D.3a we plot where this approximation predicts haplont, diplont or haploid-diploid life cycles to evolve for comparison to the explicit multi-locus simulation (described above).

Above, as in previous work, we consider the average dominance and selection coefficients (h, s_d and s_h). We can approximate the effect of small amounts of variation (and covariation)

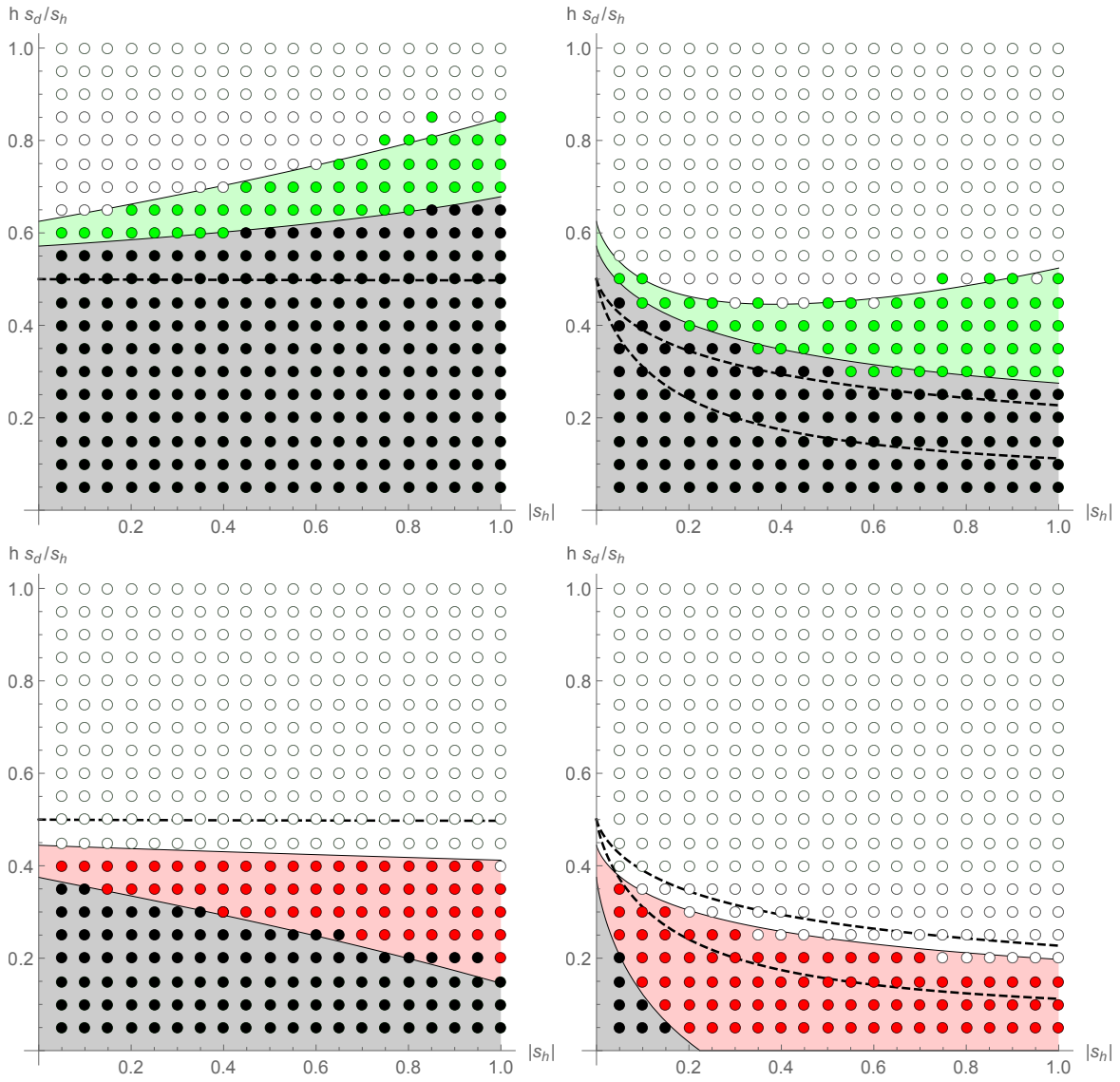

Figure D.2: Parameter space where haplont, diplont and haploid-diploid life cycles are favoured where the strength of selection against deleterious mutations ($|s_h|$) and effective dominance hs_d/s_h is varied. Background colors: prediction from the two-locus stability analysis extrapolated to multiple loci. Circles: multilocus simulation results starting from three different initial haploidy rates (t_{init} = 0.01, 0.5, or 0.99), with population size 20,000. White: evolution toward haplonty. Green: convergence stable haploid-diploid life cycles. Red: either haplonty or diplonty is favoured, with a repelling state in between. Black and gray: evolution toward diplonty. (a) and (b): diploids have higher intrinsic fitness ($S_h = 0$, $S_d = 0.025$) (c) and (d): haploids have higher intrinsic fitness ($S_h = 0.025$, $S_d = 0$). Map length: $R = 100$ ((a) and (c)) and $R = 0.35$ ((b) and (d)). The dashed lines show where haplonty (above dashed lines) and diplonty (below dashed lines) are favoured when there is no difference in intrinsic fitness ($S_h = S_d = 0$). In (b) and (d), there is a repelling point between the dashed lines. Mutants change the life cycle by a small amount ($|t_{Mm} - t_{MM}| = 0.001$) and the genome-wide haploid mutation rate, $U = 0.1$.

among loci in these coefficients by performing a Taylor expansion, as described in Lynch and Walsh (1998), Appendix 1 (see *Mathematica* file for details). Because we have assumed that deleterious mutations are rare, s_d is always found with h and we consider variation in s_h and the compound parameter hs_d. Assuming that deviations between coefficients and their mean value are of order ϵ and that selection is weak (as assumed in equation D.2), yields

$$\begin{aligned}\lambda_{net} \approx & 1 + (t_{Mm} - t_{MM})\Bigg(\sigma_h - \sigma_d + 2(-s_h)L\hat{q}_a\left(\frac{hs_d}{s_h} - \frac{1}{2}\right) \\ & + \frac{(1+t_{MM})L\hat{q}_a(-s_h)}{\mu^2}\Bigg((1-t_{MM})\left(\frac{hs_d}{s_h}\operatorname{Cov}(hs_d, s_h) - \operatorname{Var}(hs_d)\right) \\ & + t_{MM}\left(\frac{hs_d}{s_h}\operatorname{Var}(s_h) - \operatorname{Cov}(hs_d, s_h)\right)\Bigg)\Bigg) + O(\epsilon^3)\end{aligned} \tag{D.6}$$

Based on this analysis, variation in s_h generally makes haplonty more stable to invasion (reduces λ_{net} for $t_{MM} = 1, t_{Mm} < 1$). Similarly, variation in hs_d makes diplonty more stable to invasion (where $t_{MM} = 0, t_{Mm} > 0$). Positive covariation between hs_d and s_h has the opposite effect. Yeast deletion data indicate that the heterozygous effects of deleterious mutations may be much less variable than their homozygous effects, due to a negative correlation between h and s (Phadnis and Fry 2005, Agrawal and Whitlock 2011, Manna et al. 2011). Even if s_d and s_h are on average the same, it may thus be that the variance of hs_d is much lower than the variance of s_h.

Beneficial Mutations

Previous models of life cycle evolution in sexual organisms have used numerical approaches to show that, in the absence of intrinsic fitness differences, haplonty or diplonty can be favoured during sweeps of beneficial mutations (Otto 1994, Orr and Otto 1994). Here, we demonstrate that beneficial mutations interact with intrinsic fitnesses in a similar way to deleterious mutations and thus generate similar selection on the life cycle. We obtain analytical results using a quasi-linkage equilibrium (QLE) approximation, in which selection is assumed to be weak relative to recombination so that linkage disequilibrium ($D = x_1x_4 - x_2x_3$) equilibrates quickly relative to the rate of change of allele frequencies ($p_A = x_1 + x_3$ and $p_M = x_1 + x_2$) Assuming weak selection, $O(\epsilon)$, and low mutation rates, $O(\epsilon^2)$, the leading order term for the quasi-equilibrium value of linkage disequilibrium ($\hat{D}_Q$) is given by

$$\hat{D}_Q \approx \delta_t \frac{s_h}{r} p_M(1-p_M)p_A(1-p_A)\left(1 - p_A\frac{hs_d}{s_h} - (1-p_A)(1-h)\frac{s_d}{s_h}\right) + O(\epsilon^2), \tag{D.7}$$

where $\delta_t = (p_M(t_{Mm} - t_{MM}) + (1 - p_M)(t_{mm} - t_{Mm}))$ is the effect of the modifier on the length of the haploid phase (δ_t is positive if m increases the haploid phase with $t_{mm} > t_{Mm} > t_{MM}$ and negative if $t_{mm} < t_{Mm} < t_{MM}$).

