Genetic and evolutionary analysis of diversification and reproductive isolation in yeast

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Gönensin Ozan Bozdag
Plön, December, 2015

Contents

Summary of the dissertation

In this thesis, I examine the mechanisms of evolution at different levels, from evolutionary conflict between selfish genes within a single individual (Chapter 1), through social evolution acting within a species (Chapter 2), to genetic divergence and incompatibility between closely related species (Chapters 4 & 5). The thesis therefore investigates how tiny genetic differences occurring in individuals accumulate and produce discontinuous groups.

The first chapter explores an interesting form of natural selection, acting independently on different genomes within the same cell. Natural selection can act at the level of individual genes: an allele that promotes its own transmission can increase in frequency despite reducing the fitness of the rest of the genome (Dawkins 1978). This phenomenon, known as intra-genomic conflict (Hurst 1992), has long been hypothesized to drive evolution, forcing different lineages to adapt to the genes within their own genomes and therefore causing their genomes to diverge, and potentially, to become incompatible types. Here I test whether intra-genomic conflict drives evolutionary change by evolving yeast populations in the laboratory, to see if intra-genomic conflicts would lead genomes in independent populations to become incompatible. After allowing populations to evolve under two treatments of strict vertical transmission of mitochondria, or mixed horizontal/vertical transmission, I tested the evolutionary changes in interactions between mitochondrial and nuclear genomes in the continuum of mutualism and selfishness. As predicted, increasing the independence of mitochondria from their hosts (by increasing outbreeding) reduced the evolved fitness benefit that mitochondria provided to their un-evolved hosts. The results presented in this chapter hint that intra-genomic conflicts can speed up the evolution of cyto-nuclear reproductive isolation between allopatric populations.

The second chapter also looks at whether conflict, this time between individuals in a population rather than between genes within an individual, can lead to diversification, not just in the form of single nucleotide replacements but at the under-examined form of copy number variation. The sharing of the secreted enzyme invertase (encoded by *SUC* genes) by yeast cells is a well-established laboratory model used to test social conflict models. Moreover, yeast populations vary in *SUC* gene copy numbers. The observed copy number variation has been suggested to be the result of natural selection acting at the level of social conflict. However, genetic variation might instead be explained by adaptation of different populations to different local availabilities of sucrose, the substrate for the *SUC* gene product. Here, I

provide evidence showing that the variation observed in natural populations is better explained by the environmental adaptation hypothesis rather than the social conflict hypothesis (Bozdag & Greig 2014).

The final chapters take a different approach: rather than at bottom-up approach testing how natural selection (intra-genomic conflict, social conflict and environmental adaptation) may drive diversification or divergence into different types, I take a top-down approach, testing which genetic changes are responsible for the discontinuities between already established types (between two species of yeast, *S. cerevisiae* and *S. paradoxus*).

In chapter three, I look at how nucleotide sequence variation can accumulate to such an extent that it prevents the segregation of diverged chromosomes, causing sexual incompatibilities between established types (different species). Here, I have genetically manipulated interspecific hybrids with the aim of inducing crossovers between their diverged chromosomes. This manipulation increased recombination rates significantly compared to un-manipulated hybrids. Increased recombination caused a remarkable increase in the fertility of the yeast hybrids, from 0.5% viable gametes to over 30% viable gametes. I conclude that the reduced recombination in interspecific hybrids is responsible for at least one third of the hybrid gamete death.

And finally in chapter four, I determine how individual genetic changes can cause incompatibility, potentially preventing certain individuals from breeding together and therefore allowing the accumulation of further genetic changes. Here I assayed a hybrid strain for two-locus incompatibilities (Bateson-Dobzhansky-Muller genic incompatibilities) between the two parental yeast species. If such genic incompatibilities exist, the proportion of viable offspring bearing the hybrid combination for a pair of loci should be significantly lower than the proportion bearing the *non*-hybrid (i.e. parental) combination. To check this, I exploited the improved viability of interspecific hybrids obtained in the chapter three. As a result, I present seven putative BDMI regions between the two sibling species of yeast.

Zusammenfassung der Dissertation

Diese Dissertation behandelt Evolutionsmechanismen auf drei verschiedenen Ebenen. Im ersten Kapitel untersuche ich evolutionäre Konflikte zwischen egoistischen Genen desselben Genoms, im zweiten Kapitel untersuche ich Aspekte der sozialen Evolution innerhalb einer Art, und im vierten und fünften Kapitel analysiere ich den Einfluss von genetischer Divergenz auf die Inkompatibilität zwischen zwei nah verwandten Arten. Zusammenfassend zeigt diese Arbeit, wie kleine Unterschiede innerhalb des Genoms sich zu klar definierten Arten akkumulieren können.

Das erste Kapitel beleuchtet den Einfluss natürlicher Selektion auf die verschiedenen Genome innerhalb einer Zelle. Zum Beispiel kann ein Allel, das seine eigene Vererbung fördert, seine Frequenz in der Population erhöhen, obwohl es gleichzeitig die Fitness des restlichen Genoms reduziert (Dawkins 1978). Es wird schon seit langem angenommen, dass dieses Phänomen, auch inner-genomischer Konflikt genannt (Hurst 1992), die Evolution antreibt, und dazu führt, dass das Genom sich seinen eigenen Genen anpassen muss und so potenziell mit sich selbst inkompatibel wird. Ich benutze Bäckerhefe (*Saccharomyces sp.*), um zu testen, ob inner-genomische Konflikte zu evolutionären Veränderungen und zu genetischen Inkompatibilitäten von Populationen führen können. Dazu unterziehe ich Populationen entweder einer rein vertikalen Vererbung von Mitochondrien, oder einer gemischt vertikalen und horizontalen Vererbung, und analysiere die Veränderungen in der Interaktion zwischen Mitochondrien und dem nuklearen Genom in Bezug auf Mutualismus und Eigennutz („selfishness') der Gene. Wie erwartet nimmt der Fitnessvorteil, den die evolvierten Mitochondrien für ein nicht evolviertes Wirtsgenom haben, mit der Auskreuzungsrate („outbreeding rate") ab. Meine Ergebnisse deuten darauf hin, dass inner-genomische Konflikte die Evolution von cyto-nuklearer, reproduktiver Isolation zwischen allopatrischen Populationen beschleunigen können.

Das zweite Kapitel untersucht ebenfalls, ob Konflikte (diesmal zwischen Individuen derselben Population und nicht zwischen Genen desselben Individuums) zu Diversifizierung führen können, in Form von Punktmutationen, aber auch in Form von Kopienzahlvariation ("copy number variation") von Genen.

Die gemeinsame Benutzung von Enzyminvertasen (von SUC-Genen kodiert) in Hefepopulationen ist ein etabliertes System zum Testen von sozialen Evolutionsmodellen. Hefepopulationen variieren in der Anzahl der Kopien an SUC-Genen. Es wird angenommen,

dass diese Kopienzahlvariation ein Ergebnis natürlicher Selektion aufgrund von sozialem Konflikt ist. Alternativ dazu könnte diese Variation jedoch auch ein Ergebnis von Adaption an die unterschiedliche Verfügbarkeit von Saccharose in verschiedenen Umwelten sein. Ich zeige hier, dass die Kopienzahlvariation besser mit einer Umweltanpassung als mit sozialem Konflikt übereinstimmt (Bozdag & Greig 2014).

In den letzten beiden Kapiteln verwende ich eine andere Methode. Anstelle einer bottom-up Analyse, wo ich den Einfluss der natürlichen Selektion (inner-genomische und soziale Konflikte und Umweltanpassung) auf die Divergenz von Populationen teste, verwende ich hier eine top-down Analyse, mit der ich teste, welche genetischen Veränderungen für die Diskontinuität zwischen bereits etablierten Arten (*S. cerevisiae* und *S. paradoxus*) verantwortlich sind.

Im dritten Kapitel untersuche ich wie die Nukleotidsequenzen zweier Arten zu einem solchen Ausmaß divergieren können, dass ein ordnungsgemäßes Segregieren der Chromosomen in der Meiose verhindert wird, was zu sexueller Inkompatibilität führen kann. Hierfür habe ich zwischenartliche Hybride aus *S. cerevisiae* und *S. paradoxus* so genetisch manipuliert, dass die Chromosomen der beiden Arten wieder Cross-over miteinander bilden können. Diese Manipulation hat im Vergleich zu nicht manipulierten Hybriden die Rekombinationsrate signifikant erhöht, und die erhöhte Rekombinationsrate führte zu einem bemerkenswerten Anstieg in der Fertilität der Hybriden, von 0.5% lebensfähigen Gameten bis zu über 30%. Dies bedeutet, dass die reduzierte Rekombinationsrate in zwischenartlichen Hybriden der Grund für mindestens ein Drittel der Hybridensterblichkeit ist.

Im vierten Kapitel untersuche ich wie individuelle genetische Veränderungen zu Inkompatibilitäten führen, die bestimmte Individuen davon abhalten sich miteinander fortzupflanzen, und damit zu einer Ansammlung von weiteren genetischen Unterschieden führen. Hierfür überprüfe ich eine zwischenartliche Hybridkreuzung nach zwei-Merkmals-Inkompatibilitäten („two-locus Bateson-Dobzhansky-Muller incompatibilities"). Die Vorhersage ist, falls solche Inkompatibilitäten bestehen, dass der Anteil an lebensfähigen Nachkommen mit ebendieser Hybrid-Kombination an Merkmalen signifikant niedriger ist als der Anteil an Nachkommen mit einer Nicht-Hybrid-(also einer parentalen) Kombination. Um dies zu überprüfen, nutzte ich die verbesserte Lebensfähigkeit der Hybriden aus Kapitel 3 und konnte zeigen, dass die beiden Hefearten sieben mögliche Bateson-Dobzhansky-Muller Regionen aufweisen.

All life on earth has a single common ancestor. From this single ancestor, a huge diversity of different types of organisms has evolved. These types are not continuously varying, but fall into discontinuous groups. This thesis looks at how selection acts at different levels (genomes within an individual, between genomes of members of the same species, and between genomes of different species) and how small genetic differences occurring in individuals accumulate and thus produce discontinuous groups. Understanding these arguments requires an introduction to both how genomes vary, and how different events of evolution act on this variation. Within this framework, the following pages firstly summarize types of mutations that introduce evolutionarily significant variation into populations. This is necessary because variation studied in this thesis looks at different types of mutations (i.e., single nucleotide replacements, copy number changes). The introduction then moves to the evolutionary events (i.e. drift and selection) that fix these mutations within populations. At that point, it is necessary to look at different forms of natural selection (i.e. intra-genomic conflicts, social conflicts, effect of physical environment) acting on these mutations, because the thesis examines the mechanisms of evolution at different levels, from evolutionary conflict between selfish genes within a single individual (Chapter 1), through social evolution acting within a species (Chapter 2). Then, the introduction will summarize the forms of reproductive barriers that prevent inter-breeding between diverged species, as this study also looks at how fixed genetic differences function in forming reproductive barriers between species (Chapters 3 & 4). The study investigates all these concepts using two yeast species: *S. cerevisiae* and *S. paradoxus*. Therefore the general introduction finally revisits some of these concepts (i.e. types of mutations and their adaptive importance, and the models of reproductive barriers) very briefly in the context of the two yeast species at the end of this introduction.

Sources and types of genetic variation

The ultimate source of genetic variation is replication error. Alterations to DNA that are not repaired can escape into newly born individuals. These alterations arise mostly as polymerase errors during DNA replication or, rarely, as chemical degradation before replication (e.g. Preston *et al*. 2010). Mutations, which have different rates in different parts of the genome and in the genomes of different species, are essential for the evolutionary change to occur (Ayala & Fitch 1997; Sniegowski *et al*. 2000).

These errors are random events with respect to sequence but definitely not to function. Most of them have neutral or deleterious effects on individuals (reviewed in Eyre-Walker & Keightley 2007). A very low proportion of these random mutations are beneficial for a given environmental condition (e.g. Peris *et al*. 2010; Hall & Joseph 2010; but also see Rutter *et al*. 2010). A mutation may promote the transmission of an allele at the level of that allele, the genome it resides in, the organism that genome resides in, or the population that organism resides in. The term beneficial (or deleterious) depends on which level we are looking at. Different mutations can affect the likelihood that they will be transmitted to the next generation. The adaptive genotypes for that novel environmental condition will be selected from that preexisting variation already found in the population (Luria & Delbrück 1943; Lederberg & Lederberg 1952; Zhang *et al*. 2011). Mutation rates vary across different species (Lynch 2010). Moreover, rate of mutations can evolve in an adaptive way (reviewed in Sniegowski *et al*. 2000).

Three types of mutations are subject to this thesis study: Single nucleotide mutations, copy number changes, and chromosomal rearrangements (CRs).

Single nucleotide mutations: Single nucleotide variations are differences at specific nucleotide positions between genomes (or haplotypes within a heterozygous individual). Natural populations usually contain high numbers of single nucleotide polymorphisms (Avise 1994). For example, between two randomly chosen humans there is an average of one single nucleotide substitution in every 1000-2000 nucleotides (Sachidanandam *et al*. 2001). Single nucleotide substitutions can alter the function and/or abundance of the proteins. Therefore, the introduction of single nucleotide mutations into populations is very important for evolutionary change. These simple mutations can alter the fitness (survival and reproductive success) of individuals providing raw material for evolutionary adaptation (Hall & Joseph 2010; Peris *et al*. 2010). When gene flow between populations is limited, different variants at a single nucleotide position can fix in each population, resulting in genetic divergence between populations (e.g. Colosimo *et al*. 2005).

Copy number variation: Replication and repair errors that occur during sexual or asexual reproduction (i.e. unequal crossover events) or by the activity of retrotransposons can increase the copy number of a gene and as a result can be important for evolutionary change (Zhang 2003; Hurles 2004). Duplication of a gene may cause a fitness change as it can increase the

quantity of a gene product, increasing fitness if that product is limiting, or decreasing fitness due to the metabolic cost of overproduction or by disrupting stoichiometry (Kondrashov & Kondrashov 2006). Certain stress conditions (i.e. heavy metal stress) have been shown to select for individuals with increased copies of resistance genes (i.e. heavy metal transporter gene duplications) that confer resistance to such stressors (Yang *et al*. 2010; Chow *et al*. 2012). Increasing the copy number of a gene, apart from potentially conferring rapid adaptation (reviewed in Kondrashov 2012), can also be important for evolutionary innovation and further diversification. Since duplication relaxes the purifying selection either on the novel or on the ancestral copy of a gene, one of these copies can accumulate mutations at a faster rate, which may in the end evolve proteins with completely novel functions (e.g. Zhang *et al*. 1998). This is especially important because evolution of novel gene families is mostly driven by such neo-functionalization events of duplicated gene copies (reviewed in De Grassi *et al*. 2008). Additionally, if populations divide into different subpopulations after the duplication event, and both populations lose the function of the duplicate gene at different loci, then hybrids may suffer from reduced fitness (Bikard *et al*. 2009; Maclean & Greig 2010). Described further below.

Chromosomal variation: Many closely related species exhibit differences in the arrangement of their chromosomes. That is, the colinearity of chromosomal sections can vary between populations (White 1978). Chromosomal rearrangements (CRs) such as i) inversions and ii) translocations are important sources of genetic variation and are critical for the evolutionary diversification (e.g. Kirkpatrick & Barton 2006).

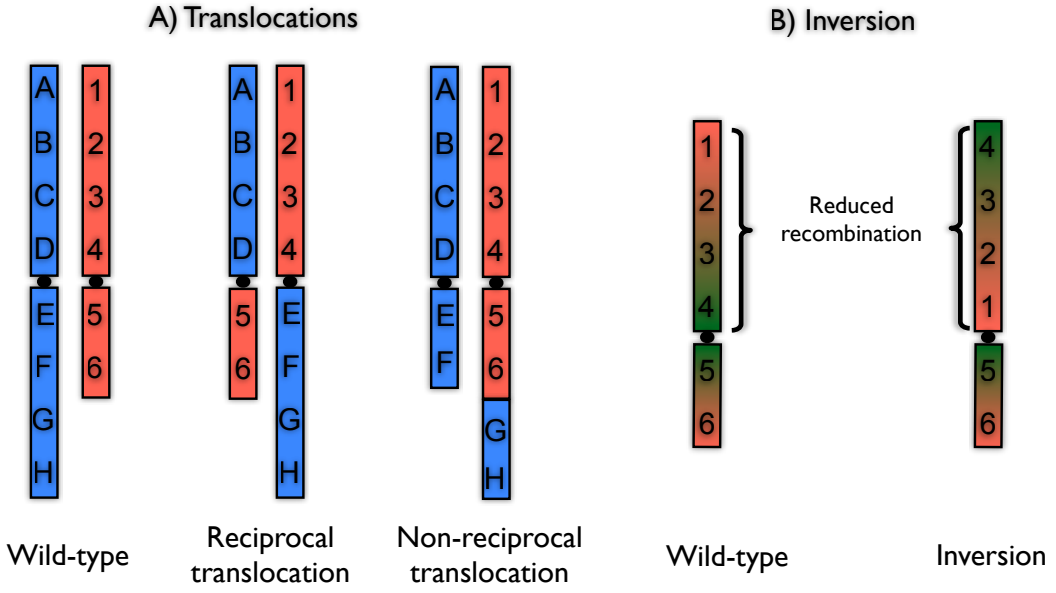


Figure 1. Chromosomal rearrangements. A) Translocations can be reciprocal (bidirectional) or non-reciprocal (unidirectional). B) Inversions are defined as the inverted order of genes located along a chromosome. Failure of synapses formation between inverted positions of heterozygous chromosomes reduces the recombination rates (Kirkpatrick 2010). Inverted regions can be fixed for different alleles in different populations. Those alleles can be important in conferring local adaptation. Therefore, a reduction in recombination between inverted regions at population level would prevent maladaptive gene flow between populations and keep co-adapted genes together.

i) Translocations are reciprocal (bidirectional) or non-reciprocal (unidirectional) movement of large chromosome segments to any other position within a genome (Figure 1A). Since translocations can lead to intrinsic problems during meiosis (see below), they are likely to be selected against. However chromosomal rearrangements may have adaptive roles that increase their spread in populations (Dunn *et al*. 2013).

ii) 'Inversions' reverse the order of genes along a chromosome (see Figure 1B; reviewed in Kirkpatrick 2010). Genes at the break points or within inverted regions can have altered expression (Laayouni *et al*. 2007). Such expression differences can cause genetic diseases (Castermans *et al*. 2007). More relevant to the divergence of different populations, inverted regions reduce crossing-over events in heterozygotes (i.e. when one haplotype contains the inversion and the other does not). If the inversion has fixed in one population, low levels of gene flow prevents recombination to break up locally adapted combinations as they would without the inversion (e.g. Joron *et al*. 2011; Wadsworth *et al*. 2015). Similarly inversions can reduce the gene flow *between* closely related species. This also keeps locally adapted alleles of a species together (Noor *et al*. 2001). These locally adapted alleles located within inverted

regions can increase fitness via epistasis or via additive effects (Charlesworth & Charlesworth 1979; Kirkpatrick & Barton 2006). Thus species that have adapted to different environmental conditions do not suffer from maladaptive gene flow (Mayr 1970). The *Drosophila pseudoobscura* populations that are fixed for diverse chromosome-III right-arm inversions provide evidence supporting the importance of inversions at local adaptation (Dobzhansky 1949; Wallece *et al*. 2013; Fuller *et al*. 2014). Chromosomal inversions that confer local adaptation have been shown in ecotypes of yellow monkeyflower (Lowry & Willis 2010); different inversions conferred local adaptation to different levels of water availability (Lowry & Wills 2010).

Population genetic models analyzing fixation dynamics of chromosomal mutations (e.g. inversions) are different from models analyzing fixation dynamics of other types of mutations (e.g. beneficial single nucleotide replacements). An individual carrying a beneficial single nucleotide mutation would mate with any other individual found in that population with no reduction in reproductive fitness. The situation is different for population dynamics under chromosomal mutation models. Mating between two individuals with different karyotypes can be problematic as gametes they produce through meiosis may suffer from gene imbalances (e.g. missing or additional genes can be deleterious) (White 1978; see Fig. 2 below). An individual with a *de novo* chromosomal mutation will encounter individuals that have wild-type karyotypes in its mating attempts (Bengtsson & Bodmer 1976). The heterozygotes formed from such outbreeding events may have reduced reproductive fitness in comparison to homozygotes (Bengtsson & Bodmer 1976; Kirkpatrick & Barton 2006). This initial reproductive fitness cost makes the fixation of chromosomal mutations in the population less likely, even if such chromosomal mutations are significantly beneficial outside of meiosis. Different models are proposed to explain the population dynamics of such chromosomal evolution events (reviewed in Rieseberg 2001; Hoffmann & Rieseberg 2008). Early models suggested that the fixation is mainly achieved by formation of new populations via bottleneck events (Lande 1979; Hedrick 1981; Rieseberg 2001), increasing the likelihood of mating between individuals with the same karyotype. Recent models offer selection-based scenarios those require no aid from genetic drift (Hedrick 1981; Kirkpatrick & Barton 2006; Hoffmann & Rieseberg 2008). Chromosomal mutations, i.e. inversions, can spread through the population when the selective advantage of locally-adapted alleles within the inversion is strong enough to counter the reproductive cost (Kirkpatrick & Barton 2006). Further, mixed effect models, where drift and natural selection are acting in combination, can also aid the initial increase of a chromosomal mutation in frequency through the population (Hedrick

1981; Stathos & Fishman 2014). Once the frequency of a chromosomal mutation surpasses to 50% in a population, then likelihood of fixation is highly possible since most mates will have the same karyotype and the reproductive burden shifts to the wildtype arrangement (Bengtsson & Bodmer 1976). Understanding the spread of chromosomal mutations through a population is still an active area of research (Hoffmann & Rieseberg 2008; Faria & Navarro 2010; Stathos & Fishman 2014).

In summary, chromosomal rearrangements can alter gene expression levels of nearby genes, and more importantly, keep locally adapted gene combinations intact and thus become very important for ecological adaptation of populations (Navarro & Barton 2003; Kirkpatrick & Barton 2006; Pérez-Ortîn *et al*. 2002). As a result, CRs directly contribute to accumulating evolutionary diversification of populations (reviewed in Brown & O'Neill 2010).

The causes of evolutionary change

Mutations inevitably happen and provide raw material for evolutionary change. Then random and non-random events work on these variation and lead evolutionary change to occur. Different mutations can affect the likelihood that they will be transmitted to the next generation. When mutations are responsible for the change in the frequency of an allele, we call this effect "natural selection" – a non-random event. When alleles change in frequency without regard to the effect of the mutation, we call this process "genetic drift" – a random event. Further, crossing-over and random segregation of chromosomes are also important sources of randomness that affect the outcome of evolutionary change. Through the action of all these evolutionary events, variants in different populations can reach to fixation, and as a result populations diverge at the genetic level.

i) Random events

Evolutionary events other than natural selection are classified as random events of evolution (Barton *et al*. 2007). The ones will be summarized here are crossing-over, random segregation, and genetic drift (Barton *et al*. 2007).

a) Crossing-over and random segregation

Crossing-over greatly increases the adaptive potential of sexual populations (Mayr 1970). Crossovers are induced by double-strand breaks made at different nucleotide sequences that are distributed through out the genomes of sexual organisms. Multiple breaks are made on every chromosome but only a few of these breaks resolve as crossovers. The remaining

breaks resolve as gene-conversion events (Mancera *et al*. 2008). The random decision made at this point brings different parental alleles together (also see: Bishop & Zickler 2007; Whitby 2005; Baudat & de Massy 2007). When two loci occur on different chromosomes, random segregation causes all combinations of parental alleles to be inherited with the same likelihood. However, when two loci occur on the same chromosome, parental types will co-segregate unless a crossover event occurs between the two loci, in which case the offspring will inherit a new combination of alleles. The two alleles forming a novel interaction may then increase the fitness of the offspring significantly in comparison to the fitness values achieved by the interaction of old parental combinations (Otto & Barton 1997). Thus resolution decision more or less random in nature is likely to have significant outcomes for the evolutionary change/adaptation (Felsenstein 1974; but see: Chen *et al*. 2008; Rattray *et al*. 2014; Kauppi *et al*. 2004).

Segregation of chromosomes is an important random event in evolutionary change. To put it simply, two such loci exemplified above may be found at different chromosomes. In this case, segregation of the two chromosomes into the progeny is also a random event ('Law of independent assortment' of Mendel. The random assortment can form novel inter-chromosomal allelic-combinations those were not found in the early generations (Monaghan & Corcos 1984). Combination of such unique sets of genes located at different chromosomes can form favorable variants that will be important for evolutionary change.

In sum, random decision of crossover resolution and following event, random segregation of that recombined (or non-recombined) chromosomes into the progeny, by bringing favorable allelic interactions together, may increase the adaptive potentials of populations (Barton & Charlesworth 1998). Nevertheless, it is important to note that the population models and evolutionary significance of these two random processes are more elaborate than the summary presented here (see: Charlesworth *et al*. 1977; Barton 1995; Otto & Lenormand 2002).

b) Random genetic drift

Random sampling of alleles from a gene pool into subsequent generations is designated as random genetic drift (Herron & Freeman 2013). Drift is a random event in evolution since its effect on allele frequencies is independent from the relative fitness of those alleles (Barton *et al.* 2007). Drift tends to decrease variation within populations (Eimes *et al*. 2011); since rare alleles are less likely to be sampled, they are often lost under drift. However, given the stochastic nature of drift, in some cases initially rare alleles can increase in frequency and

even fix due to drift alone. Given that these changes are entirely random, they are unlikely to be replicated across populations and as a result, drift tends to increase variation between populations (Li 1997, Barton *et al*. 2007).

The effect of drift in reducing within population variation has been shown in laboratory experiments (e.g. Li & Roossinck 2004). Drastic reduction in population size of the northern elephant seal in the early 20th century is a well-known example on the effects of drift on the genetic variation (Stewart *et al*. 1994); the population bottleneck has drastically reduced the genetic variation of the elephant seals (Hedrick 1995). Similarly, establishment of a new population by the dispersal of a few individuals can rapidly fix genetic differences between new (founder) and old populations. For instance, founder events in *Teleogryllus* cricket populations led to the evolution of cricket populations with no singing ability (Tinghitella *et al*. 2011).

ii) Non-random event: Natural selection and its forms

a) Natural selection allows adaptation and divergence

The probability that a mutation occurs is entirely random with respect to the effect of the mutation in fitness. Given that random changes are more likely to impair rather than improve function, most mutations are likely to be deleterious or neutral. Since deleterious mutations have, by definition, a reduced probability of being transmitted to the next generation, they will likely be lost from the population. In rare cases, a mutation may be beneficial and will therefore, by definition, be more likely to be transmitted to the next generation (Li 1997; Otto & Whitlock 2005; Charlesworth 2015). Relative differences in fitness of genotypes allow the process of natural selection (Haldane 1927; Gerrish & Lenski 1998; Orr 2009). As beneficial alleles increase in frequency and deleterious alleles decrease in frequency, this process of natural selection increases adaptation (fitness in a given environment) over time (Herron & Freeman 2013).

The higher probability of survival and reproductive success of certain genotypes *in a given environment* makes natural selection a non-random event. However, since the environmental conditions constantly change thorough time, evolution by natural selection would not prepare populations for future conditions unless those conditions are not predictable (Herron & Freeman 2013). Furthermore, environmental conditions vary through space. Sub-populations adapting to diverse physical conditions may diverge from each another in time. Hence, a certain genotype adapted to a certain environment could be maladaptive in a different

environment (reviewed in Sgro & Hoffmann 2004). Varying fitness performance of a genotype in different environments is known as the genotype-by-environment (GxE) interactions. For instance, bacteria populations that have high fitness in low temperature environment have been shown to perform worse on a high temperature (Bennett & Lenski 2007). Similarly, *Drosophila melanogaster* populations adapted and increased in fitness in a novel stress-environment had a reduction/trade-off in its fitness in ancestral non-stress environment (Shirley & Sibly 1999). Therefore, it is common to observe populations those adapting to diverse external physical environments to diverge from each on the fitness landscape.

b) Selection occurs at different levels

Natural selection requires three components: Phenotypic variation between the reproductive units (e.g. individuals or replicators), inheritance of that variation from one generation to the next, and differential reproductive success (Godfrey-Smith 2009). Any reproducing unit that has these three components can be a potential unit or target of the natural selection (Godfrey-Smith 2009). The unit of selection is argued to be above (e.g. groups) or below (e.g. genes) the 'individual level' (see Box 3 in Pigliucci 2008). If the units of natural selection can take diverse forms (e.g. genes vs. individuals), as a consequence, the forms that natural selection act on those levels themselves can also be diverse.

c) Selection at different levels causes intra-genomic conflict

For instance 'genes as the units of selection' argument proposes that natural selection can target genes within an individual (Hurst 1992). A particular gene, by "self-promoting" its transmission, can spread through a population in a non-Mendelian fashion (e.g. Goddard & Burt 1999). This means that, such a gene, instead of equally segregating into the progeny relative to its alternative form (e.g. 50% Mendelian inheritance), can over-replicate and thus get into more than 50 % of the progeny (Burt & Trivers 2006). Its increase in frequency in the population does not represent a Darwinian process in which genes increase in frequency in populations simply because they are beneficial in the context of organismal fitness (Orgel & Crick 1980). If not neutral, the spread of such a gene through the population may be negative for the fitness of the individual because synthesizing non-beneficial molecules found in such high copy numbers would be costly (e.g. Pasyukova *et al*. 2004; Harrison *et al*. 2012). Although selfish genetic elements can harm their hosts by reducing their fitness, these replicators can still spread through a population (e.g. Goddard *et al*. 2001; reviewed in Hurst & Werren 2001). Further, existence of such a harmful genetic element would directly

decrease the fitness of all the other genes that are found within the same host (e.g. individual or genome) (reviewed in Werren 2011). As a result, genes at other loci within that individual are likely to be in conflict with such genetic elements (Hurst 1992). Such an intra-genomic conflict can lead to and "arms race" between selfish, harmful-to-the-host mutations, and compensatory mutations (Kleene 2005). Thus natural selection would increase the frequency of genes that counteract the harmful effects of the selfish genes (Hurst 1992). Such a perpetual arms race seen between genes within an individual can fix co-evolving mutations within that population (reviewed in Chou & Leu 2015). Following this, since allopatric populations would accumulate co-evolving mutations independent from one another, such intra-genomic conflicts are suggested to diverge populations from each other (see Chapter 1). This different view of natural selection argues that such perpetual intra-genomic conflicts taking place internally within the individuals are very important in evolutionary adaptation (Dawkins 1978). Under this view, the target of selection is the genes (or replicators) rather than individuals (Pigliucci 2008; Godfrey-Smith 2009). Moreover, this view suggests that the importance of external environment is less significant in natural selection than the internal environment of an individual where the conflicts take place and these conflicts are suggested to be more decisive in diversification (Hurst 1992; see for some criticism: Kaplan & Pigliucci 2001; Godfrey-Smith 2009).

d) Selection at different levels causes social conflict

The social conflict model of evolution can drive diversification within populations (e.g. Eldar 2011). Social interactions may take place between members of a population in a way that can determine the relative fitness of individuals in a dynamic fashion. Such interactions would then be the target of natural selection when the fitness of a genotype depends on the frequency of competing genotypes in the same population (Bell 2008). As a result, natural selection, normally acting on varying genotype-by-environment interactions, this time would also act on varying social interactions seen between individuals of a population. These interactions are known as the genotype-by-genotype (GxG) interactions. This idea of seeing social interactions as the evolutionary background is similar to the classical model of natural selection that takes the external physical environment as the background of evolutionary change. For instance, Smith *et al*. has shown that the fitness of a plant genotype (recipient) has been shown to be affected negatively or positively depending onto the genotypes (actors) of its neighbors (1970). If the recipient genotype coexists with a neighbor that has a positive affect on its fitness, then natural selection will increase the frequency of the recipient (reviewed in Mitri & Foster 2013). On the contrary, if the neighbor has a negative effect on

the fitness of the recipient, then natural selection is expected to decrease the frequency of the recipient genotype through time. Therefore, coexistence of different genotypes with diverse social interactions can evolve in populations through distinct outcomes and thus increase the diversity within and between populations (Herron & Doebeli 2013). These social interactions can take diverse forms (Mitri & Foster 2013). For instance, coexistence of two microbial genotypes can increase the fitness of both types, in comparison to their fitness values when both are found in monocultures. This interaction can retain diverse genotypes within the population instead of fixing one type to 100 % (Mitri & Foster 2013). Further, interacting genotypes can be in a one-sided conflict, in which one party benefits but the other party suffers from that interaction (e.g. Vos & Velicer 2009). But still, the relative frequencies of these conflicting parties can be under oscillation or under stable coexistence (discussed in Velicer 2003). Such social conflicts orchestrated by natural selection are speculated to be shaping the genetic makeup (i.e. at the gene copy number level) of microbial populations (Greig & Travisano 2004). Albeit, these social interactions taking place between individuals of populations do not take place in a vacuum. The physical environment is also immensely important in determining the stability of coexistence (discussed in: West *et al*. 2006; Mitri & Foster 2013). While diverse types can coexist under certain environmental conditions, perturbations of the environmental conditions can drive the extinction of one type and therefore decrease the diversity (Williams & Lenton 2008). The outcome of such genotype-by-genotype interactions is then strongly affected by the environmental conditions in which these social interactions take place (Thompson 2013). In sum, social interactions can be important targets of diversification by natural selection. Yet, since genotype-genotype interactions are directly affected by the external environment, putative social traits that are thought to have evolved as a mere product of social evolution should be carefully examined in order to distinguish them from the social traits those have evolved as a by-product of environmental adaptation (see Chapter 2).

To sum up, intra-genomic conflicts or social conflicts are important backgrounds for the evolutionary process. However, the external environment in which populations exist is extremely important in shaping adaptation of these populations. Further, these forms are not mutually exclusive and all are dynamically inter-connected in shaping the relative fitness values of genotypes, and thus they all in combination affect the ever-continuing evolutionary change of populations (Michod 1999, Preface; Bell 2008; Thompson 2013).

Formation of independently evolving populations

Departure of the continuous variation between two diverging populations to bimodal "discontinuities" results in "the single most important event in evolution": the origin of species (Mayr 1963, p.11). The question of how continuous variation between populations transforms into discontinuous types known as species is a very critical one in evolutionary biology (Coyne 1992). We know that both natural selection and genetic drift fix mutations within populations (Barton *et al*. 2007). Over time accumulating genetic differences and very limited gene flow result in the formation of independently evolving populations (e.g. Platt *et al*. 2010). Estimates suggest that there are about 1.1-6.6 billion such independently evolving populations on earth (Hughes *et al*. 1997). The number is astounding in itself although the diversity of microbes, nematodes, and fungi are not included in this estimate (Hughes *et al*. 1997). These independently evolving populations are likely candidates that would evolve into discontinuous groups. However evolution of these groups into discontinuous types is completed only after the evolution of mechanisms (or barriers) that prevent genetic exchange between the groups (Dobzhansky 1937). Otherwise gene flow between such groups can homogenize allele frequencies between populations fusing discontinuous types back into a single continuously varying type. Thus strong reproductive barriers complete the evolution of irrevocable discontinuous types (Barton & Bengtsson 1986). To explain the evolution of unique types it is therefore necessary to look at the forms of reproductive barriers (Coyne & Orr 2004; Mallet 2006).

Reproductive isolation (RI) can take many forms and all these forms can evolve as a by-product of divergence of populations in allopatry (Mayr 1970; Turelli *et al*. 2001). These forms, in combination, prevent successful breeding between groups of organisms (e.g. Jennings *et al*. 2014). Depending on the complexity of the breeding system of the groups under study, reproductive isolation can act at many levels (Mayr 1970): i) premating isolation, ii) postmating-prezygotic isolation, and iii) postzygotic isolation.

Forms of reproductive isolation

a) Premating isolation

Premating barriers prevent or reduce mating between individuals from different populations (Mayr 1970). Temporal differences in gamete production or discrimination against mating signals are the most common examples of premating barriers (for some examples: Sobel *et al*. 2010).

Coral species of *Montastraea annularis* and *M. franksi* are isolated by temporal premating isolation (Knowlton *et al*. 1997). Spawning time difference (about one to two hours) between the two species is a barrier to interspecific gene flow (Knowlton *et al*. 1997). Some populations of the Japanese winter moth are also temporally isolated: seasonal difference in the breeding of the two moth species limit gene flow between them (Yamamoto & Sota 2009). Thus the two groups diverge as a result of temporal isolation under sympatry (Alexander & Bigelow 1960). Another example for temporal premating isolation has been shown between the plant species *Gelsemium tankinii* and *G. sempervirens*: although both plants share the same pollinator, flowering times were shown to be different within the observed years (Pascarella 2007). This temporal difference was suggested to be the primary cause of reproductive isolation between the two species (Pascarella 2007).

Premating barriers may also arise through the discrimination of (visual, vocal, or chemical) mating signals. As a result, individuals prefer to mate with individuals that secrete/signal compatible/familiar mating cues. Coyne *et al*. has shown that male flies preferentially court females who secrete conspecific pheromones (1994). Diverged coevolution may also reduce the gene flow between populations. Color of the male cichlid fishes or wing patterns of the *Pieris* butterflies are examples of visual premating barriers (Wiernasz & Kingsolver 1992; van der Sluijs 2008).

Premating barriers can evolve as a by-product of evolutionary divergence of loci involved in mate discrimination (Muller 1942; e.g. Dodd 1989; Vines & Schluter 2006), but they can evolve to reduce the hybridization of diverged groups that would otherwise lead diverged groups to waste their reproductive resources on such unsuccessful mating events (Coyne & Orr 1989). Therefore natural selection is suggested to directly involve in the evolution of mate discrimination between species that are already isolated, for instance, at the post-zygotic level (e.g. Sætre *et al*. 1997; Höbel & Gerhardt 2003).

b) Postmating-prezygotic isolation

Even if individuals from different populations successfully copulate, the next stage - the fertilization of gametes - may present barriers known as postmating-prezygotic isolation (Coyne & Orr 2004). Sperm-egg incompatibilities or pollen-stigma incompatibilities are important postmating-prezygotic barriers seen between animal or plant species (Mayr 1970).

In animal populations, biochemical incompatibilities between the male ejaculate and the

female reproductive tract have been shown to form barriers to fertilization. Rapid evolution (following gene duplication events) of female reproductive tract protease genes has been suggested to reduce inter-specific fertilization between some *Drosophila* species (Kelleher *et al*. 2007; Lawniczak & Begun 2007). Prezygotic isolation has been observed between *Drosophila virilis* and *D. americana:* fertilization almost completely fails between these two species, while conspecific crosses of the two show high fertilization/hatching success rates (Sweigart 2010). Moreover, proteins localized on the external coat of egg or spermatids may be involved in prezygotic isolation. For instance, sperm competition and/or female choice have been suggested to cause such prezygotic barriers between mammalian species (Wyckoff *et al*. 2000; Swanson *et al*. 2001).

c) Postzygotic isolation

Gametes from different species can sometimes fuse without experiencing any prezygotic barriers. However such interspecific hybrids may have reduced viability or fertility as a result of the postzygotic isolation. Postzygotic isolation is classified under two forms: 'extrinsic isolation' or 'intrinsic isolation' (Coyne & Orr 2004).

Extrinsic postzygotic isolation: Maladaptive genotype-by-environment interactions of hybrid crosses can cause isolation in parental habitats - this is also known as the 'ecological inviability' (reviewed in Seehausen *et al*. 2014). Extrinsic postzygotic isolation is a by-product of divergent ecological adaptation of parents to different habitats. A hybrid with intermediate trait values of its parents may do worse in either environment in comparison to the local parents (Schluter & Conte 2009), and therefore be inviable or unfit, reducing gene flow between parental populations.

Populations of the leaf beetle *Neochlamisus bebbianae* represent a clear example of extrinsic postzygotic isolation. The population that has adapted to the maple does better on a maple tree compared to another population that has adapted to the willow tree - and *vice versa* (Egan & Funk 2009). However the hybrids of these two species perform worse on both parental habitats.

