Heterosis in yeast hybrids

Dissertation

in fulfillment of the requirements for the degree
Doctor rerum naturalium

of the Faculty of Mathematics and Natural Sciences at the
Christian Albrechts University of Kiel

Submitted by

Joana P. Bernardes

Group of Experimental Evolution
Max Planck Institute for Evolutionary Biology

Kiel, January 2018

First referee: Dr. Duncan Greig

Second referee: Prof. Dr. Eva Stukenbrock

Third referee: Dr. Arne Nolte

Chairperson: Prof. Dr. Mathias Leippe

Date of oral examination: 05/03/2018

Table of Contents

Summary

Hybridisation occurs when individuals of divergent populations or individuals of different species mate and produce viable hybrid offspring. Hybridisation is an important factor in the evolutionary history of several populations, because it brings together alleles that have not been exposed to selection. Thus hybrid adaptation to ancestral or new environmental conditions depends on the background and interaction of the parental genotypes. I am particularly interested in the performance of the first hybrid generation (F1 hybrids), especially when hybrids are viable and able to outperform one or both parents under different environmental conditions, a phenomenon known as heterosis.

The aim of my thesis was to understand and identify mechanisms underlying heterosis. I used *Saccharomyces* yeasts as a model system due to their laboratory practicality, their ability to form viable hybrids and reliable fitness measurements. First, I competed a range of different F1 hybrids with their wild or domesticated parental populations; I identified prevalent heterosis for crosses between domesticated and wild populations of different yeast species but not for crosses between wild populations of the same yeast species. Thus the environment from where parental strains were isolated seems to affect heterosis, and F1 hybrids with a domesticated background display more extensive heterosis. Domesticated yeasts are characterized by being highly heterozygous, which can potentially mask recessive deleterious alleles in the genome. When yeasts strains are brought to the laboratory they undergo an extreme form of inbreeding that induces a haploid spore to grow vegetatively, and allows mate-type switching followed by mating within the same haploid colony. This inbreeding process creates a complete homozygous monosporic clone with recessive deleterious alleles exposed. By crossing two domesticated monosporic clones derived from divergent populations several recessive deleterious alleles might be complemented and the F1 hybrid would have an advantage in comparison to its monosporic parents. Using monosporic clones as parental strains in heterosis studies may inflate heterosis measurements due to parental disadvantage and not the F1 hybrid advantage. Thus I compared asexual fitness or growth of heterozygous yeast isolates with homozygous monosporic clones for both domesticated and wild yeast populations; I found that the monosporic cloning might explain some, but not all, of

the heterosis seen, potentially accounting for the difference in heterosis between domesticated and wild yeast strains. Thus heterosis is not solely explained by complementation of recessive deleterious alleles, and other mechanisms might affect the F1 hybrid advantage. I focus on heterosis at the transcriptome level and analysed the transcription of a representative heterotic F1 hybrid relative to its parents in environments that favoured one or the other parent. Hybrid transcription was varied and resembled the fitter parent in specific environments. Thus at the transcriptome level, the F1 hybrid may repress potential deleterious alleles, making them recessive, and induce more advantageous alleles, making them dominant, by differentially transcribing its parental alleles. For the first time to our knowledge, multigenic heterosis at a transcriptome level was identified, which render the F1 hybrid better adapted than its parents to different environmental conditions.

Heterosis studies in *Saccharomyces* yeasts, due to their simplicity, can evidence characteristics with an impact on heterosis while also tracing the evolutionary history of divergent populations. These types of studies have interesting applications agriculture sector where hybridisation has been used for centuries to make higher yield crops and bigger cattle to fulfil human consumption needs.

Zusammenfassung

Hybridisierungen sind Kreuzungen zwischen Individuen verschiedener Populationen oder verschiedener Arten. Hybridisierung spielt eine wichtige Rolle bei der Rekonstruktion der evolutionären Geschichte von Populationen, doch auch die Hybriden selbst könnten ein Schlüssel für die Lösung und das Verständnis biologischer Konzepte sein. Hybriden sind das Produkt einer Kreuzung von bereits getrennten, parentalen Genotypen, und die Anpassungsfähigkeit von Hybriden wird maßgeblich vom genetischen Hintergrund der Elterngeneration bestimmt. Ich interessiere mich besonders für die Leistung der ersten Generation von Hybriden (F1 Hybriden), insbesondere wenn die F1 Hybriden die Fähigkeit besitzen in unterschiedlichen Umgebungen besser zu wachsen, als ihre Eltern. Dieses Phänomen wird als Heterosis bezeichnet.

Das Ziel meiner Arbeit ist es, den der Heterosis zugrundeliegenden Mechanismus zu identifizieren und zu verstehen. Ich habe Hefen als Modell-System gewählt, weil sie die Fähigkeit besitzen, Hybriden zwischen genetisch divergenten Populationen (intraspezifisch) oder zwischen genetisch divergenten Arten (interspezifisch) innerhalb der Gattung *Saccharomyces* zu bilden. Zusätzlich sind Hefen ein perfektes Modell-System, um asexuelle Fitness zu messen, da sie einfach in großer Anzahl und unter kontrollierbaren und wiederholbaren Bedingungen zu kultivieren sind. Bei unterschiedlichen Vergleichen von F1 Hybriden mit ihren Eltern konnte ich Heterosis, sowohl auf der genetischen, als auch der transkriptiven Ebene finden.