Linkage disequilibrium is a measure of whether certain genotypes are found together more often than expected by chance. When $D > 0$, alleles A and M are more often found together, as are alleles a and m. When $s_h = s_d$ and $0 < h < 1$, as assumed in Otto (1994) and Orr and Otto (1994), equation (D.7) shows that m alleles that increase the length of the haploid phase ($\delta_t > 0$) are associated with the beneficial mutation, a ($\hat{D}_Q > 0$). This association should be retained for longer when the recombination rate is low. Hence lower recombination rates should favour haplonty, as found by Otto (1994) and Orr and Otto (1994).

The change in the frequency of the modifier allele, m (Δq_m) can then be expressed as a function of linkage disequilibrium ($\hat{D}_Q$) and allele frequencies, p_A and p_M. Assuming that selection is weak and mutation rates are low, the leading order term of Δq_m is given by

$$\Delta q_m \approx \delta_t p_M (1 - p_M) \left(\sigma_h - \sigma_d + s_h (1 - p_A) \left(1 - 2 p_A \frac{h s_d}{s_h} - (1 - p_A) \frac{s_d}{s_h} \right) \right) + O(\epsilon^2). \quad \text{(D.8)}$$

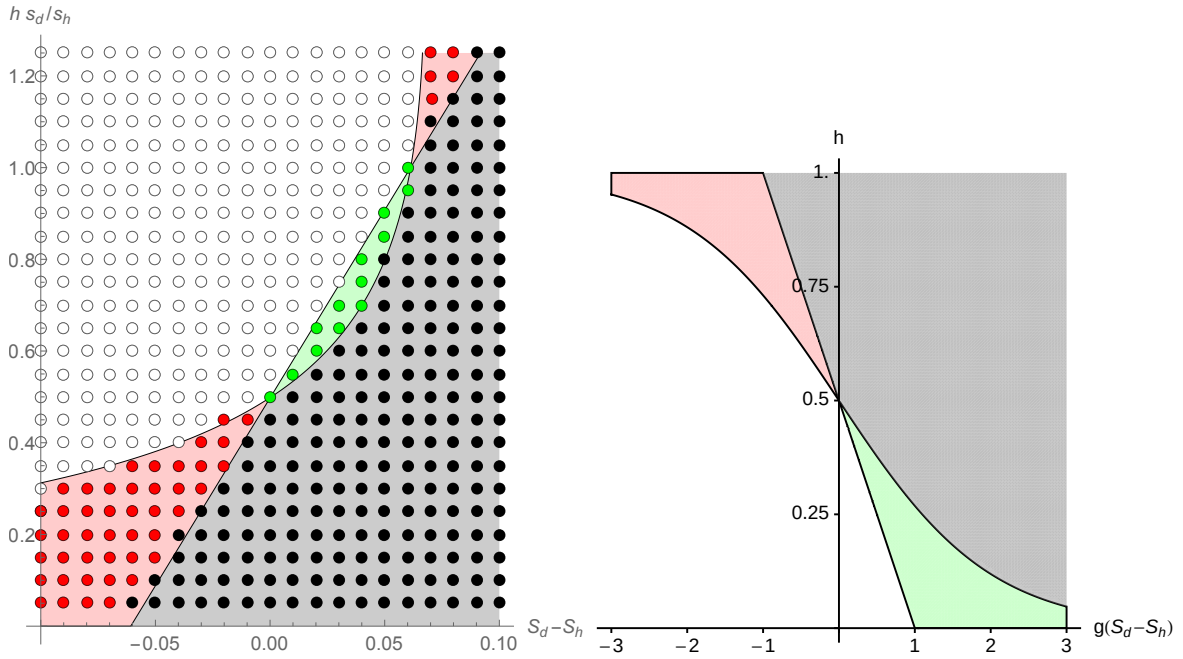

Figure D.3: Parameter space for which (a) deleterious mutations and (b) beneficial mutations favour haplont, diplont and haploid-diploid life cycles as a function of the difference in intrinsic fitness between haploids and diploids ($S_d - S_h$). (a) Shows the effective dominance of deleterious mutations (hs_d/s_h) against intrinsic fitness differences ($S_d - S_h$), parameters and symbols as in figures D.2a and D.2c with $|s_h| = 0.4$. (b) Regions in which particular life cycles are favoured in the presence of beneficial mutations, evaluated using equation D.11. g is the number of generations between fixation events. Population size, N, is 20000.

Unlike deleterious mutations, beneficial mutations reach high frequencies in the population, so the dynamics of the modifier depend on the fitness of both heterozygous and homozygous mutants. Equation (D.8) shows that, when fixed ($p_A = 0$), a beneficial mutation with a different effect size in haploids and diploids ($s_d \neq s_h$) affects life cycle evolution in a similar manner to intrinsic fitness differences (σ_d and σ_h). However, there is also transient selection on the life cycle that occurs during the fixation of a beneficial mutation. We isolate the transient selection on the life cycle from the effect on intrinsic fitnesses by considering the case where $s_d = s_h = s$ so that

$$\Delta q_m \approx \delta_t p_M(1-p_M)(\sigma_h - \sigma_d + 2p_A(1-p_A)(1/2-h)s) + O(\epsilon^2). \tag{D.9}$$

Equation (D.9) demonstrates that, in the absence of intrinsic fitness differences ($\sigma_d = \sigma_h$), haplonty is favoured during sweeps of partially recessive ($h < 1/2$) beneficial mutations and diplonty is favoured during sweeps of partially dominant ($h > 1/2$) beneficial mutations (as found numerically by Orr and Otto 1994).

Whether life cycle evolution is dominated by differences in intrinsic fitness or transient selection occurring generated by beneficial mutations depends on the rate at which beneficial mutations occur and how long they segregate in the population. The fixation time of beneficial mutations is different for different life cycles (longer when diploid phases are longer). We assume that the mutant life cycle allele is rare or similar enough to that of the resident that the time taken to fix a beneficial mutation depends on the life cycle of the resident and then measure the transient selection on the modifier over the entire time course of the sweep using

$$\int p_M(1-p_M)2p_A(1-p_A)p_A(1/2-h)s\,\mathrm{d}t. \tag{D.10}$$

This integral can then be evaluated assuming that a beneficial mutation will initially be found at frequency $1/N$, where N is the population size.

Assuming that the rate of adaptation is limited by the rate of environmental change so that a beneficial mutation fixes every g generations and considering selection on the life cycle from all L loci, the average invasion fitness of a rare life cycle modifier per generation is

$$\begin{aligned}\Delta\bar{q}_m \approx & \delta_t p_M(1-p_M)\Bigg((S_h - S_d) \\ & - \frac{1}{g}\ln\left[\frac{1}{N} + \frac{(N-1)(h(1-t_{MM})+t_{MM})}{N(1-h(1-t_{MM}))}\right]/(1-t_{MM})\Bigg),\end{aligned} \tag{D.11}$$

where the last term accounts for the fact that the beneficial mutations occur only once every g generations.

As with deleterious mutations, there can be haploid-diploid life cycles ($0 < t_{MM} < 1$) that are evolutionarily singular strategies. Assuming that the population size is large, mutants that increase the length of the haploid phase ($\delta_t > 0$) can only invade a resident population that has a short haploid phase ($t_{MM} = 0$) if beneficial mutations are partially recessive ($0 < h < 1/2$). Similarly, mutants that decrease the length of the haploid phase ($\delta_t < 0$) can only invade a resident population that has a long haploid phase ($t_{MM} \approx 1$) if beneficial mutations are partially recessive ($0 < h < 1/2$). Therefore, a haploid-diploid life cycle can only be convergence stable when $0 < h < 1/2$ (green in figure D.3b). Figure D.3b also shows the region in which both haplonty and diplonty cannot be invaded by small life cycle modifiers, in which case the singular strategy represents a repelling point (red).

When the rate of adaptation is not limited by the rate of environmental change, but by the rate of fixation of beneficial mutations, the time between fixation events depends on the occurrence of beneficial mutations ($1/g$) and their fixation probability (P_{fix}), which is given by $2s(t_{MM} + (1 - t_{MM})h)$. Fixation probability decreases when the diploid phase is longer because beneficial mutations are partially hidden by the extra chromosomal copy in diploids. Under mutation-limited adaptation g can be replaced in equation (D.11) by g/P_{fix}. In this case, haploid-diploid life cycles are never maintained by selection. Thus, beneficial mutations can only favour haplont or diplont life cycles if the rate of adaptation is not mutation-limited.