Apart from 'ecological inviability', hybrids may also suffer from 'behavioral sterility' (Coyne & Orr 2004 p.249). Here, again, the problem is that hybrids show intermediate trait values. For instance, gametes of the wolf spider hybrids experience a healthy developmental period. However, male hybrids fail to attract non-hybrid females as a result of intermediate courtship

behavior (Stratton & Uetz 1986).

Intrinsic postzygotic isolation: Postzygotic isolation can occur *independent* of the environment hybrid individuals find themselves in. Therefore hybrids are maladapted to any possible environment (Coyne & Orr 2004 p. 253). Unlike the extrinsic isolation, hybrids experiencing intrinsic postzygotic isolation do not show phenotypes that are intermediate between their parent species, but are instead worse than either parent.

There are two general and well-accepted genetic mechanisms of intrinsic postzygotic reproductive isolation (Coyne & Orr 2004): a) chromosomal incompatibility (including anti-recombination) and b) genic incompatibility.

a) Chromosomal incompatibility is a result of fixed chromosomal rearrangements (i.e. inversions or translocations) that form barriers to gamete production (Faria & Navarro 2010). Fixed chromosomal mutations (e.g. translocations) can reduce the hybrid fertility (White 1978 p. 54). This is because the random assortment of chromosomes may produce hybrid sex cells lacking essential genes (Figure 2). Therefore chromosomal rearrangements, when fixed differentially between populations, can cause postzygotic isolation (Zeyl 2014).

Indeed such fixed chromosomal mutations are very common between closely related species (White 1978). For example, *Drosophila pseudoobscura* populations are fixed for inversions that keep adaptive gene pools of each population intact by reducing recombination between populations (Dobzhansky 1949; Wallece *et al*. 2013; Fuller *et al*. 2014). Humans and chimpanzees differ in their karyotype by one fusion and thirty-three inversions (Feuk *et al*. 2005). It has been suggested that these rearrangements between human and chimps could have promoted the accumulation of further divergence between the two lineages (Navarro & Barton 2003; discussed in Rieseberg & Livingstone 2003). However, further evidence is needed to understand if these inversions were directly involved in the early phases of speciation (Kirkpatrick & Barton 2006; Ayala & Coluzzi 2005).

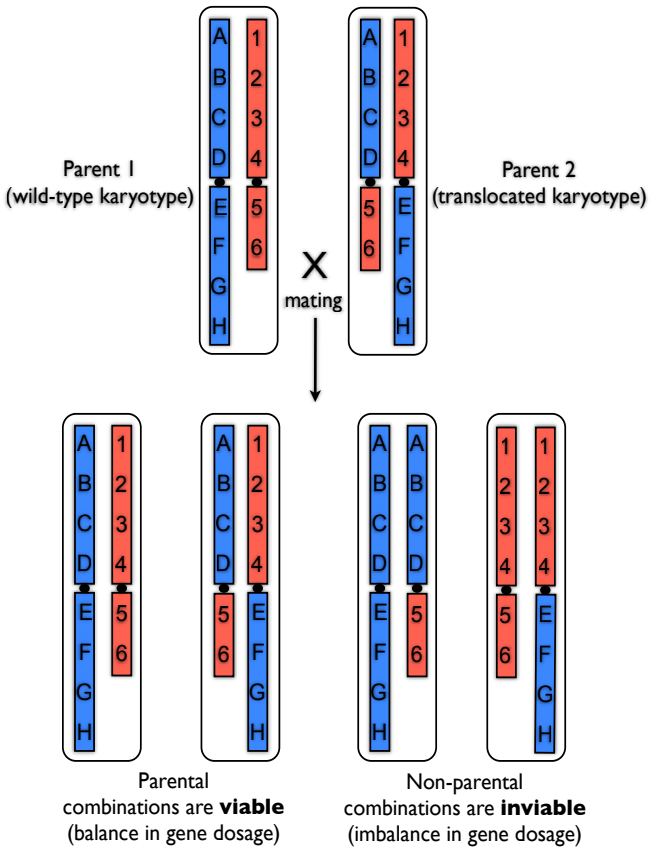


Figure 2. Chromosomal mutations (e.g. translocations) can reduce fertility as a result of gene imbalances. A parent with wild-type chromosomal karyotype and a parent with a trans-located chromosomal karyotype mates (on top). Even though the diploid hybrid is viable, half of the progeny they produce will be inviable (bottom-right). Offspring that have non-parental combinations may lack essential genes. Even if reciprocal translocations do not involve essential genes, still, extra genes seen in the non-parental offspring can be lethal too (Zeyl 2014). Such fixed chromosomal translocations are commonly observed between sibling species and are important in reproductive isolation observed between such heterospecific crosses (White 1978).

Another form of the reproduction isolation that can be mentioned under the 'chromosomal isolation' model is the 'anti-recombination' (AR) barrier (Greig 2009). Extensive single nucleotide divergence spread throughout the genomes of populations can generate reproductive isolation even in the absence of chromosomal rearrangements. This is because divergence in nucleotide sequence inhibits recombination. Since faithful chromosome segregation relies on at least one crossover event per chromosome (e.g. Hillers & Villeneuve 2003) any reduction in the number the crossovers is likely to reduce sexual fertility in eukaryotes. Accumulation of single nucleotide mutations to the extent that they prevent successful breeding of diverging groups can be effective in the reproductive isolation (Greig 2009). However level of single nucleotide divergence that would significantly reduce the crossing-over activity between diverged types is still under-studied (Sniegowski 1998; Fraser *et al*. 2007). And the only direct evidence of the model acting as the reproductive barrier comes from yeast crosses (Chambers *et al*. 1996; Hunter *et al*. 1996).

b) Genic incompatibilities (the Bateson-Dobzhansky-Muller incompatibility model or BDMI) are the most studied model of intrinsic postzygotic isolation (Bateson 1909; Dobzhansky 1937; Muller 1942). The model explains post-zygotic isolation as a by-product of simple genetic divergence in allopatry (Orr & Turelli 2001). The genic incompatibilities arise in

some hybrid backgrounds when genes located at two or more loci interact in *negative epistasis* (Orr 1995). Two loci model of the genic incompatibility can be explained as follows (e.g. Coyne & Orr 1998; Orr 1995, etc.): an ancestral population with the genotype of aabb at two-locus splits into two. In one population a *de novo* mutation, 'A', appears and substitutes for the 'a'. Eventually, either though drift or from some benefit of this mutation, this population becomes fixed for the AAbb genotype. In the other population, this time, the other allele is fixed for the aaBB genotype. Since natural selection "tests" the compatibility of these novel genetic interactions, AAbb and aaBB genotypes can spread through the populations as a result of the independent co-evolution (i.e. the A allele is compatible with the b allele and the B allele is compatible with the a allele). When these two allopatric populations reunite under the same genetic background by hybridization, since 'A' and 'B' were never tested together by natural selection, this novel allelic combination may not function well together. Thus the F_1 hybrids (AaBb – if incompatible alleles A and B are dominant over compatible alleles a and b) or only the F2 hybrids (AABB – if they are recessive) may be inviable or sterile (Coyne & Orr 2004). Such epistatic interactions do not need to reduce the fitness drastically to promote reproductive isolation (Orr & Turelli 2001). Small decreases in relative fitness of hybrids, can also be important, as multiple such BDM-pairs in total can have a significant effect in the reproductive isolation (Orr 1995). This is a relaxed model in the sense that mutations causing incompatibility can arise in multiple evolutionary scenarios, such as the neutral or the positive evolution (Coyne & Orr 1998). Such incompatibility regions/genes have been detected in many taxa showing the universal importance of the model (reviewed in: Rieseberg & Willis 2007; Presgraves 2010; Maheshwari & Barbash 2011).

BDM regions/genes are best documented between the members of *Drosophila* genus (Noor & Feder 2006). Evidence of a two-locus BDM incompatibility (*HMR* and *LHR*), that cause male lethality, has been identified between *Drosophila simulans* and *D. melanogaster* (Brideau *et al*. 2006). BDM genes/regions have also been documented between the mouse species, *Mus m. musculus* and *Mus m. domesticus* (Mihola *et al*. 2009). Moreover, recent evidence showed that multiple two-locus BDM regions exist between the natural populations of swordtail fish: 150 pairs (at FDR = 2%) of genetic incompatibility regions were defined by linkage analysis (Schumer *et al*. 2014). BDM regions have also been identified between plant species: interspecific crosses of *Oryza* genus or between-strain crosses of the *Arabidopsis thaliana* (Chen *et al*. 2014; Bikard *et al*. 2009). It has been suggested that higher-order (including three or more loci) BDMIs are easier to evolve in comparison to simple two-locus BDMIs (Cabot *et al*. 1994; Orr 1995). Indeed examples of the higher-order BDMI regions have started to be

documented. For instance, many complex multi-locus BDM sterility regions have been identified between the hybrids of *Mus musculus* genus (Turner & Harr 2014).

Reproductive isolation between closely related groups can be the result of multiple barriers acting with different strengths (e.g.: Ramsey *et al*. 2003; Lowry *et al*. 2008). Therefore multiple barriers (e.g. premating and postmating in combination) can act together and prevent interbreeding. After the evolution of these irrevocable barriers, it is then possible that, independent populations in very long periods of time would diverge further and form discontinuous types seen on earth (Barton & Bengtsson 1986).

The yeast model system to study speciation

Model organisms are very useful tools to understand the evolution of reproductive barriers and speciation (Fry 2009). In this respect, *Saccharomyces* yeast - being one of the most studied eukaryote in the molecular biology and genetics - is a useful system to study the genetics of reproductive isolation in depth (Delneri *et al*. 2003; Dettman *et al*. 2007; Anderson *et al*., 2010; Dujon 2010; Kvitek & Sherlock 2011; Hou & Schacherer 2015). Yeast is especially interesting because, although it is a unicellular eukaryote, different species belonging to this genus are sexually isolated (Greig 2009). Reproductive barriers between these species act at postzygotic level (Hunter *et al*. 1996; Delneri *et al*. 2007; Liti *et al*. 2006; Charron *et al*. 2014; Hou *et al*. 2015).

Most of speciation research in yeast has been done between the laboratory model *S. cerevisiae* and its closest relative *S. paradoxus*. Haploids of these two species fertilize to form a hybrid diploid; there is little, if any, prezygotic isolation (e.g. Rogers *et al*. 2015, but also see Murphy *et al*. 2006). Hybrid diploids can reproduce successfully via mitosis. However, when the hybrid diploid undergoes meiosis, almost all (99 %) of the resulting gametes are inviable across all environments (Liti *et al*. 2006). Therefore these two species are *intrinsically* isolated at the postzygotic stage (see "Intrinsic Postzygotic Isolation" above). However we still do not have a conclusive understanding of yeast hybrid sterility (Greig 2009). Current models of sexual sterility between the two species can only explain ~10 % of hybrid inviability (see Chapter 3). Therefore genetic models of the post-zygotic isolation mentioned above should be tested in the laboratory to explain remaining of the hybrid gamete death (~89%).

Following is the brief information on the yeast life-cycle (a), genetic differences between the

two yeast species (b), and models of reproductive isolation (c).

i) The yeast life-cycle

Saccharomyces yeasts reproduce both sexually and asexually. They grow predominantly as diploids via mitosis (Fig. 3-1). However, under starvation conditions, cells switch from asexual to sexual reproduction (Fig. 3-2) (reviewed in Neiman 2011). Diploid parent undergo meiotic division to produce four haploid gametes (ascospores). Two of these are 'a' mating-type and the other two are 'α' mating-type. Sexual progeny, at this stage, are kept inside the ascus wall. The ascus wall protects haploid spores from environmental stress (Coluccio *et al*., 2008).

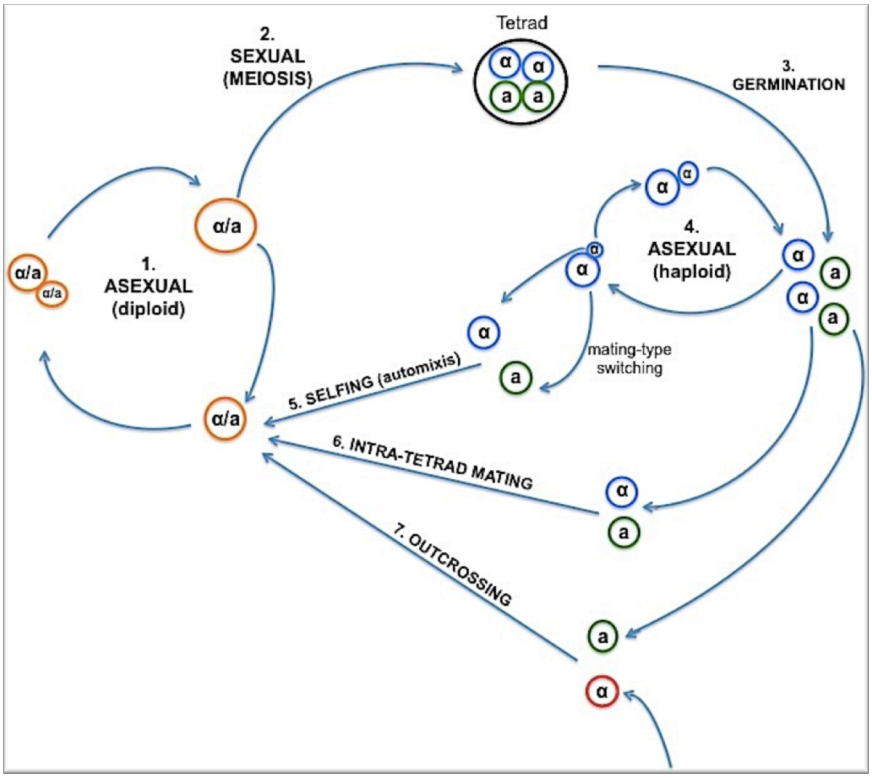


Figure 3. Yeast life cycle has two important modes of reproduction: sexual and asexual. In nature, yeast almost always (~99% of cell divisions) reproduces asexually (Tsai *et al*. 2008). Diploid cells are formed via the fertilization of 'a' and 'α' cells. Mitochondria inheritance at the fertilization stage is bi-parental: both mating-types contribute to the mitochondria pool of the newly formed diploid.

When external conditions become benign, haploid gametes germinate and reproduce mitotically (Fig. 3-4). At this stage there are three possibilities to form a diploid again. i) Haploid spores within a tetrad may fertilize with its opposite sex sibling (Fig. 3-6). This is known as the intra-tetrad mating ii) A haploid cell produces a daughter cell. Then the mother switches its mating type and fuses with its own daughter (Nasmyth 1987; Fig. 3-5). This extreme form of inbreeding is known as the auto-diploidization. iii) Two "unrelated" haploids (Murphy & Zeyl 2010) of opposite mating type coming from different tetrads fuse (inter-tetrad mating) and form a heterozygous diploid (Fig. 3-7). Both mating-types contribute to mitochondria pool of newly formed diploid cells.

ii) Genetic differences between *S. cerevisiae* and *S. paradoxus* genomes

a) Single nucleotide differences: The genomes of *S. cerevisiae* and *S. paradoxus* are extensively (~14%) diverged at the single nucleotide level (Cliften *et al*. 2001; Liti *et al*. 2006). Molecular evolution tests comparing populations of *S. cerevisiae* and *S. paradoxus* have shown that a high fraction of single nucleotide differences between the two species are neutral and deleterious mutations (Doniger *et al.* 2008). Frequency of the detected positively selected sites (as base substitutions) within- and between-species is low (Doniger *et al*. 2008; Will *et al*. 2010). However adaptive single nucleotide mutations have been regularly documented in yeast populations (Li *et al*. 2009; Trindale *et al*. 2009; Kvitek & Sherlock 2011; Kohn & Anderson 2014).

b) Gene copy number differences: Gene copy number variation within the *S. cerevisiae* is significantly higher than gene copy number variation within the *S. paradoxus* isolates (Bergström *et al*. 2014). 32% of the sub-telomeric loci within 18 *S. cerevisiae* isolates were subject to copy number expansion, while only 9.3% of the sub-telomeric loci were polymorphic within 19 *S. paradoxus* isolates. Gene duplications in yeast have been shown to be important in adaptation to novel environmental conditions (Fogel & Welch 1982; Dunham *et al*. 2002; Brown *et al*. 1998; Kvitek & Sherlock 2011). Moreover gene duplication events promoted evolution of functional novelty in yeast (Papp *et al*. 2003; Hittinger & Carroll 2007).

c) Chromosomal differences: Genomes of both species are collinear (Liti *et al*. 2006; Fischer *et al*. 2000). Comparison of the two genomes shows no large chromosomal rearrangements (Kellis *et al*. 2003). Only four small inversions and three segmental duplications have been found between the tested strains of two species (Kellis *et al*. 2003). On the other hand, cases of adaptive chromosomal rearrangement, from laboratory and natural isolates, have been detected in yeast (Dunham *et al*. 2002; Pérez-Ortîn *et al*. 2002; Rancati *et al*. 2008; Yona *et al*. 2012; Chang *et al*. 2013; Dunn *et al*. 2013; Hou *et al*. 2014; Clowers *et al*. 2015).

d) Differences between mitochondrial genomes: Both species have 8 protein coding mitochondrial genes (Procházka *et al*. 2012). Differences between mitochondrial sequences are: Gene order differences at the mtDNA, high divergence of inter-genic sequences, and an addition of active *ori*-elements within the *S. paradoxus* mtDNA (Procházka *et al*. 2012).

iii) Models of postzygotic reproductive isolation in yeast

S. cerevisiae and S. paradoxus fulfill the biological species concept (BSC) criteria (Naumov 1996). Haploids from different species mate efficiently - the pheromones and receptors required for fertilization are compatible between species (Rogers *et al*. 2015). However the heterozygous hybrids are sterile at the postzygotic level. Hybrid diploids undergoing meiosis produce only ~ 1% viable gametes, whereas conspecific diploids of both species produce over 90% viable sexual gametes from meiosis (Liti *et al*. 2006).

Three genetic models (all discussed above) have been suggested to explain postzygotic sterility of *Saccharomyces* hybrids (Kao *et al*. 2010; Delneri *et al*. 2003; Hunter *et al*. 1996).

a) Chromosomal model of reproductive isolation in yeast: As mentioned above, fixed chromosomal mutations (e.g. translocations) can reduce the hybrid fertility (White 1978). The model has been tested for the crosses of the two species, and also within populations of *S. cerevisiae* (Delneri 2003, Liti *et al.* 2006, Hou *et al*. 2014). Translocations can reduce spore viability in *intraspecific* crosses of these two yeast species (Clowers *et al*. 20015; Charron *et al*. 2014). However, since there are collinear genomes between the two species, it suggests that rearrangements did not cause the split (but arose later) (Liti *et al*. 2006). As a summary, chromosomal speciation model is unlikely to explain the reproductive isolation between these two species.

b) Anti-recombination model of reproductive isolation in yeast: Genome wide single nucleotide divergence reduces the crossing-over frequency dramatically (Kao *et al*. 2010). Crossing-over is essential for faithful chromosome segregation (reviewed in Roeder 1997; see for an exception: Davis & Smith 2003). If genomes accumulate single nucleotide differences to a level that reduces crossing-over rates and thus impairs the chromosome segregation, then such fixed substitution differences can form sexual barriers between the species. This is because any F_1 hybrid spore lacking one chromosome in a set of 16 chromosomes will be inviable (Figure 4).

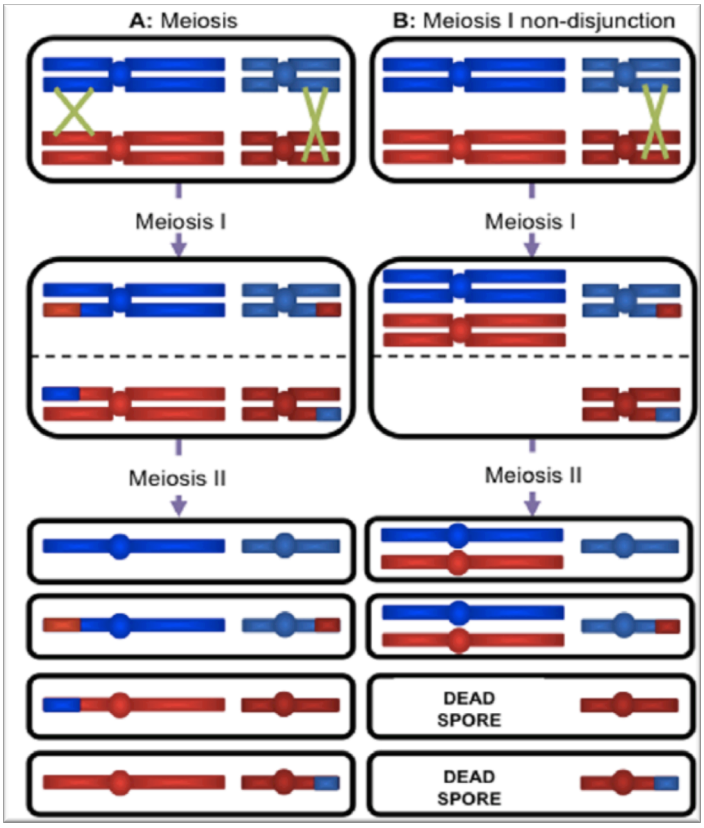


Figure 4. Anti-recombination can prevent faithful segregation by reducing crossing-over between homeologous chromosomes and thus can render hybrids sterile. A) Crossing-over between homologous chromosomes (blue and red at the same length) secures segregation into sexual progeny in yeast. B) Meiosis-I non-disjunction event for one pair of chromosomes can result in two dead spores within four gametes produced. If genomes are extensively diverged at single nucleotide level (e.g. homeology seen between the two yeast species), then crossing-over is reduced or inhibited. As a consequence, homeologs can be pulled into the same pole at meiosis-I. Figure is adapted from Amin *et al*. (2010).

Anti-recombination has been shown to contribute to the postzygotic sterility between the two species (Hunter *et al*. 1996). Furthermore, Hunter *et al*. (1996) showed that the anti-recombination was responsible for 10 % of the hybrid gamete death. However absolute strength of the anti-recombination barrier between the two species is still unknown (Chapter 3).

c) The genic incompatibility model of reproductive isolation in yeast: This model has been tested several times in *S. cerevisiae* and *S. paradoxus* hybrids (Greig 2007; Kao *et al*. 2010; Xu & He 2011). Yet, so far, there is no evidence for any two-locus BDMI regions between these species (Li *et al*. 2013).

Recently, Hou *et al*. (2015) identified the first example of a two-locus BDM incompatibility in yeast. This evidence is from intraspecific crosses of *S. cerevisiae* (2015). Hou *et al*. (2015) crossed a laboratory strain and a clinical isolate of *S. cerevisiae*. Meiotic progeny were viable on the regular laboratory media. Interestingly though, Hou *et al*. (2015) observed that one fourth of the spores from this cross were inviable on a respiratory carbon source (glycerol). They then identified a two-locus BDM incompatibility region causing the respiratory deficiency in F_1 spore of the inter-population cross. Discovery of this conditional

incompatibility, between the two relatively close (~0.35% divergence) populations, suggests that such incompatibilities would also exist between far more divergent (~14%) crosses of the two species that are subject to this thesis study.

As a result, although the two-locus incompatibilities have not been detected between the two species, this model is still a strong candidate to explain the remaining gamete death. Therefore it should be re-tested using larger sample size that allows the detection of any potential "weak" incompatibilities (Li *et al*. 2013; see Chapter 4).

Overview

All types of mutations mentioned in the beginning of the general introduction are important in evolutionary divergence and diversification. These genetic mutations become fixed between populations by the action of random drift and natural selection. However, the forms that natural selection can take are as diverse as the forms of mutations that they act on. Therefore collecting evidence to understand the relative importance of these forms in increasing genetic divergence between populations is in central importance in evolutionary biology research.

In this regard, in chapter one, I examine intra-genomic conflicts that occur within a cell. Using yeast, I test if conflicts between the nuclear and mitochondrial genomes can speed up the divergence of populations that have evolved in isolation. To vary the amount of conflict between the two genomes, I manipulated the mode of sexual breeding of different populations. When conflict between the genomes was increased, I found that mitochondria provided less benefit to their host. I suggest that the intra-genomic conflicts can be effective in the evolution of negative cyto-nuclear interactions between populations.

In the second chapter, I question the social-conflict hypothesis that was suggested to be the cause of the *SUC* gene copy number variation detected between yeast populations. Apart from the social conflict hypothesis, an alternative hypothesis points to the importance of the external physical environment's role (the environmental adaptation hypothesis) in explaining the evolution of copy number variation. I genotyped wild yeast populations isolated from multiple environments to see whether the *SUC* gene copies varied within- or between-environments. I provide evidence favoring the environmental adaptation hypothesis over the social conflict hypothesis.

All these forms of evolution (i.e. intra-genomic conflicts, social conflicts and environmental

adaptation) that act at different levels (i.e. replicators and/or individuals) are expected to diverge populations towards becoming isolated entities. Following this by taking a top-down approach, in the next two chapters I investigate two different genetic models of postzygotic reproductive isolation that are suggested to be effective in preventing inter-breeding of already established yeast species.

In chapter three, I investigate the significance of overall single nucleotide divergence that has accumulated throughout the genomes of two yeast species (*S. cerevisiae* and *S. paradoxus*) in forming barriers to mating. I attempted to overcome the segregation problem (e.g. the anti-recombination barrier) that has been shown to be important during the meiosis of the hybrid diploids. By inducing crossovers between the chromosomes of two diverged species, I provide novel evidence on the relative role of the anti-recombination barrier in causing reproductive isolation between the two species.

Chapter four also looks at the importance of genetic divergence in preventing healthy gamete production between these two sibling species. Instead of looking at the problems that arise *during* meiosis due to the overall single nucleotide divergence, in this final chapter, I investigate the role of divergence observed between individual loci of both species in causing sterility *after* the completion of meiosis in hybrids. This chapter attempts to answer whether strong two-locus genic incompatibilities (BDMIs) cause reproductive isolation between the two species, in addition to the role of the model documented in the chapter three.

In summary, in the following chapters, I investigate how evolution of discontinuous forms could become possible, firstly by looking at how yeast populations can diverge at initial stages of population division by the action of diverse forms of natural selection (Chapters 1&2) and then by looking at what prevents the genomes of already established yeast species from mixing (Chapters 3&4).

Mode of mitochondrial transmission determines intracellular conflict in yeast

Introduction

Mitochondria are endosymbionts that were once free-living alpha-proteo-bacteria (Yang *et al*. 1985). This mutualism started with the acquisition of bacteria into an archaea-like host, and has given rise to the evolution of true complex life on earth (Grosberg & Strathmann 2007; Spang *et al*. 2015). Although the mutualism between mitochondria and host has evolved for over a billion years (Margulis 1970; Knoll *et al*. 2006), even such a long-term collaboration is prone to conflicts. Mutations in mitochondrial genomes may be selected in the short term, even when they can cause genetic diseases and therefore reduce organismal fitness (Wallace 1999).

Saccharomyces yeast, unlike most eukaryotes, do not require mitochondria for survival. This is because yeast cells that have no functional mitochondrial genome can still produce energy via fermentation (De Deken 1966). Mitochondria are only necessary for respiration. Another distinguishing feature of yeast is the bi-parental inheritance of its mitochondria (Nunnari *et al*. 1997). Both yeast mating-types, 'a' and 'α' cells, contribute to the mitochondria pools of the newly formed diploid cell. This is very different to the uni-parental inheritance seen in most animal and plant species where only one sex contributes to the mitochondria pool (Burt & Trivers 2006).

Since yeast cells with no functional mitochondria are viable, nonfunctional mitochondria can replicate within the host, even though they no longer provide any benefit (Burt & Trivers 2006). Such mitochondria can be thought of as selfish mitochondria. Haploid yeast host about 25-50 mitochondria per cell, and therefore, cells can be polymorphic for the mitochondrial genomes they contain in (Burt & Trivers 2006). If some of the mitochondria within a pool are selfish but the rest are mutualistic, then such a "heteroplasmic" yeast cell will be able to respire. However, some yeast cells carry only non-functional mitochondria. These respiratory deficient yeasts are known as the petite mutants (ρ^-) (Goldbring *et al*. 1970). Yeast petite mutants are formed at a frequency of ~1 % per cell division (Goldbring *et al*. 1970; Baruffini *et al*. 2007).

The selfish nature of the mitochondria in petite yeast cells is a result of several types of mutations. Large genomic deletions can give a replication advantage to selfish types over wild-types, as can duplications of DNA replication origins which enable faster reproduction (Petersen *et al*. 2002). Therefore, selfish mutants can replicate faster than wild-type competitors and spread through the population of hosts (MacAlpine *et al*. 2001) even if such selfish mutants are costly to the host especially under sexual reproduction (Burt & Trivers 2006). On the other hand, fixation of petites in a population of hosts is not possible since sexual reproduction (meiosis) relies on a mitochondrial energy source (Jambhekar & Amon 2008). Therefore non-functional mutant mitochondria can only increase in frequency by hitchhiking along functional mitochondria.
Nevertheless, the spread of selfish mitochondria relies of sexual reproduction; their higher replication rates cause them to be over-represented in offspring relative to wild-type mitochondria. When a heteroplasmic cell is crossed to a wild type, it produces 18-37% petite offspring despite both parents being capable of respiration (Harrison *et al*. 2014). Respiratory deficient mutants (petites) do not increase in frequency under asexual reproduction, as they provide no benefit to the host (Harrison *et al*. 2014).

We can determine the behavior of mutant mitochondria by crossing a yeast cell carrying fully functional mitochondria to a yeast cell carrying selfish mitochondria (Williamson 2002). The mitochondria that replicates faster will be over represented in the progeny of that cross. When a mitochondrion conferring a petite phenotype is overrepresented in the host's progeny compared to the wild-type, it is known as "suppressive" petite (Petersen *et al*. 2002). "Hyper-suppressive" petite mitochondria (Burt & Trivers 2006) replicate so much faster than wild-type mitochondria that over ~95 % of the offspring of crosses with wild type yeast are themselves petite (Petersen *et al*. 2002). Such hyper suppressive petite mitochondria carry multiplied origin of replication regions within their genomes giving them a decisive over-replication advantage (Blanc & Dujon 1980; Zamaroczy *et al*. 1981).

Here, there is an interesting questions requiring investigation: Does the transmission mode, i.e. *vertical* or *horizontal* transmission of mitochondria, affect host-endosymbiont interactions within the continuum of *mutualism* and *parasitism*? For instance, do horizontally transmitted mitochondria evolve to be less mutualistic or more selfish when compared to vertically transmitted mitochondria?

Data on infectious diseases have suggested that the transmission mode of parasites is key to the evolution of virulence (Herre 1993; Lipsitch *et al*. 1996; Frank 1996). For instance, the "continuum hypothesis" (Ewald 1987) predicts evolution of lower virulence under the vertical transmission, where parasites are transferred from parents to their offspring (Fine 1975; Kover & Clay 1998). This is because a highly virulent vertically-transmitted pathogen that reduces host fitness will in turn reduce its own fitness by reducing the number of new hosts (i.e. offspring) it can infect. Since a vertically transmitted parasite is imprisoned within the cellular lineage it has formed in and the only way it can spread is through the offspring of that host, reducing host survival and reproduction rates will reduce the parasite's own chances of spreading through the population. Therefore under vertical transmission selection against virulence should be strong. On the contrary, under horizontal transmission, since parasites can move (or escape) to other individuals within a population, a relatively higher virulence might evolve (e.g. Magalon *et al*. 2010; Pagán *et al*. 2014). The reason is that a parasite reducing host fitness, differently from the vertically transmitted parasite, is not imprisoned within that host, and thus more virulent parasites still have their chances to remain in the population via moving to other hosts. Therefore selection against more virulent parasites is low. Moreover, since different parasite genotypes are mixed more frequently within common hosts under the horizontal transmission, such a polymorphism would increase within-host among-parasite selection, and therefore select for faster replicating virulent types. The theory then suggests a trade-off between virulence and these two modes of transmissions (Anderson & May 1982).

By extending the continuum hypothesis from the host-parasite interactions to the intra-genomic host-endosymbiont interactions, we can predict that the mode of transmission is likely to effect the interaction between mitochondria and their host for three reasons:

First, varying the association between mitochondria offspring and host offspring affects correlation between host fitness and mitochondria fitness. Under strictly vertical transmission, the fitness of host and mitochondria are correlated because they are locked together in the same cellular lineage. Hosts cannot acquire new mitochondria, and mitochondria cannot get into new hosts. Since the fate of both parties are tied together, it is expected that cells where both parties provide increased services to each other (i.e. mutualism) will have higher fitness and are therefore should increase in frequency through the population. Further, a higher among-lineage-selection is expected under a more vertical selection generating higher selection pressure against host lineages carrying relatively more selfish mitochondria (Taylor *et al*. 2002). Thus changing transmission to be more vertical should select for more

cooperation, less selfishness, and more dependence. Changing transmission to be more horizontal will have the opposite effect. A host bearing more selfish mitochondria suffers from a reduction in fitness and thus is likely to be purged from the population under more horizontal transmission. But such a selfish mitochondria can get into a different host (e.g. by outbreeding or by sexual reproduction) instead of just by getting into host's offspring. This then breaks the correlation between host fitness and mitochondria fitness (and dependence) more frequently under a more horizontal transmission.

Second, varying the association between mitochondria offspring and host offspring affects co-evolution. This is really a corollary of point one, and it just means that more vertical transmission allows more co-evolution. For example, imagine that a mutation increases mitochondrial exploitation of the host, lowering host fitness. Under horizontal transmission a compensatory mutation is of little benefit, because the hosts' offspring will receive new mitochondria, which may not have the selfish mutation (or may have a different selfish mutation). It is then less likely for nuclear DNA to accumulate compensatory mutations that directly target mitochondrial mutations to neutralize harmful effects of more selfish or less mutualistic mitochondria (Chou & Leu 2015). Thus changing transmission to be more vertical should select for more co-evolution and changing transmission to be more horizontal will have the opposite effect.

Third, varying horizontal transmission affects within-host competition between mitochondria. Under strictly vertical transmission, variation among mitochondria can only increase due to *de novo* mutation. But under horizontal transmission, mitochondria can also come in from other lineages. Thus changing transmission to be more horizontal will increase within-host competition among mitochondria, selecting for mitochondria that have increased within host competitiveness, potentially reducing host fitness. Similarly, relatedness of mitochondria within a host is lower under more horizontal transmission (Hastings 1999). This increases selection for faster-replicating (or over-replicating) mutant mitochondria - a result of increased within-cell-selection (Taylor *et al*. 2002). Changing transmission to be more vertical will have the opposite effect.

In this study, I explore fitness effects of host-mitochondria evolution in relation to two different modes of transmission using *Saccharomyces* yeast as a model. With this purpose, I evolved yeast populations in the laboratory for 15 sexual cycles under two treatments (see Figure 1). In the 'horizontal transmission' treatment, I increased the amount of horizontal

mitochondria transmission by only allowing mating of gametes from different tetrads (see the 'Yeast Life-Cycle' under the General Introduction for more details on yeast mating). I removed the opportunity for inbreeding by intra-tetrad mating or by auto-diploidization following mating-type switching. In the 'vertical transmission' treatment, I removed the opportunity for mating between or within tetrads, only allowing mating between mother cells and their clonal daughters in the form of auto-diploidization. In, what I call, horizontal transmission, mitochondria moves into a different genetic background, by merging of two different cellular lineages (inter-tetrad mating), since intra-tetrad mating is prevented. Although, vertical transmission lines also have same amount of sex, and includes fusion of cells, this time, mother and daughter mates (strict intra-tetrad mating), thus mitochondria stay within the same cellular lineage. All populations of both lines also underwent frequent mitotic growth in which mitochondria were transmitted vertically. Therefore manipulating the mating system either increased or decreased the amount of horizontal transmission only during the mating (e.g. fertilization) stage. And then I measured: 1. The overall adaptation of co-evolved nuclei and mitochondria, 2. The differences in the dependence on mitochondria of nuclei evolved under the two treatments, 3. The difference in the services provided by the mitochondria evolved under the two treatments. For the first, I expect that there would be less overall adaptation in the 'horizontal transmission' treatment because co-evolution is disrupted more frequently under horizontal transmission. For the second, I expect nuclei to evolve less dependence on to their mitochondria under the 'horizontal transmission' treatment. On the contrary, under the 'vertical transmission' treatment, I expect mitochondria and nuclei to evolve in a more inter-dependent manner. Thus populations of the 'vertical transmission' treatment are expected to have a greater reduction in fitness when their functional mitochondria are removed. For the third, I expect mitochondrial evolution to be less mutualistic or more selfish under the 'horizontal transmission treatment in comparison to mitochondrial evolution under the 'vertical transmission' treatment.

Methods

Experimental evolution:

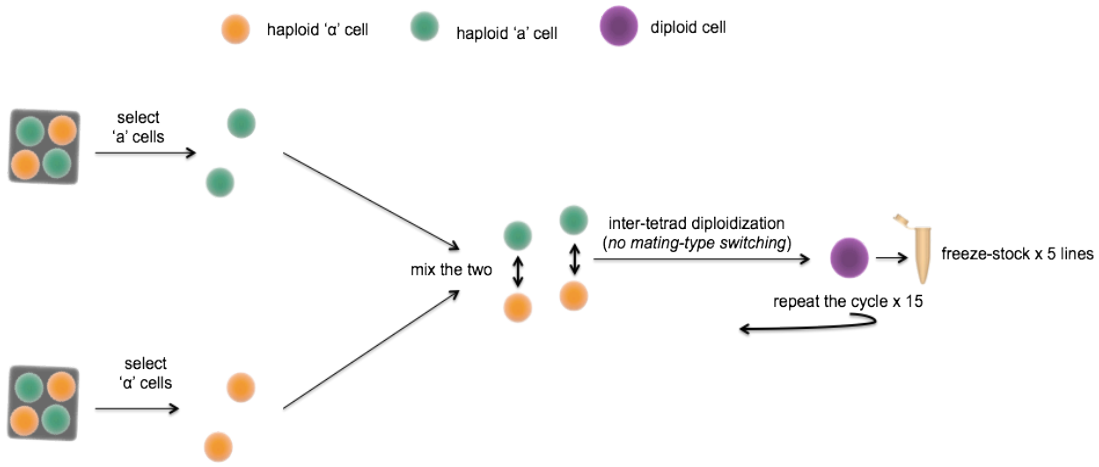
I started each cycle of evolution transfer by culturing *Saccharomyces cerevisiae* (Y55) cells in 5 ml KAc (2% potassium acetate) at room temperature for 4-5 days (Fig. 1-A/B). This induces sex (meiosis and sporulation). I confirmed completion of meiosis observing tetrads under the microscope. I aliquoted 2 ml of the KAc culture. Then I applied 1% NaOH to kill the cells that have not fertilized and gone through meiosis (Coluccio *et al*. 2008).

To select for ‘a’ or ‘α’ cells (Fig. 1-A/B), next, I transferred 100 µl of the tetrads into two separate 3 ml of synthetic media, either without histidine or without leucine amino acids, and grew two separate cultures for 1.5 days shaking at 30 °C.