Im ersten Kapitel dieser Arbeit habe ich das Wachstum von F1 Hybriden unter Konkurrenz im direkten Vergleich mit beiden Elternpopulationen gemessen. Die F1 Hybriden wurden durch Kreuzungen von *S. cerevisiae* und *S. paradoxus* (interspezifische Hybriden) und durch Kreuzungen von *S. paradoxus* mit wilden Populationen (intraspezifische Hybriden) erzeugt. Ich fand, dass das Wachstum der F1 Hybriden relativ zum Durchschnitt der Eltern, einem Wert für die elterliche Heterosis, mit dem Wachstumsunterschied zwischen den Eltern und den F1 Hybriden, einem Wert für phänotypische Unterschiede, korreliert war. Dies deutet darauf hin, dass Allele mit einer geringen Fitness vom einen Elternteil, durch Allele mit einer

hohen Fitness vom anderen Elternteil ausgeglichen werden. Die interspezifischen F1 Hybriden zeigten eine stärkere Heterosis, als intraspezifische F1 Hybriden, was darauf hindeutet, dass die genetische Divergenz in einem Zusammenhang mit der Gesamt-Heterosis stehen könnte. Um die elterliche genetische Divergenz unabhängig vom Genotypen zu manipulieren, habe ich außerdem das Wachstum unter Konkurrenz von einzelnen interspezifischen F1 Hybriden relativ zu beiden Eltern in zwölf unterschiedlichen Umgebungen gemessen. Hierbei habe ich nicht nur den zuvor erwähnten Zusammenhang zwischen der phänotypischen Divergenz der Eltern und der mittleren parentalen Heterosis, sondern auch eine schwache Beziehung zwischen der phänotypischen Divergenz und der „best-parent" Heterosis. Dies deutet darauf hin, dass die Komplementierung von schadhaften Allelen nicht der einzige Grund für interspezifische Heterosis sein könnte. Ich vermute, dass die reziproke Komplementierung der weniger fitten Allele zwischen den Eltern einen fitteren Hybriden hervorbringt.

Im zweiten Kapitel habe ich den interspezifischen F1 Hybriden genutzt und sein Transkriptionsprofil mit dem der jeweiligen Eltern in zwei unterschiedlichen Umgebungen verglichen. Die Bedingungen innerhalb dieser Umgebungen wurden so gewählt, dass der F1 Hybrid immer einen Vorteil gegenüber beiden Eltern haben würde und dass beide Eltern in je einer der beiden Umgebungen einen Vorteil im Vergleich mit der anderen Elternpopulation haben würden. Meine Ergebnisse zeigen, dass das Transkriptionsprofil der F1 Hybriden dem von *S. paradoxus* in der Umgebung ähnlicher waren, in der der parentale *S. paradoxus* einen Vorteil gegenüber *S. cerevisiae* hatte und dem Transkritionsprofil von *S. cerevisiae* in der Umgebung mehr ähnelte, in der der parentale *S. cerevisiae* einen Vorteil gegenüber *S. paradoxus* hatte. Des Weiteren fand ich, dass die Transkription des F1 Hybriden eine hohes Level an Dominanz der jeweils fitteren Elternpopulation in der jeweiligen Umgebung zeigte. Zudem war eine allelspezifische Expression bei den F1 Hybriden häufig zu beobachten. Demnach trat die *cis*-regulierte Transkription häufiger auf, als die *trans*-regulierte Transkription, und diese Eigenschaft ist konsistent mit der Komplementierung von weniger fitten Allelen des einen Elternteils durch Allele mit einer höheren Fitness vom anderen Elternteil. Interessanterweise kann der F1 Hybrid nicht nur seine Transkription so beeinflussen, dass das jeweils fittere der beiden

parentalen Allele in der jeweiligen Umgebung induziert wird, sondern hat auch die Fähigkeit seine Transkription auf einen Umweltreiz hin zu modifizieren. In meiner Arbeit habe ich Spuren von Heterosis in der Transkription von F1 Hybriden gefunden, was ich damit erkläre, dass der F1 Hybrid die Fähigkeit hat beide parentale Allele zu regulieren und bevorzugt das vorteilhaftere Allel an mehreren Loci zu transkribieren.