Discussion

Empirical evidence suggests that the fitness of haploid and diploid types may not be equal and that the fitness effects of new mutations are not generally the same (Thornber 2006, Gerstein 2013, Zörgö et al. 2013), leading to selection in favour of one ploidy type. Large differences in intrinsic fitnesses favour expansion of the phase with higher fitness (Jenkins and Kirkpatrick 1994). On the other hand, without differences in intrinsic fitness, life cycle evolution depends on the dominance of mutations (e.g., Perrot et al. 1991). In this study, we show how life cycles are expected to evolve when both of these selective forces act. Notably, we find that haploid-diploid life cycles involving an 'alternation of generations' can evolve under certain conditions, discussed below.

For haploid-diploid life cycles to evolve by selection, individuals with longer diploid phases must be favoured in predominantly haploid populations and individuals with longer haploid phases must be favoured in predominantly diploid populations. Previous models predicting the evolution of biphasic haploid-diploid life cycles have posited indirect benefits from decreasing senescence by reducing phase-specific generation time (Jenkins 1993), reducing the frequency

of sexual reproduction (Richerd et al. 1993), or exploiting more ecological niches (Bell 1997, Hughes and Otto 1999, Rescan et al. 2016). However, haploid-diploid life cycles are not a unique way of accessing these benefits. For example, diplont or haplont species can reduce generation times or the frequency of sexual reproduction without evolving haploid-diploid life cycles. Similarly, differentiated life cycle stages (Steenstrup alternations), phenotypic plasticity or genetic polymorphism can allow diplontic or haplontic species to exploit multiple ecological niches without tying growth form to the sexual cycle. Here, we use a population genetic model to show that haploid-diploid life cycles can evolve as a direct consequence of ploidy if the intrinsic fitness of haploids and diploids is not equal.

Deleterious and beneficial mutations affect life cycle evolution because changing the life cycle changes the amount of selection heterozygous zygotes will subsequently experience as heterozygous diploids versus as the component haploid genotypes. Heterozygous diploids have higher fitness than the average of the two component haploids when deleterious mutations are effectively partially recessive ($0 < hs_d/s_h < 1/2$) or when beneficial mutations are partially dominant ($1/2 < hs_d/s_h < 1$). Consequently, diplonty is favoured when partially recessive deleterious alleles are present at mutation-selection balance and during sweeps of partially dominant beneficial alleles. Conversely, partially dominant deleterious alleles or recessive beneficial alleles favour haplonty. The strength of this selection on the life cycle (caused by masking alleles) depends on the equilibrium frequency of deleterious alleles or the time taken for beneficial alleles to reach fixation, both of which are greater when the diploid phase is longer (assuming $0 < hs_d/s_h < 1$; when $hs_d/s_h > 1$ increasing the length of the diploid phase decreases the equilibrium frequency of deleterious alleles and haploid-diploid life cycles do not evolve).

When diploids have higher intrinsic fitness ($S_d > S_h$), a haploid-diploid life cycle can evolve if deleterious or beneficial mutations favour haplonty (figure D.3). In this case, the strength of selection in favour of haplonty is strong when the diploid phase is longer (because deleterious mutations reach higher frequencies and sweeps of beneficial mutations take longer) and can outweigh the intrinsic fitness differences. When the diploid phase is short, intrinsic fitness differences dominate, favouring a longer diploid phase. This combination ensures that evolution converges towards a haploid-diploid life cycle.

When haploids have higher intrinsic fitness ($S_h > S_d$), either haplonty or diplonty is always favoured. Even if selection due to deleterious or beneficial mutations favours diplonty and is of a similar strength to differences in intrinsic fitness, this can only lead to a repelling point, such that either haplonty or diplonty evolves. For these parameters, selection in favour of diplonty is stronger when the diploid phase is longer, favouring even longer diploid phases (because the benefits of masking deleterious mutations or revealing beneficial mutations are greater). Conversely, intrinsic fitness differences dominate when the diploid phase is short,

favouring longer haploid phases. Thus haplonty and diplonty can both be stable strategies (figure D.3).

Therefore, a strong condition for haploid-diploid life cycles to evolve in the model presented here is that diploids have higher intrinsic fitness than haploids. In theory, a diploid intrinsic fitness advantage may be particularly likely due to several previously proposed hypotheses. Firstly, Orr (1995*b*) showed that diplonty can protect organisms from partially recessive somatic mutations (e.g., masking potentially cancerous mutations that arise during development). Although Orr (1995*b*) did not explicitly explore whether haploid-diploid life cycles could evolve, considering somatic mutations that are partially recessive in his model generates a diploid advantage of the type considered here (see *Mathematica* file). Secondly, Haig and Wilczek (2006) proposed that, when diploid growth is partly provisioned by the female haploid (e.g., if diploids grow on haploids), paternally expressed genes will favour greater female allocation to his diploid offspring, improving the fitness of that phase. Finally, the presence of a homologous chromosome copy may serve as a template for more efficient DNA repair (Michod and Wojciechowski 1994).

Diploids also appear to have an intrinsic fitness advantage based on some of the mortality rates observed in natural populations of macroscopic algae, which is the fitness measure most closely related to the intrinsic fitness considered in this study. For example, diploids of the red algae *Mazzaella flaccida* and *Chondrus crispus* have moderately increased survivorship relative to haploids ($S_d - S_h \approx 0.1$, Bhattacharya 1985, Thornber and Gaines 2004). Other studies have found no difference in survivorship, perhaps because there is limited power to detect smaller differences in mortality rates (e.g., Engel et al. 2001, Thornber and Gaines 2004).

The model we present here evaluates the balance between selection due to intrinsic fitness differences and the effect of masking/revealing mutations. If either force is much stronger than the other, then it will dominate, as suggested by Jenkins and Kirkpatrick (1994). For example, figure D.3A shows how life cycles are expected to evolve when the deleterious mutation rate per haploid genome (U) is 0.1, approximately equal to estimates of the deleterious mutation rate in *Amsinckia* and *Arabidopsis* plants (Schoen 2005, Halligan and Keightley 2009). Figure D.3A suggests that these forces are of similar strength when the intrinsic fitness difference between haploids and diploids ($S_d - S_h$) is between 1% and 6%. For deleterious mutation rates that are a factor f larger, the scale of the x-axis on this figure can be multiplied by f to determine when selection on the life cycle due to deleterious mutations should be approximately the same strength as selection due to differences in intrinsic fitness.

Given that deleterious mutations are typically partially recessive (Simmons and Crow 1977, Agrawal and Whitlock 2011, Manna et al. 2011), the region in which a haploid-diploid life

cycle evolves is unlikely to be commonly encountered, except in two circumstances. First, if mutations are more deleterious in homozygous diploids than in haploids ($s_d > s_h$), haploid-diploid life cycles can be favoured when deleterious mutations are partially recessive (figure D.2a). Second, when recombination rates are low, the region in which haploid-diploid life cycles moves into the zone where deleterious mutations are partially recessive (figure D.2b).

A previous investigation by Otto and Marks (1996) found that haploidy was also favoured by recessive deleterious mutations when selfing, asexual reproduction or assortative mating is common (similar to low recombination). These results were interpreted via the fact that these mating schemes partly cause the effective recombination rate to be reduced, e.g., recombination has no impact in a selfed, homozygous individual. However, this analysis assumed that homozygotes and haploids have equal fitness, thus increased homozygosity had no direct impact on fitness. Here, we show that, when haploids and diploids have unequal fitness and/or when new mutations occur during the life cycle (e.g., at meiosis), the net effect of selfing can favour haploidy or diploidy (Appendix). We also note that the frequency of deleterious mutations, and thus their relative impact on life cycle evolution, is also decreased with increased selfing because they are exposed to selection in the homozygous state (Appendix). Thus, if the fitness of haploids and homozygous diploids differs, we caution against generally predicting that haplont and haploid-diploid life cycles should be more common in species where selfing, asexual reproduction and assortative mating are frequent. For example, this may explain why a survey by Mable and Otto (1998) found no correlation between haploidy and the estimated degree of sexuality in protists or green algae.