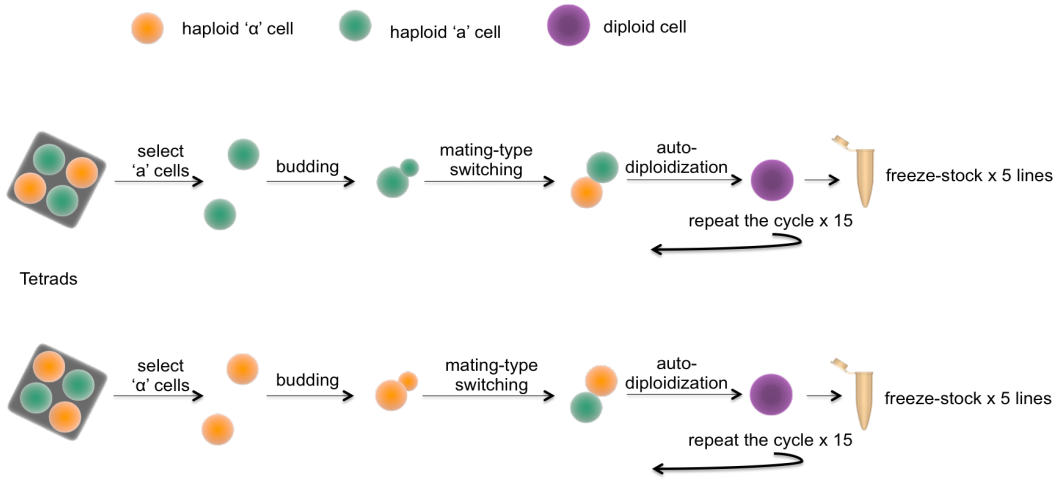
To induce fertilization of haploid cells and apply laboratory selection, I transferred 200 µl of drop-out cultures into 5ml of COM media (complete amino acid mixture, 2% glucose, yeast-nitrogen base) - see below Horizontal Evolution Regime (Figure 1-A) and Vertical Evolution Regime (Figure 1-B) for the differences at this stage. I kept the cells un-shaken within the COM liquid media for 5 days of extended growth at 30 °C. The rationale behind an extended growth for 5 days is to induce yeast cells to switch from fermentation to mitochondrial respiration under nutrient depletion. This stage is the critical point in all evolution experiment: Fertilization of haploids makes the difference between *horizontal* (A) or *vertical* (B) mitochondria transmissions (Fig. 1-A/B). However, still, in both evolution regimes most of the cell divisions took place as mitosis where mitochondria are vertically transmitted. Thus horizontal transmission is only induced in-between many vertical transmission stages. However for the simplicity, I call this evolution regime (Figure 1-A) as the ‘horizontal evolution regime’ and strictly vertical transmission regime (Figure 1-B) as the ‘vertical evolution regime. This stage is also the point when I applied environmental selection.

To repeat the same selection regime, after the completion of between-tetrad-mating (Fig. 1-A, horizontal transmission) *or* within-tetrad-mating (Fig. 1-B, vertical transmission) that is followed by an extended growth of laboratory selection in COM-liquid culture, I restarted the whole cycle beginning from sporulation/meiosis stage in KAc media. These three stages of transfers have been repeated for 15 cycles for both horizontal and vertical evolution lines.

It is important to note that a yeast cell can undergo sporulation and meiosis only if that cell contains respiring mitochondria (Jambhekar & Amon 2008). Therefore, each sexual cycle (meiosis and sporulation) in my selective regime eliminates cells carrying hyper-suppressive mitochondria. Still, since each cytosol carries a pool of mitochondria, my selective regime allows nonfunctional “selfish” mitochondria to hitchhike to the next stages with the functional mitochondria.



A)



B)

Figure 1. A) Horizontal transmission using strains unable to switch mating-types. Thus every mating event creates a horizontal transmission chance. It is crucial to note that, apart from the horizontal mitochondria transmission at the fertilization stage ('inter-tetrad diploidization'), most of the cell divisions in these selection regimes are mitotic divisions. Horizontal mitochondria transmission could be rarely induced within a mostly vertical mitochondria transmission mode. This line is therefore horizontal and vertical transmission but for the simplicity, in the subsequent part of the text I will refer to it as the horizontal evolution regime. **B)** Vertical selection using strains those contain functional mating-type switching gene (*HO*). Thus the endosymbiont is transferred only by vertical (from mother to daughter) transmission. This line has a strictly vertical transmission of mitochondria with no horizontal transmission.

A) Horizontal transmission regime:

Horizontal transmission forces cells from different tetrads to mate - known as the inter-tetrad-mating. To achieve this I used mating-type switching gene (*HO*) knock-out mutants to prevent clonal mating (Fig. 1-A).

To induce the horizontal mitochondria transmission, after selecting only 'a' or only 'α' cells in his-drop-out or leu-drop-out media (Fig. 1-A), I mixed two cultures within the same COM-liquid media. This critical difference increases the horizontal endosymbiont transmission by enforcing 'a' and 'α' cells that were in different tetrads. Thus mitochondria pools explore more cytosolic hosts. Then, similar to vertical lines I applied environmental selection for an extended growth in liquid COM cultures.

I ran five parallel replicates in total. I stored evolving population with regular intervals as frozen stocks (Fig. 1-A).

B) Vertical transmission regime:

Vertical transmission means that a mother cell buds to generate a daughter cell; next, the mother switches its mating-type and mates with its daughter (Nasmyth 1987). Thus, fertilization occurs only between identical clones (Fig. 1-B), and any "selfish" mutant is imprisoned within the cellular lineage it has formed in.

To induce the vertical mitochondria transmission, after selecting 'a' or 'α' cells, I did not mix two cultures within the same COM-liquid media. Five lines of 'a' mating-type and five lines of 'α' mating-type ran separately giving in total 10 replicates for vertical evolution regime. That is why in all comparisons there are two times more data points for vertically evolved lines in comparison to horizontal evolution regime.

Functional mating-type switching gene (*HO*) in this regime allowed separate cultures with different mating-types to form diploids via mother-daughter fertilization. This extreme form of within-tetrad mating ensures vertical transmission of the mitochondria pools within populations. I stored evolving populations at – 80 °C.

Assay for overall co-adaptation by measuring growth yields:

After completion of 15 transfers, I measured fitness of all 15 lineages (five lines of vertical-'a' selection, five lines of inbred-'α' selection, and five lines of horizontal selection) by quantifying the overall growth yield. This is to see if the selection lines adapted to the

experimental media (extended growth in COM-liquid) after 15 transfers of evolution. "Overall" adaptation in this context means that I have measured the adaptation of cells containing co-evolved nuclei and mitochondria. Cells from evolved populations carrying functional mitochondria and nuclei were grown for five days in the original selection environment (COM-liquid). I also grew ancestral clones as controls. I transferred 150 µl of cells into 96-well plates as five technical replicates. Then, I measured the optical density (OD_{600}) of each line as three reads per plate (15 data points each).

To visualize the overall adaptation of each line more simply, I subtracted mean growth yield results of ancestors from mean growth yield results of their corresponding evolved populations, and obtained the bar graphics in Figure 2. However the statistical tests used to determine if each evolved line has adapted significantly in comparison to their ancestral clones were done using 15 raw data points in comparison to corresponding ancestral control. Dunnett's ANOVA was used to calculate multiple comparison corrected p-values. Finally, to test if overall adaptation differed between vertical and horizontal evolution groups, I treated all data points from each treatment as one population, and performed an unpaired t-test with Welch's correction for unequal variances (equality rejected via F-test).

Testing the fitness of the evolved strains in the absence of respiratory services from co-evolved mitochondria:

To determine how the two transmission treatments affected the dependence of the evolving nuclear genomes on the services provided by their co-evolving mitochondria, I compared the fitness of the evolved cells after they had been cured of functioning mitochondria (i.e. petites). A petite of the strain YDG851 (ura3::KANMX/) was used as a marked common-competitor strain. Three clones per petite were used in independent fitness assays. I confirmed the petite nature of the evolved and competitor clones by replica plating colonies onto respiratory carbon source (glycerol). Thus I could pick clones that were unable to grow on glycerol.

For each biological replicate of the three fitness assays, I grew the marked-competitor and evolved petite clones separately in 3 ml of COM liquid shaking at 30 °C overnight. Next day, I mixed 300 µl of each evolved line and 700 µl of YDG851 common competitor in a 1.5 ml Eppendorf tube. The mixture was plated onto YEPD plates as 10^{-4}-fold serial dilutions. About 200-500 colonies were observed on each plate, I replica plated initial YEPD colonies onto YEPD-G418 and YEPD-only plates. By counting the number of drug-resistant colonies

relative to sensitive colonies, I recorded the time zero (T0) relative frequencies of the two types. From that initial mixture of competitor and evolved petites, I transferred 100 µl of the culture into a fresh 5 ml COM-liquid (original selection environment) and incubated the mixture steady at 30 °C for five days. After five days of growth I plated the mixture of the two types onto YEPD plates as 100 µl of 10^{-4}-fold serial dilutions as three technical replicates. Then I replica plated those colonies onto YEPD-only and YEPD-G418 plates. Finally I counted colonies of both genotypes as final relative frequency of both types at time T1.

Fitness of evolved lines was calculated using the Malthusian parameters (Lenski *et al*. 1991). I also log transformed (at base 10) the relative fitness values before the final statistical calculations. To compare the horizontal and the vertical evolution groups as two populations of data points, I performed a Mann-Whitney test.

Testing the fitness of ancestral strains containing evolved mitochondria:

I compared the fitness effect of horizontally or vertically evolved mitochondria in a common un-evolved host. This tests if mitochondria pools evolved via different transmission modes differ in their final fitness effect over the same un-evolved host. I transferred evolved mitochondria of each line into a common un-evolved host with no functional mitochondria (Y55 strain, YDG949 p^-, kar1, can1, cyh2).

I confirmed petite nature of control strain (YDG949) by replica plating onto the glycerol media. Mitochondria recipient un-evolved host was a kar1 mutant. This point mutation prevents the fusion of the nuclei during fertilization (see Lancashire & Mattoon 1979).

All evolved strains were diploid. This prevents them to undergo fertilization with the haploid recipient. Thus I initially grew the evolved diploids in 3 ml KAc cultures for 4 days to induce sporulation/meiosis. I could obtain four haploid gametes. I grew the recipient strain (YDG949 p^-) overnight in liquid YEPD before transferring their mitochondria. Next I mixed 100 µl of the un-evolved recipient host and 250 µl of the evolved mitochondria donors. I spun down the mixture, washed off the media, and after vortexing the recipient and donors, I plated a 15 µl of each mixture as a patch onto the YEPD plates. This initiates fertilization of un-evolved recipient and evolved hosts. After seven hours of growth, I suspended the mixture in 5 ml sterile water, and plated a 400-700 µl of mixtures on selective media. Here, glycerol selects for the functional mitochondria of the evolved donors. Two drugs (canavanine and cycloheximide) select for the genomic DNA of the un-evolved host and select against the

genomic contents of the evolved lines. With this method, I could obtain a handful of colonies on selective media after 4-5 days of growth at 30 °C. I picked seven biological replicates for each mitochondrial-transfer events.

For the Malthusian fitness assays I applied the same method described in the section "Testing the fitness of the evolved strains in the absence of respiratory services from co-evolved mitochondria". However initial mixture of the test clones and the drug resistance competitor YDG963 (un-evolved common competitor strain containing its native mitochondria and ura3::*KANMX* marker) was in equal volumes. For T0 and T1, I plated 80-100 µl 10^{-4}-fold serial dilutions of cultures (gave 100-300 colonies on each plate). To compare the horizontal and the vertical evolution groups as two populations of data points, I performed a Mann-Whitney test.

Results

Overall adaptation of evolved cells containing co-evolved mitochondria:

Following 15 transfers of selection, I quantified the overall adaptation by measuring growth rates of each line in comparison to growth rates of the ancestral clones (Figure 2). Overall adaptation is a measure of the growth rate increase of each line in comparison to corresponding ancestors as a result of the laboratory selection applied as an extended growth in COM-liquid cultures. 'Overall' means that I am looking at the combined effects of both nuclear and mitochondrial evolution, instead of looking at growth rate increase as a result of selection acting on each genome separately.

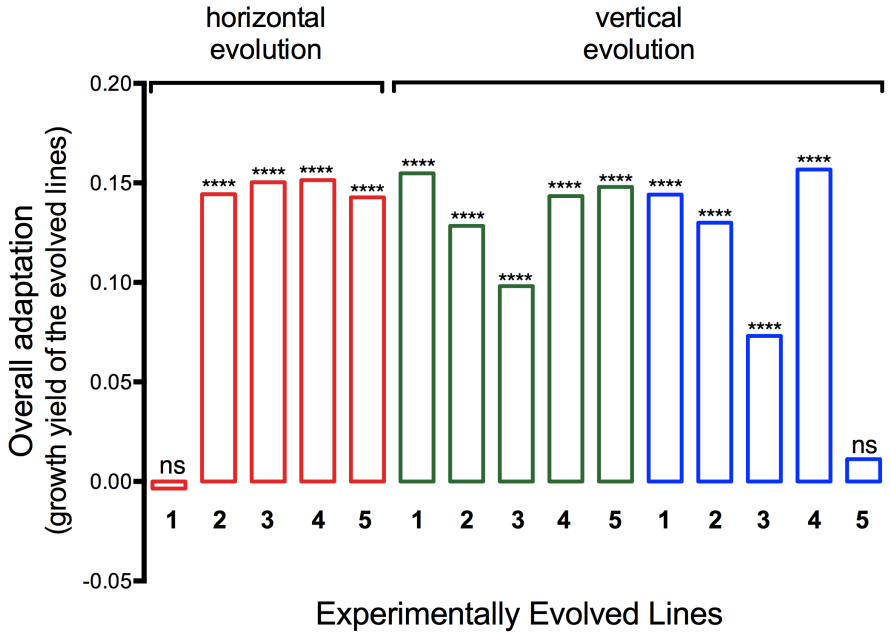


Figure 2. Overall adaptation of horizontally and vertically selected yeast lines. Each bar represents the growth yield of an evolved line measured in the selection environment at the end of 15 transfers. Growth values are the difference between the ancestor and the growth yields of each evolved line. Stars above the bars represent statistically significant ($p<0.0001$) differences between an evolved lined against its corresponding ancestor (Ordinary one-way ANOVA, Dunnett's multiple-comparison corrected). Red bars show the fitness improvement of replicate populations that evolved under horizontal transmission, green ('a' mating-type selection) and blue bars (''α' mating-type selection) show the fitness improvement of replicate populations that evolved under vertical transmission (see Methods). There was *no difference* in overall adaptation between horizontal and vertical transmission regimes ($p=0.0937$ in Welch's t-test).

Most replicate lines, except two outliers (see 'ns' above the bars), evolved to grow significantly more than the corresponding ancestors in the original selection environment (Fig. 2). To see if the transmission regimes of the mitochondria (horizontal vs. vertical) have an effect on the overall adaptation rates of yeast, I compared the growth yield of horizontally- and vertically-evolved lines as two groups of data points (red bars vs. green bars in Figure 2) However there was no significant difference in overall adaptation between horizontally and vertically selected groups ($p=0.0937$ in Welch's t-test).

Fitness of the evolved strains in the absence of respiratory services from co-evolved mitochondria:

Next I wanted to see if the nuclear genome evolved to have different fitness values as a result

of differences in horizontal and vertical mitochondrial transmission modes. Therefore I performed a competition assay by mixing each evolved line separately with its corresponding ancestor, and measuring fitness of each evolved population relative to its ancestor in the selection environment. To compare only the evolution of the nuclear genomes of both horizontal and vertical evolution regimes, I obtained evolved and ancestral cells that had no functional mitochondria ('mitochondria-free evolved cells').

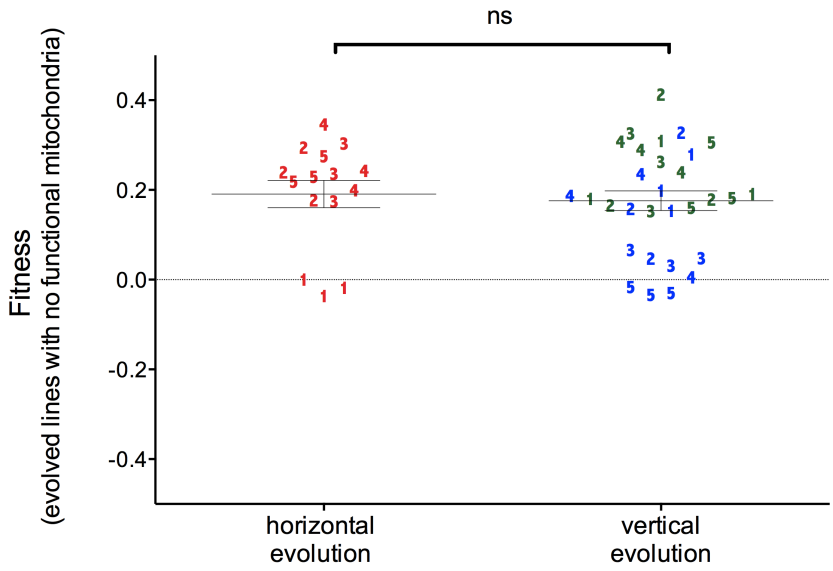


Figure 3. Testing the fitness improvement of the evolved cells in the absence of functional mitochondria. Both horizontal and vertical transmission lines evolved to perform better against their ancestors as a population (see mean values for both horizontal-evolution and vertical-evolution groups in comparison to the base-line '0' depicting no fitness change). However there was no significant difference between horizontal and vertical evolution groups (p=0.5740 in Mann-Whitney test). Red data points show horizontal transmission lines, green data points show vertical transmission lines in 'a' mating-type selection and blue data points show vertical evolution lines in ''α' mating-type selection (see Methods). Numbers shown as data points represent their experimental evolution labeling. The horizontally-selected lines have half the data points of vertically-selected lines due to the experimental setup (Fig. 1, A vs. B). I tested three biological replicates for each line. Error bars represent SEM over and below the mean.

Although both groups evolved to perform better against their ancestors as a whole (Fig. 3), there was no significant difference in the fitness of evolved nuclear genomes between horizontally and vertically evolved groups (p=0.5740 in Mann-Whitney test). The fitness increase of each line with no functional mitochondrial genome (Fig. 3) showed a similar trend observed in overall adaptation (Fig. 2). For instance, horizontal-transmission line-1 had no difference from its ancestor as a cell containing its functional mitochondria (Fig. 2) and the

same line also performed poorly against its ancestor in fitness assays this time as a cell that did not contain any functional mitochondria (Fig. 3).

Fitness of ancestral strains containing evolved mitochondria:

My main motivation in the study was to see if evolved mitochondria would have a different effect on the fitness of the host depending on whether the mitochondrial evolved under horizontal or vertical transmission. This will allow me to judge if the horizontal evolution lines evolved less mutualistic or more selfish mitochondria pools in comparison to the vertical evolution lines as predicted by theory. To test this I transferred each evolved mitochondria pool separately into the same un-evolved host that had no functional mitochondria. Then I performed relative fitness assays for each line by mixing the un-evolved host bearing evolved-mitochondria pools against the un-evolved host bearing its native/un-evolved mitochondria under the original selection conditions (5 days of growth in COM-liquid media).

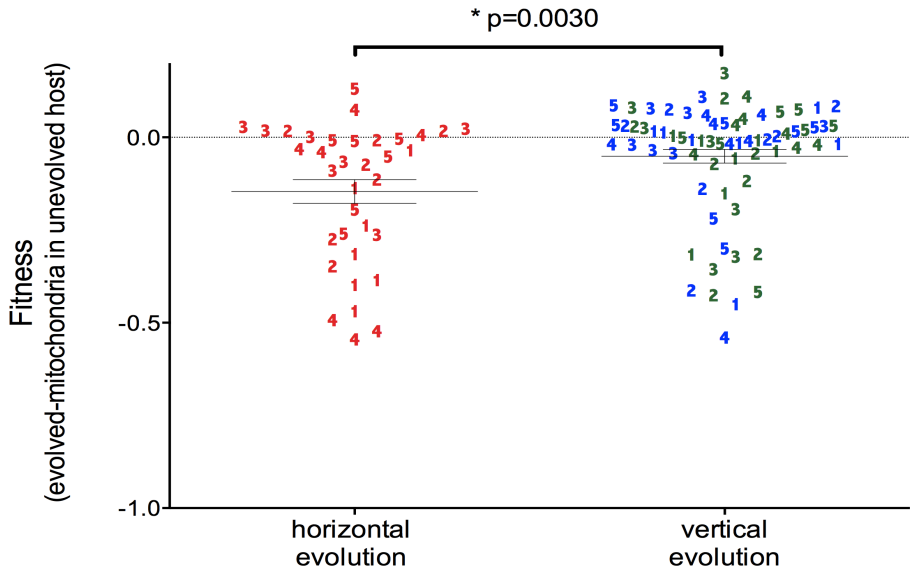


Figure 4. Fitness effect of evolved mitochondria pools in an un-evolved host against the same un-evolved host bearing its native mitochondria. Mitochondria evolved under horizontal and vertical transmission had significantly different effects on the same un-evolved host: The un-evolved host carrying evolved mitochondria from horizontal transmission as a group had lower fitness than the un-evolved host carrying evolved mitochondria from vertical transmission as a group (p=0.0030 in Mann-Whitney test). Red data points show horizontal transmission lines, green data points show vertical evolution lines in ‘a' mating-type selection and blue data points show vertical evolution lines in 'α' mating-type selection (see Methods). Numbers shown as data points represent their experimental evolution labeling. The horizontally-selected lines have half the data points of vertically-selected lines due to the experimental setup (Fig. 1, A vs. B). Fitness assays were repeated as seven biological replicates of mitochondria transfer clones. Error bars represent SEM around the mean.

Results for evolved mitochondria pools differed between horizontal and vertical evolution regimes. Fitness of the un-evolved host that carried the horizontally evolved mitochondria pools was significantly lower than the fitness of the same un-evolved host that carried the vertically evolved mitochondria pools ($p= 0.0030$ in Mann-Whitney test, Fig. 4).

Discussion

As predicted by theory, I found that the mitochondria that had evolved under greater horizontal transmission were less beneficial when transferred back to the ancestor than those that evolved under strictly vertical transmission (Figure 4). This might be because they had evolved to be more selfish since they depend less on the host genome for their fitness than vertically transmitted mitochondria do. However, the two treatments did not significantly affect the overall rate of adaptation under experimental evolution: populations under both treatments had adapted approximately equally, despite the mitochondria under the horizontal transmission treatment being less mutualistic or more selfish. Presumably, this occurred because the nuclear genomes that had co-evolved with more selfish or less mutualistic mitochondria evolved compromising mutations, in an arms race. Therefore, the reduced fitness of ancestral yeast strains carrying mitochondria that evolved under horizontal transmission can be seen as a form of the Bateson-Dobzhansky-Muller (BDM) incompatibility, because these mitochondria do not reduce the fitness of their co-evolved parents.

No difference in overall adaptation under vertical and horizontal transmission:

Here my expectation was to detect less overall adaptation in the horizontal transmission treatment. This is firstly because selfish mitochondria might reduce the host fitness more because it may spread more throughout the host population under this treatment. Second, from the host's point of view, mutations that allow nuclear genome to compensate ('compensatory mutations') harmful effects of mitochondria are more likely to coevolve under vertical transmission treatment rather than horizontal transmission treatment (Chou & Leu 2015). Mutations in the nuclear genome would be of more effective in counteracting mitochondrial mutations since mitochondria inside the cells are less polymorphic under a more vertical transmission. Compensatory mutations in the vertical treatment would be more effective as these mutations can efficiently compensate against a single mitochondrial genotype under vertical transmission treatment. On the contrary, under horizontal transmission treatment, mutations that can protect the interests of the host against the selfish

mitochondria might be ineffective and therefore costly because of a higher polymorphism of mitochondria pools of hosts.

Contrary to predictions, there was no fitness difference between the two transmission groups: mean values of growth rate differences in overall adaptation of co-evolved nuclear and mitochondrial genomes did not differ significantly between horizontal and vertical evolution groups. This might be because even though the evolved mitochondria became more selfish under a more horizontal transmission and reduced overall fitness of the host and the endosymbiont, nuclei could have adapted more rapidly under this treatment to compromise harmful effects of selfish types efficiently. The reason can be that under the horizontal transmission outcrossed sex would increase the adaptive potential of the host itself in comparison to the other treatment that has more clonal reproduction. Recombination of beneficial mutations under the same nuclear host via inter-breeding can increase host fitness more rapidly. Under strict vertical transmission, since recombination of diverse nuclear genomes is prevented, beneficial mutations competing under independent cellular lineages can slow down the rate of adaptation (e.g. clonal interference). As a result, although a more mutualistic co-evolution is predicted under the vertical transmission treatment, more outcrossed sex might increase host fitness more effectively and rapidly under the horizontal transmission treatment, and this difference can eventually balance out overall adaptation levels between the two treatment populations.

No difference in the dependence on mitochondria of nuclei evolved under the two treatments:
Here the expectation was to see a difference in their dependence on mitochondria of nuclei evolved under the two treatments. I expected nuclei of the populations from the vertical transmission treatment to evolve greater dependence on their mitochondria. This is predicted to be so because under vertical transmission a cellular lineage can increase in frequency more efficiently within a population if its nuclei and mitochondria cooperate more. Such a co-evolution would also allow accumulation of more efficient compensatory mutations in nuclear genomes that neutralize mitochondrial mutations under the vertical transmission treatment. The prediction then is that if I take away mitochondria from vertical transmission populations, hosts would suffer more from a reduction in fitness. Expectation for the nuclei evolution under the horizontal transmission populations was opposite. They were expected to depend less onto their mitochondria. Since mutations in nuclear genomes are less likely to be effective in neutralizing negative effects of more polymorphic mitochondria, instead of compensating mutations, ‘compromising mutations’ that increase the fitness of the host

independent of mitochondrial mutations are likely to be selected, rendering nuclei more independent from its endosymbiont partners. Then the prediction for this treatment is that even if I take away mitochondria from horizontal evolution populations, host would suffer less from a reduction in fitness, in comparison to the other treatment.

Contradicting the expectation, removal of functional mitochondria did not lead to a difference in fitness between two evolved groups of host populations (Figure 3). There was no statistically significant reduction in host fitness in the vertical transmission populations. This was calculated by comparing the relative fitness differences between the vertical transmission treatment and the horizontal transmission treatment before and after the removal of functional mitochondria. The reason for this lack of difference might be that the length of the experiment was too short to detect the evolution of such a trend. As another option, maybe the co-evolution between nuclei and mitochondria of the vertical transmission treatment was not robust enough to significantly increase the fitness of this system via a positive epistasis between the nuclei and mitochondria (e.g. Zeyl *et al*. 2005). Thus under both modes of transmission most of the evolutionary adaptation could be a result of nuclear genome evolution and therefore, removal of mitochondria from both population of hosts will have no significant reduction in their fitness levels.

Adaptation is mostly due to selection acting on the nuclear DNA:

In line with the final statement above, it is known that most of the functional genes in eukaryotes are encoded from the nuclei - in comparison to only eight protein-coding genes encoded from mitochondria that function in respiration (Procházka *et al*. 2012). Thus, most of the adaptation observed in both selection lines (Figure 2) could be the result of selection acting mainly on the nuclei rather than the interaction between nuclei and mitochondria. Results are in favor of that prediction. I see that individual lines follow similar patterns in their fitness levels comparing fitness levels of each line looking at both nuclei and mitochondria and fitness levels of each line after taking away functional mitochondria. For instance, in the horizontal transmission treatment, line-1 showed no fitness increase in comparison to the ancestor, with or without its functional mitochondria (between Fig. 2 and Fig. 3). Similarly, in vertical transmission treatment, line-5 (‘α’ mating-type in blue in Fig.2 and Fig. 3) shows no fitness increase in comparison to the ancestor, with or without its functional mitochondria. Concordance seen in fitness patterns of individual lines with or without functional mitochondria suggests that populations of both treatment regimes have adapted to the selection environment mostly due to the selection that has acted on the nuclei.

Horizontal transmission selected for more selfish behavior between mitochondria and their hosts:

The central goal of this study was to investigate whether a difference in the services provided by the mitochondria would evolve due to differences in their mode of transmission. The prediction is that the horizontal transmission treatment would evolve more selfish or less mutualistic mitochondria in comparison to mitochondria evolved under the vertical transmission treatment populations. The reason was that vertically transmitted mitochondria would spread in the population by improving the fitness of its host (mutualism) or by not harming its host (neutral evolution). On the contrary, under horizontal transmission, mitochondria can increase its own fitness by moving into different hosts without depending on a certain host. Therefore, under horizontal transmission it is more likely for mitochondria to evolve by accumulating "selfish mutations" allowing them to over-replicate, and therefore have less mutualistic or more selfish effects on their hosts. This then could be detected as different fitness effects of mitochondria from two treatments in a common host.

To see if this was true, I transferred horizontally or vertically evolved mitochondria into a common host with no functional mitochondria, and performed relative fitness assays by mixing these transfer lines with the same host that carries mitochondria. Confirming the prediction, mitochondria that evolved under a greater horizontal transmission were less beneficial when transferred back to the un-evolved host, whereas mitochondria that evolved under strict vertical transmission treatment had more or less a neutral effect on the fitness of the un-evolved host (Figure 4). This result suggests that mode of transmission can indeed evolve interactions between endosymbionts and their hosts towards significantly different directions of fitness outcomes.

Nonetheless, to have a more clear understanding on the nature of host-endosymbiont interactions evolved, frequencies of selfish mitochondria (e.g. petites) formation rates within a population of hosts should also be quantified comparing two treatment populations. Further, measuring the copy number of mitochondrial genome per host per population would be informative to underpin the nature of interactions between host and endosymbiont.

The Red-Queen Dynamics and population divergence:

One form of reproductive barrier that is observed between closely related species is a result of incompatible interspecific interactions seen between organelle genomes and nuclear genomes.

Known as the cyto-nuclear incompatibilities, as in the form of Bateson-Dobzhansky-Muller (BDM) model, such interactions can cause hybrid breakdown (Chou & Leu 2015). There is growing evidence from diverse taxa suggesting that cyto-nuclear BDM incompatibilities are common in forming barriers to hybridization (Burton *et al*. 2013; Greiner *et al*. 2011). We can explain common observations of cyto-nuclear incompatibilities by setting up a hypothesis in the context of Red-Queen dynamics (Daugherty & Malik 2012) based on the data presented in Figure 4. Since mutations in mitochondrial DNA can easily evolve selfish types that would spread in populations, compensatory mutations in the nuclear DNA can evolve to restore the organismal fitness (Chou & Leu 2015). However, every time the nuclear genome compensates the mitochondrial mutations and restores cellular fitness, novel mutations within the mitochondrial genomes would constantly allow the mitochondrial DNA to escape. Therefore nuclear DNA will need to re-evolve compensatory mutations. This constant red queen dynamics seen in an intra-genomic level may result in allopatric populations diverging more rapidly from each other (see Hurst 1992). When independently co-evolved populations reunite as hybrids of nuclear and mitochondrial genomes, such intra-genomic hybrids would suffer from incompatibilities (Paliwal *et al*. 2014). Evolved-mitochondria and un-evolved host of horizontal lines (Figure 4) may be a result of the evolution of derived-ancestral type of negative epistasis between evolved-mitochondria and un-evolved-nuclei. Hence, the reduced fitness of ancestral yeast strains carrying mitochondria that evolved under horizontal transmission can be seen as a form of the Bateson-Dobzhansky-Muller (BDM) incompatibility, because these mitochondria do not reduce the fitness of their co-evolved parents. However to see if this interpretation holds true, mitochondria transfers should be done between evolved mitochondria pools and evolved hosts in a long-term experimental evolution study.

Conclusion

There is growing evidence showing that horizontal transmission of selfish genetic elements allow their spread through the population of hosts despite their harmful effects on hosts. Sexual reproduction (analogous to the horizontal transmission) allows the spread of selfish genetic elements as a result of intra-genomic conflicts whereas asexual reproduction (analogous to the vertical transmission) limits the spread of such elements. This has been shown to be true for homing-endonuclease genes (Goddard *et al*. 2001), 2μ plasmids (Futcher *et al*. 1988), transposable elements (Zeyl *et al*. 1996), and recently for the selfish mitochondria of yeast (Harrison *et al*. 2014). This study provides more direct (e.g. fitness cost over the host – Figure 4) evidence to that growing body of research. I show that mitochondria

can evolve in different directions in the continuum of mutualism and selfishness in their interactions with the host due to the differences the mode of transmission. From a more general perspective, the findings represented here suggest that the extension of the continuum hypothesis to a host-endosymbiont context is possible.

Finally, before finishing, it is important to point out one limitation of my experimental setup. Although I could manipulate endosymbiont transmission between two extremes of horizontal and vertical transmissions during the fertilization steps (Figure 1-A/B), still most of the cell divisions in my evolution experiments were mitotic divisions. Thus, in the horizontal evolution regime the majority of the endosymbiont inheritance was via vertical transmission (see also Lipsitch *et al*. 1996). This is simply a limitation of studying these interactions in a non-obligate sexual microbe: Sexual cycles can only infrequently be induced within a large number of asexual divisions in this system. As clearly pointed by Hastings (1999), low frequency of sexual divisions (e.g. horizontal transmission) in between frequent asexual divisions (e.g. vertical transmission) can greatly reduce the evolution of more selfish interactions. Nevertheless it is encouraging to see that the final result presented here supports theoretical predictions in spite of such a conservative experimental setup.

MOLECULAR ECOLOGY

Molecular Ecology (2014) 23, 5061–5071 doi: 10.1111/mec.12904

The genetics of a putative social trait in natural populations of yeast

G. O. BOZDAG* and D. GREIG* †

*Max Planck Institute for Evolutionary Biology, August Thienemann Strasse 2, Plön 24306, Germany, †The Galton Laboratory, Department of Genetics, Evolution, and Environment, University College London, Gower Street, London WC1E 6BT, UK

Abstract

The sharing of secreted invertase by yeast cells is a well-established laboratory model for cooperation, but the only evidence that such cooperation occurs in nature is that the *SUC* loci, which encode invertase, vary in number and functionality. Genotypes that do not produce invertase can act as 'cheats' in laboratory experiments, growing on the glucose that is released when invertase producers, or 'cooperators', digest sucrose. However, genetic variation for invertase production might instead be explained by adaptation of different populations to different local availabilities of sucrose, the substrate for invertase. Here we find that 110 wild yeast strains isolated from natural habitats, and all contained a single *SUC* locus and produced invertase; none were 'cheats'. The only genetic variants we found were three strains isolated instead from sucrose-rich nectar, which produced higher levels of invertase from three additional *SUC* loci at their subtelomeres. We argue that the pattern of *SUC* gene variation is better explained by local adaptation than by social conflict.

Keywords: cheating, cooperation, copy number variation, droplet digital PCR, *Saccharomyces*, *SUC*

Received 21 January 2014; revision received 21 August 2014; accepted 25 August 2014

Introduction

In contrast to other eukaryotes, the genome of *Saccharomyces cerevisiae* is compact, containing few redundant genes or pseudogenes (Goffeau et al. 1996; Lafontaine et al. 2004). The *SUC* genes, which encode the extracellular enzyme invertase, are exceptional. There are nine known loci for *SUC* genes: *SUC1* –*SUC5* and *SUC7* – *SUC10* (Naumov & Naumova 2010). *SUC2*, the ancestral locus, is located in the left arm of chromosome IX, but the other copies are all found at subtelomeric regions (Carlson & Botstein 1983; Carlson et al. 1985; Naumov & Naumova 2010). Most strains contain only a single *SUC2* gene, but some contain one or more of the subtelomeric *SUC* loci in addition to *SUC2*, and others have a *suc2* pseudogene and produce no invertase (Carlson & Botstein 1983; Naumov et al. 1996; Denayrolles et al. 1997).

The variation in *SUC* genotypes can be explained using social evolution theory (Greig & Travisano 2004). The invertase produced from *SUC* genes is secreted to digest extracellular sucrose into the preferred sugars glucose and fructose, which can be taken up by the cell and metabolized. Sugars diffuse readily, so cells that cannot produce invertase themselves because they lack any functional *SUC* genes can use the glucose and fructose produced by those that do (Gore et al. 2009). Thus, invertase production is analogous to public goods cooperation with nonproducers as cheats that can exploit and invade populations of cooperators (Greig & Travisano 2004). The phenomenon of telomeric silencing (Wyrick et al. 1999) has been used to explain subtelomeric *SUC* loci. If *suc2* cheats retain an unexpressed but functional copy of *SUC*, then having invaded a colony of cooperators and depleted the public good, they could regain the ability to produce invertase from a silent subtelomeric 'backup' copy (Greig & Travisano 2004). Consistent with this social theory model, laboratory experiments find that the relative fitness of nonproducers can be higher or lower than that of producers, depending on factors such as density (Greig & Travisano 2004), frequency (Gore et al. 2009; Damore & Gore

Correspondence: Gonensin O. Bozdag, Fax: +49 4522 763-260; E-mail: ozan.bozdag@gmail.com

2012) and sucrose concentration (Koschwanez et al. 2011). However, a recent experiment found that mixed cultures of producers and nonproducers had higher mean fitness than monocultures of producers, inconsistent with the model of nonproducers as cheats (MacLean et al. 2010).

An alternative explanation for SUC variation is that different SUC genotypes have adapted to environments with different availabilities of sucrose. For thousands of years, humans have used yeast to make alcohol, and more recently, to raise bread, to flavour foods, to study genetics and to secrete bio-engineered products such as insulin (Thim et al. 1986; Botstein & Fink 1988; Porro et al. 1995). A survey of the drinks available in a typical bar reveals some of the diverse substrates that domesticated yeast strains are grown on. Yeast produces invertase constitutively, even in the absence of sucrose, although high levels of glucose can suppress invertase production (MacLean et al. 2010). Substrates low in sucrose might favour the loss of costly invertase production and the selection of suc2 null mutants. Conversely, substrates rich in sucrose might select for additional subtelomeric copies of SUC if they were not completely silenced and could therefore contributed to increased invertase production (Denayrolles et al. 1997; Batista et al. 2004). Thus, the observed diversity in SUC genotypes may simply be due to domestication in different environments (Libkind et al. 2011). Similar increases in diversity are seen in other domesticated species, for example domesticated dogs have much greater morphological variation than wolves, their wild ancestors (Wayne 1986) indeed the range of body sizes in among breeds of this single domesticated species exceeds the range of all other wild canid species (Lindblad-Toh et al. 2005).

These two competing hypotheses can be tested by examining how individuals vary within and between habitats. The social conflict hypothesis predicts that different strains isolated from the same habitat will differ in their SUC genotypes, because some will be cheats and others will be cooperators. The sucrose adaptation hypothesis predicts that different strains from the same habitat will have the same SUC genotype, but that strains from environments with different sucrose availabilities will differ. Naumov et al. (1996) surveyed SUC gene variation in 91 strains isolated from many different environments, finding eleven invertase nonproducing strains that contained only a nonfunctional suc2 allele. Five of these came from olive processing (Vaughan & Martini 1987), and two came from human faeces (Naumov et al. 1990), environments that are low in sucrose (Marsilio et al. 2001), which is consistent with the sucrose adaptation hypothesis. But two came from wine, an environment that also provided many invertase producers, consistent with the social conflict hypothesis. The remaining nonproducer (GM51) has unknown origins (Naumov et al. 1996). Ten strains had multiple SUC genes, and all came from sucrose-rich environments (strawberry, grape, ginger wine, billi wine and palm wine, Naumov et al. 1993; Basson et al. 2010; Kim & Lee 2006) or from fermentations that are artificially supplemented with sucrose (distilling and champagne making, Naumov et al. 1996), supporting the sucrose adaptation hypothesis. These results are difficult to interpret because different lineages of S. cerevisiae are often genetically mixed (Liti et al. 2009), perhaps by the process of human domestication (Libkind et al. 2011), and because many of the strains were not systematically isolated and their origins are unclear.

We therefore decided to determine the frequency of invertase nonproducers in S. paradoxus, the wild relative to S. cerevisiae. S. paradoxus has several advantages over S. cerevisiae for this study. The most important is that S. paradoxus is not used in human fermentations and instead has a well-established and well-sampled natural habitat extending to several continents on oak trees (Naumov et al. 1998; Johnson et al. 2004), and on Canadian maple trees (Charron et al. 2014). Unlike S. cerevisiae, whose populations show little geographical structure and high gene flow, perhaps because humans move strains around the world and mix them (Liti et al. 2009), S. paradoxus populations have strong geographical structure, with little mixing between lineages from different places (Kuehne et al. 2007; Liti et al. 2009). These properties mean that any S. paradoxus strain is likely to have evolved in the environment from which it was isolated, and is very unlikely to be a recent immigrant adapted to a different environment or to contain genetic material from such an immigrant. To test the hypothesis that social conflict should produce SUC variation among individuals within a single type of habitat, we determined the invertase production, the SUC loci and the SUC gene copy number of a set of 80 S. paradoxus strains: 65 isolated from oak trees and 15 isolated from maple trees. We did not have similarly large sets of S. cerevisiae strains from well-defined habitats and cannot exclude the possibility that some wild-caught strains might originate from human fermentations, or be related to such feral escapees. Nevertheless, we also tested 30 S. cerevisiae strains we could find that were isolated from apparently natural sources, including 15 recent isolates from primeval forests, which form distinct lineages compared to all the other S. cerevisiae strains identified so far (Wang et al. 2012). Finally, we tested whether strains with SUC2 deleted produced invertase from their subtelomeric SUC loci or whether they were 'silent'.

Materials and methods

All strains, their original strain numbers, references, details of their origins and inclusion in genome sequencing projects are described in Appendix S1 (Supporting information).

To determine whether SUC genotypic variation occurs within (rather than between) wild populations, it is necessary to have multiple examples of wild strains isolated from a single well-defined habitat. S. paradoxus is ideal for this because it is not domesticated and many strains have been systematically isolated from oak trees. We tested all the oak-associated strains that we could access, including 29 that we isolated ourselves in Germany, 25 from the United Kingdom, 7 from Russia, 3 from North America and 1 from Japan (see Appendix S1, Supporting information for details). More recently, Canadian maple trees have been identified as a habitat for S. paradoxus, and we included 15 strains of S. paradoxus isolated from Canadian maple trees (Charron et al. 2014). We tested all S. paradoxus strains that we could acquire, but we excluded single strains isolated from unique or poorly described habitats and those from insect vectors which might have fed on unknown substrates.

We were concerned that any S. cerevisiae strains we tested might have recently escaped from human fermentations, or might have been crossed to such feral strains. Further, the natural habitat of S. cerevisiae is less well established than that of S. paradoxus. We therefore focused primarily on S. paradoxus. However, a large set of S. cerevisiae has recently been isolated from primeval forests in China, far from human influence, and there is good evidence they represent a truly wild population (Wang et al. 2012). We were able to get hold of 15 of these strains to test (7 from rotten wood, 2 from soil, 2 from oak, 2 from beech and one each from persimmon and oriental raisin trees). The majority of other available S. cerevisiae strains have been isolated from human fermentations or associated places, such as vineyards and food processing facilities. However, we were able to find 15 additional S. cerevisiae strains from a variety of apparently natural habitats: 5 from oak, 3 from soil, 3 from Bertram palm nectar and one each from cactus, cactus fruit, fig and cocoa.

We also tested various S. cerevisiae strains as controls and for comparison purposes. Our standard control strains were C.Lab.1. and C.Lab.1. suc2::KANMX , are isogenic with strains that have been used as a 'cooperators' and 'cheats', respectively, in previous laboratory studies on cooperation (Greig & Travisano 2004; MacLean & Brandon 2008; Gore et al. 2009; MacLean et al. 2010). We included two domesticated strains, C.Ginger.wine and C.Billi.wine, as positive controls with known multiple SUC copies. Finally, we knocked out the SUC2 loci from these two strains as well as from the three wild strains that turned out to have multiple SUC copies, creating five new strains: C.Ginger.wine. suc2::NATMX, C.Billi.wine. suc2::NATMX , C.Nectar.1. suc2::NATMX , C.Nectar.2. suc2::NATMX and C.Nectar.3. suc2::NATMX .

Screening wild strains for invertase nonproducers

To determine which of our strains produced invertase, we used the Glucose (HK) Assay Kit (Sigma-Aldrich, St. Louis, MO, USA), which produces a colorimetric reaction in response to glucose. We calibrated the assay using known dilutions of purified invertase (Sigma-Aldrich). Twenty microlitre of each dilution was combined with 100 l L sodium acetate buffer (0.2 M, pH = 5.2), and 50 l L of 0.5 M sucrose added. The reaction was incubated at 37 °C for 20 min, then stopped by adding 300 l L of 0.2 M K_2HPO_4 and heating at 100 °C for 5 min. One-hundred and fifty microlitre of this reaction mixture was added to 1 mL glucose assay reagent provided by the kit, and the optical absorbance at 340 nm was determined after following the kit instructions. We found the assay gave a linear response between absorbances of 0.11 and 0.78 (Appendix S4, Supporting information).

We optimized the assay conditions using a laboratory strain, C.Lab.1., which produces invertase from a single SUC2 locus and has been used as a 'cooperator' in previous work on sociality (Greig & Travisano 2004; MacLean & Brandon 2008; Gore et al. 2009; MacLean et al. 2010). Each strain was grown in 2 mL of YEPD (1% yeast extract, 2% peptone, 2% dextrose) overnight at 30 °C. We spun down 1 mL of the culture, washed it with 1 mL of sterile water and centrifuged again. The pellet was resuspended in 1 mL of 0.9% sterile saline, and 5 l L of the cell suspension was spotted onto a YEPS plate (1% yeast extract, 2% peptone, 2% sucrose and 2.5% agar) and incubated it for 2 days at 30 °C. The resulting colony was then resuspended in 5 mL of sterile water, and a 100 l L sample was spun down and washed twice, then resuspended in 50 l L of sterile water, combined with 100 l L sodium acetate buffer and 50 l L of 0.5 M sucrose, incubated at 37 °C for 20 min and stopped with 300 l L of 0.2 M K_2HPO_4 heating as described above. One-hundred microlitre of the reaction mixture was added to 1 mL glucose assay reagent, and the absorbance was read. We used the same method on the isogenic 'cheat' strain C.Lab.1. suc2::KANMX . Pilot experiments indicated that wild strains produced so much more invertase than the laboratory 'cooperator'

strain C.Lab.1 that they saturated the assay, so we reduced the volume of the resuspended cells from 100 to 20 l L, making the suspension up to 100 l L with 80 l L of the nonproducer C.Lab.1. suc2::KANMX prepared in the same way. Measurements were then multiplied by five to correct for this dilution. We screened the invertase production of all 110 wild strains in this way (see Appendix S1, Supporting information).

SUC alleles in whole genome sequences

Whole genome sequences were available for 29 S. paradoxus and 8 S. cerevisiae strains. For details of which strains had sequences, and where the sequences can be accessed, please see Appendix S1 (Supporting information). These genome sequences were used to determine whether a strain contained intact SUC open reading frames or suc pseudogenes. The nucleotide sequences of SUC genes that were identified in this way are listed in Appendices S5 and S6 (Supporting information).

Southern blots for SUC loci

Whole genome sequences were not available for most of our strains, and even for the 29 S. paradoxus and 8 S. cerevisiae that had been sequenced, we could not reliably infer the SUC loci or copy numbers from the sequences because of the short reads and low sequencing coverage. Subtelomeric SUC genes are embedded in highly repetitive DNA which may not be properly assembled in genome sequencing projects. To determine the SUC loci in our wild strains, we therefore made Southern blots of whole-chromosome pulsed-field gels, and probed them with labelled SUC2 fragments. We also included controls on the pulsed-field gels: the C.Lab.1. suc2::KANMX as nonproducer containing no known SUC genes, C.Lab.1 as a producer containing a single SUC2 gene and the domesticated strains C.Ginger.wine and C.Billi.wine as positive controls previously identified as containing multiple SUC loci (Naumov et al. 1996). All strains are described in Appendix S1 (Supporting information).

We prepared chromosomal DNA plugs according to Carle & Olson (1985). S. cerevisiae CHEF DNA size standard (YNN295 strain) was used in all pulse-field gel electrophoresis runs (Bio-Rad, Hercules, CA, USA). After the pulse-field gel electrophoresis (0.5 9 TBE, 14 °C, 200 V for 15 h with 60-s switching time, and for 8 h with a 90-s switching time), DNA was transferred to positively charged nitrocellulose membrane (GE Healthcare, Buckinghamshire, UK).

The number and chromosomal location of each SUC locus were determined by probing the membrane with DIG-labelled probes (Eurofins, Ebersberg, Germany). S. paradoxus and S. cerevisiae probes were designed according to the most conserved 5 ^{0} regions of SUC2 gene.

Hybridization and detection reactions were carried out according to the Roche's DIG High Prime DNA Labeling and Detection Starter Kit 1 (Roche, Mann-Heim, Germany).

S. paradoxus SUC2 probe sequence:

CGTCTGGGGTACGCCATTGTATTGGGGCCATGCT
ACTTCCGATGATTTGACCCACTGGCAAGACGAA
CCCATTGCTATTG

S. cerevisiae SUC2 probe sequence:

ATGACAAACGAAACTAGCGATAGACCTTTGGTC
CACTTCACACCCAACAAGGGCTGGATGAATGAT
CCAAATGG

ddPCR for SUC copy number

The Southern blots of whole-chromosome pulsed-field gels could detect SUC loci in addition to SUC2 . But because each chromosome has two telomeres, and because different chromosomal bands can colocalize on the gel, it cannot be used to precisely determine SUC copy number in strains that have subtelomeric copies of SUC in addition to SUC2 . We therefore used droplet digital PCR (Bio-Rad QX100 system) to determine SUC copy number in the strains that had been determined by Southern blotting to contain multiple SUC loci, as well as in the 15 Chinese S. cerevisiae strains which we received most recently (as an alternative to Southern blotting). ddPCR uses simultaneous duplex reactions for target and reference genes within a single tube that contains about 20 000 reaction microdroplets, which are individually scored as positive or negative for the presence of amplicons by TaqMan fluorescence (see Huggett et al. 2013 for an introduction to the digital PCR technology). We used prevalidated TaqMan gene probes and primers designed by Life Technologies (CA, USA) for SUC2 (VIC, Sc04134115_s1) and two reference genes RPN5 (FAM, Sc04107686_s1) and MNN1 (FAM, Sc04117288_s1).

We isolated genomic DNA (MasterPureTM Yeast DNA Purification Kit, Epicentre Biotechnologies) from C.Lab.1 as a single-copy control, C.Lab.2. suc2::KANMX as a zero copy control, the two strains identified by a previous study as having multiple SUC loci (C.Ginger.wine and C.Billi.wine; Naumov et al. 1996) as positive controls, as well as the wild strains to be tested (Please see Appendix S1, Supporting information). Genomic DNA was restricted with Hin dIII, as this enzyme has a conserved cut site within the SUC2, RPN5 and MNN1 open reading frames, but outside the

We also confirmed that SUC2 was deleted from these five strains using the CHEF Southern blot (see Fig. S1, Supporting information). We then performed quantitative invertase production assays using the conditions described above (under 'Variation in invertase production') on the five wild-type strains (C.Nectar.1, C.Nectar.2 and C.Nectar.3; C.Ginger.wine and C.Billi.wine) and the five SUC2 knockouts derived from them (C.Nectar.1. suc2::NATMX, C.Nectar.2. suc2::NATMX, C.Nectar.3. suc2::NATMX, C.Ginger.wine. suc2::NATMX, C.Billi.wine. suc2::NATMX, respectively). As mentioned before, three independent replicates were made of each assay, allowing quantitative comparisons to be made (raw data are on Appendix S3, Supporting information).

Results

No wild strains were invertase nonproducers

Figure 1 shows that all the 110 wild strains that we tested produce more invertase than the standard invertase nonproducer or 'cheater' used in several previous experiments about cooperation (Greig & Travisano 2004; MacLean & Brandon 2008; Gore et al. 2009; MacLean et al. 2010). None of the 110 wild strains produced invertase at a level low enough to fall within the 95% confidence interval around the residual invertase activity of the standard nonproducer laboratory strain, C.Lab.1.-suc2::KANMX. In fact, all the wild strains also had higher invertase activity than the 95% confidence interval around the invertase activity of the standard laboratory producer strain, C.Lab.1. We applied a Tukey post-hoc test to a one-way ANOVA on these three groups and found that the 110 wild strains, as a group, produced significantly more invertase than both nonproducer and producer laboratory strains ($F_{2,113} = 125.5$, $P < 0.0001$).

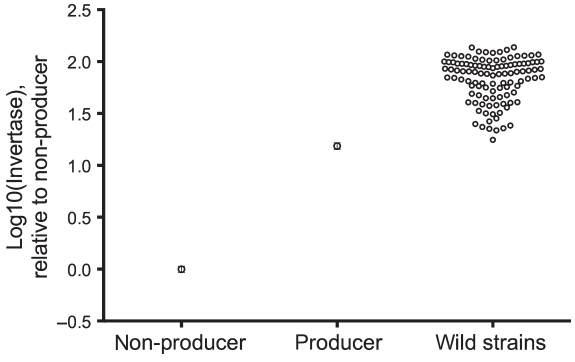


Fig. 1 Screening of wild strains for invertase nonproducers. The invertase production of the 110 wild strains screened, as well as the standard laboratory invertase-producer strain C.Lab.1, is shown, relative to the production of the standard laboratory invertase nonproducer strain C.Lab.1. suc2::KANMX. All strains are described in Appendix S1 (Supporting information), and all data are listed in Appendix S2 (Supporting information).

No suc2 pseudogenes were detected in wild strains

Whole genome sequences existed for 29 S. paradoxus strains (Liti et al. 2009; Bergström et al. 2014). Consistent with their ability to produce invertase, we found intact open reading frames (ORFs) homologous to the reference S. cerevisiae strain (s288c/C.Lab.1) in all these strains. The length of the ORF was identical among all 29 S. paradoxus strains. Also for 8 S. cerevisiae strains, we found intact ORFs homologous to the reference strain (SGRP1: Liti et al. 2009; SGRP2: Bergström et al. 2014). There were no frameshift or nonsense mutations in any of the wild strains for which sequence was available (see Appendices S5 and S6, Supporting information for the SUC2 nucleotide sequences identified in the wild strains used in this study).

Three S. cerevisiae strains contained additional SUC genes

Our Southern blots showed that all the wild S. paradoxus strains isolated from oak and maple trees contained just a single SUC locus, SUC2, located on chromosome IX. All 27 S. cerevisiae strains isolated from nature also contained SUC2 on chromosome IX, but three S. cerevisiae strains (C.Nectar.1, C.Nectar.2 and C.Nectar.3) contained additional SUC loci on chromosome II (SUC3), on chromosome X (SUC8) and on chromosome XIV (SUC9) (Figs S1 and S2, Supporting information). ddPCR (Fig. 2) shows that the SUC copy number of the three wild strains with multiple loci is closest to four, corresponding to one SUC open reading frame for each chromosome with a SUC locus (SUC2, plus the extra loci SUC3, SUC8 and SUC9). All three of these wild strains were isolated from the same environment: Bertam palm (Eugeissona tristis) nectars in West Malaysia (Liti et al. 2009).

Producers vary in their invertase production

We found that 11 different S. cerevisiae strains isolated from nine different domestic and wild environments varied significantly in their invertase production (Fig. 3; $F_{10,22} = 39.92$, $P < 0.0001$). Post-hoc Tukey tests (letter above the bars in Fig. 3) found that some, but not all, strains with four SUC copies produced significantly more invertase than strains with a single copy; some strains also produced significantly more invertase than other strains that had the same number of SUC copies.

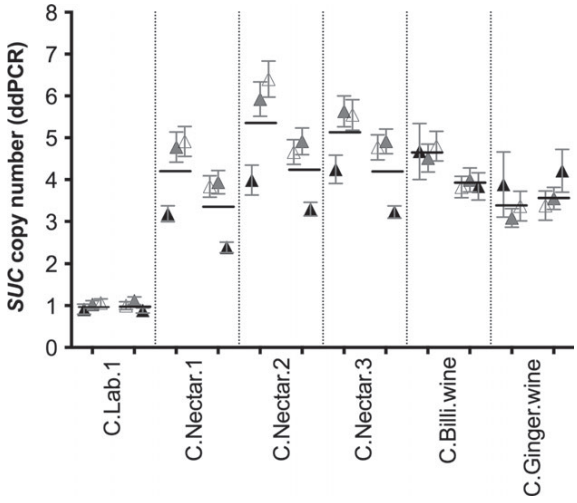


Fig. 2 SUC gene copy number detection using droplet digital PCR (ddPCR) in five multilocus strains, normalized to a known single-copy control C. Lab.1 (first column). Three different symbol tones (dark, grey and empty) represent three different biological replicates. Copy number estimates calculated against RPN5 reference probe are on the left-hand side of each column, and copy number estimates calculated against MNN1 are on the right-hand side of each column. Black bars show the means of each set of three biological replicates.

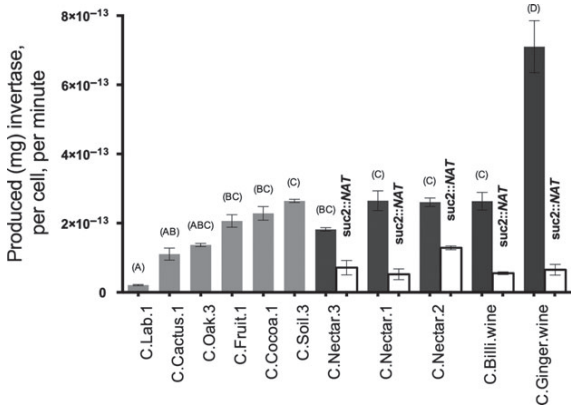


Fig. 3 Light grey bars show the mean invertase production per cell for the single-copy standard producer and 5 other single-copy S. cerevisiae strains from different wild habitats. Dark grey bars show the production for the three strains with subtelomeric SUC loci isolated from Bertram palm nectar, and two strains from domesticated origins with subtelomeric loci as controls. Letters above the filled bars indicate which strains differ from each other with respect to their wild-type invertase production: strains with a letter in common are not significantly different. Open bars show the residual production of invertase from subtelomeric loci after SUC2 was knocked out. Three replicate assays were made for each strain; error bars show the standard error of the mean.

When grouped by number of SUC copies, the five strains with multiple SUC copies produced significantly more invertase than the six strains containing only SUC2 (Student's t-test, $P = 0.0023$, $t = 3.32$, $DF = 31$), but this difference was driven by two strains (laboratory strain C.Lab.1 and domesticated strain C.Ginger.wine): when the analysis was repeated on the wild strains alone, no significant difference was detected between the strains with 1 SUC copy and the strains with four copies (Student's t-test, $P = 0.0713$, $t = 1.8951$, $DF = 22$). Thus, it was unclear whether or not additional subtelomeric copies of SUC contributed to the variation in invertase production, or whether it was caused simply by variation in SUC2 expression. We therefore decided to test directly, by knocking out SUC2 , whether the additional subtelomeric copies of SUC are expressed or whether they function only as silent backup copies for 'cheats'.

Subtelomeric SUC copies are not silent

Figure 3 shows that SUC2 contributes much more to total invertase production than subtelomeric copies of SUC . Knocking out SUC2 in the five strains with multiple copies reduces invertase production in every case, a statistically significant effect ($P = 0.0312$, paired sign test). The average reduction in invertase when SUC2 was deleted was 64%, suggesting that each of three subtelomeric SUC genes contributes only about 12% to total invertase production. But subtelomeric copies are far from silent: the SUC2 knockouts all produce more invertase than the standard laboratory producer strain C.Lab.1 (Fig. 3).

Discussion

Our results do not support the hypothesis that natural variation in SUC genes is caused by social conflict (Greig & Travisano 2004). It is more likely that different SUC genotypes are selected by habitats with different availabilities of sucrose (Naumov et al. 1996), but our survey does not contain enough variation for us to be certain.

Invertase nonproducers

Our main aim was to determine whether invertase nonproducers existed in the same natural habitats as producers, which would be required in order for nonproducers to cheat. We found no nonproducers among the 65 oak-associated S. paradoxus strains that we tested nor among the 15 strains from maple trees. Unfortunately, the other habitats included in the survey had only a few strains available from each, so we might not find both producers and nonproducers cooccurring

in the same type of habitat even if they were there. Nevertheless, we found no nonproducers at all among a total of 110 different wild strains (Fig. 1). There is therefore no evidence to support the idea that nonproducing cheats occur among wild strains.

Our results stand in contrast to Naumov et al.'s (1996) finding of 11 nonproducers among a sample of 91 S. cerevisiae strains. One explanation is that Naumov et al. (1996) surveyed strains from a wider range of environments, which might select for or against the production of invertase according to sucrose adaptation hypothesis. Another is that most of Naumov's strains were associated with humans, whereas ours came only from natural sources. Artificial selection on domesticated species can increase diversity (Vila et al. 1999), as well as allowing loss of functions that would be important for survival in the wild (e.g. loss of pigmentation in domestic pigs and horses, Andersson & Georges 2004). It is therefore possible that invertase nonproducing mutants that would be eliminated by natural selection in the wild can persist by drift or even be selected in anthropogenic environments that are abundant in sugars other than sucrose or which lack producers as competitors. Thus, the variation observed in human-associated strains may be due to changes in environment, demography and population structure resulting from domestication. It is also possible that some domesticated environments produce conditions that allow cheating, for example by increasing population densities or environmental stability, compared to those conditions that would exist naturally, and thus, variation in domesticated strains could be due to the social conflict hypothesis. Because the evolutionary history of human-associated strains is obscure, it would be difficult to disentangle these explanations for the variation among domesticated yeast, but as there is no evidence for nonproducers and producers occupying the same habitat and abundant evidence for variability in sucrose availability, we, like Naumov et al. (1996), favour the sucrose adaptation hypothesis for domesticated strains as well as for the wild strains we describe here.

Copy number variation

A secondary aim of the project was to determine whether variation in SUC copy number was consistent with social evolution.

According to the social conflict hypothesis as originally formulated (Greig & Travisano 2004), subtelomeric SUC loci could act as transcriptionally silenced backups which can be stochastically de-repressed (Gottschling et al. 1990; Louis 1995) or which could restore function to a suc2 pseudogene by gene conversion (analogous to mating-type switching using silent telomeric copies of the hidden mating-type, HM, loci) (Naumov & Tolstorukov 1973; Hicks & Herskowitz 1977). Silent copies of SUC could allow cheats to switch back to invertase production when there are no cooperators to exploit, a form of 'facultative cheating' (Gore et al. 2009). This part of the social conflict hypothesis is now much less plausible because subsequent research has shown that subtelomeric silencing is predominantly a haploid phenomenon (Mercier et al. 2005). Indeed, the three wild strains we found with multiple SUC copies produced invertase at a high level, and they continued to do so even when SUC2 was knocked out, showing that the remaining subtelomeric SUC loci are transcriptionally active and are not silenced backup copies (Fig. 3). Further, all the strains with subtelomeric copies came from the same environment, Bertram palm nectar, and all strains from this environment contained three subtelomeric SUC alleles in addition to SUC2 : there was no genotypic variation within the environment as predicted by the social conflict hypothesis. This could simply be because our tiny sample contained only three strains, but it is also most consistent with the sucrose adaptation hypothesis. Sucrose is the major carbon source in most plant nectars (Corbet 2003; Pacini et al. 2003; Dupont et al. 2004; Wiens et al. 2006; Peay et al. 2012), and Bertam palm nectars contain high and stable concentrations of sucrose (~10%; Wiens et al. 2006). However, these strains are very closely related: C. Nectar.1 differs from C.Nectar.2 by just 0.0059% of nucleotides across the whole genome, and from C. Nectar.3 by 0.019%; C.Nectar.2 and C.Nectar.3 differ by 0.012% (Liti et al. 2009). Given the small sample size, the high genetic relatedness and the likelihood that all three strains inherited their subtelomeric SUC genes by common descent, we cannot exclude the possibility that the expansion of the SUC gene family in this environment is due to neither social evolution nor environmental selection, but simply genetic drift.

The expression of invertase from strains with subtelomeric SUC loci shows that they are not 'cheats'. However, a social model could still be used to explain their evolution if the originally proposed roles of cooperator and cheat were reversed. If strains with more SUC copies produce more invertase, they could be considered cooperators instead of cheats, and they could feed other, cheating, strains that have only SUC2 and produce less. As under the original social explanation for SUC genetic variation, we would predict that cheats and cooperators should occur in the same environment. Whilst we might not expect to detect such copy number variation among only three strains from Bertram palm nectar, we would expect to find variation within the well-sampled oak-tree and maple-tree habitats, but we did not. Instead, we find copy number variation

between (but not within) environments that differ in sucrose availability. Whilst we must be cautious not to overgeneralize from just three closely related strains, the little copy number variation we do find in our survey is clearly better explained by the sucrose adaptation hypothesis than by the social conflict hypothesis.

Is invertase production a cooperative trait?

We previously proposed the social conflict hypothesis to explain variation in *SUC* genotypes among *S. cerevisiae* strains (Greig & Travisano 2004). But because *S. cerevisiae* is domesticated, and isolates came from many different sources, it was difficult to know whether different genotypes evolved in a common environment that would permit social cheating. In this survey of wild strains, we find very little variation of *SUC* genotypes, and the limited variation we do find occurs between, and not within, environments. The genetic variation is therefore better explained by adaptation to different environmental levels of sucrose than by social conflict. However, given the lack of variation in our samples, we have very limited power to differentiate between the two hypotheses. The ideal survey would test the invertase production and the *SUC* genotype of multiple strains isolated from at last two different natural habitats that differed in their sucrose availability. Such a design would have the best chance of being able to definitively distinguish the difference between the two hypotheses explaining variation for *SUC*. If different *SUC* genotypes are selected by the local availability of sucrose, then the two environments will be fixed for different genotypes. If social conflict produces variation, then we would expect more variation within the high-sucrose environment than within the low-sucrose environment. Unfortunately, such well-sampled natural habitats differing in sucrose availability do not exist, but we hope that as research in yeast natural history progresses, such a survey may be possible in the future.

Authors have previously cited the variation in *SUC* genotypes as evidence that cheating occurs in nature (Greig & Travisano 2004; MacLean & Brandon 2008; Gore *et al.* 2009), but here we show that the evidence has been misinterpreted. This has significant consequences for the use of invertase production as an experimental model of cooperation. Cooperative traits are properly defined not merely as those traits that benefit others, which would be nonsensically overinclusive, but those traits that evolved because of the benefits they convey to others (West *et al.* 2007). Thus, it is important to show that cooperation occurs in the environment in which a putative cooperative trait evolved, and the existence of natural genetic variation was presented as evidence that invertase production evolved in nature as a cooperative trait. It is worth noting as an aside, though, that the existence of natural cheats is not sufficient to prove a trait as cooperative: we would not consider scatter-hoarding of nuts by squirrels to be a cooperative trait, even though hoarded nuts are often eaten by scroungers and not by the squirrel that buried them (Stapanian & Smith 1978). To prove that invertase production evolved as a cooperative trait, one would need to show that not only that social conflict over invertase sharing occurs in nature, but also that invertase sharing was actually selected. Surveys like ours cannot therefore determine whether or not invertase production is a cooperative trait. Even if the natural variation for *SUC* copy number is not caused by social conflict, social conflict may nonetheless underlie other forms of genetic variation for invertase production (for example, Fig. 3 shows there is considerable and significant variation in invertase production even among strains containing only *SUC2*). And even if social conflict does not cause any natural genetic variation in invertase production, it is still possible that invertase production evolved as a cooperative trait in nature. And even if it did not evolve as cooperative trait in nature, invertase sharing in an experimental setting could still be a useful model for cooperation. We are mindful, though, of the words of G.C. Williams: 'Adaptation should be attributed to no higher a level of organization than is demanded by the evidence' (Williams 1996). In our opinion, a trait should not be called cooperative until more parsimonious explanations for its evolution have been rejected.

Acknowledgements

We thank Leonid Kruglyak for four natural *S. cerevisiae* strains, Feng-Yan Bai for 15 wild *S. cerevisiae* strains isolated from China, Christian R. Landry for 15 *S. paradoxus* strains isolated from maple trees in Québec (Canada) and Eric L. Miller for providing the Plön (Germany) *S. paradoxus* strains and for his comments on the manuscript. We thank David W. Rogers for his recommendations to improve our invertase assays, Mathias Wegner for helping interpret our Southern blot analysis, and Zeljka Pezer for helping us with the ddPCR. Jai Denton and Laurence Hurst read an early draft of the manuscript, and the latter recommended Molecular Ecology. The study was greatly improved thanks to positive criticism from three anonymous reviewers. This work was supported by the Max Planck Society.

References

Andersson L, Georges M (2004) Domestic-animal genomics: deciphering the genetics of complex traits. *Nature Reviews Genetics*, 5, 202–212.

Basson CE, Groenewald J-H, Kossmann J, Cronjé C, Bauer R (2010) Sugar and acid-related quality attributes and enzyme

activities in strawberry fruits: invertase is the main sucrose hydrolysing enzyme. Food Chemistry, 121, 1156–1162.
Batista AS, Miletti LC, Stambuk BU (2004) Sucrose fermentation by Saccharomyces cerevisiaelacking hexose transport. Journal of Molecular Microbiology and Biotechnology, 8, 26–33.
Bergström A, Simpson JT, Salinas F et al. (2014) A High-definition view off unctional genetic variation from natural yeast genomes. Molecular Biology and Evolution, 31, 872–888.
Botstein D, Fink G (1988) Yeast: an experimental organism for modern biology. Science, 240, 1439–1443.
Carle GF, Olson MV (1985) An electrophoretic karyotype for yeast. Proceedings of the National Academy of Sciences, USA, 82, 3756.
Carlson M, Botstein D (1983) Organization of the SUC Gene Family in Saccharomyces Molecular and Cellular Biology, 3, 351–359.
Carlson M, Celenza JL, Eng FJ (1985) Evolution of the dispersed SUC gene family of Saccharomycesby rearrangements of chromosome telomeres. Molecular and Cellular Biology, 5, 2894–2902.
Charron G, Leducq J-P, Bertin C, Dube AK, Landry CR (2014) Exploring the northern limit of the distribution of Saccharomyces cerevisiaeand Sacchromyces paradoxusin North America. FEMS Yeast Research , 14, 281–288.
Corbet SA (2003) Nectar sugar content: estimating standing crop and secretion rate in the field. Apidologie, 34, 1–10.
Damore JA, Gore J (2012) Understanding microbial cooperation. Journal of Theoretical Biology, 299, 31–41.
Denayrolles M, de Villechenon E, Funel LA, Aigle M (1997) Incidence of SUC-RTM telomeric repeated genes in brewing and wild wine strains of Saccharomyces Current Genetics, 31, 457–461.
Dupont Y, Hansen D, Rasmussen J, Olesen J (2004) Evolutionary changes in nectar sugar composition associated with switches between bird and insect pollination: the Canarian bird-flower element revisited. Functional Ecology, 18, 670–676.
Goffeau A, Barrell BG, Bussey H et al. (1996) Life with 6000 genes. Science, 274, 546–567.
Gore J, Youk H, Van Oudenaarden A (2009) Snowdrift game dynamics and facultative cheating in yeast. Nature, 459, 253–256.
Gottschling DE, Aparicio OM, Billington BL, Zakian VA (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell, 63, 751–762.
Greig D, Travisano M (2004) The Prisoner's Dilemma and polymorphism in yeast SUC genes. Proceedings. Biological sciences/The Royal Society, 271, S25–S26.
Hicks JB, Herskowitz I (1977) Interconversion of yeast mating types II. Restoration of mating ability to sterile mutants in homothallic and heterothallic strains. Genetics, 85, 373–393.
Huggett JF, Foy CA, Benes V et al. (2013) Minimum information for publication of quantitative digital PCR experiments. Clinical Chemistry , 59, 892–902.
Johnson LJ, Koufopanou V, Goddard MR, Hetherington R, Schafer SM, Burt A (2004) Population genetics of the wild yeast Saccharomyces paradoxusGenetics, 166, 43.
Kim D-H, Lee Y-C (2006) Changes in some quality factors off rozen ginger as affected by the freezing storage conditions. Journal of the Science of Food and Agriculture, 86, 1439–1445.
Koschwanez JH, Foster KR, Murray AW (2011) Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. PLoS Biology, 9, e1001122.
Kuehne H, Murphy H, Francis C, Sniegowski P (2007) Allopatric divergence, secondary contact, and genetic isolation in wild yeast populations. Current Biology, 17, 407–411.
Lafontaine I, Fischer G, Talla E, Dujon B (2004) Gene relics in the genome of the yeast Saccharomyces cerevisiae. Gene, 335, 1–17.
Libkind D, Hittinger CT, Valério E et al. (2011) Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. PNAS , 108, 14539–14544.
Lindblad-Toh K, Wade CM, Mikkelsen TS et al. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature, 438, 803–819.
Liti G, Carter DM, Moses AM et al. (2009) Population genomics of domestic and wild yeasts. Nature, 458, 337–341.
Louis EJ (1995) The chromosome ends of Saccharomyces cerevisiae Yeast, 11, 1553–1573.
MacLean CR, Brandon C (2008) Stable public goods cooperation and dynamic social interactions in yeast. Journal of Evolutionary Biology, 21, 1836–1843.
MacLean RC, Fuentes-Hernandez A, Greig D, Hurst LD, Gudelj I (2010) A mixture of "Cheats" and 'Co-Operators' can enable maximal group benefit. PLoS Biology, 8, e1000486.
Marsilio V, Campestre C, Lanza B, De Angelis M (2001) Sugar and polyol compositions of some European olive fruit varieties (Olea europaeaL.) suitable for table olive purposes. Food Chemistry, 72, 485–490.
Mercier G, Berthault N, Touleimat N et al. (2005) A haploid-specific transcriptional response to irradiation in Saccharomyces cerevisiae Nucleic Acids Research, 33, 6635–6643.
Naumov G, Naumova E (2010) Polygenic control for fermentation of b-fructosides in the yeast Saccharomyces cerevisiae: new genes SUC9 and SUC10 . Microbiology, 79, 160–166.
Naumov GI, Tolstorukov II (1973) Comparative genetics of yeast. X. Reidentification of mutators of mating types in Saccharomyces Genetika, 9, 81–91.
Naumov G, Turakainen H, Naumova E, Aho S, Korhola M (1990) A new family of polymorphic genes in Saccharomyces cerevisiae: alpha-galactosidase genes MEL1-MEL7. Molecular & General Genetics: MGG , 224, 119–128.
Naumov G, Naumova E, Sancho E, Korhola M (1993) Taxogenetics of the Saccharomycessensu stricto yeasts from western and South Africa. Cryptogamie. Mycologie, 4, 263–270.
Naumov GI, Naumova ES, Sancho ED, Korhbla MP (1996) Polymeric SUC genes in natural populations of Saccharomyces cerevisiae FEMS Microbiology Letters , 135, 31–35.
Naumov GI, Naumova ES, Sniegowski PD (1998) Saccharomyces paradoxus and Saccharomyces cerevisiaeare associated with exudates of North American oaks. Canadian Journal of Microbiology, 44, 1045–1050.
Pacini E, Nepi M, Vesprini J (2003) Nectar biodiversity: a short review. Plant Systematics and Evolution, 238, 7–21.
Peay KG, Belisle M, Fukami T (2012) Phylogenetic relatedness predicts priority effects in nectar yeast communities. Proceedings. Biological sciences/The Royal Society, 279, 749–758.
Porro D, Brambilla L, Ranzi B, Martegani E, Alberghina L (1995) Development of metabolically engineered Saccharomy-

ces cerevisiaecells for the production of Lactic-Acid. Biotechnology Progress, 11, 294–298.
Stapanian MA, Smith CC (1978) A model for seed scatterhoarding: coevolution off ox squirrels and black walnuts. Ecology, 59, 884–896.
Thim L, Hansen MT, Norris K et al. (1986) Secretion and processing ofi nsulin precursors in yeast. Proceedings of the National Academy of Sciences, USA, 83, 6766–6770.
Vaughan MA, Martini A (1987) Three newly delimited species of Saccharomyces sensu stricto. Antonie van Leeuwenhoek, 53, 77–84.
Vila C, Maldonado JE, Wayne RK (1999) Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. Journal of Heredity, 90, 71–77.
Wach A (1996) PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in S. cerevisiae. Yeast, 12, 259–265.
Wang Q-M, Liu W-Q, Liti G, Wang S-A, Bai F-Y (2012) Surprisingly diverged populations of Saccharomyces cerevisiaein natural environments remote from human activity. Molecular Ecology, 21, 5404–5417.
Wayne RK (1986) Limb morphology of domestic and wild canids: the influence of development on morphologic change. Journal of morphology187, 301–319.
West SA, Griffin AS, Gardner A (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. Journal of Evolutionary Biology, 20, 415–432.
Wiens F, Zitzmann A, Hussein NA (2006) Fast food for slow lorises: is low metabolism related to secondary compounds in high-energy plant diet? Journal of Mammalogy 87, 790–798.
Williams GC (1996) Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought. Princeton University Press, Princeton, New Jersey.
Wyrick JJ, Holstege FC, Jennings EG et al. (1999) Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. Nature, 402, 418–421.

D.G. designed the study. G.O.B. performed the research. D.G. and G.O.B. analysed data and wrote the paper.

Data accessibility

The available SUC2 nucleotide sequences, in FASTA format, for 29 S. paradoxus and 8 S. cerevisiae strains are accessible on Appendices S5 and S6 (Supporting information), respectively. Raw data for all invertase assay experiments, including absorbance reads and calculations, are accessible on Appendices S2, S3 and S4 (Supporting information).

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 . List of all strains used in this study with strain names and details.

Appendix S2 . Invertase producer screening results for each wild strain, producer strain, and nonproducer strain. Wild strains were tested once; control strains were tested as 3 replicates.

Appendix S3 . Invertase assay result for 11 strains and 5 suc2 knockout strains. Three measurements were performed for each strain.

Appendix S4 . Invertase assay results performed using commercial invertase enzyme to determine the linear response range and to calculate the invertase production using the equation of the linear regression ($R^2 = 0.948$, six data points, three replicates each). None of our measurements gave more glucose reading over the top point of this linear regression line.

Appendix S5 . S. paradoxus SUC2 nucleotide sequences available for 29 strains.

Appendix S6 . S. cerevisiae SUC2 nucleotide sequences available for eight strains. Sequence files include ~130 base pairs upstream and 200 base pairs downstream of the start and the stop codons, respectively.

Fig. S1 Southern blot confirmation of SUC2 knockouts in the multiple-copied strains, and common laboratory strain C.Lab.1 (S288c).

Fig. S2 CHEF gel and corresponding Southern blot assay showing four different chromosomal locations of SUC genes in C.Nectar.1, C.Nectar.2. and C.Nectar.3 strains. This CHEF gel was ran longer (0.5 9 TBE, 14 °C, 200 V for 30 h with 60-s switching time, and for 12.5 h with a 90-s switching time) to further separate chromosome II, XIV, X bands.

Increasing recombination increases hybrid fertility 70-fold

Introduction

Sexual reproduction comprises the processes of meiosis, the production of gametes, and fertilization, the fusion of gametes. These sexual processes generate novel diploid allelic combinations by crossing-over and segregation. Meiosis reduces the genome content from diploid to haploid by two divisions. The first meiotic division (Meiosis I) separates parental centromeres into two cells. The second meiotic division (Meiosis II) simply separates sib-centromeres in a process that is similar to haploid mitosis (Wilkins & Holliday 2009). Recombination by chromosomal crossing-over during Meiosis I can increase the genetic variation due to sexual reproduction and increase the evolutionary potential of populations (Mayr 1970).

During Meiosis I the pairing of homologous chromosome sets inherited from the parents is initiated by the formation of the double-strand breaks (DSBs). DSBs expose DNA strands by forming three-prime overhangs. This initiates the homology search (Fig. 1). When homologs match and align by the complementary base pairing then the crossing over (CO) initiates (reviewed in McKee 2004). These two molecular mechanisms, pairing and crossing over (2 & 5B in Fig. 1), secure the segregation of exactly one of each chromosome into the newly formed haploid gametes (i.e. Hillers & Villeneuve 2003; Fledel-Alon *et al*. 2009).