When the balance between intrinsic fitness differences and the effect of mutations favours convergence on haploid-diploid strategies, disruptive selection then arises such that polymorphisms can evolve with alternative alleles coding for longer haploid and longer diploid phases (i.e., a polymorphic strategy of specialists). In our simulations, a single modifier locus is able to confer fully haplont or diplont life cycles, polymorphism at this locus therefore means that these specialists life cycles can be relatively common (along with the life cycle of the heterozygote at the modifier locus). If genetic control of the life-cycle instead involves many loci at the modifier loci, each of which was limited to a having a small effect on the length of the haploid phase, a higher proportion of intermediate phenotypes would be observed in a population experiencing disruptive selection due to mating and recombination. This is especially true when modifier loci are loosely linked because associations between alleles at different loci (linkage disequilibria) are small when recombination is large relative to selection (e.g, Otto 2007, equation 9.45). Disruptive selection was also observed in a density-dependent model where haploids and diploids occupy different niches with or without deleterious mutations (Rescan et al. 2016). Temporal variability of niche sizes can, however, stabilize obligatory alternation between phases (Rescan et al. 2016). Thus, for haploid-diploid life cycles to be

favoured over a polymorphic population of specialist haploids and diploids appears to require constraints on the genetic architecture underlying life cycle variation or external variability.

It is intuitively and empirically reasonable that haploids and diploids should both differ in intrinsic fitness and in the extent to which new mutations are masked/revealed to selection. Here, we find the conditions under which these selective forces are approximately balanced and show that this suggests a new hypothesis for the evolution of haploid-diploid life cycles. A significant strength of this hypothesis is that haploid-diploid life cycles evolve in species undergoing an alternation of haploids and diploid phases without positing any extrinsic benefits.

Acknowledgements

We sincerely thank Sally Otto (SPO) for many helpful discussions, guidance and comments on the manuscript and thank Denis Roze for help with multilocus extrapolation analysis, multilocus simulations and helpful comments. Funding for this project was provided by the University of British Columbia and by a Natural Sciences and Engineering Research Council of Canada grant to SPO.

Appendix

We consider four models: two continuous selection models and two discrete selection models with mutations occurring at either meiosis or gamete production. We allow selfing to occur among gametes at rate σ, following Otto and Marks (1996). In the main text, we primarily discuss the continuous selection model with mutations at meiosis where $\sigma = 0$. We denote the genotypes MA, Ma, mA and ma by indices 1 to 4, the frequency of these genotypes in the next generation x_1', x_2', x_3' and x_4') are given by

$$\begin{aligned} x_1' = (1-\mu)\big((1-\sigma)\big(&x_1^2 w_{11,A} + x_1 x_2 w_{12,A} + x_1 x_3 w_{13,A} + x_1 x_4 w_{14,A} - rD w_{14,A}\big) \\ &+\sigma x_1 w_{11,A}\big)/\overline{W} \end{aligned} \tag{D.12a}$$

$$\begin{aligned} x_2' = \big((1-\sigma)\big(&x_2 x_1 w_{12,a} + x_2^2 w_{22,a} + x_2 x_3 w_{23,a} + x_2 x_4 w_{24,a} + rD w_{14,a}\big) \\ &+\sigma x_2 w_{22,a} \\ +\mu\big((1-\sigma)\big(&x_1^2 w_{11,A\mu} + x_1 x_2 w_{12,A\mu} + x_1 x_3 w_{13,A\mu} + x_1 x_4 w_{14,A\mu} - rD w_{14,A\mu}\big) \\ &+\sigma x_1 w_{11,A\mu}\big)\big)/\overline{W} \end{aligned} \tag{D.12b}$$

$$\begin{aligned} x_3' = (1-\mu)\big((1-\sigma)\big(&x_3 x_1 w_{13,A} + x_3 x_2 w_{23,A} + x_3^2 w_{33,A} + x_3 x_4 w_{34,A} - rD w_{14,A}\big) \\ &+\sigma x_3 w_{33,A}\big)/\overline{W} \end{aligned} \tag{D.12c}$$

$$\begin{aligned} x_4' = \big((1-\sigma)\big(&x_4 x_1 w_{14,a} + x_4 x_2 w_{24,a} + x_4 x_3 w_{34,a} + x_4^2 w_{44,a} + rD w_{14,a}\big) \\ &+\sigma x_4 w_{44,a} \\ +\mu\big((1-\sigma)\big(&x_3 x_1 w_{13,A\mu} + x_3 x_2 w_{23,A\mu} + x_3^2 w_{33,A\mu} + x_3 x_4 w_{34,A\mu} - rD w_{14,A\mu}\big) \\ &+\sigma x_3 w_{33,A\mu}\big)\big)/\overline{W} \end{aligned} \tag{D.12d}$$

where $D = x_1 x_4 - x_2 x_3$ and $\overline{W}$ is the sum of the numerators. The notation $w_{ij,k}$ refers to the fitness of a zygote formed by gametes with indices i and j that produces a haploid of type k without mutation, $w_{ij,k\mu}$ is similar but where the k haploid produced by meiosis mutates. These fitnesses for the discrete and continuous selection models are given in table D.2. When mutations occur at gamete production, mutation does not affect fitness and $w_{ij,A\mu} = w_{ij,A}$. The fitness values where mutations occur at meiosis are given in table D.3.

We then calculate the frequency of the a allele ($\hat{q}_a$) when the modifier locus is fixed for a resident allele, M, which is given by

$$\hat{q}_a = \frac{\mu w_{11,A\mu}}{w_{11,A} - (1-\sigma) w_{12,a} - \sigma w_{22,a}}, \tag{D.13}$$

Table D.2: Fitnesses in discrete and continuous selection models.

Fitness	Continuous selection	Discrete selection
$w_{11,A}$	$w_{AA}(t_{MM})w_A(t_{MM})$	$w_{AA}d_{MM} + w_A(1-d_{MM})$
$w_{12,A}$	$w_{Aa}(t_{MM})w_A(t_{MM})$	$w_{Aa}d_{MM} + w_A(1-d_{MM})$
$w_{12,a}$	$w_{Aa}(t_{MM})w_a(t_{MM})$	$w_{Aa}d_{MM} + w_a(1-d_{MM})$
$w_{13,A}$	$w_{AA}(t_{Mm})w_A(t_{Mm})$	$w_{AA}d_{Mm} + w_A(1-d_{Mm})$
$w_{14,A} = w_{23,A}$	$w_{Aa}(t_{Mm})w_A(t_{Mm})$	$w_{Aa}d_{Mm} + w_A(1-d_{Mm})$
$w_{14,a} = w_{23,a}$	$w_{Aa}(t_{Mm})w_a(t_{Mm})$	$w_{Aa}d_{Mm} + w_a(1-d_{Mm})$
$w_{22,a}$	$w_{aa}(t_{MM})w_a(t_{MM})$	$w_{aa}d_{MM} + w_a(1-d_{MM})$
$w_{24,a}$	$w_{aa}(t_{Mm})w_a(t_{Mm})$	$w_{aa}d_{Mm} + w_a(1-d_{Mm})$
$w_{33,A}$	$w_{AA}(t_{mm})w_A(t_{mm})$	$w_{AA}d_{mm} + w_A(1-d_{mm})$
$w_{34,A}$	$w_{Aa}(t_{mm})w_A(t_{mm})$	$w_{Aa}d_{mm} + w_A(1-d_{mm})$
$w_{34,a}$	$w_{Aa}(t_{mm})w_a(t_{mm})$	$w_{Aa}d_{mm} + w_a(1-d_{mm})$
$w_{44,a}$	$w_{aa}(t_{mm})w_a(t_{mm})$	$w_{aa}d_{mm} + w_a(1-d_{mm})$

Table D.3: Fitnesses of mutated types when mutations occur at meiosis.

Fitness	Continuous selection	Discrete selection
$w_{11,A\mu}$	$w_{AA}(t_{MM})w_a(t_{MM})$	$w_{AA}d_{MM} + w_a(1-d_{MM})$
$w_{12,A\mu}$	$w_{Aa}(t_{MM})w_a(t_{MM})$	$w_{Aa}d_{MM} + w_a(1-d_{MM})$
$w_{13,A\mu}$	$w_{AA}(t_{Mm})w_a(t_{Mm})$	$w_{AA}d_{Mm} + w_a(1-d_{Mm})$
$w_{14,A\mu} = w_{23,A\mu}$	$w_{Aa}(t_{Mm})w_a(t_{Mm})$	$w_{Aa}d_{Mm} + w_a(1-d_{Mm})$
$w_{33,A\mu}$	$w_{AA}(t_{mm})w_a(t_{mm})$	$w_{AA}d_{mm} + w_a(1-d_{mm})$
$w_{34,A\mu}$	$w_{Aa}(t_{mm})w_a(t_{mm})$	$w_{Aa}d_{mm} + w_a(1-d_{mm})$

where we ignore terms on the order of μ^2. For the continuous selection model with mutations at meiosis and $\sigma = 0$, this is equivalent to equation (D.1). As in the main text, we then evaluate the spread of a rare modifier using the leading eigenvalue (λ_l) of the system described by equations D.12c and D.12d. Full expressions of λ_l for each of the life cycles considered can be found in the supplementary *Mathematica* notebook.