However, sequential and complex nature of segregation makes it susceptible to various problems. For instance, homology search and pairing (Step-2, Fig. 1) can be problematic due to sequence divergence between partner chromosomes as this step relies on complementary base pairing (Boumil *et al*. 2003; Tessé *et al*. 2003; Gerton & Howley 2005). High single nucleotide divergence between chromosomes decreases the efficiency of the homology search and pairing, reducing the ability of CO to form which consequently decreases segregation efficiency. Even if pairing is successful between diverged homologs, CO events might be blocked by the action of the anti-recombination proteins such as the mismatch repair system (Hunter *et al*. 1996). These proteins reduce crossing over rates by blocking or dissociating

invading strands (Step 3, Fig. 1) (Martini *et al*. 2011). This helps maintain genomic stability by reducing ectopic recombination between non-homologues. As an example of this, it has been observed that intraspecific yeast crosses had reduced CO levels as a result of anti-recombination activity (Datta *et al*. 1997; Martini *et al*. 2011).

For this reason, segregation can fail in F_1 hybrids between species with highly diverged chromosomes. Indeed, extensive (~14%) single nucleotide divergence between *S. cerevisiae* and *S. paradoxus* homeologous (i.e. homologous but highly diverged) chromosomes results in inhibition of COs and reduces successful segregation in F1 hybrids, causing hybrid sterility (Kao *et al*. 2010). This is because every chromosome pair requires at least a single crossover event to ensure perfect segregation (e.g. Hillers & Villeneuve 2003). If segregation fails, then the resulting sex cells will suffer from missing chromosomes and/or from gain of additional chromosomes. This leads to the reduced fitness or lethality (Torres *et al*. 2008). Yeast cells cannot survive if any of its chromosomes is missing. Therefore, inhibition of crossing-over in hybrid crosses between yeast species results in the sex cell inviability as a result of aneuploidies (Kao *et al*. 2010). In summary, single nucleotide divergence induced anti-recombination can form a barrier to gene flow between diverged populations of microbial species (Matic *et al*. 1995; Rayssiguier *et al*. 1989; Vulic *et al*. 1997; Datta *et al*. 1997).

However reduction or total inhibition of recombination (anti-recombination) is not the sole genetic cause of the hybrid sterility in yeast. Another cause of hybrid sterility can be a result of gross chromosomal rearrangements (GCRs) between diverging species (Coyne & Orr 2004). When GCRs fix between incipient or completely isolated species, F_1 hybrids formed between them may produce inviable gametes due to chromosome imbalances causing missing essential genes (see Figure 1 in the General Introduction). There is evidence from a range of plant and animal taxa showing that GCRs contributing to reproductive isolation between diverged populations (Baker & Bickham 1986; Searle 1993; Ramsey & Schemske 2002; Hauffe *et al*. 2011; Nei & Nozawa 2011; Dion-Côté *et al*. 2015). Fixed GCRs between some *Saccharomyces* populations have also been shown to contribute to the hybrid sterility as they reduce the F_1 spore viability (Delneri *et al*. 2003; Charron *et al*. 2014; Hou *et al*. 2014). However these are examples from intraspecific yeast crosses but not from interspecific crosses of *S. cerevisiae* and *S. paradoxus* parents. Genomes of both species are collinear with no GCRs except for four small inversions which contain no essential genes rejecting the potential role of GCR model in the split of these two species (Kellis *et al*. 2003).

Another model of post-zygotic reproductive isolation (RI) between species invokes the Bateson-Dobzhansky-Muller genetic incompatibilities (BDMI) (Dobzhansky 1937; Muller 1942; Orr 1995). The BDMI model suggests that genes at different loci within a population co-evolve (Coyne & Orr 2004). However genes evolved independently in allopatric populations were never tested by the natural selection. Thus when diverged populations reunite to form hybrids, a gene from one species cannot function well with another gene from another species as a result of between-loci negative epistasis. This can lead to the post-zygotic isolation (see the General Introduction for a detailed explanation of the BDMI model). This form of RI has extensive theoretical and empirical support, especially from *Drosophila* research (Phadnis & Orr 2008; Presgraves *et al*. 2003; Tang & Presgraves 2009; Barbash *et al*. 2003; Brideau *et al*. 2006; Sawamura & Yamamoto 1997; Ferree & Barbash 2009).

Although tested in different studies there is no proof for BDMI regions those causing post-zygotic spore inviability between *S. cerevisiae* and *S. paradoxus* hybrids (Greig 2007; Kao *et al*. 2010; also see Chapter 4). On the other hand, it is technically challenging to test whether BDMIs play a role in reproductive isolation between inter-specific yeast hybrids. This is because only 1% of *S. cerevisiae* and *S. paradoxus* hybrid spores survive after the completion of meiosis (Liti *et al*. 2006). And even if these 1% F_1 spores are recovered by massive random plating, those haploids would be aneuploids due to anti-recombination (Kao *et al*. 2010). Aneuploidy, then, would mask recessive BDMIs enabling spores that carry them to survive.

This study tries to reduce the anti-recombination barrier observed between *S. cerevisiae* and *S. paradoxus* hybrids. This is important for two reasons: First, from the speciation research perspective, understanding the relative role of the anti-recombination barrier in F_1 spore death is valuable for the speciation research itself (Coyne & Orr 2004). Second, a significant induction in hybrid spore viability will allow us to collect pure haploid gametes. This will become a useful tool to study the other potential models of reproductive isolation (e.g. the BDMI model). To achieve this, study follows a two-step strategy.

i) In order to initiate potential crossovers in the sites where crossovers are most likely to be successful, i.e. in chromosomal regions with the highest identity between the two species, I devised a way to induce DSBs at any chromosomal site of my choice in meiosis. I used the natural homing endonuclease *VDE1,* a selfish genetic element which cleaves at a specific unique DNA sequence and thus forms DSBs (Nogami *et al*. 2002). Since *VDE1* is naturally located in the nucleus only during the meiosis and is known to recruit meiosis toolbox

proteins those required to repair formed DSBs, this enzyme and its specific cut-site is a promising candidate to increase the likelihood of crossing-over rates (Fukuda *et al*. 2003). By inserting the VRS cleavage site into chromosomal locations flanked with highly identical sequences between the two species, I tried to reduce the mismatch protein activity around artificially introduced hotspots. Thus DBSs formed via *VDE1* activity would more likely to be repaired via crossing-over resolution. The initial aim was to test the crossover induction and aneuploidy reduction effect of VRS initially at a single chromosome (chromosome-III of *S. cerevisiae* parent) and if it was successful, then, to expand the system, inserting VRS sites into conserved regions of every chromosome, improving global segregation and hence F_1 spore viability.

ii) The second step of my strategy was to repress two potential anti-recombination proteins in order to increase the likelihood of crossover resolution between all 16 homeologous chromosome pairs of the two species (Fig. 1, Step-3). a) Sgs1 was suggested to resolve strand invasion events as non-crossovers instead of crossovers, when the interacting sequences are low in identity (De Muyt *et al*. 2012). I therefore aimed to increase crossover rates over non-crossover rates between the homeologous chromosomes of hybrids by repressing the meiotic expression of this protein (Myung *et al.* 2001a; Spell & Kinks-Robertson 2004; Amin *et al*. 2010). b) I also repressed meiotic expression of the *MSH2* gene. *Msh2p* has already been demonstrated to play a role in anti recombination in yeast hybrids. Hunter *et al*. (1996) artificially increased the CO rates between these species by deleting *MSH2* and therefore improved the hybrid spore viability (i.e. fertility) up to 10.2 % (compared to 1.2% before manipulation). Thus by inhibiting the meiotic activities of Sgs1 (this study) and Msh2 (Hunter *et al*. 1996) in combination, I tried to overcome the anti-recombination barrier between the two species.

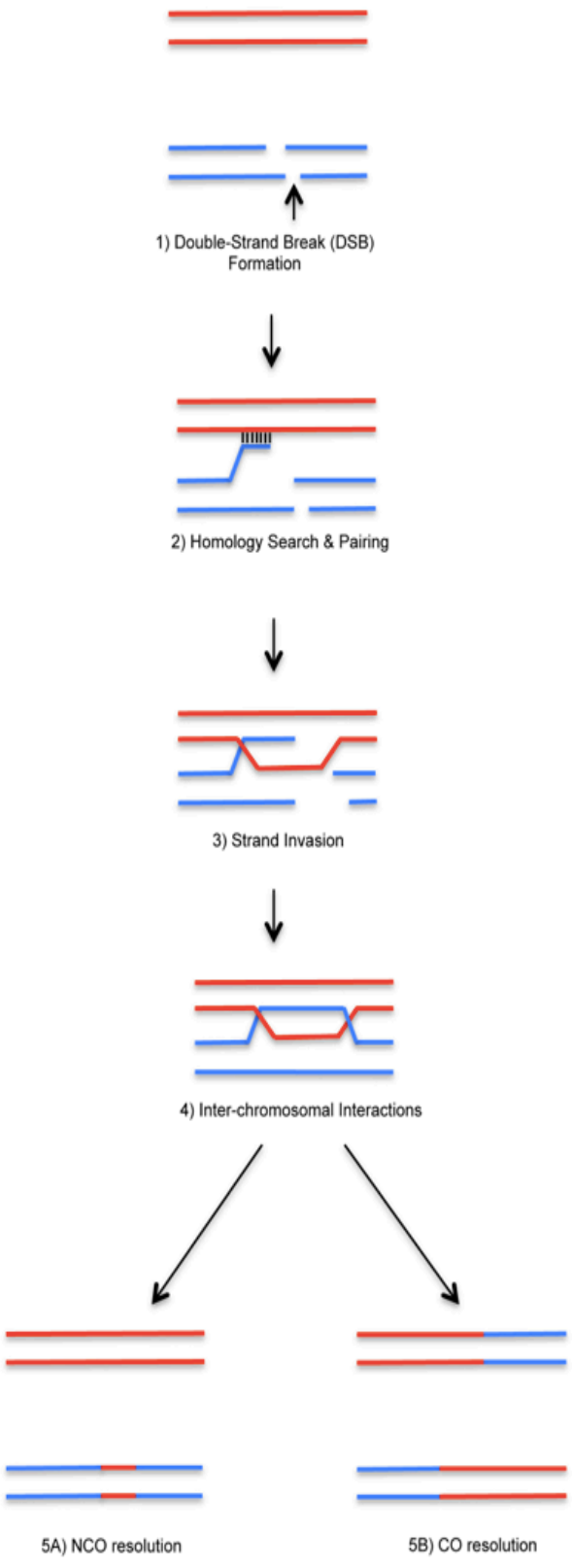


Fig. 1. A simplified-model for homologous chromosome pairing and double-strand break (DSB) repair. 1. DSBs are formed at multiple regions of the genome. 2. This initiates the homology search by freeing DNA as stretches of 3-prime overhangs. Homologs match through complementary base pairing (Tessé *et al*. 2003). This is the step when Msh2 mismatch protein shows its anti-recombination activity. 3. When homologs align efficiently, strand invasion begins. 4. Next, inter-chromosomal interactions, regulated by the meiosis toolbox proteins, takes place. At this point, the decision over the resolution of the DBS repair is regulated. Sgs1 helicase functions as an anti-recombination protein at this stage by preventing stable strand invasion and as a result resolving invading-strands prematurely. This activity leads to NCO resolution if interacting chromosomes are diverged at single nucleotide level (Chen & Jinks-Robertson 1999; Welz-Voegele & Jinks-Robertson 2008). 5A) Most of the DBSs are repaired by non-crossover (NCO) resolution. However, in theory, NCO-resolution does not efficiently secure the faithful segregation of the chromosomes into the newly formed sexual progenies. 5B) Crossing-over (CO) resolution, on the other hand, is necessary for many species to secure the segregation of the chromosomes. Thus, at least, one CO per chromosome secures evenly segregation of each chromosome into the haploid progeny (Hillers and Villeneuve 2003).

Methods

Yeast strains and genetic manipulations:

I used W303 strain background as the *S. cerevisiae* parent and N17 strain background as the *S. paradoxus* parent. Haploid *S. cerevisiae* and *S. paradoxus* cells were mated on YEPD plates (2% glucose, yeast extract, peptone, agar) to obtain hybrid diploids.

Gene replacements, to detect chromosome-III recombination and aneuploidy rates, and VRS double-strand break (DSB) site insertions were done on the *S. cerevisiae* parent. I followed the PCR-mediated gene replacement protocol (Wach 1996) to replace promoters, to insert chromosome-III markers, and to insert the VRS cut-site. I used the LiAc method for transformations (Gietz & Schiestl 2007).

All strains used are listed in the Supplementary Table 1 and all primers used are listed in the Supplementary Table 2.

Chromosome-III segregation assay system:

I marked chromosome-III of the *S. cerevisiae* parent on two arms by deleting essential amino acid metabolism genes (*LEU2* and *THR4*) using drug resistance markers (Figure 2). Therefore I could score hybrid gametes for aneuploidy and crossing-over events.

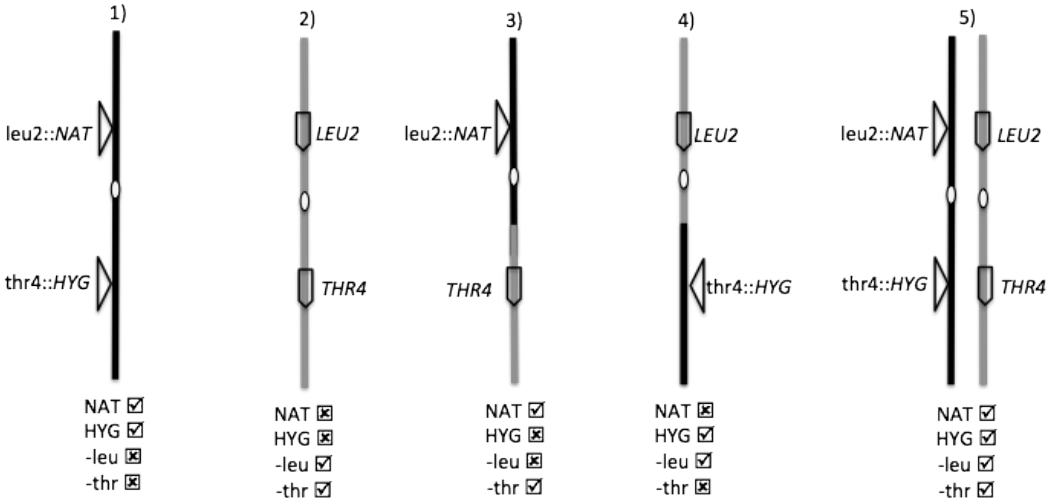


Figure 2. Markers on chromosome-III to quantify crossing-over and aneuploidy rates in the F_1 hybrid spores. I replaced *LEU2* and *THR4* genes with drug resistance markers *NAT* and *HYG* on *S. cerevisiae* parent (black). Thus simply by replica-plating viable F_1 spores, I could score each colony as crossover-negative-euploid, crossover-positive-euploid, or as chromosome-III aneuploid. (1) and (2) in figure show outcomes for crossover-negative euploid. (3) and (4) show outcomes for crossover positive euploid spores. And (5) shows hybrid spores that are aneuploid for chromosome-III, pointing the segregation failure. Genetic markers on both chromosomes allow the detection of crossovers only within the *LEU2-THR4* interval. This interval spans the one-third of the physical distance of the chromosome.

DSB target location in chromosome-III assay system:

I introduced VRS cut-site into the *FEN1* locus of the chromosome-III. Insert site shared ~70

base pairs of 100 % identical sequences at both directions between two species. Short but perfect identity between the two species around the artificial DSB site makes it more likely region to allow crossover resolution.

To introduce the hotspot, I designed a p-Bluescript (+) plasmid construct containing two 31 bp VRS cut-sites around the *URA3* auxotrophy marker (p_uniVRSURA3VRS_uni). I amplified *VRS-URA3-VRS* construct from the plasmid, and transformed (Walch 1996) into the *S. cerevisiae* parents targeting the *FEN1* gene (Fig.3).

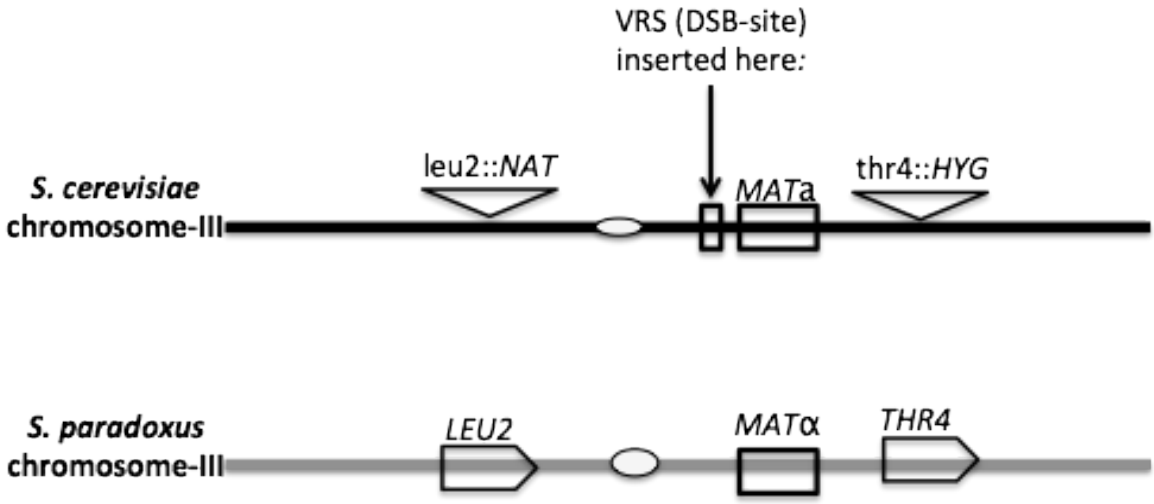


Figure 3. Position of the inserted DSB site (VRS) relative to the chromosome-III markers. Artificial recombination hotspot (*VRS* cut-site) is inserted at the *FEN1* locus of *S. cerevisiae*. ~70 base pairs of up- and down-stream 100 % identity between two species flank the DSB site to increase the likelihood of crossovers. If *VDE1* cleaves the cut-site at the *S. cerevisiae* chromosome-III during meiosis-I, *the DSB would be repaired using the sequence information of the complementary S. paradoxus* region. This region is shorter in *S. paradoxus* because it lacks the cut-site and the primers that were used to insert the cut-site. Thus the decrease at the length of the *S. cerevisiae* region following the repair event would allow the detection of the *VDE1* homing efficiency simply by performing PCR reactions on the F_1 hybrid spores. Formed DSB can be repaired as crossovers or non-crossovers (Figure 1 5A or 5B).

After confirming the insertion via diagnostic PCRs, I replica-plated the uracil prototroph colonies onto the 5-Fluoroorotic acid plates (5-FOA, ammonium sulfate, yeast nitrogen base without amino acids, dextrose, necessary amino acid supplements, agar). Next I picked 5-FOA resistant colonies, and confirmed the *URA3* looping-out event via diagnostic PCR. Finally I confirmed the intactness of the 31 base pairs cut-site via Sanger sequencing.

Initiation of meiosis and DSBs in targeted location and scoring chromosome-III genotypes:

To initiate double strand break formation at the target site on chromosome-III, I cultured hybrid diploid in 3ml of liquid KAc sporulation medium (2% potassium acetate). This induces meiosis in hybrid diploids. I incubated the cultures at the room temperature by shaking at 225 rpms for 4 to 5 days for the completion of the sporulation. DSBs made by the *VDE1* activity are formed at this KAc culturing step. During this step, repairing of the VDE cleaved DSBs by meiosis toolbox proteins lead to crossover or non-crossover repairs (Figure 1 5A or 5B).

Random Spore Analysis:

After the completion of meiosis/sporulation I plated F_1 hybrids onto synthetic complete media containing canavanine and cycloheximide drugs. Canavanine (recessive drug resistance gene located on the chromosome V of *S. cerevisiae* parent) and cycloheximide (recessive drug resistance gene located on the chromosome VII of *S. paradoxus* parent) allow me to select for random haploid spores. This is because diploids plated onto the double drug would be sensitive as a result of the heterozygosity at Can1 and Cyh2 loci. Haploid hybrid spores that inherit chromosome V only from *S. cerevisiae* parent and chromosome VII only from *S. paradoxus* parent would be able to grow as random colonies. I streaked out these viable colonies of different treatments onto the YEPD plates. Then I replica-plated random hybrid spores onto appropriate plates to detect whether and F_1 hybrid spore was aneuploid, crossover positive, or euploid for chromosome-III (Fig. 2).

Detecting *VDE1* homing efficiency at the target location:

The *S. cerevisiae* (w303) VRS insertion site includes 31-bp of recognition sequence and an additional 40-bp of universal primers. *S. paradoxus* (n17) counterpart at the same position is shorter and is detectable on an agarose gel (see Figure 3). If *VDE1* cleaves the cut-site at *S. cerevisiae* chromosome-III, the DSB would be repaired from complementary sequence of the *S. paradoxus FEN1* gene which is shorter. I could detect homing-efficiency by performing PCR reactions on crossover-positive euploid spores, crossover-negative euploid spores, and aneuploids.

SGS1 and *MSH2* strain construction:

To increase the likelihood of crossover resolution over non-crossover resolution at DSB sites all over the homeologous chromosomes in hybrid diploids, I replaced the native promoters of

SGS1 and *MSH2* with meiotically-repressed *CLB2* promoter (Grandin & Reed 1993). Thus I could inhibit the potential anti-recombination activity of both proteins during the meiosis. The promoter replacement construct was designed by Lee & Amon (2003). I amplified the *CLB2* promoter and *KANMX* drug resistance marker from a construct that was already inserted in another laboratory strain of *S. cerevisiae* (SK1 strain kindly provided by Neil Hunter - see Oh *et al*. 2008). After amplifying the constructs (*KANMX*-p*CLB2* targeting *SGS1* promoter or *MSH2* promoter), I transformed the amplicon into the haploids of both parental species. I selected positive colonies on YEPD-G418 plates for *KANMX* activity and confirmed the position of the replacements via diagnostic PCRs.

I obtained either single-mutant (p*CLB2SGS1*) or double-mutant (p*CLB2SGS1* and p*CLB2MSH2*) *S. cerevisiae* (a) and *S. paradoxus* (@) haploid parents. Single mutant is to investigate whether the Sgs1 meiotic repression would increase the crossing-over levels and reduce the segregation problems when alone. Double mutation aimed to achieve the ultimate goal of this study to detect the relative role of the anti-recombination in hybrid spore viability via recovering sexual fertility of full hybrids. I crossed single or double-mutant parental species to obtain hybrid diploids for further tests.

To initiate meiosis and to detect chromosome-III crossover and aneuploidy levels in F_1 hybrid spores, I applied the protocols mentioned under the "Initiation of meiosis and DSBs in targeted location and scoring chromosome-III genotypes" of this section.

Tetrad assay system:

I mainly used F_1 random spore analysis to detect crossover and aneuploidy rates since the alternative method, tetrad-dissection, was not feasible due to low spore viability. Although random spore analysis is reliable to recover haploid progeny by segregation of recessive drug resistant markers, canavanine and cycloheximide (Hopper & Hall 1974), it can nevertheless cause overestimations of aneuploids. This is may be because double drug media will fail to kill heterozygous diploids at a 100 % rate leading to 'diploid leakage' on analysis. The increase in viable spore production in Sgs1-only (Sgs1-m) and Sgs1 + Msh2 double mutants were efficient thus I could use the alternative method, tetrad dissection, to measure spore viability and aneuploidy rates for these genotypes. Tetrad analysis ensures to score haploid spores by allowing specifically picking haploid spores under the microscope.

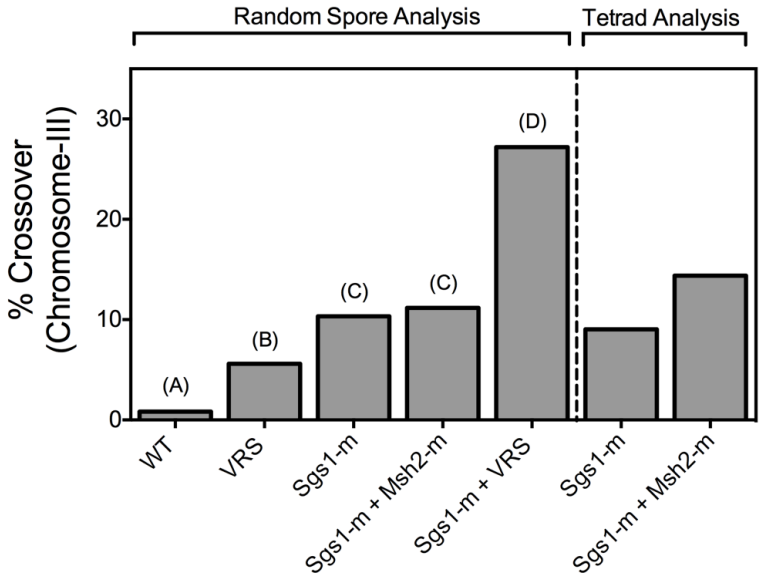
To dissect haploid hybrid spores, I first digested the ascus walls of the tetrads from the mutant

hybrids (Sgs1-m single mutant and Sgs1-Msh2 double-mutant) by incubating ascospores in 20 µl of Zymolyase (3 units / 100 µl, Zymo Research, USA) for 30 minutes. Next I placed individual F_1 hybrid spores onto YEPD plates using a micro-manipulator microscope (MSM 400, Singer Instruments, UK). I incubated plates at 30 °C for two days for germination and colony formation. Then I performed the same replica plating protocol to detect chromosome-III crossover and aneuploidy rates (Figure 2).

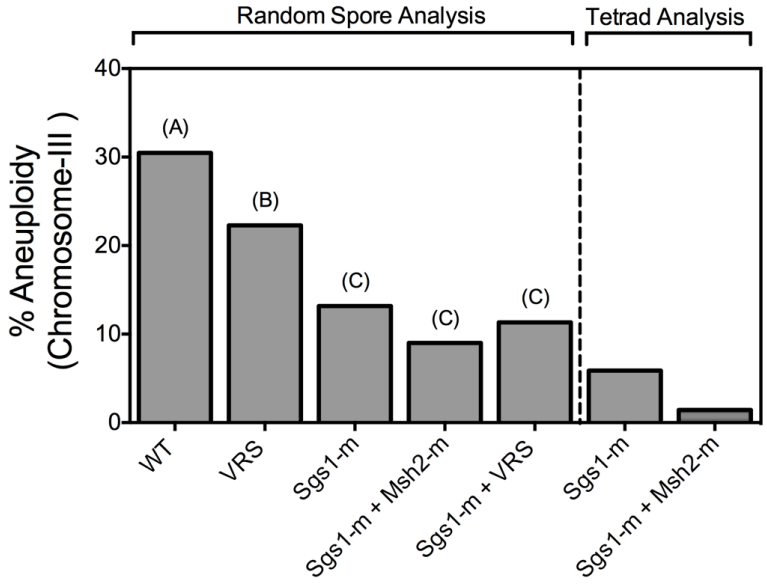
I have also detected the viability of the hybrid spores by dissecting the tetrads with the same protocol and counting the number of viable colonies after two days incubation for colony formation. The viability scores were obtained for eight genotypes, including hybrid treatment genotypes and non-hybrid control genotypes at least by dissecting 250 tetrads (1000 spores) or 100 tetrads (400 spores), for hybrids and non-hybrids, respectively (Supplementary_Table_1).

Results

Crossover increase and segregation rescue for chromosome-III in random spores



A.



B.

Figure 4. (A) Chromosome-III crossover percentages and **(B)** Chromosome-III aneuploidy percentages in F_1 hybrids. **Left-panels** in both (A) and (B) show random spore values including all tested genotypes. About 600 spores for each treatment were scored *in random spore analysis*. Letters above the bars represent statistically significant (common letters) and insignificant (different letters) results of the pair-wise comparisons calculated only for the random spore analysis. I applied Fisher's exact tests for all 10 possible pairwise comparisons of the treatment genotypes. Raw numbers are used for statistical tests while percentage values are used for the graphics. p = 0.005 was taken as a significance cut-off after correcting for the 10 pairwise comparisons. **Right-panels** show crossover (A) and aneuploidy (B) percentages for the two genotypes (Sgs1-m single-mutant and Sgs1-m + Msh2-m double-mutant) that had a high fertility and therefore could be analyzed by *tetrad analysis* to get more

precise values. I scored 443 F_1 hybrid spores in Sgs1-m meiotic repression treatment and scored 487 F_1 hybrid spores in Sgs1-m and Msh2-m meiotic repression treatment.

Crossover rate induction was effective for all treatment genotypes (Figure 4-A left-panel).

Wild-type hybrids had only 0.83 % crossover positive F_1 spores (n=597). Introduction of DBS cut-site (VRS) induced chromosome-III crossing-over rates up-to 5.7 % ($p<0.0001$ in Fisher's exact test, Fig. 4-A left-panel). The aneuploidy levels as a result decreased from 30.4% levels of wild-type hybrids down to 22.4 % in the DSB introduced hybrids ($p=0.0021$ in Fisher's exact test, Fig. 4-B left-panel).

Although VRS treatment was effective, yet the crossover induction and the aneuploidy reduction were far from ideal levels seen in non-hybrids (Jessop *et al*. 2006). I was curious about the homing efficiency of the *VDE1* at the artificial target site. Therefore I investigated the *VDE1* homing (see "Detecting *VDE1* homing efficiency at the target location" under the Methods). I detected that *VDE1* homing at the DSB site was very effective as all tested crossover-negative hybrid spores (n=53) were positive for the homing. Yet the DSBs were mostly resolved via non-crossover repair pathway (Supplementary Figure 1 and Figure 1 5A vs. 5B).

Next I took a genome-wide approach by repressing the meiotic activity of a candidate protein, Sgs1p, to achieve a better improvement in segregation success in hybrids. Repression of Sgs1p during meiosis aimed to increase the likelihood of crossing-over propagation of DSBs throughout all chromosomes in the hybrids. Sgs1-m treatment induced the crossing-over levels up-to 10.3 % ($p<0.0001$ against WT values in Fisher-exact test, Sgs1-m in Figure 4-A left-panel). This in turn reduced the aneuploidy level from 30.4 % of the WT-hybrids down to 13.16 % in the Sgs1-m hybrids ($p<0.0001$ against WT values in Fisher-exact test, Figure 4-B left-panel). Induction in crossovers and reduction in aneuploids were greater in Sgs1-m treatment, also, in comparison to DSB-site (VRS) treatment ($p=0.0049$ for CO and $p<0.0001$ for aneuploidies in Fisher's exact test).

When I combined Sgs1 (this study) and Msh2 (Hunter *et al*. 1996) meiotic repressions under the same hybrid genetic background to overcome the anti-recombination problem, I could increase the chromosome-III crossing-over percentage up-to 11.1 % and decrease aneuploidy levels down to 9 % (Fig. 4-A/B, left-panels). However both crossover and aneuploidy levels

were no different than Sgs1-m alone (Fig. 4-A/B).

I also tested Sgs1-m repression together with the DSB site. Combined treatment of Sgs1-m and DSB site increased crossing-over levels up to 27.1 % (Fig.4-A left-panel). This increase was significantly higher in comparison to crossover rates induced in all the other treatments (see letters above the bars of Figure 4-A left-panel). Interestingly though, aneuploidy rate, 11.3 %, was higher than expected (Fig. 4-B, left panel).

Segregation improvement was greater when spores analyzed via tetrad analysis:

There was no correlation between the increase in crossovers and the decrease in aneuploidy when Sgs1-m + VRS treatment applied (Fig. 4-B left-panel). Indeed aneuploidy decrease leveled off for the three treatments (Sgs1-m, Sgs1-m + Msh2-m, and Sgs1-m + VRS) (Supplementary Figure 3). Therefore, I applied tetrad analysis for 443 F_1 spores of Sgs1 single-mutant and for 487 F_1 spores of Sgs1-and-Msh2 double mutant treatments, to obtain more precise measures of crossing-over and aneuploidy levels, in addition to earlier random spore analysis (see the 'Tetrad Assay System' under the Methods).

As predicted, I detected a greater rescue of segregation for the samples of these two treatments by using the tetrad analysis, (Fig. 4-B right-panel). Sgs1-m treatment had 5.8 % aneuploids - this was 13.7 % in random spore analysis. Sgs1-m and Msh2-m double-mutant treatment had only 1.4 % aneuploids in tetrad analysis.

Remarkable increase in hybrid fertility as a result of segregation rescue:

Finally, I detected F_1 spore viability of the hybrids by dissecting hybrid tetrads. Results show that the Sgs1-m treatment increased hybrid fertility up to 20.8 % - about a 45-fold improvement in comparison to the wild-type hybrids (Figure 5). Double-mutation treatment had even a higher induction in the hybrid fertility: 32.6 % of the dissected spores were viable.

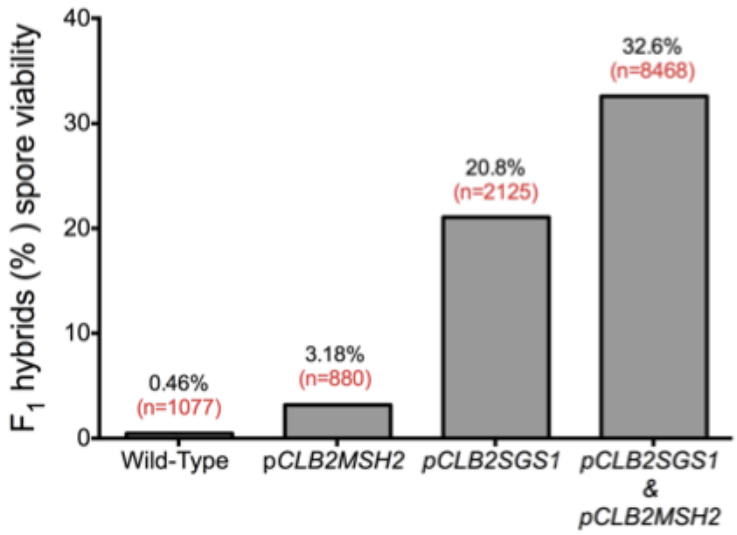


Figure 5. Increase in the hybrid fertility as a result of crossover induction. Percentage values over each bar represent viability levels and red values represent the dissected spore number (n) for each genotype. See the Supplementary Figure 2 for a visual sample of the viable colony growth and the Supplementary Figure 3 to see the results for the control genotypes scored for viability.

Discussion

Potential mechanisms (AR, BDMI, GCRs) contributing to reproductive isolation (RI) between the two yeast species have been investigated intensely (Greig 2009). However, research focusing on different mechanisms of the yeast RI is far from being conclusive. This study focused on understanding relative role of the AR model in forming barrier to breeding between two yeast siblings. Additionally, this research aimed to increase the fertility of hybrids to enable research on other reproductive barriers (e.g. BDMI model) possible.

Artificial DSB-site approach to induce recombination:

Double-strand breaks (DSBs) are readily formed at multiple regions in diploid hybrids of *S.cerevisiae* and *S. paradoxus* (Boumil *et al*. 2003). But since genomes of both species are extensively diverged at single nucleotide level, crossovers are inhibited in hybrids. This is because DSBs are repaired via non-crossovers (e.g. gene conversion) instead of crossovers. And non-crossover (NCO) resolution, unlike crossover resolution, does not secure chromosome segregation (Roeder 1997). Therefore DSBs should be made at sites with highly identical flanking sequences. That can increase the likelihood of crossover propagation. Therefore, I inserted an artificial DSB site to a region on chromosome III that had 70 base pairs of 100 % identity between the two species, predicting an induction in crossovers and

reduction in aneuploidies for the chromosome III.

The artificial DSB site effectively increased crossing-over rates from 0.83 % (WT) to 5.7 % (VRS) (Fig. 4-A). However, the induction in recombination was still much less than the expected levels seen in non-hybrids – non-hybrid level is around 25-30 % (Jessop *et al*. 2006). Further, although crossing-over was improved by 6.7-fold, the reduction in aneuploids was not proportional as it was only a 1.3-fold reduction (Fig. 4-B). This could have been because DSBs at the target site were inefficiently formed. This idea was rejected: Formation of the DSBs was efficient (Supplementary Fig. 1). Even crossover-negative spores tested (n=57) were positive for *VDE1* homing. This result suggested that, even though DSB site had identical flanking sequences between the two species, the length of identical region (~70 bp) was not long enough to propagate crossovers. Therefore DSBs were mostly repaired via the NCO pathway instead of the CO repair pathway (Fig.1 5A vs. 5B). NCO resolution here hinted a possible anti-recombination protein activity.

Anti-recombination protein Sgs1 inhibits homeologous crossovers in full-hybrids:

Nucleotide mismatches between short DNA segments or between homeologous chromosomes are known to invoke the anti-recombination (AR) protein activity (Shen & Huang 1989; Zahrt & Maloy 1997; Vulić *et al*. 1997; Selva *et al*. 1995; Datta *et al*. 1997). In yeast, AR proteins canalize the CO repair to the NCO repair (Welz-Voegele & Jinks-Robertson 2008; Myung *et al*. 2001a). AR protein activity could explain biased NCO repair levels seen in the DSB-site introduced hybrids. To test this further, and following the aim of recovering hybrid fertility, next I repressed meiotic activity of a candidate helicase, Sgs1p (Watt *et al*. 1996). My expectation was to induce crossovers and reduce aneuploidies more efficiently.

Sgs1 meiotic repression treatment induced crossing-over levels effectively by 12-fold (Sgs1-m in Figure 4-A/B, left-panel). As a consequence, aneuploidy level was reduced down to a 5.8 % (Fig. 4B, right-panel). This shows that anti-recombination activity of a single protein can greatly inhibit crossing-over between collinear but extensively diverged yeast chromosomes.

Effect of Msh2 and Sgs1 double-mutant treatment in aneuploidy reduction:

Multiple checkpoints have evolved to keep genomes stable (Myung *et al*. 2001b). Activities of the anti-recombination proteins are dispersed through time and space: MMR protein family member Msh2p takes an earlier role at blocking the homeologous recombination. Sgs1p on

the other hand takes a later role at blocking the recombination by unwinding inter-chromosomal interactions at *S. cerevisiae* mitosis (Spell & Kinks-Robertson 2004). Therefore I repressed Sgs1 and Msh2 proteins to be able to induce crossovers even further. Homeologous crossing-over levels increased about 13-fold when the meiotic expression of both proteins are repressed (Fig. 4-A, left-panel and right-panel). The rescue of the segregation was remarkable: Only 1.4% of the 483 dissected F_1 spores were aneuploid for that chromosome (Fig. 4-B right-panel).

As mentioned earlier, the artificial DSB site did not induce crossovers by itself. This was suggested to be a result of anti-recombination protein activity. I tested the DBS site introduction and the Sgs1 meiotic repression in combination to see if this was true. I expected a higher increase in crossovers in this combined treatment compared to when each treatment applied alone. Crossing-over levels indeed increased very effectively: From 0.83 % to 27.7 % - a 33-fold induction (Fig. 4A).

However, the segregation rescue under this combined treatment did not correlate with crossover induction (Supp. Fig. 3). The VRS + Sgs1-m double-mutant did not have any reduction in the chromosome-III aneuploidy level as expected, in comparison to the Sgs1-only and the Sgs1 + Msh2 double mutant treatments (Fig. 4-b, left-panel). Aneuploidy rates leveled off around 10 % level in these three treatments (Supplementary Figure 3). Insensitivity of the random spore analysis could have caused over-estimation of real aneuploidy rates (see Methods). Therefore I tested Sgs1-m single-mutant and Sgs1-m and Msh2-m double-mutant aneuploidy and crossover rates via *tetrad analysis* in order to obtain more precise measures. This was proven to be right as dissected tetrads showed a correlated response between induction in crossover and reduction in aneuploidy (Fig. 4-B, right-panel). Only 1.4 % of the dissected double-mutation treatment spores were aneuploid for the chromosome-III.

The anti-recombination amounts to at least one third of hybrid sterility in yeast:

The next obvious thing to check is measuring the changes in hybrid fertility. The rescue in segregation is expected to be throughout all 16 homeologous chromosomes of hybrids in Sgs1 and Msh2 treatments. This then should induce the hybrid fertility as predicted by theory. This was proven to be true. The F_1 hybrid gamete viability increased remarkably to 32.6% in double mutant treatment (Fig. 5). The increase corresponds to about a 70-fold induction in viability in comparison to the wild-type fertility percentage. In addition, Sgs1-m and Msh2-m

double-mutation had a viability cost as a result of induced genome instability (Myung *et al*. 2001a). Non-hybrid double-mutant controls had 73% (*S. cerevisiae*) and 88% (*S. paradoxus*) spore viability (Supplementary Table 3). These fertility levels of double-mutant non-hybrids are upper bound for possible fertility induction. Therefore it can be claimed that the anti-recombination leads to at least one third of the sexual sterility between the two species.

What can be the cause of the remaining spore death?

Although the hybrid fertility has increased up-to 32.6 % by reducing the anti-recombination effect, most of the hybrid spores are still inviable. The aneuploidy level I have detected for the chromosome-III (1.4 %) suggests that the segregation rescue was almost completely successful (Fig.4-B right-panel). Further, Kao *et al*. (2010) showed that they could fix segregation problem remarkably by deleting the *MSH2* gene completely. Aneuploidy they measured was 0.29 per viable hybrid spore as a mean of all chromosomes (Kao *et al*. 2010). These two data sets can create a biased impression suggesting that the anti-recombination problem is fully solved in these hybrids. The problem is that the aneuploidy rates in inviable spores are still unknown. Therefore it is possible that the segregation problem has not yet fully solved and the anti-recombination is still the major cause of the hybrid spore death for the remaining fraction.

Yet, if we assume that the anti-recombination problem has been fixed completely by the double mutation treatment in hybrids, then we need to think about the other potential genetic models of reproductive isolation to fully understand the sterility observed between the two species. For instance, the genic incompatibility model (BDMIs) is the most likely candidate that is responsible for the remaining hybrid spore death. Thus it would be interesting to calculate how many such two-locus BDMI regions can exist between the two species as a cause of the remaining sterility between the two species.

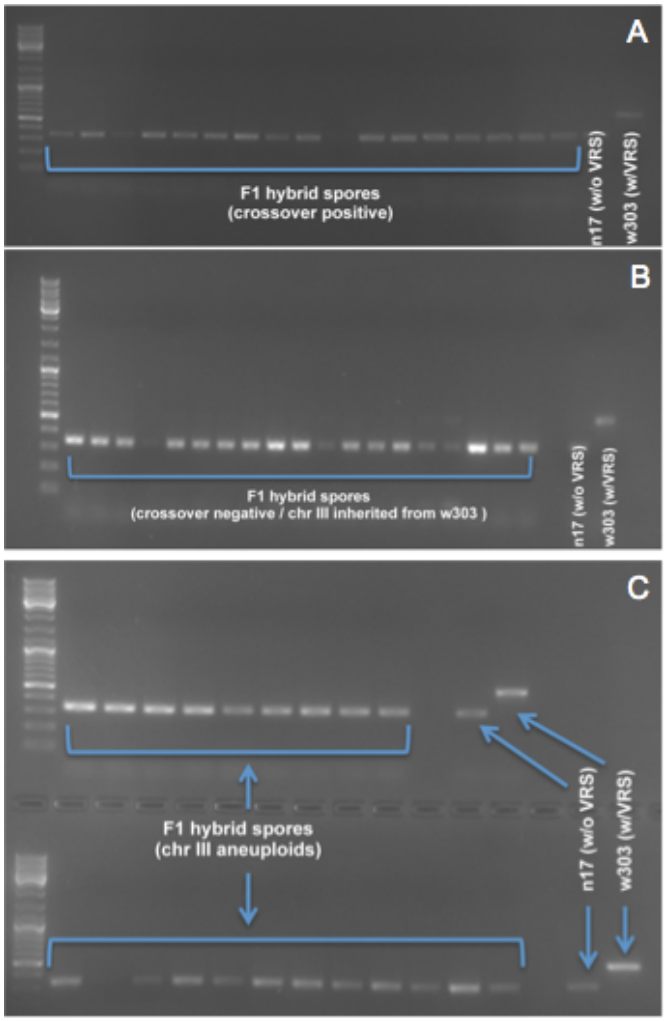
To calculate the potential number of two-locus BDMI regions, the equation formulated by Li *et al*. (2013) would be useful (see Supplementary 2 for the details). The equation assumes that there is three causes of hybrid spore death: 1) Random spore death rate ('R') introduced here by the mutations on *MSH2* and *SGS1*. 2) Spore death as a result of the anti-recombination related aneuploidy ('U'). 3) Two-locus genic incompatibilities with different killing-strengths ('I'). Multiplication of values for these three independent effects should give the wild-type hybrid spore survival rate ('T').

From the data produced in this study, the survival rate 'T' is 0.0046 (0.46 %). Random spore death probability 'R' due to the double mutation is 0.2683 (26.83 %). The probability of spore death 'U' due to the anti-recombination -assuming that it has been fixed fully- is 0.326 (32.6 %) – data obtained from the Supplementary Table 1. Then the number of potential two-locus BDM incompatibility regions ('N') can be estimated using varying values of killing-strengths ('I') for pairs of BDMI regions. For instance, if killing-probability ('I') of every pair of BDMI region is 0.9 then there can be about 18 BDMI regions between the two species. For a killing-probability of 0.5 there is room for 35 BDMI pairs, and for a killing-probability of 0.1 there is room for 185 BDMI pairs between these yeast species. All in all, experimental designs that can deal with these varying killing-strengths should be designed and applied to detect the potential BDMI regions between these two species. Such a research is even more promising especially after the detection of the very first two-locus incompatible region between a cross of two highly related (only ~0.35 % diverged) *S. cerevisiae* populations.

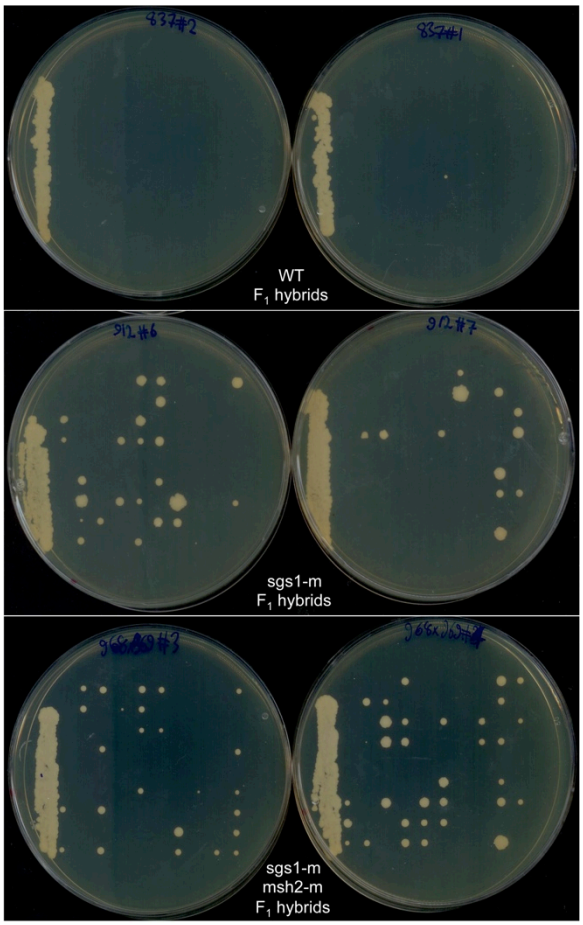
Conclusion

From speciation research perspective, this study presents evidence that the homeology driven anti-recombination barrier amounts to at least one third of the gamete death in F_1 hybrids between the two species. Here increase in hybrid spore viability was remarkably (70-fold). Yet improvement of fertility up-to 32.6 % level still left a large "room": ~ 51 %. It would be very intriguing to search for the genic incompatibility regions (the BDMIs) that most likely exist and therefore responsible for spore death between the hybrids of the two species using improved hybrid viability obtained.

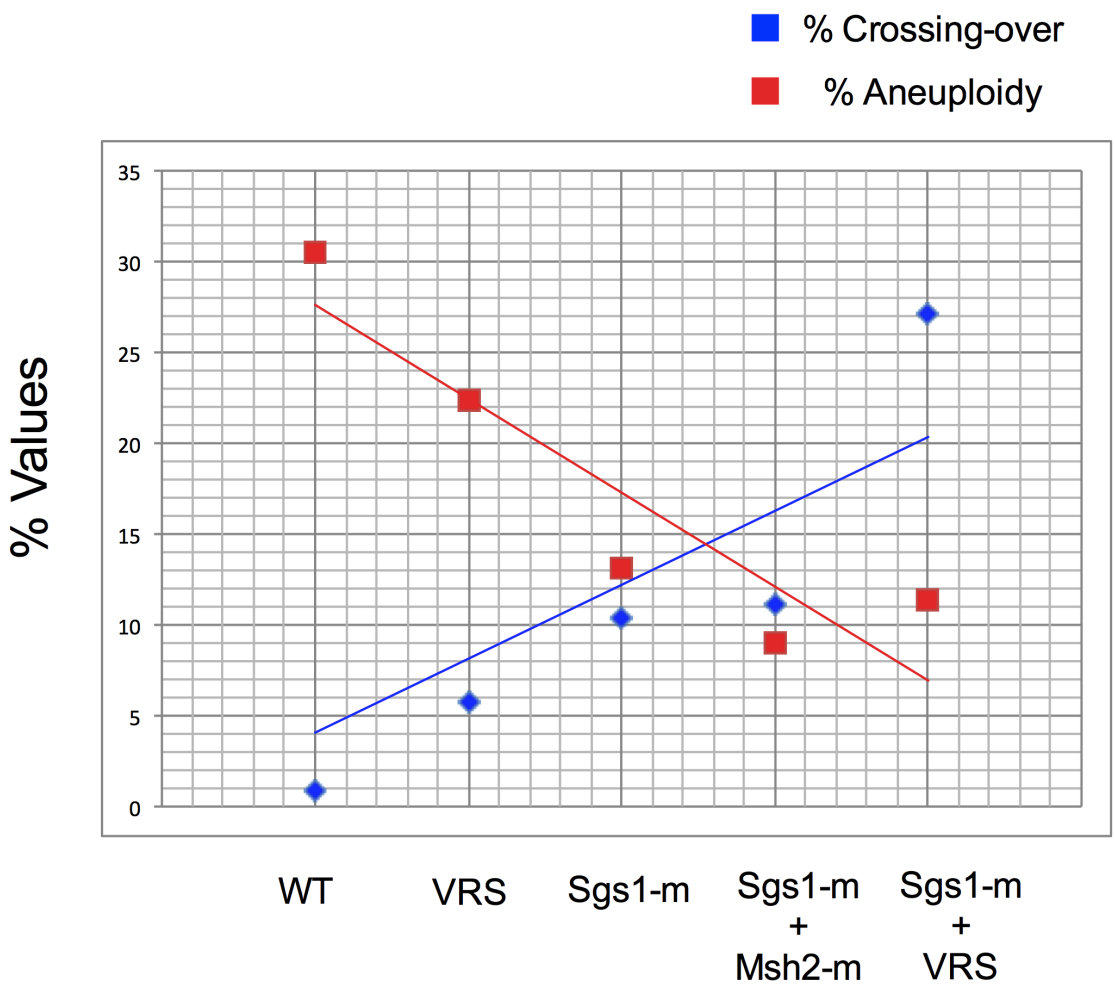
Supplementary Data



Supplementary_Figure_1. VDE homing was 100% in all tested F_1 hybrid spores. In all PCR reactions, I included parental haploid strains of n17 and w303 as controls – visible in the last two lanes of each agarose gel. **A)** VDE homing in crossing over positive spores (n=16). **B)** VDE homing in crossing over negative euploid spores. Here, I only analyzed spores those inherited chromosome III of *S. cerevisiae* parent, since it has the cut-site (n=17). **C)** VDE homing in chromosome III aneuploid spores (n=20). See *VRS homing detection on artificial cut-site* under Materials and Methods section, for details.



Supplementary_Figure_2. Examples of dissected colony morphologies for three hybrid genotypes. For each F_1 hybrid genotype two sample plates are represented. Each plate has 80 dissected spores coming from 20 4-spore-tetrads. Top: WT hybrid spores; middle: p*CLB2SGS1* hybrid spores; bottom: p*CLB2SGS1* and p*CLB2MSH2* double-mutant hybrid spores. Plates were kept in 30 °C for 2 days and then scanned.



Supplementary_Figure_3. Change in the crossing-over (blue) and aneuploidy (red) levels through the treatments. Expectation is that with an increase in crossover percentages between the treatments a correlated decrease in the aneuploidy levels should be seen. However, especially the significant increase in crossovers seen visible moving from Sgs1-m and Msh2-m to Sgs1-m and VRS treatment (blue-dots), does not follow with a decrease in aneuploidy levels when compared between the same two treatments (red-dots).

Genotype	Strain	Percentage spore viability	Viable spores / dissected spores
hybrid, wild-type	N17 x W303	0.46	5/1077
hybrid, p*CLB2MSH2*	N17 x W303	3.18	28/880
hybrid, p*CLB2SGS1*	N17 x W303	20.8	443/2125
hybrid, p*CLB2SGS1* & pCLB2MSH2	N17 x W303	32.6	2763/8468
S.p. p*CLB2SGS1*	N17	87.75	351/400
S.c. p*CLB2SGS1*	W303	83.47	323/396
S.p., p*CLB2SGS1* & *pCLB2MSH2*	N17	88.5	354/400
S.c., p*CLB2SGS1* & *pCLB2MSH2*	W303	73.17	281/384

Supplementary_Table_1 Spore viability levels of tested for hybrids with different treatment genotypes and for non-hybrid control genotypes. The double mutation of Sgs1-m and Msh2-m can be costly as when under *S. cerevisiae* background only 73.17 % and when under *S. paradoxus* background only 88.5 % of the spores survive. The upper-bound of 73.17 % viability of the *S. cerevisiae* non-hybrid is accepted as the highest possible viability we can get in a full hybrid double-mutant, and this value has been taken as the random spore death probability ('R') in the BDMI calculations.

Equation (Li *et al*. 2013) T=(1-R)(1-U)(0.75+0.25(1-I))^N	
(T) observed survival rate of the wild-type hybrid	0.0046
(U) probability of spore death due to the segregation problem	0.326
(R) probability of spore death due to the double mutation	0.2683
(I) killing-stregth of a pair of BDMI interaction	(N) number of asymmetrical incompatibilities required to explain the observed spore death
0.99	16
0.9	18
0.8	21
0.7	24
0.6	29
0.5	35
0.4	44
0.3	60
0.2	91
0.1	185

Supplementary_Table_2. Number of possible two-locus Bateson-Dobzhansky-Muller incompatibilities that can be responsible for the remaining hybrid spore death. Assuming that segregation problem is fixed via the double mutation treatment, then, the BDMI two-locus model would be the best candidate to explain remaining hybrid spore death. Assuming the simplest segregation model (i.e. no centromere linkage) for every two loci inter-chromosomal region, we can predict the number of such pairs of genic incompatibilities between the two species. To do this I use the asymmetrical incompatibility model: For instance only the 'Ab' (big-A and little-b) inter-specific combination have the potential to be lethal and the other inter-specific combination 'aB' (little-a and big-B) is non-lethal. Such an inter-chromosomal two-locus interaction ('Ab') will segregate in one fourth (1:4 or 0.25) of the spores (every meiosis produces equally likely fractions of the genotypes AB, ab, Ab, and aB). Thus the potential number of inter-chromosomal BDM regions to explain the remaining hybrid spore death can be calculated using the formula initially introduced by Li *et al*. (2013) (see the table above). All two-locus BDM pairs are assumed to be independent for simplicity.

CHAPTER FOUR

A screen for Bateson-Dobzhansky-Muller incompatibility regions in yeast

Introduction

Speciation, at the genetic level, is a result of accumulating genetic differences between splitting lineages. Since genomes are subject to different types of mutations, from single nucleotide polymorphisms to large chromosomal rearrangements, the *combination* of all types of mutations may contribute to the evolution of reproductive isolation between populations (Wu & Ting 2004; Seehausen *et al*. 2014). Based on different mutational histories, genetic models of speciation focused mainly on two models of intrinsic postzygotic isolation: 'chromosomal' and 'genic incompatibility' models (reviewed in: Coyne & Orr 2004).

The genic incompatibility (the Bateson-Dobzhansky-Muller or BDMI) model of speciation has received great empirical and theoretical attention since its early formalization at 1930s (Bateson 1909; Dobzhansky1937; Muller 1942; Orr 1996). The BDMI model presents a simple and solid idea that explains the speciation event as an inevitable by-product of genetic divergence in geographic isolation (i.e. allopatry). Further, speciation by the evolution of such incompatibilities is possible through all different models of molecular evolution (i.e. neutral theory or positive selection) (Barton & Charlesworth 1984; Palopoli & Wu 1994; Coyne & Orr 1998; Presgraves *et al*. 2003; Brideau *et al*. 2006). The idea is as follows. Genes (and inter-genic sequences that control genes) at different positions within a population co-evolve. Mutations at different positions (loci) are constantly tested by natural selection within a population, i.e. genes within a population co-evolve. On the other hand, different gene pools of geographically isolated populations evolve independently (no coevolution) from one another. As a result, when members of such allopatric populations interbreed after a period of divergence, between-loci negative epistasis may cause formation of unfit, inviable, or sterile hybrids or hybrid offspring (see also: Wagner *et al*. 1994; Corbett-Detig *et al*. 2013). Thus sequence divergence that affects gene function may act at the postzygotic stage to cause reproductive isolation between populations (reviewed in: Welch 2004; Wu & Ting 2004; Johnson 2010; Maheshwari & Barbash 2011).

Most of the evidence collected for the BDMI model of speciation has been obtained from *Drosophila* species (Noor & Feder 2006; Presgraves 2010). Genetic mapping identified many BDMI *regions* in *Drosophila*, but only a handful of *genes* are identified. Some examples of *Drosophila* genes causing sterility or inviability are: *OVD*, *NUP96*, *NUP160*, *HMR*, *LHR*, and *ZHR* (Phadnis & Orr 2008; Presgraves *et al*. 2003; Tang & Presgraves 2009; Barbash *et al*. 2003; Brideau *et al*. 2006; Sawamura & Yamamoto 1997; Ferree & Barbash 2009). Examples of BDMI genes discovered in different taxa also exist (e.g. Seidel *et al*. 2008; Mihola *et al*. 2009; Chen *et al*. 2014; Chae *et al*. 2014). For instance, incompatibility between *HPA1* and *HPA2* that leads to the embryo lethality in inter-strain crosses of *Arabidopsis thaliana* is an example of the evolution of genic incompatibility which has evolved as a result of divergent evolution in duplicate gene copies (Bikard *et al*. 2009). *XMRK2* is another hybrid incompatibility gene resulting in lethality due to tumor formation in hybrids of the fish species *X. maculatus* and *X. helleri* (Walter & Kazianis 2001).

Technical advances of genomic era have not uncovered the molecular mechanisms underlying BDMI. One reason is BDMI interactions leading to postzygotic inviability might involve more complex interactions than are apparent from crossing experiments (see Wu & Palopoli 1994). Multiple-two-locus or high-order (including at least three loci) interactions are more likely explanations for the evolution of hybrid inviability or sterility (i.e. easier to evolve according to the fitness landscape models of population genetics - see Cabot *et al*. 1994 and Orr 1995). Detecting those higher order interactions, then, is a matter of statistical power (Li *et al*. 2013). Thus, applying high-throughput genotyping techniques on larger samples to extend the evidence for BDMI to different taxa is much needed for the speciation genetics research (Noor & Feder 2006; Li *et al*. 2013).

The yeast species *Saccharomyces cerevisiae* and *Saccharomyces paradoxus*, are good candidates for studying the genetics of hybrid sterility. Haploid cells of these two species mate efficiently to form healthy diploids. However, when the resulting F_1 hybrids undergo meiosis, only ≤1% of the F_1 hybrid gametes are viable, which is classified as intrinsic post-zygotic isolation (Greig 2009). Anti-recombination has been shown to contribute a large part of this hybrid sterility (Chapter 3), however the possibility remains that other mechanisms also play a role. Apart from four small inversions and three segmental duplications, genomes of both species are collinear (Kellis *et al*. 2003). This eliminates gross chromosomal rearrangements (GCRs) as having significant effects on post-zygotic isolation. On the other hand, it is quite possible that the BDMI model can be a significant cause of reproductive

isolation between the two species, since genomes of both are diverged extensively at the single nucleotide level (~14%). Even within a single species, *S. cerevisiae*, crosses whose parents are only diverged about 0.35% at the single nucleotide level have been shown to have incompatible genes (Hou *et al*. 2015, also see: Heck *et al*. 2006). To date, no BDMIs have been detected that contribute to hybrid sterility in these two species (Greig 2007; Kao *et al*. 2010; Xu & He 2011), however previous methods have suffered from limitations. Greig substituted whole chromosomes from *S. paradoxus* into *S. cerevisiae*, and found that they were compatible, however this method would not detect weak BDMIs (2007). Kao *et al*. 2010 screened the rare (<1%) viable gametes produced by F_1 meiosis for under-represented combinations of loci, however such surviving gametes are highly aneuploid due to the action of AR during hybrid meiosis, so recessive BDMIs would be masked by complementation. Here, we present a method that solves both of these problems.

In this study, to discover BDMIs in yeast, we firstly reduced the anti-recombination barrier by repressing the meiotic expression of genes *MSH2* and *SGS1* (see Chapter 3). Repression of these proteins, promoted recombination between the two species chromosomes, and therefore increased the viability of *S. cerevisiae* and *S. paradoxus* hybrids, from 0.5% up-to 32,6%. Improvement of the fertility still left room for the BDMIs as being the other alternative that could as well be responsible for the rest of the hybrid spore death (see Supplementary Table 3 in Chapter 3). Secondly, we collected 336 viable gametes from 84 F_1 hybrid meiosis events, selecting only gametes from tetrads in which all four spores were viable. This ensured that they were all haploid overcoming the Kao *et al*.'s problem of aneuploid gametes (2010). We genotyped viable hybrid spores and tested for underrepresented genotypes that would indicate BDMIs.

Methods

Experimental Procedure

Generation of four spore viable F_1 hybrids for genotyping:

To obtain haploid F_1 hybrids for sequencing, we initially crossed *S. paradoxus* strain n17 (YDG980) and *S. cerevisiae* strain w303 (YDG981) and hence obtained a diploid hybrid (YDG982). Expressions of *MSH2* and *SGS1* anti-recombination genes were repressed during meiosis (see Chapter 3 for details of genetic modifications). We induced sporulation/meiosis by incubating the hybrid diploid (YDG982) in 3 ml KAc (2% potassium acetate sporulation media) by shaking the cultures for four days in room temperature. We confirmed completion

of meiosis by observing the tetrads under the microscope. We dissected tetrads that had four visible spores. To digest the ascus wall of the hybrid spores, we incubated spores in 1unit (per 10 µl) zymolyase (Zymo Research EU, Freiburg, Germany) for 30 minutes. After enzymatic digestion of the ascus wall, we replaced four spores of tetrads onto the YEPD (2% glucose, yeast extract, peptone, agar) plates using the MSM400 tetrad dissection microscope (Sigma Instruments, UK). Spores were kept on the YEPD plates for two days for germination and colony formation at 30 °C incubator.

DNA extractions:

To isolate the genomic DNA, we picked only the four-spore-viable tetrads. We excluded any F_1 spores with three, two, or one spore viable spores. We picked in total 93 four-spore-viable samples (372 F_1 haploid hybrids). However, we reduced our sample size in our final analysis down to 336 spores (84 tetrads), because sequencing results of 36 spores (nine tetrads) were unclear. We also sampled four F_1 spores of both *S. cerevisiae* and *S. paradoxus* non-hybrids, and four samples of hybrid diploids as controls. Thus we initially isolated genomic DNA from 384 samples (including 36 samples that were excluded from further analysis later) using MasterPure ^{TM} Yeast DNA Purification Kit (Epicentre, Biozyme Biotech, Oldendorf). We validated the quantity and quality of the DNA extracts by using NanoDrop (Thermo Scientific Inc.) and by running samples on 1 % agarose gels. We diluted all DNA samples down to 50 nanograms (per µl) for library preparations.

Library preparation using RAD-tagged doubled digest method:

To prepare the samples for sequencing, we used double digestion based RAD-tag library preparation method (Etter *et al*. 2011; Hohenlohe *et al*. 2010; Peterson *et al*. 2012). We used restriction enzymes Csp6I and PstI to double digested the genomes of all samples. Separately for each sample, we mixed ligase, two restriction enzymes (PstI and Csp6I), 50 ng of sample DNA, and adapters (adapterX_TagY_fq and adapterX_TagY_rv) to a final volume of 20 µl. Reaction mixtures were incubated at 37 °C for two hours. After the ligation reaction, we cleaned up the excess adapters, enzymes, and fragments smaller than 300bp by using Ampure beads at a 1:1 ratio. Next we mixed Phusion Hot Start II High-Fidelity DNA Polymerase (2U/µl), adding P5 and P7 primers at 10 mM concentration, dNTPs (2mM per dNTP), and 5X Phusion HF Buffer to amplify the target regions (Acinas *et al*. 2005; Etter *et al*. 2011). 30 µl PCR mixture was amplified as initial 98 °C incubation for 30s, followed by 25 cycles of 98 °C 10s, 68 °C 15s, 72 °C 30s, and with a final extension at 72 °C for 5 mins. We used fluorescence NanoDrop to check the quantity and quality of the samples.

Genotyping the samples:

To sequence the tagged samples, we mixed all tagged samples in one pool. All samples were multiplexed using combinations of unique barcodes therefore reads from a single sequencing reaction would have unique reverse and forward tags that will help us to distinguish all samples after obtaining the pool of MiSeq reads. We used MiSeq platform to obtain 300 bp paired-end reads.

Data Analysis

Mapping the reads:

To map the reads, we assembled two reference genomes using the gene order of *S. cerevisiae* (strain w303) parent. This did not create any problems since genomes of both species are almost perfectly collinear. Nevertheless, we removed open reading frames that were absent at the same position of either one of the species from our reference genomes. We also removed intergenic regions from our reference genomes. Some reads were clearly mapped onto one species and did not map onto the other species. Nevertheless, some reads mapped to both parents in regions where there was high sequence identity between the sequences of the two genomes. This was not an issue though. There was a clear difference at the alignment scores for such reads. The parental state was assigned to the parent on to which most reads mapped. We used the GATK software for mapping the reads.

Visualization of the genotypes:

To visualize the genotyped samples, we used freely available PlotTetradSeg component of the Recombine package (Anderson *et al*. 2011) on R (http://www.R-project.org). For the visual map of the pairwise BDMIs we used the Circos software (Krzywinski *et al*. 2009). Supplementary Figure 1 shows a sample of genotype of the first four spores. Supplementary Figure 2 shows the crossover rate in relation to the length of each chromosome. Supplementary Table 2 shows the raw numbers for chromosome length and rates of recombination.

Segmentation of chromosomes according to the crossover breakpoints:

To identify the position of crossover break points for each chromosome in each sample, we took mid-points at positions where switches in mapping took place from one species to the other species. Chromosomes were divided into segments using the information of crossover breakpoints using all break point information obtained from all samples (see Kao *et al*. 2010).

This gave us 1208 segments in total. This means that, starting from chromosome I to chromosome XVI, in all our samples, crossovers divided all of the samples into 1208 segments in total. See Supplementary Table 1 for the number of segments per chromosome (in a total of 336 hybrid spores).

Calculating two-locus incompatibility regions by generating two-by-two Chi-square tables:

To calculate for two-locus interactions, we performed Chi square analysis using segmentation data from all 336 samples. Pairwise comparisons were made only between two segments located at different chromosomes (e.g. inter-chromosomal). Thus, any intra-chromosomal comparisons were excluded from analysis. These inter-chromosomal two-locus comparisons were performed for all possible pairs across all 16 chromosomes.

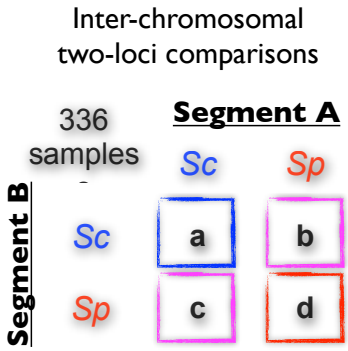


Figure 1. Two-by-two tables to identify incompatibilities between two-locus of *S. cerevisiae* (Sc) and *S. paradoxus* (Sp) genomes. Inter-chromosomal two-locus comparisons were performed for all pairwise interactions across 16 chromosomes. This gave 730236 two-by-two tables. Chi squared statistics were computed for all of the tables. If any two loci are incompatible between two species and therefore likely to kill off the spores when they co-segregate into that sample as in hybrid combination, then, we expect deviation from random segregation for such two-locus interactions. For instance, on the figure, values for 'a' and 'd' non-hybrid segregants will be over-represented and 'b' and 'c' hybrid segregants will in comparison be under-represented in viable samples. This is because hybrid combinations have a probability (e.g. killing effect) of causing inviability when they segregate into the same spore (see Li *et al*. 2013).

Four segregants (Sc-Sc, Sp-Sp, Sc-Sp, and Sp-Sc) are formed for two loci found at different chromosomes (Figure 1). However, if any two loci show BDMI interaction, hybrid combinations (Sc-Sp, and Sp-Sc) are less likely to survive due to negative effect of epistasis, and therefore scores of samples in two-by-two table deviate from the neutral expectation (that is 25% for each genotype categories), and thus increase the Chi square scores. Spore numbers

for all four categories were entered into two-by-two tables for Chi square statistical analysis. All in all, we generated 730236 of such two-by-two tables to detect any two-locus negative-epistasis between the two species. From 730236 tables, we arbitrarily selected top 330 segment pairs that appeared on top of our Chi squared tables as potential BDMI interactions. All these interactions had a Chi squared score above 20.0.

Calculating the false discovery rate (FDR) for the top 330 putative BDMIs:

To calculate the false discovery rate of among the 330 putative BDMIs with the highest chi-squre values, we simulated the meiosis to produce a null distribution of genotypes for comparison. We redistributed all randomized segments into 84 tetrads (336 samples) randomly, and repeated the simulation for 80 times. Simulations would give interactions with varying p-values just by chance due to enormous number of pairwise comparisons. Thus, for the Chi square value threshold of 20.0, we also recorded number of pairwise interactions appeared just by random chance above that threshold. We divided the number of seemingly significant interactions obtained from the simulations to the number of significant interactions obtained from the analysis of the real data (i.e. 330).

Determining the number of interactions with a potential BDMI strength:

To discover the two-locus incompatibilities by exploiting big data using different methods, apart from two-by-two Chi squared table analysis, we also performed a different analysis by comparing proportions of tetrad class that were parental ditype to tetrad class that were non-parental ditype for all possible two-locus (inter-chromosomal) pairwise comparisons.

To better explain the parental/non-parental proportion analysis, segregation of two loci found on different chromosomes should be explained in the context of yeast meiosis. Yeast meiosis produces four spores at each division. This forms three classes of tetrads for two independent loci. Imagine ‘A’ and ‘B’ are the genotypes of one parental species and ‘a’ and ‘b’ are the genotypes of the other parental species. Even though these two loci are located on different chromosomes, random segregation of them would reproduce i) parental ditype (PD) class, ii) non-parental ditype (NPD) class, and iii) tetratype (TT) class of four gametes. PD tetrads contain two AB and two ab genotypes. NPD tetrads contain two Ab and two aB genotypes. TT tetrads contain one of each possible genotype: AB, ab, Ab, and aB. Random assortment gives a ratio of 1:1:4 for the three classes of tetrads (PD:NPD:TT). If one of these two loci (either A/a or B/b) are not centromere linked then this 1:1:4 (PD:NPD:TT) ratio is secured. However, if both loci have centromere linkage, depending onto the distance of the more

distant loci, there is a reduction in the TT-type proportion. PD:NPD ratio is not affected by centromere linkage and therefore stays at 1:1, if interactions are neutral (i.e. no BDMIs) (Sherman 2002). Following the BDMI model, it is expected that the two-locus incompatibilities would lead to a deviation of the PD/NPD ratio non-randomly. That is absence or reduction in hybrid two-locus combinations (NPDs) would increase the ratio above the expected random threshold.

To check for this, we initially determined PD, NPD, and TT classes of every two-locus interaction those have segregated into 84 four-spores. We excluded TT tetrad counts and calculated PD to NPD ratio for two-locus regions. If any two-locus are perfectly compatible between two species, the expected mean of PD to NPD ratio would be 1. However data will deviate from 1.0 randomly in both directions. After calculating the PD/NPD ratio distribution, by summing the three standard deviation (one standard deviation= 0.396) and median value of PD/NPD (median=1.079), we calculated a cut-off value of 2.226. Then we detected the number of two-locus interactions that deviated from the cut-off value as to get a rough estimate of two-locus interactions that may involve in negative epistasis with different killing-strengths.

Results

Seven putative BDMI interactions detected between the two species:

We obtained 1208 segments across 16 chromosomes in a total of 336 samples (Supplementary Table 1). We took each segmental region as the shortest individual loci of our dataset. Next, we tested segregation ratios of segments found on different chromosomes (two-locus interactions) to discover BDMI regions between the two species. To do that, we generated two-by-two tables and entered the number of samples for every segregation categories of every possible two-locus pairwise comparison. We generated in total 730236 two-by-two tables (Figure 1). We computed Chi-square statistics to detect interactions between segments from different chromosomes. We have chosen top 330 pairs of interactions (Chi squared value > 20.0) as putative BDMI regions (see *Methods*).

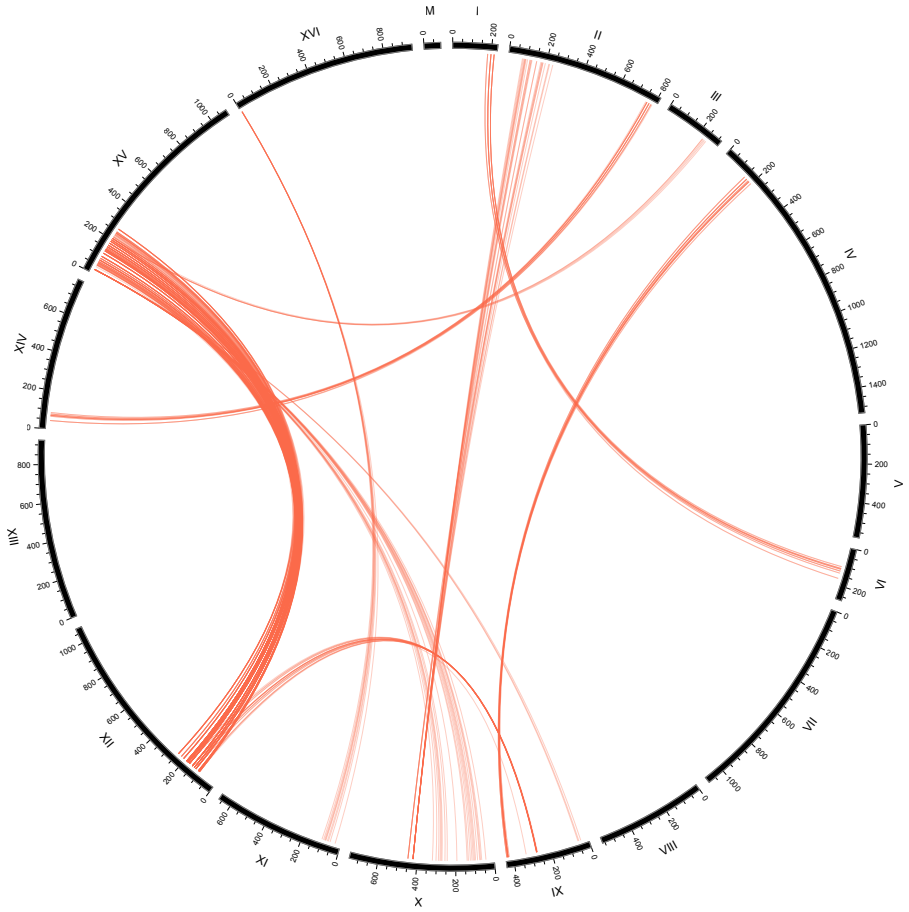


Figure 2. Interaction-map of the strongest two-locus interactions that deviate from expected segregation values. We have chosen 330 segments from our top Chi-squared categories of a total of 730236 tables. These fall into 10 categories of two-locus interactions. Seven of these interactions are putative BDMIs. False discovery rate for these 330 pairs of interactions is 89%.

These top 330 pairs of segments fell into 10 categories (Figure 2). Seven of these interactions are in the negative direction expected, which represent putative BDMIs. For all of the categories of these seven putative BDMI regions, samples were biased towards containing more of the non-hybrid segregants in surviving spores. Interaction between chromosome XII and chromosome XV is the strongest candidate, as it includes 201 pairs of segments of these 330 top interacting pairs. All of these 201 pairs of segments are underrepresented in viable samples for hybrid combinations of both loci.

We detected three interactions (in Figure 2) within those top ten categories of interactions that had a positive direction (Table 1). This means that the hybrid combinations of the three regions were more likely to survive compared to non-hybrid combinations of all three two

loci interactions. Chi squared level below which we start to observe interactions in both directions (i.e. hybrid combinations underrepresented and hybrid combinations overrepresented), can be a good cut-off point to select for BDMIs with relatively stronger killing strengths.

False discoveries due to massive multiple testing was possible within these 330 pairs. To detect the false discovery rate (FDR), we simulated the meiosis by randomizing segmented dataset to produce a null distribution of genotypes (see Methods). By dividing the number of pairs of interactions appeared in our random simulations to the number of interactions found in experimental data (using the same Chi squared level threshold), we calculated a false discovery rate of 89%.

9061 two-locus segment pairs with potential BDMI-strength:

Finally, we computed the ratio of parental ditype tetrad samples to non-parental ditype tetrad samples for every pairwise interaction (see *Methods*). This informs us on the number of two-locus interactions that can involve BDMIs with varying but mostly very weak killing-probabilities. Proportion of the parental to non-parental tetrad classes in our samples (PD/NPD ratio) should be equal (or around equality) if interaction between two loci is neutral (PD/NPD = 1.0). However, non-random deviation from the ratio of 1.0 is possible when there is negative epistatic interaction between any two loci, which decreases the number of non-parental type tetrad survival that inherits such two loci as in hybrid combination. If there are BDMI interactions we expect the PD types to be over-represented in comparison to the NPD types in surviving spores. The data can be exploited to get a rough estimate of two-locus regions that may involve in hybrid spore killing at various killing-rates by looking at the data distribution of PD/NPD ratio for all two loci interactions. As a result, we found that 9061 two-locus pairs had potential BDMI interactions with very weak strengths, all of which appeared at three standard deviation away from the mean (Figure 3).

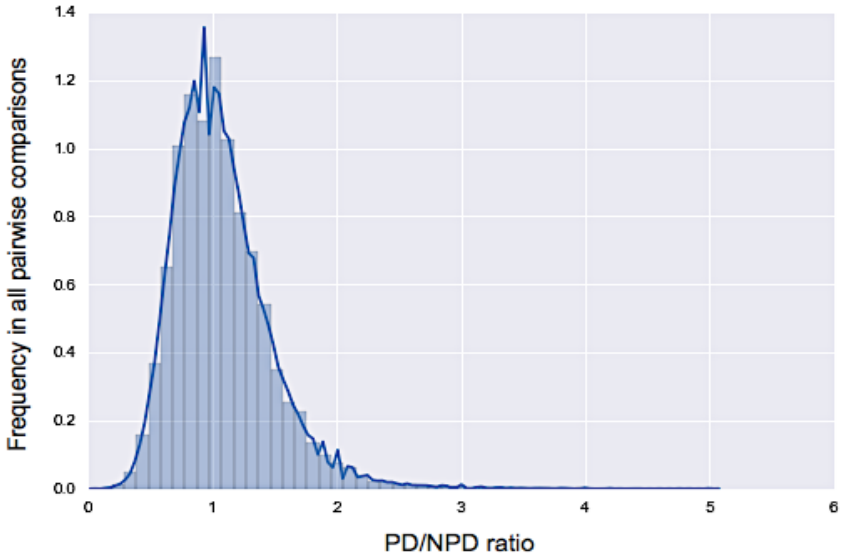


Figure 3. Parental ditype (PD) to non-parental ditype (NPD) ratio of samples for all pairwise inter-chromosomal interactions. The expectation for two loci interactions that are neutral would give 1.0 for the PD/NPD ratio. Median is 1.079, and standard deviation (S.D.) is 0.396. Three-sigma rule gives a cut-off of 2.226. 9061 two-locus interactions with potential BDMI strength were found. Parental type (non-hybrid) segregants were relatively higher than non-parental (hybrid) segregants in our sample of tetrads (see *Methods*).