In the models in which mutations occur at gamete production, and assuming that the fitnesses of A haploids and AA diploids are equal (such that $w_{11,A} = w_{13,A} = w_{33,A} = 1$), invasion occurs ($\lambda_l > 1$) if

$$\begin{aligned} 0 <& \sigma(w_{22,a} - w_{44,a})(w_{12,A} - w_{14,A}(1-r)) \\ &+ r(1-\sigma)(w_{12,A}w_{14,a} + w_{14,A}(w_{12,a} - 2w_{14,a}) \\ &+ (w_{12,A} - w_{14,A})(1 - w_{14,a}(1-\sigma) - w_{22,a}\sigma). \end{aligned} \tag{D.14}$$

Increased selfing can either increase or decrease the parameter range over which this inequality

is satisfied unless it is further assumed that the fitness of a haploids and aa diploids are equal (such that $w_{22,a} = w_{44,a}$ and the first term in D.14 is 0).

When the fitnesses of haploids and homozygous diploids are equal and mutations occur at gamete production, Otto and Marks (1996) showed that haploidy is always favoured over a larger parameter space when selfing is higher in the discrete selection model. Similarly, in the continuous selection model, where we also assume that modifiers have a small effect, $t_{Mm} - t_{MM} = \delta_{tMm}$ is of order μ, modifiers that increase the length of the haploid phase ($\delta_{tMm} > 0$) invade if

$$\begin{aligned} h(w_{AA}(t_{MM})w_A(t_{MM}) - (1-\sigma)w_{Aa}(t_{MM})w_a(t_{MM}) - \sigma w_{aa}(t_{MM})w_a(t_{MM})) \\ > r(1-\sigma)(1-2h)w_a(t_{MM})w_{AA}(t_{MM}). \end{aligned} \tag{D.15}$$

This condition is always met when $h > 1/2$ and is always satisfied for a greater parameter range with higher selfing rates (higher σ) if $h < 1/2$.

In the continuous selection model with mutations at meiosis, however, the impact of selfing is not so simple. Even when we assume the fitnesses of haploids and homozygous diploids is equal ($s_h = s_d$ and $\sigma_d = \sigma_h = 0$) and modifiers have a small effect ($t_{mm} - t_{MM} = \delta_{tmm}$ and $t_{Mm} - t_{MM} = h_m\delta_{tmm}$, where δ_{tmm} is of order μ and terms of $O(\mu^2)$ are discarded) and make the further assumption that recombination is free ($r = 1/2$), haploidy is favoured when

$$h > \frac{1-(1-h_m)(1-\sigma)(1+\sigma w_a(t_{MM})w_{Aa}(t_{MM})/K_1)}{2h_m}, \tag{D.16}$$

where $K_1 = w_{AA}(t_{MM})w_A(t_{MM}) - \sigma w_{aa}(t_{MM})w_a(t_{MM})$. For dominant modifiers ($h_m = 1$), this condition is satisfied if and only if $h > 1/2$, such that selfing has no effect on whether haploidy or diploidy is favoured. When $0 < h_m < 1$, increased selfing increases the right hand side of inequality (D.16). Therefore, increased selfing decreases, rather than increases, the parameter range under which haploidy is favoured. Although selfing can facilitate the evolution of haploidy when $r < 1/2$ (presumably because the impact of disequilibrium is greater), our overall finding is that when mutations occur at meiosis, selfing does not uniformly favour haploidy even when we assume that the fitness of haploids and homozygous diploids are equal.

In addition, the convergence properties of discrete and continuous selection models differ. For example, Hall (2000) found that, without selfing or intrinsic fitness differences, haploid-diploid life cycles can evolve in the discrete selection model where mutations occur at meiosis. However, in the main text we show that haploid-diploid life cycles do not evolve in the continuous

selection model where mutations occur at meiosis without intrinsic fitness differences. For the purposes of this study, one important distinction between models is whether haploid-diploid life cycles evolve for recessive deleterious mutations with selfing and loose linkage ($\sigma > 0$, $r = 1/2$). In figure D.4, we show a numerical example of life cycle evolution with selfing, loose linkage, and $s_d = s_h$. For these parameters, haploid-diploid life cycles evolve for low h in the discrete selection model but not in the continuous selection model (where mutations occur at gamete production in both cases). Thus in both the case considered by Hall (2000) (mutations at meiosis with no selfing) and in figure D.4 (mutations at gamete production with selfing), life-cycle models in which selection occurs continously (figure D.1b) favour haploid-diploid life cycles less often than discrete life cycle models (figure D.1a)

Finally, we clarify how selfing affects the disequilibrium between the M and A loci, which was discussed in Otto and Marks (1996). Using the same model and assumptions as Otto and Marks (1996), where $w_{AA} = w_A = 1$, $w_{Aa} = 1 - hs$, and $w_a = w_{aa} = 1 - s$ we find that the disequilibrium, $D = x_1x_4 - x_2x_3$ during invasion of a modifier is given by

$$D = \frac{(d_{Mm} - d_{mm})(1-h)\mu(1-\sigma)}{K_5(1 - d_{MM}(1-h)(1-\sigma))} \tag{D.17}$$

where $K_5 = r(1-\sigma) + s(1-d_{Mm})(1-h)(1-r) + hs(1-r)(1-\sigma) + \sigma s$ is strictly positive. Thus, disequilibrium has the same sign as $(d_{Mm} - d_{MM})$ and is positive for modifiers that increase the the diploid phase (modifiers associated with the less fit allele) and negative for modifiers that increase the haploid phase, as found by Otto and Marks (1996). However, the magnitude of this disequilibrium decreases with increasing selfing, contrary to the result stated in Otto and Marks (1996). In the supplementary *Mathematica* file we show that the magnitude of the disequilibrium increases with increasing selfing if $\hat{q}_a$ is held constant but because selfing also helps purging and reduces $\hat{q}_a$, the net effect on disequilibrium is opposite.

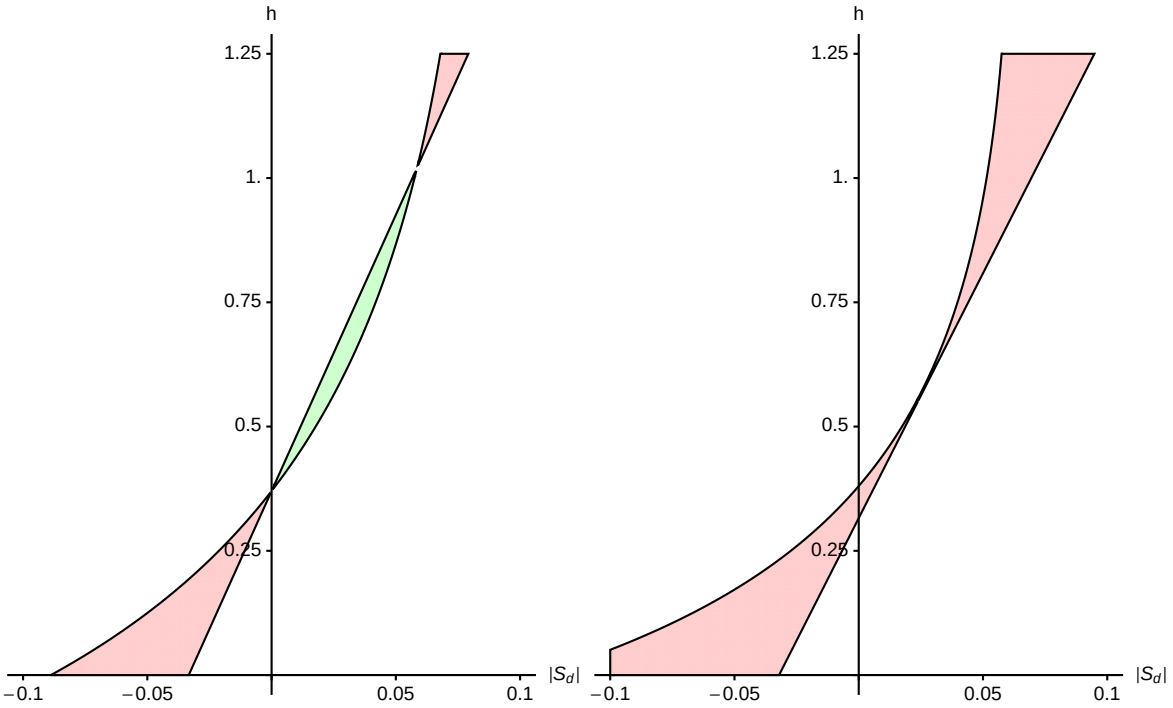

Figure D.4: Here we plot whether haplont, diplont, or haploid-diploid life cycles are favoured when there is selfing among gametes as a function of the intrinsic fitness of diploids (S_d) for (a) the discrete selection model with mutations at gamete production and (b) the continuous selection model with mutations at gamete production. To evaluate expected life cycle evolution we evaluated the stability of pure haplont ($d_{MM} = 0$, $t_{MM} = 1$) or diplont ($d_{MM} = 1$, $t_{MM} = 0$) strategies using equation (D.5) with the full expression of λ_l where terms on the order of μ^2 are discarded, which can be found in the supplementary *Mathematica* file. In both plots $\sigma = 0.4$, $r = 1/2$, $s_d = s_h = -0.3$, $U = 0.1$, $L = 1000$, $S_h = 0$, and modifiers have a small and dominant effect ($t_{mm} = t_{Mm}$, $|t_{Mm} - t_{MM}| = 1/10,000$, $d_{mm} = d_{Mm}$, $|d_{Mm} - d_{MM}| = 1/10,000$).