Discussion

Our study is designed to discover BDMIs between the two siblings of *Saccharomyces* yeast. Here, we show that there are seven putative two-locus BDMI regions between the two species. However, we calculated a very high false discovery rate for our top categories of interactions.

First weak putative BDMI regions documented between the two species:

Previous studies that were designed to discover BDMIs in yeast did not found any such interactions (e.g. Greig 2007; Kao *et al*. 2010). As stated by Kao *et al*., this could mean that, rather than strong effect BDMIs (those kill hybrids with a 100% killing probability), most of the BDMIs that exist between the two species could have weak effects (2010). In line with this, Kachroo *et al*. (2015) could recover functions of almost half (47%) of the yeast genes that were essential for survival, by complementing their copies with orthologs from human genome. This shows that two-locus interactions between siblings of yeasts with strong killing effect may evolve rarely. However, previous studies had their limitations (Greig 2007; Kao *et al*. 2010). Kao *et al*. screened BDMIs in the rare (<1%) viable gametes those had extra chromosomes. Since BDMIs in yeast are likely to be recessive (Greig *et al*. 2002), aneuploidy

can mask those interactions. To discover weak two-locus BDMIs, it is necessary to use large samples (≥400) with no aneuploidy (Li *et al*. 2013). We designed our study to solve aneuploidy issue by screening only four spore viable gametes, as they are guaranteed not to have any extra or missing chromosomes.

As a result, we could detect putative BDMI regions in yeast. We found 330 pairs of regions located at different chromosomes that deviated from the expected neutral segregation rates. These 330 pairs of segments fell into ten categories of two-locus interactions (Figure 2). We show that seven of these two-locus interactions are putative BDMI. Within these seven, direction of the interactions was negative where hybrid combinations of them were less likely to survive, which is an expected outcome for BDMI interactions. However, their killing effects are low (see below). Each of those interactions, which fell into ten general categories, has consistent results where deviation from expected neutral segregation ratios in viable spores was always in the same direction. This means that, for instance, all pairs of segments (201 in total), that make up the interaction between chromosome XII and chromosome XV, are underrepresented for hybrid combination of segregants, and none of the 201 segment pairs are enriched for hybrid combinations.

Unexpectedly, within these top ten categories, we found three interactions those had a positive direction in their segregation ratios, where hybrid combinations were more likely to survive (Table 1). A similar result has been observed in the study of Kao *et al*. (2010). They have identified three chromosomes were less likely to be inherited from the same species (Kao *et al*. 2010).

Finally, it is very critical to point that, in our random meiosis simulations, we calculated a very high false discovery rate for the 330 pairs of segments. An FDR of 89% prevents us to be certain on the true BDMI regions within our top biased two-locus segregants. The interaction between chromosome XII and chromosome XV contains significantly more segment pairs (201 pairs in total) within the top 10 interactions (Supplementary Table 2). Therefore, it is the candidate to contain the strongest BDMI interaction between the two species.

Killing strength of putative BDMIs detected in this study:

It is intriguing to test the relative spore killing effect of BDMIs for yeast. We could quantify total spore killing effect of the putative BDMIs by using very simplified assumptions. To do

this, we detected the deviation of each category of interactions from the expected ratio (25%) of neutral segregation. Next, we took spore-killing effect of seven negative interactions (putative BDMIs) and three positive interactions, and summed the deviations of all interactions. Finally, since such spore killing interactions will segregate into one fourth of the segregants, we divided the total spore killing effect by four.

Interaction	Number of segments involved	% Killing Effect (mean for involved segments)	Direction (negatives are putative BDMIs)	Individual killing effect (%) divided by four (probability of co-segregation)	Locus 1 (Start-End position)	Locus 2 (Start-End position)
1	201	-14.08434	-	-3.521085051	chr12 98687 - 267200	chr15 30795 - 240219
2	24	-13.51687	-	-3.37921627	chr02 71863 - 251063	chr10 419763 - 465996
3	21	-14.68254	-	-3.670634921	chr10 25663 - 323458	chr15 16405 - 26368
4	19	-13.78446	-	-3.446115288	chr09 269759 - 278951	chr12 98687 - 196103
5	18	-13.5582	-	-3.389550265	chr04 142920 - 206794	chr09 418734 - 442905
6	15	13.452381	+	3.363095238	chr02 754283 - 791315	chr14 24397 - 88103
7	15	13.412698	+	3.353174603	chr01 177342 -247982	chr06 98086 - 173664
8	9	14.153439	+	3.538359788	chr11 26923 - 103253	chr16 280 - 8755
9	5	-13.09524	-	-3.273809524	chr03 231320 - 257445	chr15 200961 - 218156
10	3	-15.47619	-	-3.869047619	chr09 32385 - 56467	chr15 16405 - 26386
			Total killing effect	14.29%		

Table 1. Individual and total killing-strength of putative BDMIs detected. Two loci segregate into four categories of genotypes. If interactions between any two loci are neutral, expectation for each category is 25%. By subtracting deviation of each negative (seven putative BDMIs) and positive interaction from their expected rate of 25%, we can try calculating individual killing-effect of each interaction. Next, by summing killing-effects of all ten interactions, and dividing the sum by four (assuming interactions are asymmetric and thus segregates into one fourth of gametes), we calculated killing-strength of top categories of interactions to be about 14%.

As a result, we calculated a spore killing effect for BDMI model to be about 14% (Table 1). We know that about 32% hybrid gametes are inviable due to the anti-recombination problem (Chapter 3). This leaves about 53% room of remaining hybrid inviability between the two species. The remaining spore death could be a result of anti-recombination barrier, as we cannot be sure if we have fixed segregation problems between the two species (Chapter 3). Anti-recombination model is still the major candidate responsible for remaining spore death. However, BDMIs with very weak spore killing effect (i.e. Figure 3) can also have a role in remaining sterility. Nevertheless, our assumption of 14% spore killing effect of BDMIs should be experimentally validated. To do this every putative BDMI interactions should be tested by replacing these regions in both species' genetic background using the sequence of the other species, and then by running the partial replacement lines through meiosis, and F_1 gamete viability should be scored.

More complex interactions involving many weak-effect regions:

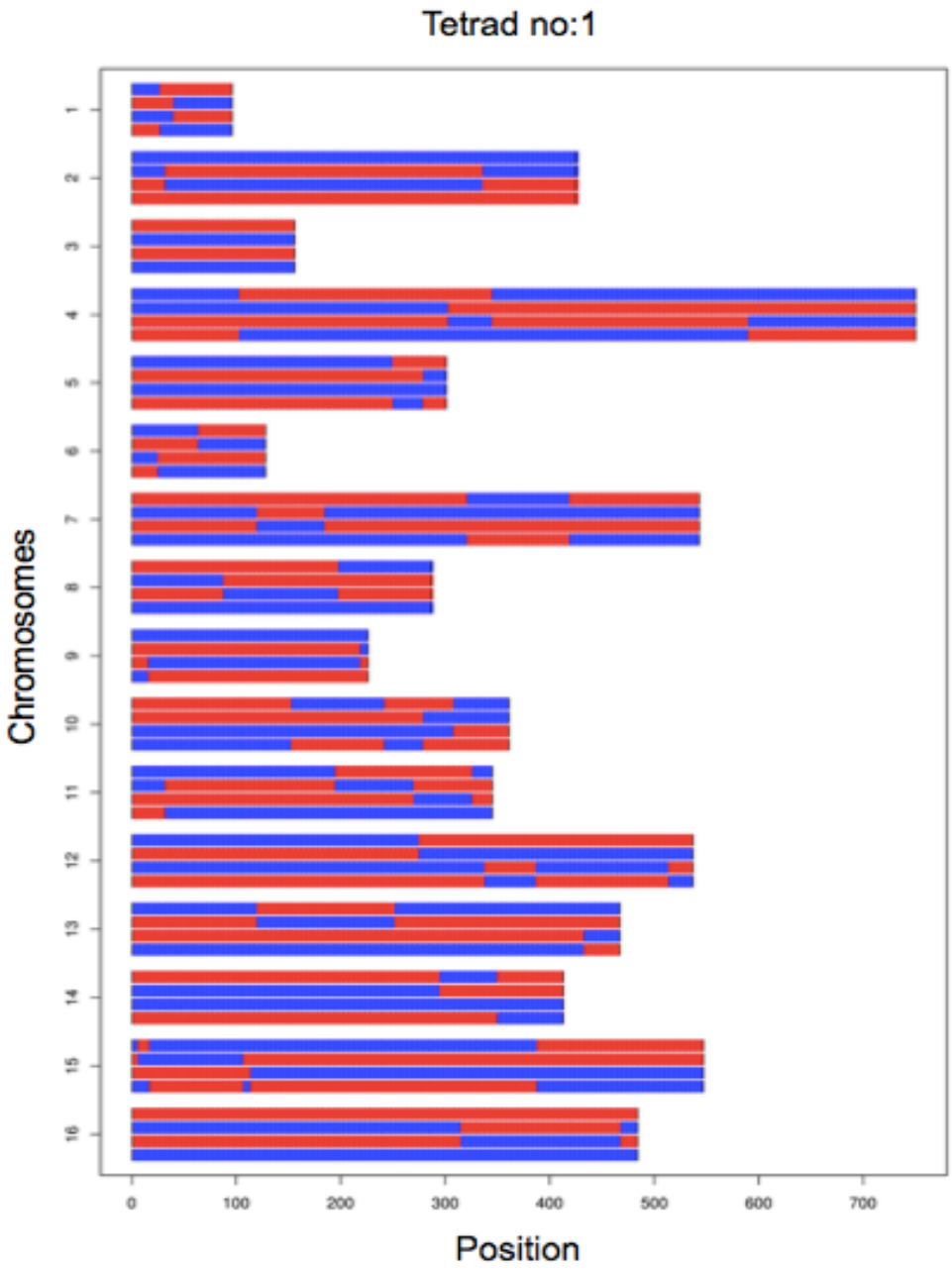
Previous studies showed that, even at very low genetic distance levels (i.e. ≤ 1% single nucleotide divergence), two-locus BDMIs exist in yeast (Heck *et al*. 2006; Hou *et al*. 2015). Moreover, experimental evolution studies lead to evolution of two-locus negative epistatic interactions between allopatric populations (Anderson *et al*. 2010; Kvitek & Sherlock 2011; Bui *et al*. 2015). Evidence then suggests that such two-locus negative epistatic interactions are likely to evolve in very short periods between yeast populations. On the other hand, genic incompatibilities which kill-off yeast hybrids seem to require at least three loci interactions (Kao *et al*. 2010). From the population genetics perspective, it is known that such higher-order incompatibilities are easier to evolve in comparison to the evolution of effective BDMI regions those contain very few interacting loci (Orr 1995). Such complex BDMIs can kill-off yeast gametes when three or more interacting loci co-segregate into the same gamete (but see: Leducq *et al*. 2012). However, it is challenging to detect weak effect BDMIs. Because, very complex interactions all with very tiny killing effect requires extremely large sample size.

Nevertheless, our data hints the existence of such complex BDMI networks. For instance, chromosome XV seems to be a central hub for multiple two-locus interactions. The long region at the arm of the chromosome XV has interactions with four other chromosomes (Figure 2). Moreover, interactions sharing identical regions between four loci in four chromosomes (III-XV-XII-IX) suggest that higher order interactions, involving four-loci are likely (Figure 2). Supporting the argument further, we detected 9061 two-locus regions that deviate from the normal PD/NPD distribution. Existence of such two-locus interactions with potential BDMI effects points that there may be many complex interactions, which have very weak effects in reducing hybrid fertility when alone (Figure 3). In sum, although our study do not have the power to test for complex BDMIs, yet networks appear in our two-locus interaction analysis hint the existence of more complex BDMIs between the two species (see for similar examples: Chae *et al*. 2014; Schumer *et al*. 2014; Turner & Harr 2014; Paixão *et al*. 2014).

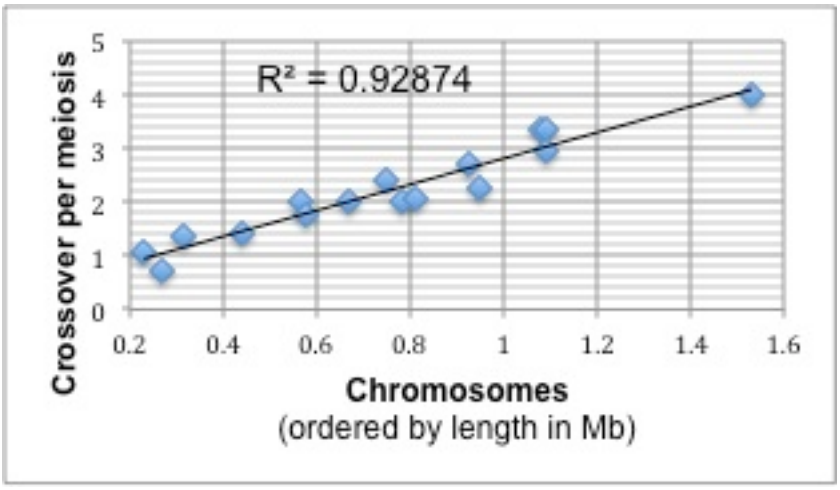
Conclusion

This study provides the first evidence of putative BDMI regions in yeast. Yet, we cannot be certain about the significance of these regions as BDMIs since our false discovery rate is very high. Nevertheless, we estimate spore-killing effect of top BDMI categories we have detected here to be about 14%. To be conclusive, experimental validation, by replacing the interacting regions in either species' backgrounds would be a necessary step to take in a further study.

Supplementary Data



Supplementary Figure 1. Genotypes of the first four spores (first tetrad) in a total of 336 samples (84 tetrads). Four genotypes obtained from four viable spores are visualized as sets of four. Red segments represent *S. cerevisiae* ancestry and blue segments represent the *S. paradoxus* parental ancestry. Every chromosome was segmented according to the shortest regions those did not had a recombination within themselves. PlotTetradSeg (Anderson *et al*. 2011) script was used on 'R' to visualize these tetrad images.



Supplementary Figure 2. Crossover events per meiosis ordered by length of every chromosome. Length of the chromosomes and the number of crossover events observed correlate (R^2=0.928).

Chromosome	Number of segments
1	26
2	86
3	53
4	119
5	69
6	26
7	92
8	69
9	52
10	74
11	73
12	96
13	87
14	74
15	107
16	105

Supplementary Table 1. Segment number per chromosome obtained using crossover breakpoint date of 336 samples in total.

Chromosome	Length (Mb)	Crossover per meiosis
1	0.230218	1.058139535
6	0.270161	0.720930233
3	0.31662	1.360465116
9	0.439888	1.430232558
8	0.562643	2
5	0.576874	1.755813953
11	0.666816	2.034883721
10	0.745751	2.418604651
14	0.784333	2.023255814
2	0.813184	2.081395349
13	0.924431	2.709302326
16	0.948066	2.26744186
12	1.08	3.360465116
15	1.09	3.360465116
7	1.09	2.965116279
4	1.53	4

Supplementary Table 2. Crossover rate per chromosome in *Sgs1* and *Msh2* double-mutant hybrid. Recombination events increased to 35.5 per meiosis observed in hybrid tetrads (17.7 recombination event per spore).

Contributions to the thesis

Chapter one

This chapter will be submitted to the Journal of Evolutionary Biology (JEB).

GOB (Gönensin Ozan Bozdag) and DG (Duncan Greig) conceived the study. GOB designed and performed the experiments. GOB performed the statistical analysis and wrote the chapter.

Chapter two

This chapter was published in the journal Molecular Ecology in October 2014.

Citation: Bozdag, G. O. and Greig, D. (2014), The genetics of a putative social trait in natural populations of yeast. Mol Ecol, 23: 5061–5071. doi:10.1111/mec.12904

GOB and DG conceived the study and the experiments. GOB performed the experiments. GOB performed the statistical analysis. GOB and DG wrote the paper.

Chapter three

This chapter will be submitted to the PLOS ONE.

GOB and DG conceived the study. GOB performed the experiments. GOB performed the statistical analysis and wrote the chapter.

Chapter four

This chapter will be submitted to the Molecular Biology and Evolution (MBE).

GOB, David W. Rogers, and DG designed the study. I performed the experiments and generated the figures. Emre Karakoc conceived and performed the data analysis. Arne Nolte designed the protocol for library preparation. I wrote the chapter.

Concluding remarks

Natural selection acts on different levels (i.e. genomes within an individual, between genomes of members of the same species, and between genomes of different species). Further, characteristics of physical environment, social interactions between individuals of a population, and intra-genomic environment of cells, all interact dynamically and thus actively shape the outcome of evolutionary diversification. In this thesis study, I used yeast to study the role of diverse forms of natural selection acting at different levels in evolutionary diversification. Furthermore, I uncovered relative importance of genetic differences at forming barriers to gene flow between *S. cerevisiae* and *S. paradoxus* yeast species.

Genomes within an individual may have different interests (e.g. intra-genomic conflicts). Evolutionary arms race that may occur between mitochondria and nuclei of yeast can drive co-evolution of two genomes. However, co-evolution can be disrupted differently in different modes of transmission (i.e. vertical or horizontal). Findings presented in the first chapter suggest that natural selection can rapidly evolve negative interactions between nuclei and mitochondria. Results also suggest that intra-genomic conflicts may have roles in evolutionary diversification by evolving negative interactions between nuclei and mitochondria of diverse populations, as in the form of BDMIs. The set-up introduced in chapter one should be tested in more detail to see if such intra-genomic interactions that arose by action of natural selection would have a direct effect in reducing fitness of hybrids formed between allopatric populations.

Genomes (or genes) between members of the same species may also have conflicting interests. Natural selection acting on the level of social interactions may drive diversification within and between populations of microbes at the gene copy number level. On the other hand, social interactions do not take place in a vacuum. Physical environment itself, on which these interactions take place, may directly affect the evolution of variation at gene copy number level, without any influence of social interactions. Findings presented in Chapter 2 suggest that putative social trait investigated has evolved as a byproduct of environmental adaptation. Therefore, while investigating evolution of social traits and genetic fingerprints of those traits in microbial populations, more parsimonious explanations of evolution (i.e. environmental adaptation) can be considered initially as null models.

Genetic drift and all forms of natural selection (Chapters 1&2) lead to diversification of life forms. As a result of random and non-random evolutionary events, species diverge to a level

that prevents interbreeding of closely related species. Final two chapters of this thesis represent novel data on relative importance of overall nucleotide divergence (Chapters 3) and local nucleotide divergence (Chapter 4) at preventing interbreeding of the two species of yeast.

One cause of reproductive isolation has already been shown to be extensive single nucleotide divergence that has spread throughout genomes of the sibling species (Hunter et al. 1996). Evidence presented in Chapter 3 shows that ~14% single nucleotide divergence is the major barrier to interspecific mating of the two yeast species. Recovery of recombination between chromosomes of both species increased hybrid gamete viability from 0.5% to about 32%. The result therefore shows that, segregation problem that occurs during meiosis of hybrids amounts to one third of reproductive barrier in *Saccharomyces* yeast.

The other potential cause of reproductive isolation between the two yeast species was suggested to be the popular model of genic incompatibilities (i.e. the Bateson-Dobzhansky-Muller incompatibilities or BDMIs). Model suggests that, alleles at different loci function well in the genomic background they have evolved with-in. However, when a locus is transferred into a different genetic background of a different species, between-loci negative epistasis may cause formation of unfit, inviable, or sterile hybrids or hybrid offspring. Yet, no evidence of such incompatible regions has been found between the two species of yeast. Final chapter of the thesis (Chapter 4) represents the very first evidence of seven putative genic incompatibility regions in yeast. However, relative killing effect of each of those seven interactions should be experimentally tested to truly prove biological significance of those suggested BDMI interactions.

Acknowledgements

I am grateful to my supervisor Duncan Greig, for for giving me the freedom to organize my work as I wish, for his guidance with my research, and for his helps on the writing of this thesis. I was especially lucky, as I could study very diverse topics of evolutionary biology under his supervision. Duncan did not hesitate for a moment and allowed me to study on those diverse topics. As a result, I could learn a lot about evolution and yeast genetics in his laboratory. My experience in his laboratory will be one of the most enjoyable experiences of my life.

I cannot thank Dave Rogers enough in here, so the following is just a summary. I have learnt a lot from him. His understanding of evolutionary biology and molecular genetics will keep inspiring me. His ideas made every part of this thesis a lot better. Moreover, his friendship made my stay in Plön enjoyable. I will always miss those moments.

I owe a lot to Emre Karakoc for his helps with the Chapter Four of this thesis. His computational skills made the study possible. We also had very interesting conversations about many topics, such as philosophy, literature, music; his friendship made life here in Nothern Germany very interesting.

Jai Denton is a great friend. Having a conversation with him was always like a calming therapy session. He provided valuable suggestions for the early versions of the text. He also provided important advise for the study presented in Chapter Three.

I have learnt a lot from Rike Stelkens. She is one of the most positive people I have met here. I enjoyed our conversations a lot. She also willingly helped me with the text of this thesis.

I thank my friend Luisa Pallares. I visited her in her office many times. We had very interesting conversations there. She also helped me with the Chapter One. I thank Luisa also for organizing our book club meetings, which included Jorge Peña and Rafik Neme. Rafik Neme made beginning of my life in Plön a lot easier. I have moved to his place, and he helped me a lot during that period. He also willingly helped me with the Chapter Three.

I thank Ellen McConnell a lot. I had a lot of interesting conversations with her outside of the institute. She has a great taste in music: we went out of Plön many times to see great music groups. I will not forget those moments. Further, she was a great office mate.

I thank Gunda Dechow-Seligmann for her helps in the lab. Gunda and Elke Bustorf prepared the library for the sequencing of DNA samples (Chapter Four). I also thank Mahesh Panchal who helped with demultiplexing the sequencing reads.
I also thank my lab mates: Primrose Boynton, especially for organizing the experimental evolution journal club. I have learnt a lot in those meetings. I also thank Joana Bernardes, her colorful character made things in the lab a lot more interesting. I thank Eric Miller for discussing about my projects at the beginning of my Ph.D.
I thank my friend Vienna Kowallik very much. Having a conversation with her is always fun and insightful. She motivated me a lot during the final months of my Ph.D. She is a great companion.

I thank my thesis committee members Mathias Wegner and Arne Nolte. Mathias provided valuable feedback especially for the second chapter. Arne designed the system for library preparation of the DNA samples used in Chapter Four.

I thank Eva Stukenbrock for very kindly accepting to review this thesis. I thank Dietrich Ober for kindly accepting to be the chair of the talk of this study. I am thankful to the Max Planck Research School and its members for funding my Ph.D.

I thank my good friend Jens Frickel. Our conversations at the institute and also the weekends with him in Kiel were always great. I will regret not having of those moments more often.

I thank Gökce Ayan and -of course- Cagdas Sarikaya for their friendship - I had curious and fun conversations with them to say the least.
I thank my long time friend Ercan S. Unlu from Turkey. His decisions for his own research life guided me towards here.

I want to specially thank Yeliz Coskun. She gave me the confidence to move here. It was always comforting to be aware of her existence. I have visited her a lot during these years, and those times with her helped me to continue my Ph.D. smoothly.

Finally I would like to thank my mother, my father, and my brother Deniz. I was very luck to have such a family; all of them have deep interest in science, arts, and literature. That environment was crucially important for developing an interest in evolutionary biology.

References

Acinas, S.G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., and Polz, M.F. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Appl. Environ. Microbiol. *71*, 8966–8969.

Alexander, R.D., and Bigelow, R.S. (1960). Allochronic Speciation in Field Crickets, and a New Species, Acheta veletis. Evolution *14*, 334.

Amin, A.D., Chaix, A.B.H., Mason, R.P., Badge, R.M., and Borts, R.H. (2010). The Roles of the Saccharomyces cerevisiae RecQ Helicase SGS1 in Meiotic Genome Surveillance. PLoS ONE *5*, e15380.

Anderson, R.M., and May, R.M. (1982). Coevolution of hosts and parasites. Parasitology *85 (Pt 2)*, 411–426.

Anderson, C.M., Chen, S.Y., Dimon, M.T., Oke, A., DeRisi, J.L., and Fung, J.C. (2011). ReCombine: A Suite of Programs for Detection and Analysis of Meiotic Recombination in Whole-Genome Datasets. PLoS ONE *6*, e25509.

Anderson, J.B., Funt, J., Thompson, D.A., Prabhu, S., Socha, A., Sirjusingh, C., Dettman, J.R., Parreiras, L., Guttman, D.S., Regev, A., et al. (2010). Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. Curr. Biol. *20*, 1383–1388.

Avise, J.C. (2004). Molecular Markers, Natural History, and Evolution (Sunderland, Mass: Sinauer Associates).

Ayala, F.J., and Coluzzi, M. (2005). Chromosome speciation: humans, Drosophila, and mosquitoes. Proc. Natl. Acad. Sci. U.S.A. *102 Suppl 1*, 6535–6542.

Ayala, F.J., and Fitch, W.M. (1997). Genetics and the origin of species: an introduction. Proc. Natl. Acad. Sci. U.S.A. *94*, 7691–7697.

Baker, R.J., and Bickham, J.W. (1986). Speciation by monobrachial centric fusions. Proc Natl Acad Sci U S A *83*, 8245–8248.

Barbash, D.A., Siino, D.F., Tarone, A.M., and Roote, J. (2003). A rapidly evolving MYB-related protein causes species isolation in Drosophila. Proc. Natl. Acad. Sci. U.S.A. *100*, 5302–5307.

Barton, N.H. (1995). A general model for the evolution of recombination. Genet. Res. *65*, 123–145.

Barton, and Charlesworth, B. (1984). Genetic Revolutions, Founder Effects, and Speciation. Annual Review of Ecology and Systematics *15*, 133–164.

Barton, N., and Bengtsson, B.O. (1986). The barrier to genetic exchange between hybridising populations. Heredity (Edinb) *57 (Pt 3)*, 357–376.

Barton, N.H., and Charlesworth, B. (1998). Why sex and recombination? Science *281*, 1986–1990.

Barton, N.H., Briggs, D.E.G., Eisen, J.A., Goldstein, D.B., and Patel, N.H. (2007). Evolution (Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press).

Baruffini, E., Lodi, T., Dallabona, C., and Foury, F. (2007). A Single Nucleotide Polymorphism in the DNA Polymerase Gamma Gene of Saccharomyces cerevisiae Laboratory Strains Is Responsible for Increased Mitochondrial DNA Mutability. Genetics *177*, 1227–1231.

Bateson, W. (1909). Heredity and Variation in Modern Lights. In Darwin and Modern Science, (Cambridge University Press),.

Baudat, F., and de Massy, B. (2007). Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. Chromosome Res. *15*, 565–577.

Bengtsson, B.O., and Bodmer, W.F. (1976). On the increase of chromosome mutations under random mating. Theoretical Population Biology *9*, 260–281.

Bennett, A.F., and Lenski, R.E. (2007). An experimental test of evolutionary trade-offs during temperature adaptation. Proc. Natl. Acad. Sci. U.S.A. *104 Suppl 1*, 8649–8654.

Bergström, A., Simpson, J.T., Salinas, F., Barré, B., Parts, L., Zia, A., Nguyen Ba, A.N., Moses, A.M., Louis, E.J., Mustonen, V., et al. (2014). A High-Definition View of Functional Genetic Variation from Natural Yeast Genomes. Mol Biol Evol *31*, 872–888.

Bikard, D., Patel, D., Le Metté, C., Giorgi, V., Camilleri, C., Bennett, M.J., and Loudet, O. (2009). Divergent evolution of duplicate genes leads to genetic incompatibilities within A. thaliana. Science *323*, 623–626.

Bishop, D.K., and Zickler, D. (2004). Early decision; meiotic crossover interference prior to stable strand exchange and synapsis. Cell *117*, 9–15.

Blanc, H., and Dujon, B. (1980). Replicator regions of the yeast mitochondrial DNA responsible for suppressiveness. Proc. Natl. Acad. Sci. U.S.A. *77*, 3942–3946.

Boumil, R.M., Kemp, B., Angelichio, M., Nilsson-Tillgren, T., and Dawson, D.S. (2003). Meiotic segregation of a homeologous chromosome pair. Molecular Genetics and Genomics *268*, 750–760.

Bozdag, G.O., and Greig, D. (2014). The genetics of a putative social trait in natural populations of yeast. Molecular Ecology *23*, 5061–5071.

Brideau, N.J., Flores, H.A., Wang, J., Maheshwari, S., Wang, X., and Barbash, D.A. (2006). Two Dobzhansky-Muller Genes Interact to Cause Hybrid Lethality in Drosophila. Science *314*, 1292–1295.

Brown, J.D., and O'Neill, R.J. (2010). Chromosomes, conflict, and epigenetics: chromosomal speciation revisited. Annu Rev Genomics Hum Genet *11*, 291–316.

Brown, C.J., Todd, K.M., and Rosenzweig, R.F. (1998). Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. Molecular Biology and Evolution *15*, 931–942.

Bui, D.T., Dine, E., Anderson, J.B., Aquadro, C.F., and Alani, E.E. (2015). A Genetic Incompatibility Accelerates Adaptation in Yeast. PLoS Genet. *11*, e1005407.

Burt, A., and Trivers, R. (2006). Genes in Conflict: The Biology of Selfish Genetic Elements (Cambridge, Mass: Belknap Press of Harvard University Press).

Burton, R.S., Pereira, R.J., and Barreto, F.S. (2013). Cytonuclear Genomic Interactions and Hybrid Breakdown. Annual Review of Ecology, Evolution, and Systematics *44*, 281–302.

Cabot, E.L., Davis, A.W., Johnson, N.A., and Wu, C.I. (1994). Genetics of reproductive isolation in the Drosophila simulans clade: complex epistasis underlying hybrid male sterility. Genetics *137*, 175–189.

Castermans, D., Vermeesch, J.R., Fryns, J.-P., Steyaert, J.G., Van de Ven, W.J.M., Creemers, J.W.M., and Devriendt, K. (2007). Identification and characterization of the TRIP8 and REEP3 genes on chromosome 10q21.3 as novel candidate genes for autism. Eur. J. Hum. Genet. *15*, 422–431.

Chae, E., Bomblies, K., Kim, S.-T., Karelina, D., Zaidem, M., Ossowski, S., Martín-Pizarro, C., Laitinen, R.A.E., Rowan, B.A., Tenenboim, H., et al. (2014). Species-wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis. Cell *159*, 1341–1351.

Chambers, S.R., Hunter, N., Louis, E.J., and Borts, R.H. (1996). The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. Mol. Cell. Biol. *16*, 6110–6120.

Chang, S.-L., Lai, H.-Y., Tung, S.-Y., and Leu, J.-Y. (2013). Dynamic large-scale chromosomal rearrangements fuel rapid adaptation in yeast populations. PLoS Genet. *9*, e1003232.

Charlesworth, B. (2015). Correction for Charlesworth, Causes of natural variation in fitness: evidence from studies of Drosophila populations. Proc. Natl. Acad. Sci. U.S.A. *112*, E1049.

Charlesworth, D., and Charlesworth, B. (1979). Selection on Recombination in Clines. Genetics *91*, 581–589.

Charlesworth, D., Charlesworth, B., and Strobeck, C. (1977). Effects of selfing on selection for recombination. Genetics *86*, 213–226.

Charron, G., Leducq, J.-B., Bertin, C., Dubé, A.K., and Landry, C.R. (2014). Exploring the northern limit of the distribution of *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* in North America. FEMS Yeast Research *14*, 281–288.

Chen, W., and Jinks-Robertson, S. (1998). Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. Mol. Cell. Biol. *18*, 6525–6537.

Chen, C., Chen, H., Lin, Y.-S., Shen, J.-B., Shan, J.-X., Qi, P., Shi, M., Zhu, M.-Z., Huang, X.-H., Feng, Q., et al. (2014). A two-locus interaction causes interspecific hybrid weakness in rice. Nat Commun *5*, 3357.

Chen, S.Y., Tsubouchi, T., Rockmill, B., Sandler, J.S., Richards, D.R., Vader, G., Hochwagen, A., Roeder, G.S., and Fung, J.C. (2008). Global Analysis of the Meiotic Crossover Landscape. Developmental Cell *15*, 401–415.

Chou, J.-Y., and Leu, J.-Y. (2015). The Red Queen in mitochondria: cyto-nuclear co-evolution, hybrid breakdown and human disease. Frontiers in Genetics *6*.

Chow, E.W.L., Morrow, C.A., Djordjevic, J.T., Wood, I.A., and Fraser, J.A. (2012). Microevolution of Cryptococcus neoformans driven by massive tandem gene amplification.

Mol. Biol. Evol. *29*, 1987–2000.

Cliften, P.F., Hillier, L.W., Fulton, L., Graves, T., Miner, T., Gish, W.R., Waterston, R.H., and Johnston, M. (2001). Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis. Genome Res. *11*, 1175–1186.

Clowers, K.J., Heilberger, J., Piotrowski, J.S., Will, J.L., and Gasch, A.P. (2015). Ecological and genetic barriers differentiate natural populations of Saccharomyces cerevisiae. Mol Biol Evol msv112.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D., and Kingsley, D.M. (2005). Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. Science *307*, 1928–1933.

Coluccio, A.E., Rodriguez, R.K., Kernan, M.J., and Neiman, A.M. (2008). The Yeast Spore Wall Enables Spores to Survive Passage through the Digestive Tract of Drosophila. PLoS ONE *3*.

Corbett-Detig, R.B., Zhou, J., Clark, A.G., Hartl, D.L., and Ayroles, J.F. (2013). Genetic incompatibilities are widespread within species. Nature *504*, 135–137.

Coyne, J.A. (1992). Genetics and speciation. Nature *355*, 511–515.

Coyne, J.A., and Orr, H.A. (1989). Patterns of Speciation in Drosophila. Evolution *43*, 362–381.

Coyne, J.A., and Orr, H.A. (1998). The evolutionary genetics of speciation. Philos Trans R Soc Lond B Biol Sci *353*, 287–305.

Coyne, J.A., and Orr, H.A. (2004). Speciation (Sunderland, Mass: Sinauer Associates, Inc.).

Coyne, J.A., Crittenden, A.P., and Mah, K. (1994). Genetics of a pheromonal difference contributing to reproductive isolation in Drosophila. Science *265*, 1461–1464.

Datta, A., Hendrix, M., Lipsitch, M., and Jinks-Robertson, S. (1997). Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proceedings of the National Academy of Sciences *94*, 9757–9762.

Daugherty, M.D., and Malik, H.S. (2012). Rules of Engagement: Molecular Insights from Host-Virus Arms Races. Annual Review of Genetics *46*, 677–700.

Davis, L., and Smith, G.R. (2003). Nonrandom homolog segregation at meiosis I in Schizosaccharomyces pombe mutants lacking recombination. Genetics *163*, 857–874.

Dawkins, R. (1978). The Selfish Gene (New York: Oxford Univ Pr).

De Deken, R.H. (1966). The Crabtree effects and its relation to the petite mutation. J. Gen. Microbiol. *44*, 157–165.

De Grassi, A., Lanave, C., and Saccone, C. (2008). Genome duplication and gene-family evolution: the case of three OXPHOS gene families. Gene *421*, 1–6.

Delneri, D., Colson, I., Grammenoudi, S., Roberts, I.N., Louis, E.J., and Oliver, S.G. (2003). Engineering evolution to study speciation in yeasts. Nature *422*, 68–72.

De Muyt, A., Jessop, L., Kolar, E., Sourirajan, A., Chen, J., Dayani, Y., and Lichten, M. (2012). BLM Helicase Ortholog Sgs1 Is a Central Regulator of Meiotic Recombination Intermediate Metabolism. Molecular Cell *46*, 43–53.

Dettman, J.R., Anderson, J.B., and Kohn, L.M. (2008). Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus Neurospora. BMC Evolutionary Biology *8*, 35.

Dion-Cote, A.-M., Symonova, R., Rab, P., and Bernatchez, L. (2015). Reproductive isolation in a nascent species pair is associated with aneuploidy in hybrid offspring. Proceedings of the Royal Society B: Biological Sciences *282*, 20142862–20142862.

Dobzhansky, T. (1949). Observations and Experiments on Natural Selection in Drosophila. Hereditas *35*, 210–224.

Dobzhansky, T. (1982). Genetics and the Origin of Species (New York: Columbia University Press).

Dodd, D.M. (1989). Reproductive isolation as a consequence of adaptive divergence in Drosophila pseudoobscura. Evolution 1308–1311.

Doniger, S.W., Kim, H.S., Swain, D., Corcuera, D., Williams, M., Yang, S.-P., and Fay, J.C. (2008). A Catalog of Neutral and Deleterious Polymorphism in Yeast. PLoS Genetics *4*, e1000183.

Dujon, B. (2010). Yeast evolutionary genomics. Nat. Rev. Genet. *11*, 512–524.

Dunham, M.J., Badrane, H., Ferea, T., Adams, J., Brown, P.O., Rosenzweig, F., and Botstein, D. (2002). Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A. *99*, 16144–16149.

Dunn, B., Paulish, T., Stanbery, A., Piotrowski, J., Koniges, G., Kroll, E., Louis, E.J., Liti, G., Sherlock, G., and Rosenzweig, F. (2013). Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. PLoS Genet. *9*, e1003366.

Egan, S.P., and Funk, D.J. (2009). Ecologically dependent postmating isolation between sympatric host forms of Neochlamisus bebbianae leaf beetles. Proceedings of the National Academy of Sciences *106*, 19426–19431.

Eimes, J.A., Bollmer, J.L., Whittingham, L.A., Johnson, J.A., VAN Oosterhout, C., and Dunn, P.O. (2011). Rapid loss of MHC class II variation in a bottlenecked population is explained by drift and loss of copy number variation. J. Evol. Biol. *24*, 1847–1856.

Eldar, A. (2011). Social conflict drives the evolutionary divergence of quorum sensing. PNAS *108*, 13635–13640.

Etter, P.D., Bassham, S., Hohenlohe, P.A., Johnson, E.A., and Cresko, W.A. (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. Methods Mol. Biol. *772*, 157–178.

Ewald, P.W. (1987). Transmission Modes and Evolution of the Parasitism-Mutualism Continuuma. Annals of the New York Academy of Sciences *503*, 295–306.

Eyre-Walker, A., and Keightley, P.D. (2007). The distribution of fitness effects of new mutations. Nat. Rev. Genet. *8*, 610–618.

Faria, R., and Navarro, A. (2010). Chromosomal speciation revisited: rearranging theory with pieces of evidence. Trends Ecol. Evol. (Amst.) *25*, 660–669.

Felsenstein, J. (1974). The evolutionary advantage of recombination. Genetics *78*, 737–756.

Ferree, P.M., and Barbash, D.A. (2009). Species-Specific Heterochromatin Prevents Mitotic Chromosome Segregation to Cause Hybrid Lethality in Drosophila. PLoS Biology *7*, e1000234.

Feuk, L., MacDonald, J.R., Tang, T., Carson, A.R., Li, M., Rao, G., Khaja, R., and Scherer, S.W. (2005). Discovery of human inversion polymorphisms by comparative analysis of human and chimpanzee DNA sequence assemblies. PLoS Genet. *1*, e56.

Fine, P.E.M. (1975). Vectors and Vertical Transmission: An Epidemiologic Perspective. Annals of the New York Academy of Sciences *266*, 173–194.

Fischer, G., James, S.A., Roberts, I.N., Oliver, S.G., and Louis, E.J. (2000). Chromosomal evolution in Saccharomyces. Nature *405*, 451–454.

Fledel-Alon, A., Wilson, D.J., Broman, K., Wen, X., Ober, C., Coop, G., and Przeworski, M. (2009). Broad-Scale Recombination Patterns Underlying Proper Disjunction in Humans. PLoS Genetics *5*, e1000658.