Appendix E

QLE analysis of the two-locus ecological model

We will first derive results for the two-locus model (modifier + selected locus), and then consider the case of the large number of loci at mutation-selection balance for deleterious alleles (assuming each deleterious allele stays at low frequency). In the two-locus model, we neglect the effect of the modifier and the selected locus on the equilibrium population size N (assuming δ and α are small), while effects of selection on population size must be accounted for in the multilocus extension.

Genetic associations

To compute recursions on allele frequency and linkage disequilibrium, we will use the following variables: X_a is an indicator variable that equals 1 if allele a is present in a given haploid at the selected locus, and 0 otherwise, while $X_{a,1}$ and $X_{a,2}$ are similar indicator variables for the two genes at the selected locus in a given diploid. Similarly, variables $X_{m,1}$ and $X_{m,2}$ equal 1 if allele m is present at the modifier locus in a haploid, or in each haplotype of a diploid. Genetic associations between genes present on different haplotypes of a diploid zygote can be defined as follows: defining $\zeta_{a,1} = X_{a,1} - p_a$ and $\zeta_{a,2} = X_{a,2} - p_a$, and similar variables $\zeta_{m,1}$, $\zeta_{m,2}$ for the modifier locus, the genetic association between genes in sets U and V, present respectively in the first and second haplotype of a diploid individual is defined as (e.g. Kirkpatrick et al. 2002):

$$D_{U,V} = E\left[\prod_{i\in U}\zeta_{i,1}\prod_{i\in V}\zeta_{i,2}\right]$$

where E stands for the average over all zygotes (since our model does not include any asymmetry between sexes, we have $D_{U,V} = D_{V,U}$ for all U, V). Because we assume random mating, associations $D_{m,m}$, $D_{a,a}$ and $D_{ma,a}$ equal zero, while $D_{ma,ma} = D_{ma}^2$, which will be negligible under the QLE assumption.

Different superscripts will be used to denote variables measured in haploids and diploids, at different stages of the life cycle: p_a^d, p_m^d and $D_{U,V}^d$ denote the frequencies of a, m and the genetic association $D_{U,V}$ before selection, within the subset of individuals that will undergo selection as diploids, while $p_a^{d'}$, $p_m^{d'}$ and $D_{U,V}^{d'}$ denote the same variables measured after selection. Similarly, p_a^h, p_m^h and $D_{U,V}^h$ represent allele frequencies and LD among haploids, before selection, while $p_a^{h'}$, $p_m^{h'}$ and $D_{U,V}^{h'}$ represent the same variables measured after selection.

Change in frequency at the selected locus

We first express the change in frequency of allele a due to selection (Δp_a). To leading order, we can neglect the effect of the modifier on Δp_a. The frequency of allele a at the next generation is given by:

$$p_a^{'} = \frac{d\overline{W_d}p_a^{d'} + (1-d)\overline{W_h}p_a^{h'}}{d\overline{W_d} + (1-d)\overline{W_h}}$$

where $p_a^{d'}$ and $p_a^{h'}$ are the frequencies of a within the diploid and haploid subpopulations after selection, and $\overline{W_h}$ and $\overline{W_h}$ the mean fitnesses of haploids and diploids. At equilibrium for population size, $d\overline{W_d} + (1-d)\overline{W_h}$ (see equation 5 in Chapter 2), which yields:

$$p_a^{'} = \overline{W_d}p_a^{d'} + (1-d)\overline{W_h}p_a^{h'}$$

Neglecting the effect of the modifier, the frequency of allele a is the same among haploids and diploids before selection: $p_a^d = p_a^h = p_a$. Then:

$$\Delta p_a = \overline{W_d}\Delta p_a^d + (1-d)\overline{W_h}\Delta p_a^h$$

where Δp_a^d and Δp_a^h are the changes in frequency of a due to selection, within the diploid and haploid subpopulations.

To compute Δp_a^d and Δp_a^h, it is useful to express the fitness (average number of zygotes produced per individual, or half the number of gametes produced) of individuals in terms of their ζ variables: $\zeta_a = X_a - p_a$ for haploids, $\zeta_a = X_a - p_a$ and $\zeta_{a,2} = X_{a,2} - p_a$ for diploids:

$$W_h = \overline{W_h} + s_h \zeta_a$$
$$W_d = \overline{W_d} + h_a^* s_d(\zeta_{a,1} + \zeta_{a,2}) + s_d(1 - 2h_a^*)\zeta_{a,1}\zeta_{a,2}$$

where $h_a^* = h + p_a(1 - 2h)$ (which is approximately h when p_a is small).

From equation 3 in the Chapter 2, we have to the first order in α:

$$s_h = -\rho\alpha N \frac{r_h}{K_h}\left[1 - d + d\gamma_{hd}\right]$$
$$s_d = -\alpha N \frac{r_d}{K_d}\left[(1 - d)\gamma_{dh} + d\right]$$

The effect of selection on allele frequencies within diploids and haploids is then given by:

$$\Delta p_a^d = \mathrm{E}_d\left[\frac{W_d}{\bar{W}_d}\frac{\zeta_{a,1} + \zeta_{a,2}}{2}\right] , \qquad \Delta p_a^h = \mathrm{E}_d\left[\frac{W_h}{\bar{W}_h}\zeta_a\right]$$

where E_d and E_h stand for the average over all diploid and haploid individuals. Using the fact that $E\left[\zeta_{a,1}^2\right] = E\left[\zeta_{a,2}^2\right] = E\left[\zeta_a^2\right] = p_a q_a$ (where $q_a = 1 - p_a$) and that $E\left[\zeta_{a,1}\zeta_{a,2}\right] = D_{a,a} = 0$, one obtains:

$$\Delta p_a^d = \frac{h_a^* s_d p_a q_a}{\overline{W_d}} , \qquad \Delta p_a^h = \frac{s_h p_a q_a}{\overline{W_h}}$$

which finally yields:

$$\Delta p_a = \left[d h_a^* s_d + (1 - d)\, s_h\right] \tag{E.1}$$

Change in frequency at the modifier locus

The change in frequency of the modifier over one generation can be written as:

$$\Delta p_m = d\overline{W_d}\left(\Delta_r p_m^d + \Delta_s p_m^d\right) + (1 - d)\,\overline{W_h}\left(\Delta_r p_m^h + \Delta_s p_m^h\right)$$

where $\Delta_r p_m^d = p_m^d - p_m$ is the difference between the frequency of m within the diploid subpopulation before selection (p_m^d) and the frequency of m within zygotes (p_m), while $\Delta_s p_m^d = p_m^{d'} - p_m^d$ is the difference between the frequency of m within the diploid subpopulation after selection (and similarly for $\Delta_r p_m^h$ and $\Delta_s p_m^h$).

To derive recursions in terms of genetic associations, it is helpful to express the rate of diploidy of a focal individual (denoted here after d^*) in terms of its $\zeta_{m,1}$, $\zeta_{m,2}$ variables:

$$d^* = \bar{d} + d_m\left(\zeta_{m,1} + \zeta_{m,2}\right) + d_{m,m}\zeta_{m,1}\zeta_{m,2}$$

where $\bar{d}$ is the average rate of diploidy in the population, $d_m = [h_m(1-p_m) + (1-h_m)p_m]$ and $d_{m,m} = \delta(1-2h)$.

The difference between the frequency of m within the diploid and the haploid subpopulation before selection is then given by

$$\Delta_r p_m^d = \mathrm{E}\left[\frac{d^*}{\bar{d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right] , \qquad \Delta_r p_m^h = \mathrm{E}\left[\frac{1-d^*}{1-\bar{d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right]$$

where E stands for the average over all zygotes (at the start of the generation). After simplification (and using the fact that $D_{m,m} = 0$), we have:

$$p_m^d = p_m + \frac{d_m}{\bar{d}}p_m q_m , \qquad p_m^h = p_m - \frac{d_m}{1-\bar{d}}p_m q_m$$

Using a similar argument, one obtains:

$$p_a^d = p_a + \frac{d_m}{\bar{d}}D_{ma} , \qquad p_a^h = p_m - \frac{d_m}{1-\bar{d}}D_{ma}$$

However, because D_{ma} is of order δ (since it should be zero when m is neutral), at QLE $d_m D_{ma}$ is of order δ^2 and can be neglected; therefore we have $p_a^d \approx p_a^h \approx p_a$. Similarly, one can show that $D_{ma}^d \approx D_{ma}$, $D_{ma}^h \approx (1-r_{ma})D_{ma}$ (the neglected terms being of order δ^2, see next section on linkage disequilibrium). The changes in frequency of m within the diploid and haploid sub-populations du to selection are given by:

$$\Delta_s p_m^d = \mathrm{E}\left[\frac{W_d}{\overline{W_d}}\frac{\zeta_{m,1}+\zeta_{m,2}}{2}\right] = \frac{h_a^* s_d}{\overline{W_d}}D_{ma}^d \approx \frac{h_a^* s_d}{\overline{W_d}}D_{ma}$$
$$\Delta_s p_m^h = \mathrm{E}\left[\frac{W_h}{\overline{W_h}}\zeta_m\right] = \frac{s_h(1-r_{ma})}{\overline{W_h}}D_{ma}^h \approx \frac{s_h(1-r_{ma})}{\overline{W_h}}D_{ma}$$

Putting everything together yields:

$$\Delta p_m = d_m\left(\overline{W_d} - \overline{W_h}\right)p_m q_m + \left(dh_a^* s_d + (1-d)(1-r_{ma})s_h\right)D_{ma}$$

Linkage disequilibrium at QLE

In the following we compute D_{ma} at quasi-linkage equilibrium, to the first order in s and to the first order in δ (one can show easily that $D_{ma} = 0$ at QLE if either $s = 0$ or $\delta = 0$). At the next generation, we have:

$$D_{ma}^{'} = d\overline{W_d}D_{ma}^{d*} + (1-d)\overline{W_h}D_{ma}^{h*}$$

where $D_{ma}^{d*} = \mathrm{E}_d^{'}\left[\left(X_m - p_m^{'}\right)\left(X_a - p_a^{'}\right)\right]$ and $D_{ma}^{h*} = \mathrm{E}_h^{'}\left[\left(X_m - p_m^{'}\right)\left(X_a - p_a^{'}\right)\right]$. $\mathrm{E}_d^{'}$ and $\mathrm{E}_h^{'}$ standing for the averages over gametes produced by diploids and by haploids, respectively. The term D_{ma}^{d*} can be written as:

$$\begin{aligned} D_{ma}^{d*} =& \mathrm{E}_h\left[\left(X_m - p_m^{d'} - (p_m^{'} - p_m^{d'})\right)\left(X_a - p_a^{d'} - (p_a^{'} - p_a^{d'})\right)\right] \\ =& D_{ma}^{d'} + \left(p_m^{'} - p_m^{d'}\right)\left(p_a^{'} - p_a^{d'}\right) \end{aligned} \tag{E.2}$$

Similarly in haploids:

$$D_{ma}^{h*} = D_{ma}^{h'} + \left(p_m^{'} - p_m^{h'}\right)\left(p_a^{'} - p_a^{h'}\right) \tag{E.3}$$

The terms $D_{ma}^{d'} = \mathrm{E}_d\left[\left(X_m - p_m^{d'}\right)\left(X_a - p_a^{d'}\right)\right]$ and $D_{ma}^{h'} = \mathrm{E}_h\left[\left(X_m - p_m^{h'}\right)\left(X_a - p_a^{h'}\right)\right]$ are the linkage disequilibrium within the population of gametes produced by diploids and haploids, respectively. The term $\left(p_m^{'} - p_m^{d'}\right)\left(p_a^{'} - p_a^{d'}\right)$ and $\left(p_m^{'} - p_m^{h'}\right)\left(p_a^{'} - p_a^{h'}\right)$ come from the fact that allele frequencies within the pool of gametes produced by diploids may differ from frequencies within gametes produced by haploids.

Before selection, the linkage disequilibrium within individuals that remain diploid is given by:

$$D_{ma}^{d} = \mathrm{E}\left[\left(X_m - p_m^{d}\right)\left(X_a - p_a^{d}\right)\right] \quad = \mathrm{E}\left[\frac{d^*}{\bar{d}}\frac{\zeta_{m,1} + \zeta_{m,2}}{2}\right] - (p_m^d - p_m)(p_a^d - p_a)$$

where E is the average over all zygotes. The term $(p_m^d - p_m)(p_a^d - p_a)$ comes from the fact that allele frequencies may change between the population of zygotes and the population of individuals that remain diploid; however, because $p_a^d - p_a$ is of order δ^2 (see above), this term can be neglected under our weak modifier assumption. Using the fact that under random mating $D_{ma,m}$ equals zero, one obtains:

$$D_{ma}^{d} = D_{ma}\left[1 + \frac{d_m}{\bar{d}}(1 - 2p_m)\right]$$

However, since D_{ma} and d_m are both of order δ, we have (to the first order in δ): $D^d_{ma} \approx D_{ma}$. Similarly, one obtains $D^d_{m,a} \approx D_{m,a} = 0$, while $D^h_{ma} \approx (1 - r_{ma})D_{ma}$.

After selection, we have:

$$\begin{aligned} D^d_{ma} &= \mathrm{E}\left[\frac{W_d}{\overline{W_d}}\frac{\zeta_{ma,1} + \zeta_{ma,2}}{2}\right] - \left(p^{d'}_m - p^d_m\right)\left(p^{d'}_a - p^d_a\right) \\ &= D^d_{ma} + h^*_a s_d\,(1 - 2p_a)\,D_{ma} - (h^*_a s_d D_{ma})\,(h^*_a s_d p_a q_a) \end{aligned}$$

However, because D_{ma}, s_d and s_h are of order α, we have to the first order in α $D^{d'}_{ma} \approx D_{ma}$ and $D^{h'}_{ma} \approx D_{ma}$. Finally, after recombination the linkage disequilibrium within gametes produced by individuals that underwent selection as diploids is multiplied by $(1 - r_{ma})$.

Therefore, to the first order in δ and α, the linkage disequilibrium within gametes produced by diploids ($D^{d'}_{ma}$) and within gametes produced by haploids ($D^{h'}_{ma}$) both equal $(1 - r_{ma})D_{ma}$. From equations E.2 and E.3 above, one obtains:

$$D'_{ma} = (1 - r_{ma})D_{ma} + d\overline{W_d}\left(p'_m - p^{d'}_m\right)\left(p'_a - p^{d'}_a\right) + (1 - d)\overline{W_h}\left(p'_m - p^{h'}_m\right)\left(p'_a - p^{h'}_a\right)$$

Therefore, to compute D_{ma} to the first order in δ and to the first order in s, it is sufficient to express $p'_m - p^{h'}_m$, $p'_m - p^{d'}_m$ to the first order in δ, when $s = 0$, and $p'_a - p^{h'}_a$, $p'_a - p^{d'}_a$ to the first order in s, when $\delta = 0$. At the equilibrium for p_a, this gives:

$$\begin{aligned} p'_m - p^{d'}_m &= \left[\left(\overline{W_d} - \overline{W_h}\right) - \frac{1}{\bar{d}}\right] d_m p_m q_m, \ , \quad p'_a - p^{d'}_a = -\frac{h^*_a s_d p_a q_a}{\overline{W_d}} \\ p'_m - p^{h'}_m &= \left[\left(\overline{W_d} - \overline{W_h}\right) + \frac{1}{1 - \bar{d}}\right] d_m p_m q_m, \ , \quad p'_a - p^{h'}_a = -\frac{s_h p_a q_a}{\overline{W_h}} \end{aligned}$$

Putting everything together finally yields:

$$D'_{ma} = (1 - r_{ma})D_{ma} + \left[(s_h - h^*_a s_d) + \left(\overline{W_d} - \overline{W_h}\right)(dha^* s_d + (1 - d)\,s_h)\right] d_m p_a q_a p_m q_m \quad \text{(E.4)}$$