Fogel, S., and Welch, J.W. (1982). Tandem gene amplification mediates copper resistance in yeast. Proceedings of the National Academy of Sciences *79*, 5342–5346.

Frank, S.A. (1996). Models of parasite virulence. Q Rev Biol *71*, 37–78.

Fraser, C., Hanage, W.P., and Spratt, B.G. (2007). Recombination and the nature of bacterial speciation. Science *315*, 476–480.

Fry, J.D. (2009). Laboratory Experiments on Speciation (Oakland: Univ California Press).

Fukuda, T., Nogami, S., and Ohya, Y. (2003). VDE-initiated intein homing in Saccharomyces cerevisiae proceeds in a meiotic recombination-like manner. Genes to Cells *8*, 587–602.

Fuller, Z.L., Haynes, G.D., Zhu, D., Batterton, M., Chao, H., Dugan, S., Javaid, M., Jayaseelan, J.C., Lee, S., Li, M., et al. (2014). Evidence for stabilizing selection on codon usage in chromosomal rearrangements of Drosophila pseudoobscura. G3 (Bethesda) *4*, 2433–2449.

Futcher, B., Reid, E., and Hickey, D.A. (1988). Maintenance of the 2 micron circle plasmid of Saccharomyces cerevisiae by sexual transmission: an example of a selfish DNA. Genetics *118*, 411–415.

Gerrish, P.J., and Lenski, R.E. (1998). The fate of competing beneficial mutations in an asexual population. Genetica *102-103*, 127–144.

Gerton, J.L., and Hawley, R.S. (2005). Homologous chromosome interactions in meiosis: diversity amidst conservation. Nat Rev Genet *6*, 477–487.

Gietz, R.D., and Schiestl, R.H. (2007). Large-scale high-efficiency yeast transformation using

the LiAc/SS carrier DNA/PEG method. Nature Protocols *2*, 38–41.

Goddard, M.R., and Burt, A. (1999). Recurrent invasion and extinction of a selfish gene. Proc. Natl. Acad. Sci. U.S.A. *96*, 13880–13885.

Goddard, M.R., Greig, D., and Burt, A. (2001). Outcrossed sex allows a selfish gene to invade yeast populations. Proc. Biol. Sci. *268*, 2537–2542.

Godfrey-Smith, P. (2009). Darwinian Populations and Natural Selection (Oxford ; New York: Oxford University Press).

Goldring, E.S., Grossman, L.I., and Marmur, J. (1971). Petite Mutation in Yeast. J Bacteriol *107*, 377–381.

Graham, B. (2008). Selection: The Mechanism of Evolution (Oxford ; New York: Oxford University Press).

Grandin, N., and Reed, S.I. (1993). Differential function and expression of Saccharomyces cerevisiae B-type cyclins in mitosis and meiosis. Mol. Cell. Biol. *13*, 2113–2125.

Greig, D. (2007). A screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast. PLoS Genet. *3*, e21.

Greig, D. (2009). Reproductive isolation in Saccharomyces. Heredity *102*, 39–44.

Greig, D., and Travisano, M. (2003). Haploid Superiority. Science *299*, 524–525.

Greig, D., Borts, R.H., Louis, E.J., and Travisano, m. (2002). Epistasis and hybrid sterility in Saccharomyces. Proceedings of the Royal Society B: Biological Sciences *269*, 1167–1171.

Greiner, S., Rauwolf, U., Meurer, J., and Herrmann, R.G. (2011). The role of plastids in plant speciation: PLASTIDS AND SPECIATION. Molecular Ecology *20*, 671–691.

Grosberg, R.K., and Strathmann, R.R. (2007). The Evolution of Multicellularity: A Minor Major Transition? Annual Review of Ecology, Evolution, and Systematics *38*, 621–654.

Haldane, J.B.S. (1927). A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation. Mathematical Proceedings of the Cambridge Philosophical Society *23*, 838–844.

Hall, D.W., and Joseph, S.B. (2010). A High Frequency of Beneficial Mutations Across Multiple Fitness Components in Saccharomyces cerevisiae. Genetics *185*, 1397–1409.

Harrison, R.G. (1993). Hybrid Zones and the Evolutionary Process (Oxford University Press).

Harrison, E., Koufopanou, V., Burt, A., and MacLean, R.C. (2012). The cost of copy number in a selfish genetic element: the 2-μm plasmid of Saccharomyces cerevisiae. J. Evol. Biol. *25*, 2348–2356.

Harrison, E., MacLean, R.C., Koufopanou, V., and Burt, A. (2014). Sex drives intracellular conflict in yeast. Journal of Evolutionary Biology *27*, 1757–1763.

Hastings (1999). The costs of sex due to deleterious intracellular parasites. Journal of Evolutionary Biology *12*, 177–183.

Hauffe, H.C., Giménez, M.D., Vega, R., White, T.A., and Searle, J.B. (2011). Properties of a hybrid zone between highly distinct chromosomal races of the house mouse (Mus musculus domesticus) in Northern Italy, and comparisons with other hybrid zones. Cytogenet. Genome Res. *134*, 191–199.

Heck, J.A., Argueso, J.L., Gemici, Z., Reeves, R.G., Bernard, A., Aquadro, C.F., and Alani, E. (2006). Negative epistasis between natural variants of the Saccharomyces cerevisiae MLH1 and PMS1 genes results in a defect in mismatch repair. Proc. Natl. Acad. Sci. U.S.A. *103*, 3256–3261.

Hedrick, P.W. (1981). The Establishment of Chromosomal Variants. Evolution *35*, 322–332.

Hedrick, P.W. (1995). Elephant seals and the estimation of a population bottleneck. J. Hered. *86*, 232–235.

Herre, E.A. (1993). Population structure and the evolution of virulence in nematode parasites of fig wasps. Science *259*, 1442–1445.

Herron, J.C., and Freeman, S. (2013). Evolutionary Analysis (San Francisco, CA: Benjamin Cummings).

Herron, M.D., and Doebeli, M. (2013). Parallel evolutionary dynamics of adaptive diversification in Escherichia coli. PLoS Biol. *11*, e1001490.

Hillers, K.J., and Villeneuve, A.M. (2003). Chromosome-wide control of meiotic crossing over in C. elegans. Current Biology *13*, 1641–1647.

Hittinger, C.T., and Carroll, S.B. (2007). Gene duplication and the adaptive evolution of a classic genetic switch. Nature *449*, 677–681.

Höbel, G., and Gerhardt, H.C. (2003). Reproductive character displacement in the acoustic communication system of green tree frogs (Hyla cinerea). Evolution *57*, 894–904.

Hoffmann, A.A., and Rieseberg, L.H. (2008). Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? Annual Review of Ecology, Evolution, and Systematics *39*, 21–42.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., and Cresko, W.A. (2010). Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. PLoS Genet *6*, e1000862.

Hopper, A.K., and Hall, B.D. (1975). Mating Type and Sporulation in Yeast I. Mutations Which Alter Mating-Type Control over Sporulation. Genetics *80*, 41–59.

Hou, J., and Schacherer, J. (2015). Negative epistasis: a route to intraspecific reproductive isolation in yeast? Curr. Genet.

Hou, J., Friedrich, A., de Montigny, J., and Schacherer, J. (2014). Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in Saccharomyces cerevisiae. Curr. Biol. *24*, 1153–1159.

Hou, J., Friedrich, A., Gounot, J.-S., and Schacherer, J. (2015). Comprehensive survey of condition-specific reproductive isolation reveals genetic incompatibility in yeast. Nat Commun *6*, 7214.

Hughes, J.B., Daily, G.C., and Ehrlich, P.R. (1997). Population Diversity: Its Extent and Extinction. Science *278*, 689–692.

Hunter, N., Chambers, S.R., Louis, E.J., and Borts, R.H. (1996). The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. The EMBO Journal *15*, 1726.

Hurles, M. (2004). Gene Duplication: The Genomic Trade in Spare Parts. PLoS Biol *2*.

Hurst, L.D. (1992). Intragenomic Conflict as an Evolutionary Force. Proceedings of the Royal Society of London B: Biological Sciences *248*, 135–140.

Hurst, G.D., and Werren, J.H. (2001). The role of selfish genetic elements in eukaryotic evolution. Nat. Rev. Genet. *2*, 597–606.

Jambhekar, A., and Amon, A. (2008). Control of meiosis by respiration. Curr. Biol. *18*, 969–975.

Jennings, J.H., Snook, R.R., and Hoikkala, A. (2014). Reproductive isolation among allopatric Drosophila montana populations. Evolution *68*, 3095–3108.

Jessop, L., Rockmill, B., Roeder, G.S., and Lichten, M. (2006). Meiotic Chromosome Synapsis-Promoting Proteins Antagonize the Anti-Crossover Activity of Sgs1. PLoS Genetics *2*, e155.

Johnson, N.A. (2010). Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet. *26*, 317–325.

Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., Whibley, A., Becuwe, M., Baxter, S.W., Ferguson, L., et al. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature *477*, 203–206.

Kachroo, A.H., Laurent, J.M., Yellman, C.M., Meyer, A.G., Wilke, C.O., and Marcotte, E.M. (2015). Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science *348*, 921–925.

Kao, K.C., Schwartz, K., and Sherlock, G. (2010). A Genome-Wide Analysis Reveals No Nuclear Dobzhansky-Muller Pairs of Determinants of Speciation between S. cerevisiae and S. paradoxus, but Suggests More Complex Incompatibilities. PLoS Genetics *6*, e1001038.

Kaplan, J.M., and Pigliucci, M. (2001). Genes `for' Phenotypes: A Modern History View. Biology & Philosophy *16*, 189–213.

Kauppi, L., Jeffreys, A.J., and Keeney, S. (2004). Where the crossovers are: recombination distributions in mammals. Nat. Rev. Genet. *5*, 413–424.

Kelleher, E.S., Swanson, W.J., and Markow, T.A. (2007). Gene Duplication and Adaptive Evolution of Digestive Proteases in Drosophila arizonae Female Reproductive Tracts. PLoS Genet *3*.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E.S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature *423*, 241–254.

Kirkpatrick, M. (2010). How and Why Chromosome Inversions Evolve. PLoS Biol *8*.

Kirkpatrick, M., and Barton, N. (2006). Chromosome Inversions, Local Adaptation and

Speciation. Genetics *173*, 419–434.

Kleene, K.C. (2005). Sexual selection, genetic conflict, selfish genes, and the atypical patterns of gene expression in spermatogenic cells. Developmental Biology *277*, 16–26.

Knoll, A.., Javaux, E.., Hewitt, D., and Cohen, P. (2006). Eukaryotic organisms in Proterozoic oceans. Philosophical Transactions of the Royal Society B: Biological Sciences *361*, 1023–1038.

Knowlton, N., Maté, J.L., Guzmán, H.M., Rowan, R., and Jara, J. (1997). Direct evidence for reproductive isolation among the three species of the Montastraea annularis complex in Central America (Panamá and Honduras). Marine Biology *127*, 705–711.

Kohn, L.M., and Anderson, J.B. (2014). The underlying structure of adaptation under strong selection in 12 experimental yeast populations. Eukaryotic Cell *13*, 1200–1206.

Kondrashov, F.A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. Proc. Biol. Sci. *279*, 5048–5057.

Kondrashov, F.A., and Kondrashov, A.S. (2006). Role of selection in fixation of gene duplications. J. Theor. Biol. *239*, 141–151.

Kover, P.X., and Clay, K. (1998). Trade-off between Virulence and Vertical Transmission and the Maintenance of a Virulent Plant Pathogen. The American Naturalist *152*, 165–175.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645.

Kvitek, D.J., and Sherlock, G. (2011). Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genet. *7*, e1002056.

Laayouni, H., García-Franco, F., Chávez-Sandoval, B.E., Trotta, V., Beltran, S., Corominas, M., and Santos, M. (2007). Thermal evolution of gene expression profiles in Drosophila subobscura. BMC Evol Biol *7*, 1–15.

Lancashire, W.E., and Mattoon, J.R. (1979). Cytoduction: a tool for mitochondrial genetic studies in yeast. Molecular and General Genetics MGG *170*, 333–344.

Lande, R. (1979). Effective Deme Sizes During Long-Term Evolution Estimated from Rates of Chromosomal Rearrangement. Evolution *33*, 234–251.

Lawniczak, M.K.N., and Begun, D.J. (2007). Molecular population genetics of female-expressed mating-induced serine proteases in Drosophila melanogaster. Mol. Biol. Evol. *24*, 1944–1951.

Lederberg, J., and Lederberg, E.M. (1952). REPLICA PLATING AND INDIRECT SELECTION OF BACTERIAL MUTANTS. J Bacteriol *63*, 399–406.

Leducq, J.-B., Charron, G., Diss, G., Gagnon-Arsenault, I., Dubé, A.K., and Landry, C.R. (2012). Evidence for the robustness of protein complexes to inter-species hybridization. PLoS Genet. *8*, e1003161.

Lee, B.H., and Amon, A. (2003). Role of Polo-like kinase CDC5 in programming meiosis I chromosome segregation. Science *300*, 482–486.

Lenski, R.E., Rose, M.R., Simpson, S.C., and Tadler, S.C. (1991). Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. American Naturalist 1315–1341.

Li, H., and Roossinck, M.J. (2004). Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. J. Virol. *78*, 10582–10587.

Li, C., Wang, Z., and Zhang, J. (2013). Toward Genome-Wide Identification of Bateson-Dobzhansky-Muller Incompatibilities in Yeast: A Simulation Study. Genome Biology and Evolution *5*, 1261–1272.

Li, Y.-D., Liang, H., Gu, Z., Lin, Z., Guan, W., Zhou, L., Li, Y.-Q., and Li, W.-H. (2009). Detecting positive selection in the budding yeast genome. J. Evol. Biol. *22*, 2430–2437.

Lipsitch, M., Siller, S., and Nowak, M.A. (1996). The Evolution of Virulence in Pathogens with Vertical and Horizontal Transmission. Evolution *50*, 1729–1741.

Liti, G., Barton, D.B.H., and Louis, E.J. (2006). Sequence Diversity, Reproductive Isolation and Species Concepts in Saccharomyces. Genetics *174*, 839–850.

Lowry, D.B., and Willis, J.H. (2010). A Widespread Chromosomal Inversion Polymorphism Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. PLoS Biology *8*, e1000500.

Lowry, D.B., Modliszewski, J.L., Wright, K.M., Wu, C.A., and Willis, J.H. (2008). The strength and genetic basis of reproductive isolating barriers in flowering plants. Philosophical Transactions of the Royal Society of London B: Biological Sciences *363*, 3009–3021.

Luria, S.E., and Delbrück, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics *28*, 491–511.

Lynch, M. (2010). Evolution of the mutation rate. Trends Genet. *26*, 345–352.

MacAlpine, D.M., Kolesar, J., Okamoto, K., Butow, R.A., and Perlman, P.S. (2001). Replication and preferential inheritance of hypersuppressive petite mitochondrial DNA. EMBO J *20*, 1807–1817.

Maclean, C.J., and Greig, D. (2011). RECIPROCAL GENE LOSS FOLLOWING EXPERIMENTAL WHOLE-GENOME DUPLICATION CAUSES REPRODUCTIVE ISOLATION IN YEAST: DIVERGENT RESOLUTION IN ALLOTETRAPLOIDS. Evolution *65*, 932–945.

Magalon, H., Nidelet, T., Martin, G., and Kaltz, O. (2010). Host Growth Conditions Influence Experimental Evolution of Life History and Virulence of a Parasite with Vertical and Horizontal Transmission. Evolution *64*, 2126–2138.

Maheshwari, S., and Barbash, D.A. (2011). The Genetics of Hybrid Incompatibilities. Annual Review of Genetics *45*, 331–355.

Mallet, J. (2006). What does Drosophila genetics tell us about speciation? Trends in Ecology & Evolution *21*, 386–393.

Mancera, E., Bourgon, R., Brozzi, A., Huber, W., and Steinmetz, L.M. (2008). High-resolution mapping of meiotic crossovers and non-crossovers in yeast. Nature *454*, 479–485.

Margulis, L. (1971). Origin of Eukaryotic Cells (New Haven: Yale University Press).

Martini, E., Borde, V., Legendre, M., Audic, S., Regnault, B., Soubigou, G., Dujon, B., and Llorente, B. (2011). Genome-Wide Analysis of Heteroduplex DNA in Mismatch Repair–Deficient Yeast Cells Reveals Novel Properties of Meiotic Recombination Pathways. PLoS Genetics *7*, e1002305.

Matic, I., Rayssiguier, C., and Radman, M. (1995). Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. Cell *80*, 507–515.

Mayr, E. (1963). Animal Species and Evolution (Cambridge, MA: Belknap Press).

Mayr, E. (1970). Populations, Species, and Evolution: An Abridgment of *Animal Species and Evolution* (Belknap Press).

McKee, B.D. (2004). Homologous pairing and chromosome dynamics in meiosis and mitosis. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression *1677*, 165–180.

Michod, R.E. (2000). Darwinian Dynamics (Princeton University Press).

Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J.C., and Forejt, J. (2009). A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science *323*, 373–375.

Mitri, S., and Foster, K.R. (2013). The genotypic view of social interactions in microbial communities. Annu. Rev. Genet. *47*, 247–273.

Monaghan, F., and Corcos, A. (1984). On the origins of the Mendelian laws. J. Hered. *75*, 67–69.

Moyle, L.C., and Nakazato, T. (2010). Hybrid Incompatibility "Snowballs" Between Solanum Species. Science *329*, 1521–1523.

Muller H. J. (1942). Isolating mechanisms, evolution and temperature. Biol. Symp. *6*, 71–125.

Murphy, H.A., and Zeyl, C.W. (2010). Yeast sex: surprisingly high rates of outcrossing between asci. PLoS ONE *5*, e10461.

Murphy, H.A., Kuehne, H.A., Francis, C.A., and Sniegowski, P.D. (2006). Mate choice assays and mating propensity differences in natural yeast populations. Biol. Lett. *2*, 553–556.

Myung (a), K., Datta, A., Chen, C., and Kolodner, R.D. (2001b). SGS1, the Saccharomyces cerevisiae homologue of BLM and WRN, suppresses genome instability and homeologous recombination. Nature Genetics *27*, 113–116.

Myung (b), K., Chen, C., and Kolodner, R.D. (2001a). Multiple pathways cooperate in the suppression of genome instability in Saccharomyces cerevisiae. Nature *411*, 1073–1076.

Nasmyth, K. (1987). The determination of mother cell-specific mating type of switching in yeast by a specific regulator of HO transcription. The EMBO Journal *6*, 243.

Naumov, G. (1996). Genetic identification of biological species in theSaccharomyces sensu stricto complex. Journal of Industrial Microbiology & Biotechnology *17*, 295–302.

Navarro, A., and Barton, N.H. (2003). Chromosomal Speciation and Molecular Divergence--Accelerated Evolution in Rearranged Chromosomes. Science *300*, 321–324.

Nei, M., and Nozawa, M. (2011). Roles of Mutation and Selection in Speciation: From Hugo de Vries to the Modern Genomic Era. Genome Biology and Evolution *3*, 812–829.

Neiman, A.M. (2011). Sporulation in the budding yeast Saccharomyces cerevisiae. Genetics *189*, 737–765.

Nogami, S., Fukuda, T., Nagai, Y., Yabe, S., Sugiura, M., Mizutani, R., Satow, Y., Anraku, Y., and Ohya, Y. (2002). Homing at an extragenic locus mediated by VDE (PI-SceI) in Saccharomyces cerevisiae. Yeast *19*, 773–782.

Noor, M.A.F., and Feder, J.L. (2006). Speciation genetics: evolving approaches. Nature Reviews Genetics *7*, 851–861.

Noor, M.A., Grams, K.L., Bertucci, L.A., and Reiland, J. (2001). Chromosomal inversions and the reproductive isolation of species. Proceedings of the National Academy of Sciences *98*, 12084–12088.

Nunnari, J., Marshall, W.F., Straight, A., Murray, A., Sedat, J.W., and Walter, P. (1997). Mitochondrial transmission during mating in Saccharomyces cerevisiae is determined by mitochondrial fusion and fission and the intramitochondrial segregation of mitochondrial DNA. Mol. Biol. Cell *8*, 1233–1242.

Oh, S.D., Lao, J.P., Taylor, A.F., Smith, G.R., and Hunter, N. (2008). RecQ Helicase, Sgs1, and XPF Family Endonuclease, Mus81-Mms4, Resolve Aberrant Joint Molecules during Meiotic Recombination. Molecular Cell *31*, 324–336.

Orgel, L.E., and Crick, F.H.C. (1980). Selfish DNA: the ultimate parasite. Nature *284*, 604–607.

Orr, H.A. (1995). The Population Genetics of Speciation: The Evolution of Hybrid Incompatibilities. Genetics *139*, 1805–1813.

Orr, H.A. (1996). Dobzhansky, Bateson, and the Genetics of Speciation. Genetics *144*, 1331–1335.

Orr, H.A. (2009). Fitness and its role in evolutionary genetics. Nat Rev Genet *10*, 531–539.

Orr, H.A., and Turelli, M. (2001). The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution *55*, 1085–1094.

Otto, S.P., and Barton, N.H. (1997). The Evolution of Recombination: Removing the Limits to Natural Selection. Genetics *147*, 879–906.

Otto, S.P., and Lenormand, T. (2002). Resolving the paradox of sex and recombination. Nat. Rev. Genet. *3*, 252–261.

Otto, S.P., and Whitlock, M.C. (2001). Fixation Probabilities and Times. In eLS, (John Wiley & Sons, Ltd),.

Pagán, I., Montes, N., Milgroom, M.G., and García-Arenal, F. (2014). Vertical Transmission Selects for Reduced Virulence in a Plant Virus and for Increased Resistance in the Host.

PLoS Pathogens *10*, e1004293.

Paixão, T., Bassler, K.E., and Azevedo, R.B.R. (2014). Emergent speciation by multiple Dobzhansky-Muller incompatibilities. bioRxiv 008268.

Paliwal, S., Fiumera, A.C., and Fiumera, H.L. (2014). Mitochondrial-Nuclear Epistasis Contributes to Phenotypic Variation and Coadaptation in Natural Isolates of Saccharomyces cerevisiae. Genetics *198*, 1251–1265.

Palopoli, M.F., and Wu, C.I. (1994). Genetics of hybrid male sterility between drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics *138*, 329–341.

Papp, B., Pál, C., and Hurst, L.D. (2003). Evolution of cis-regulatory elements in duplicated genes of yeast. Trends Genet. *19*, 417–422.

Pascarella, J.B. (2007). Mechanisms of prezygotic reproductive isolation between two sympatric species, Gelsemium rankinii and G. sempervirens (Gelsemiaceae), in the southeastern United States. Am. J. Bot. *94*, 468–476.

Pasyukova, E.G., Nuzhdin, S.V., Morozova, T.V., and Mackay, T.F.C. (2004). Accumulation of transposable elements in the genome of Drosophila melanogaster is associated with a decrease in fitness. J. Hered. *95*, 284–290.

Pérez-Ortín, J.E., Querol, A., Puig, S., and Barrio, E. (2002). Molecular Characterization of a Chromosomal Rearrangement Involved in the Adaptive Evolution of Yeast Strains. Genome Res *12*, 1533–1539.

Peris, J.B., Davis, P., Cuevas, J.M., Nebot, M.R., and Sanjuán, R. (2010). Distribution of fitness effects caused by single-nucleotide substitutions in bacteriophage f1. Genetics *185*, 603–609.

Petersen, R.F., Langkjær, R.B., Hvidtfeldt, J., Gartner, J., Palmen, W., Ussery, D.W., and Piškur, J. (2002). Inheritance and organisation of the mitochondrial genome differ between two Saccharomyces yeasts. Journal of Molecular Biology *318*, 627–636.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLoS ONE *7*, e37135.

Phadnis, N., and Orr, H.A. (2009). A Single Gene Causes Both Male Sterility and Segregation Distortion in Drosophila Hybrids. Science *323*, 376–379.

Pigliucci, M. (2008). Is evolvability evolvable? Nat Rev Genet *9*, 75–82.

Platt, A., Horton, M., Huang, Y.S., Li, Y., Anastasio, A.E., Mulyati, N.W., Agren, J., Bossdorf, O., Byers, D., Donohue, K., et al. (2010). The scale of population structure in Arabidopsis thaliana. PLoS Genet. *6*, e1000843.

Presgraves, D.C. (2010). Speciation Genetics: Search for the Missing Snowball. Current Biology *20*, R1073–R1074.

Presgraves, D.C., Balagopalan, L., Abmayr, S.M., and Orr, H.A. (2003). Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature *423*,

715–719.

Preston, B.D., Albertson, T.M., and Herr, A.J. (2010). DNA replication fidelity and cancer. Semin. Cancer Biol. *20*, 281–293.

Procházka, E., Franko, F., Poláková, S., and Sulo, P. (2012). A complete sequence of Saccharomyces paradoxus mitochondrial genome that restores the respiration in S. cerevisiae. FEMS Yeast Res. *12*, 819–830.

Ramsey, J., and Schemske, D.W. (2002). Neopolyploidy in Flowering Plants. Annual Review of Ecology and Systematics *33*, 589–639.

Ramsey, J., Bradshaw, H.D., and Schemske, D.W. (2003). Components of reproductive isolation between the monkeyflowers Mimulus lewisii and M. cardinalis (Phrymaceae). Evolution *57*, 1520–1534.

Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R., Walton, K., Perera, A., Staehling-Hampton, K., Seidel, C.W., and Li, R. (2008). Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. Cell *135*, 879–893.

Rattray, A., Santoyo, G., Shafer, B., and Strathern, J.N. (2015). Elevated mutation rate during meiosis in Saccharomyces cerevisiae. PLoS Genet. *11*, e1004910.

Rayssiguier, C., Thaler, D.S., and Radman, M. (1989). The barrier to recombination between Escherichia coli and Salmonella typhimurium is disrupted in mismatch-repair mutants. Nature *342*, 396–401.

Rieseberg, L.H. (2001). Chromosomal rearrangements and speciation. Trends in Ecology & Evolution *16*, 351–358.

Rieseberg, L.H., and Livingstone, K. (2003). Chromosomal speciation in primates. Science *300*, 267–268.

Rieseberg, L.H., and Willis, J.H. (2007). Plant speciation. Science *317*, 910–914.

Roeder, G.S. (1997). Meiotic chromosomes: it takes two to tango. Genes Dev. *11*, 2600–2621.

Rogers, D.W., Denton, J.A., McConnell, E., and Greig, D. (2015). Experimental Evolution of Species Recognition. Current Biology *25*, 1753–1758.

Rutter, M.T., Shaw, F.H., and Fenster, C.B. (2010). Spontaneous mutation parameters for Arabidopsis thaliana measured in the wild. Evolution *64*, 1825–1835.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature *409*, 928–933.

Sætre, G.-P., Moum, T., Bureš, S., Král, M., Adamjan, M., and Moreno, J. (1997). A sexually selected character displacement in flycatchers reinforces premating isolation. Nature *387*, 589–592.

Sawamura, K., and Yamamoto, M.-T. (1997). Heredity - Abstract of article: Characterization

of a reproductive isolation gene, zygotic hybrid rescue, of Drosophila melanogaster by using minichromosomes. Heredity *79*, 97–103.

Scannell, D.R., Zill, O.A., Rokas, A., Payen, C., Dunham, M.J., Eisen, M.B., Rine, J., Johnston, M., and Hittinger, C.T. (2011). The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the Saccharomyces sensu stricto Genus. G3 (Bethesda) *1*, 11–25.

Schluter, D., and Conte, G.L. (2009). Genetics and ecological speciation. PNAS *106*, 9955–9962.

Schumer, M., Cui, R., Rosenthal, G.G., and Andolfatto, P. (2015). Reproductive isolation of hybrid populations driven by genetic incompatibilities. PLoS Genet. *11*, e1005041.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G.-P., Bank, C., Brännström, Å., et al. (2014). Genomics and the origin of species. Nature Reviews Genetics *15*, 176–192.

Seidel, H.S., Rockman, M.V., and Kruglyak, L. (2008). Widespread genetic incompatibility in C. elegans maintained by balancing selection. Science *319*, 589–594.

Selva, E.M., New, L., Crouse, G.F., and Lahue, R.S. (1995). Mismatch correction acts as a barrier to homeologous recombination in Saccharomyces cerevisiae. Genetics *139*, 1175–1188.

Sgrò, C.M., and Hoffmann, A.A. (2004). Genetic correlations, tradeoffs and environmental variation. Heredity (Edinb) *93*, 241–248.

Shen, P., and Huang, H.V. (1989). Effect of base pair mismatches on recombination via the RecBCD pathway. Molecular and General Genetics MGG *218*, 358–360.

Sherman, F. (1991). Getting started with yeast. Methods Enzymol *194*, 21.

Shirley, M.D.F., and Sibly, R.M. (1999). Genetic Basis of a between-Environment Trade-off Involving Resistance to Cadmium in Drosophila melanogaster. Evolution *53*, 826–836.

van der Sluijs, I., Van Dooren, T.J.M., Hofker, K.D., van Alphen, J.J.M., Stelkens, R.B., and Seehausen, O. (2008). Female mating preference functions predict sexual selection against hybrids between sibling species of cichlid fish. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *363*, 2871–2877.

Smith, O.D., Kleese, R.A., and Stuthman, D.D. (1970). Competition among oat varieties grown in hill plots. Crop Science.

Sniegowski, P. (1998). Mismatch repair: origin of species? Curr. Biol. *8*, R59–R61.

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

Sobel, J.M., Chen, G.F., Watt, L.R., and Schemske, D.W. (2010). The biology of speciation. Evolution *64*, 295–315.

Spang, A., Saw, J.H., Jørgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., and Ettema, T.J.G. (2015). Complex archaea that bridge

the gap between prokaryotes and eukaryotes. Nature *521*, 173–179.

Spell, R.M. (2004). Examination of the Roles of Sgs1 and Srs2 Helicases in the Enforcement of Recombination Fidelity in Saccharomyces cerevisiae. Genetics *168*, 1855–1865.

Stathos, A., and Fishman, L. (2014). Chromosomal rearrangements directly cause underdominant F1 pollen sterility in Mimulus lewisii-Mimulus cardinalis hybrids. Evolution *68*, 3109–3119.

Stewart, A.D., Logsdon, J.M., and Kelley, S.E. (2005). An empirical study of the evolution of virulence under both horizontal and vertical transmission. Evolution *59*, 730–739.

Stratton, G.E., and Uetz, G.W. (1986). The Inheritance of Courtship Behavior and Its Role as a Reproductive Isolating Mechanism in Two Species of Schizocosa Wolf Spiders (Araneae; Lycosidae. Evolution *40*, 129–141.

Swanson, W.J., Yang, Z., Wolfner, M.F., and Aquadro, C.F. (2001). Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. Proc. Natl. Acad. Sci. U.S.A. *98*, 2509–2514.

Sweigart, A.L. (2010). The Genetics of Postmating, Prezygotic Reproductive Isolation Between Drosophila virilis and D. americana. Genetics *184*, 401–410.

Tang, S., and Presgraves, D.C. (2009). Evolution of the Drosophila nuclear pore complex results in multiple hybrid incompatibilities. Science *323*, 779–782.

Taylor, D.R., Zeyl, C., and Cooke, E. (2002). Conflicting levels of selection in the accumulation of mitochondrial defects in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences *99*, 3690–3694.

Tessé, S., Storlazzi, A., Kleckner, N., Gargano, S., and Zickler, D. (2003). Localization and roles of Ski8p protein in Sordaria meiosis and delineation of three mechanistically distinct steps of meiotic homolog juxtaposition. Proc. Natl. Acad. Sci. U.S.A. *100*, 12865–12870.

Thompson, J.N. (2013). Relentless Evolution (Chicago: University Of Chicago Press).

Tinghitella, R.M., Zuk, M., Beveridge, M., and Simmons, L.W. (2011). Island hopping introduces Polynesian field crickets to novel environments, genetic bottlenecks and rapid evolution. J. Evol. Biol. *24*, 1199–1211.

Torres, E.M., Williams, B.R., and Amon, A. (2008). Aneuploidy: Cells Losing Their Balance. Genetics *179*, 737–746.

Trindade, S., Sousa, A., Xavier, K.B., Dionisio, F., Ferreira, M.G., and Gordo, I. (2009). Positive epistasis drives the acquisition of multidrug resistance. PLoS Genet. *5*, e1000578.

Tsai, I.J., Bensasson, D., Burt, A., and Koufopanou, V. (2008). Population genomics of the wild yeast Saccharomyces paradoxus: Quantifying the life cycle. PNAS *105*, 4957–4962.

Turelli, M., Barton, N.H., and Coyne, J.A. (2001). Theory and speciation. Trends Ecol. Evol. (Amst.) *16*, 330–343.

Turner, L.M., and Harr, B. (2014). Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. Elife *3*.

Velicer, G.J. (2003). Social strife in the microbial world. Trends in Microbiology *11*, 330–337.

Vines, T.H., and Schluter, D. (2006). Strong assortative mating between allopatric sticklebacks as a by-product of adaptation to different environments. Proceedings of the Royal Society B: Biological Sciences *273*, 911–916.

Vos, M., and Velicer, G.J. (2009). Social conflict in centimeter-and global-scale populations of the bacterium Myxococcus xanthus. Curr. Biol. *19*, 1763–1767.

Vulić, M., Dionisio, F., Taddei, F., and Radman, M. (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. Proceedings of the National Academy of Sciences *94*, 9763–9767.

WACH, A. (1996). IYeast Functional Analysis Reports. Yeast *12*, 259–265.

Wadsworth, C.B., Li, X., and Dopman, E.B. (2015). A recombination suppressor contributes to ecological speciation in OSTRINIA moths. Heredity (Edinb) *114*, 593–600.

Wagner, A., Wagner, G.P., and Similion, P. (1994). Epistasis Can Facilitate the Evolution of Reproductive Isolation by Peak Shifts: A Two-Locus Two-Allele Model. Genetics *138*, 533–545.

Wallace, A.G., Detweiler, D., and Schaeffer, S.W. (2013a). Molecular population genetics of inversion breakpoint regions in Drosophila pseudoobscura. G3 (Bethesda) *3*, 1151–1163.

Wallace, A.G., Detweiler, D., and Schaeffer, S.W. (2013b). Molecular population genetics of inversion breakpoint regions in Drosophila pseudoobscura. G3 (Bethesda) *3*, 1151–1163.

Walter, R.B., and Kazianis, S. (2001). Xiphophorus interspecies hybrids as genetic models of induced neoplasia. ILAR J *42*, 299–321.

Watt, P.M., Hickson, I.D., Borts, R.H., and Louis, E.J. (1996). Sgs1, a Homologue of the Bloom's and Werner's Syndrome Genes, Is Required for Maintenance of Genome Stability in Saccharomyces Cerevisiae. Genetics *144*, 935–945.

Welch, J.J. (2004). Accumulating Dobzhansky-Muller Incompatibilities: Reconciling Theory and Data. Evolution *58*, 1145–1156.

Welz-Voegele, C., and Jinks-Robertson, S. (2008). Sequence Divergence Impedes Crossover More Than Noncrossover Events During Mitotic Gap Repair in Yeast. Genetics *179*, 1251–1262.

Werren, J.H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc. Natl. Acad. Sci. U.S.A. *108 Suppl 2*, 10863–10870.

West, S.A., Griffin, A.S., Gardner, A., and Diggle, S.P. (2006). Social evolution theory for microorganisms. Nat. Rev. Microbiol. *4*, 597–607.

Whitby, M.C. (2005). Making crossovers during meiosis. Biochem. Soc. Trans. *33*, 1451–1455.

White, M.J.D. (1978). Modes of Speciation (San Francisco: W.H.Freeman & Co Ltd).

Wiernasz, D.C., and Kingsolver, J.G. (1992). Wing melanin pattern mediates species

recognition in Pieris occidentalis. Animal Behaviour *43*, 89–94.

Wilkins, A.S., and Holliday, R. (2008). The Evolution of Meiosis From Mitosis. Genetics *181*, 3–12.

Will, J.L., Kim, H.S., Clarke, J., Painter, J.C., Fay, J.C., and Gasch, A.P. (2010). Incipient Balancing Selection through Adaptive Loss of Aquaporins in Natural Saccharomyces cerevisiae Populations. PLoS Genetics *6*, e1000893.

Williams, H.T.P., and Lenton, T.M. (2008). Environmental regulation in a network of simulated microbial ecosystems. Proc. Natl. Acad. Sci. U.S.A. *105*, 10432–10437.

Williamson, D. (2002). The curious history of yeast mitochondrial DNA. Nature Reviews Genetics *3*, 475–481.

Wu, C.-I., and Ting, C.-T. (2004). Genes and speciation. Nature Reviews Genetics *5*, 114–122.

Wu, C., and Palopoli, M.F. (1994). Genetics of postmating reproductive isolation in animals. Annual Review of Genetics *28*, 283–308.

Wyckoff, G.J., Wang, W., and Wu, C.-I. (2000). Rapid evolution of male reproductive genes in the descent of man. Nature *403*, 304–309.

Xu, M., and He, X. (2011). Genetic Incompatibility Dampens Hybrid Fertility More Than Hybrid Viability: Yeast as a Case Study. PLoS ONE *6*, e18341.

Yamamoto, S., and Sota, T. (2009). Incipient allochronic speciation by climatic disruption of the reproductive period. Proc. Biol. Sci. *276*, 2711–2719.

Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G.J., and Woese, C.R. (1985). Mitochondrial origins. Proceedings of the National Academy of Sciences *82*, 4443–4447.

Yang, F., Pecina, D.A., Kelly, S.D., Kim, S.-H., Kemner, K.M., Long, D.T., and Marsh, T.L. (2010). Biosequestration via cooperative binding of copper by Ralstonia pickettii. Environ Technol *31*, 1045–1060.

Yona, A.H., Manor, Y.S., Herbst, R.H., Romano, G.H., Mitchell, A., Kupiec, M., Pilpel, Y., and Dahan, O. (2012). Chromosomal duplication is a transient evolutionary solution to stress. Proc. Natl. Acad. Sci. U.S.A. *109*, 21010–21015.

Zahrt, T.C., and Maloy, S. (1997). Barriers to recombination between closely related bacteria: MutS and RecBCD inhibit recombination between Salmonella typhimurium and Salmonella typhi. Proceedings of the National Academy of Sciences *94*, 9786–9791.

de Zamaroczy, M., Marotta, R., Faugeron-Fonty, G., Goursot, R., Mangin, M., Baldacci, G., and Bernardi, G. (1981). The origins of replication of the yeast mitochondrial genome and the phenomenon of suppressivity. Nature *292*, 75–78.

Zeyl, C. (2014). Evolution: a collection of misfits. Curr. Biol. *24*, R394–R396.

Zeyl, C., Bell, G., and Green, D.M. (1996). Sex and the spread of retrotransposon Ty3 in experimental populations of Saccharomyces cerevisiae. Genetics *143*, 1567–1577.

Zeyl, C., Vanderford, T., and Carter, M. (2003). An Evolutionary Advantage of Haploidy in

Large Yeast Populations. Science *299*, 555–558.

Zhang, J. (2003). Evolution by gene duplication: an update. Trends in Ecology & Evolution *18*, 292–298.

Zhang, J., Rosenberg, H.F., and Nei, M. (1998). Positive Darwinian selection after gene duplication in primate ribonuclease genes. PNAS *95*, 3708–3713.

Zhang, M., Azad, P., and Woodruff, R.C. (2011). Adaptation of Drosophila melanogaster to increased NaCl concentration due to dominant beneficial mutations. Genetica *139*, 177–186.

Affidavit

Hiermit erkläre ich, dass die vorliegende Arbeit
- nach Inhalt und Form meine eigene ist, abgesehen von der Beratung durch meinen Betreuer Dr. Duncan Greig
- an keiner anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen hat, noch

nicht veröffentlicht ist und auch nicht zur Veröffentlichung eingereicht wurde
- unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen

Forschungsgemeinschaft entstanden ist.

Plön, den 09.12.2015

Gönensin Ozan Bozdag