Multilocus extrapolation

We now consider the case of a large number of loci at mutation-selection balance. For simplicity, deleterious alleles at all loci have the same selection and dominance coefficients, and we assume multiplicative effects of deleterious alleles at different loci on sensitivity to competition, so that in equation (3) of the main text, ω^i_h and ω^i_d become $\prod_{j=1}^{L}(1 + \rho\alpha X_j)$ and

$\prod_{j=1}^{L}(1+h\alpha(X_{j,1}+X_{j,2})+\alpha(1-2h)(X_{j,1}+X_{j,2}))$, where L is the number of loci, while X_j, $X_{j,1}$ and $X_{j,2}$ are indicator variables that equal 1 when an deleterious allele is present at locus j in a given haploid, or at the first or second haplotype of a given diploid. Through the following, we assume that deleterious alleles are mostly present in the heterozygous state, so that the sensitivity to competition of a diploid can be written as $\prod_{j=1}^{L}[1+\alpha h\,(X_{j,1}+X_{j,2})]$. Neglecting linkage disequilibria among selected loci, the average sensitivity to competition over all haploids is given by $\overline{\omega_h}=\prod_{j=1}^{L}[1+\rho\alpha p_j]$ (where p_j is the frequency of the deleterious allele at locus j), which is approximately $\exp[\rho\alpha\sum p_j]=\exp[\rho\alpha n]$, where n is the average number of mutations per haploid genome. Similarly, the average sensitivity to competition over all diploids is approximately $\overline{\omega_h}=\exp[2h\alpha n]$.

Finally, ω_h^i and ω_d^i may be written in terms of ζ_i variables. Neglecting terms involving products of ζ variables (that will generate terms involving linkage disequilibria between deleterious alleles, that we suppose negligible), we have:

$$\begin{aligned}
\omega_h^i &= \prod 1+\rho\alpha X_j{}_{j=1}^{L} \\
&= \prod 1+\rho\alpha(\zeta_j+p_j)_{j=1}^{L} \\
&\approx \prod(1+\rho\alpha p_j)_{j=1}^{L}\left(1+\rho\alpha\sum_j\zeta_j\right) \\
&\approx \mathrm{e}^{\rho\alpha n}\left(1+\rho\alpha\sum_j\zeta_j\right)
\end{aligned}$$

Similarly,

$$\omega_d^i \approx \mathrm{e}^{2h\alpha n}\left(1+h\alpha\sum_j\zeta_{j,1}+\zeta_{j,2}\right)$$

Plugging these expressions of ω_h^i, ω_d^i, ω_h and ω_d into equation 3 of the main text (Chapter 2), one obtains that the fitness of haploid individuals can be written as $W_h\approx\overline{W_h}\left[\sum_j s_h\zeta_j\right]$, with:

$$\begin{aligned}
\overline{W_h} &= 1+r_h\left[1+\frac{N}{K_h}\,\mathrm{e}^{\rho\alpha n}\left(\frac{1-d}{1-\beta+\beta\,\mathrm{e}^{\rho\alpha n}}+\frac{\gamma_{hd}d}{1-\beta+\beta\,\mathrm{e}^{2h\alpha n}}\right)\right] \\
s_h &= -\rho\alpha r_h\frac{N}{K_h}\,\mathrm{e}^{\rho\alpha n}\left(\frac{1-d}{1-\beta+\beta\,\mathrm{e}^{\rho\alpha n}}+\frac{\gamma_{hd}d}{1-\beta+\beta\,\mathrm{e}^{2h\alpha n}}\right)
\end{aligned} \tag{E.5}$$

Similarly, one obtains $W_d\approx\overline{W_d}\left[\sum_j s_d(\zeta_{j,1}+\zeta_{j,2})\right]$ with

$$\begin{aligned}\overline{W_d} &= 1 + r_d \left[1 + \frac{N}{K_d}\, \mathrm{e}^{2h\alpha n} \left(\frac{\gamma_{dh}(1-d)}{1-\beta+\beta\, \mathrm{e}^{\rho\alpha n}} + \frac{d}{1-\beta+\beta\, \mathrm{e}^{2h\alpha n}}\right)\right] \\ s_d &= -\alpha r_h \frac{N}{K_h}\, \mathrm{e}^{2h\alpha n} \left(\frac{\gamma_{dh}(1-d)}{1-\beta+\beta\, \mathrm{e}^{\rho\alpha n}} + \frac{d}{1-\beta+\beta\, \mathrm{e}^{2h\alpha n}}\right)\end{aligned} \tag{E.6}$$

Neglecting linkage disequilibria between deleterious alleles, equation E.1 still holds, so that the mean number of deleterious alleles per haplotype at equilibrium satisfies:

$$\left[dhs_d + (1-d)\, s_h\right] n + U = 0 \tag{E.7}$$

where U is the deleterious mutation rate per haplotype. Equation E.7 must be solved numerically, together with $d\overline{W_d} + (1-d)\overline{W_h} = 1$ (where $\overline{W_d}$ and $\overline{W_h}$ are given by equations E.6 and E.5 above) to obtain N and n at equilibrium.

Finally, the change in frequency of the modifier is given by:

$$\Delta p_m = d_m \left(\overline{W_d} - \overline{W_h}\right) + \sum_j \left(dhs_d + (1-d)\,(1-r_{mj})\, s_h\right) D_{mj}$$

where r_{mj} is the recombination rate between the modifier and locus j, while D_{mj} is the linkage disequilibrium between these two loci, obtained by solving equation E.4 at equilibrium.

Abstract

Sexual reproduction leads to an alternation between haploid and diploid phases, whose relative length varies widely across taxa. The proportion of the life cycle spent in the haploid and diploid phase has important consequences on a number of adaptive processes. This thesis combines theoretical approaches exploring the effect of genetic and ecological factors on the evolution of life cycles, and experimental work on the effects of ploidy on the evolution of reproductive isolation between populations.
The theoretical part consisted in integrating ecological components into genetic models for the evolution of life cycles. In particular, I explored the interplay between niche differentiation between haploids and diploids (known to favour the maintenance of biphasic life cycles, involving development in both phases) and the effect of deleterious alleles (known to favour either haploid or diploid life cyles). While niche differentiation (or more simply intrinsic fitness differences between phases) stabilizes biphasic cycles, the presence of deleterious alleles often lead to evolutionary branching and to the stable coexistence of alleles coding for haploid and diploid cycles. Branching is prevented, however, when temporal environmental fluctuations are included into the model.
The experimental part consisted in comparing the dynamics of reproductive isolation between small populations of haploid and diploid yeasts with elevated mutation rate. The results show that while haploid hybrids tend to have a lower fitness than their parents, diploid hybrids benefit from heterosis in the F1 generation, and still have a higher fitness than the diploid homozygous parents in the F2 generation. However, the variance of hybrid fitness was much higher in haploids, with the production of some highly fit genotypes.

Résumé

La reproduction sexuée conduit à l'alternance d'une phase haploïde et d'une phase diploïde, dont la durée relative est très variable entre taxons. La proportion du cycle de vie passée en phase haploïde et en phase diploïde a d'importantes conséquences sur de nombreux processus adaptatifs. Cette thèse combine des approches théoriques qui explorent l'effet de facteurs génétiques et écologiques sur l'évolution des cycles de vie, et un travail expérimental sur l'effet de la ploidy sur l'évolution de l'isolement reproducteur entre populations.
La partie théorique a consisté à intégrer des composantes écologiques dans des modèles génétiques pour l'évolution des cycles de vie. En particulier, j'ai exploré l'interaction entre la différenciation de niche entre haploïdes et diploïdes (qui favorise le maintien de cycles biphasiques, impliquant le développement des deux phases) et l'effet d'allèles délétères (qui favorisent soit l'haploïdie, soit la diploïdie). Tandis que la différentiation de niche (ou plus simplement, une différence de valeur sélective intrinsèque entre phases) stabilise les cycles intermédiaires, la présence d'allèles délétères conduit souvent à un branchement évolutif, avec la coexistence stable d'allèles codant pour l'haploïdie et la diploïdie. Cependant, des fluctuations temporelles de l'habitat permettent d'empêcher ce branchement et de stabiliser les cycles biphasiques.
La partie expérimentale a consisté à comparer la dynamique de l'isolement reproducteur entre petites populations de levure haploïdes et de diploïdes avec de taux de mutations élevés. Les résultats montrent que tandis que les hybrides haploïdes ont une valeur sélective plus faible que leurs parents, les hybrides diploïdes bénéficient du phénomène d'hétérosis en génération F1, et ont encore une valeur sélective plus élevée que leurs parents en génération F2. La variance de la valeur sélective des hybrides était cependant beaucoup plus élevée chez les haploïdes, avec la production de certains génotypes très performants.