Im dritten und letzten Kapitel habe ich natürliche Isolate von *S. cerevisiae* aus domestizierten und wilden Habitaten mit bekannter Heterozygosität genutzt. Hierbei habe ich eine positive Relation zwischen dem Anteil an rezessiven, schädlichen Allelen von domestizierten Isolaten etabliert. Ich habe die Sporenviabilität und die asexuelle Fitness von zwölf natürlichen Isolaten, sowie die relative Fitness ihrer abgeleiteten Formen gemessen. Von den natürlichen Isolaten habe ich autodiploide Formen durch die Selbstbefruchtung einer abgeleiteten Spore abgeleitet. Inzuchtformen habe ich durch die Kreuzung zweier autodiploiden Formen, von demselben natürlichen Isolat, abgeleitet. Ich habe einen starken negativen Zusammenhang zwischen der Heterozygosität der natürlichen Isolate und ihrer Fähigkeit Sporen zu bilden gefunden. Zudem konnte ich einen stark negativen Zusammenhang zwischen der Heterozygosität und der relativen Fitness der autodiploiden Formen feststellen. Dies deutete darauf hin, dass ich bei der Eliminierung der Heterozygosität durch Autodiploidation viele rezessive, schädliche Allele freigesetzt habe, die die relative Fitness der autodiploiden Formen im Vergleich zu ihren natürlichen Isolaten senken. Im nächsten Schritt, bei der Generierung der Inzuchtformen durch die Kreuzung der beiden autodiploiden Formen, wurde die Heterozygosität jedoch wiederhergestellt und die rezessiven, schadhaften Allele wurden wieder ergänzt. Da die Autodiploidation von parentalen Stämmen gängige Praxis in Heterosisstudien ist und domestizierte, natürliche Isolate dazu tendieren eine hohe Heterozygosität aufzuweisen, während wilde, natürliche Isolate zu einer niedrigere Heterozygosität tendieren, gibt es einen Effekt in der Gesamt-Fitness der parentalen Stämme. Daher könnten Heterosis Studien, in denen domestizierte Eltern-Stämme genutzt wurden, dazu neigen rezessive, schadhafte Allele freizulegen und so das Maß an Heterosis zu überhöhen. Meine Ergebnisse deuten darauf hin, dass wenn der F1 Hybrid mit seinem ursprünglichen, natürlichen Isolat verglichen würde und nicht mit seinem autodiploiden Vorfahren, das Maß an

Heterosis gesenkt würde oder sie sogar vollständig verschwände.

Meine Arbeit beleuchtet Schlüsselaspekte der Heterosis. Sie unterstützt die Idee, dass Genotyp und Habitat der Elternpopulationen einen großen Einfluss auf die Heterosis des F1 Hybriden haben. Zusätzlich und nach unserem Wissen zum ersten Mal, habe ich multigene Heterosis auf dem Level des Transkriptoms nachgewiesen, die es dem F1 Hybriden ermöglicht besser als seine beiden Elternpopulationen an verschiedene Umgebungen angepasst zu sein. Zum Abschluss der Arbeit analysiere ich die bestehende Literatur zum Thema Heterosis in Hefen und schlage eine einfache Erklärung für die divergenten Ergebnisse innerhalb eines Rahmenkonzepts vor – nämlich, dass die Eigenschaften der parentalen natürlichen Isolate einen entscheidenden Einfluss auf die Heterosis des F1 Hybriden haben.

Introduction

1. Hybridisation & Heterosis

1.1. Hybridisation

Hybridisation describes sexual crosses between individuals of divergent populations or between individuals of different species (Barton & Hewitt, 1985). The outcome of hybridisation is often seen as a dichotomy among biologists; on the one hand, classic zoologists see hybrids as infertile evolutionary dead-ends (Dobzhansky, 1940), while on the other hand, plant breeders believe that hybridisation has been a major contributor to speciation and adaptation of plant populations, because it promotes new combinations of previously separated alleles (Arnold & Hodges, 1995; Rieseberg *et al*., 1996).

If we were to cross individuals from two mammalian species with no pre-zygotic barriers, like a horse and a donkey, even though they have a different number of chromosomes (horses carrying 64 chromosomes and donkeys 62 chromosomes) there would be no physical reproductive barriers between the two species, thus the first hybrid generation (F1 hybrid) would be a viable but infertile mule carrying 63 chromosomes. Mules are infertile because of the impaired formation of their sexual gametes, due to prevention of crossing-over and disjunction events between homologous chromosomes. These problems in meiosis unable the chromosomes to separate properly and lead unviable aneuploid oocytes (Anderson, 1939), however there is still a remote (one in four million, 2^{32}) chance for a correct chromosomal segregation (if we assume no recombination) where the female mule produces a perfect horse or a perfect donkey oocyte (Rong *et al*., 1988).

If we were to cross individuals from two plant species, like two divergent salvias, two distinct tomato plants, or two different strains of maize, the outcome would be similar to the mammalian species. The F1 hybrids may be viable but may not produce viable seeds, however, from a human consumption perspective, the plant F1 hybrids would have other attractive characteristics like big seedless fruits. Hybridisation between two plant populations can be beneficial due to the novel characteristics of the F1 hybrids (Shull, 1908), moreover the ability of many plants to reproduce asexually can

allow the production of clones from the original sterile F1 hybrids (Arnold & Hodges, 1995).

These different distinct views on hybridisation have their roots on the infertility of the F1 hybrid by problems in meiosis. Infertility might render the F1 hybrid with zero fitness, if it depends solely on sexual reproduction, but if the F1 hybrid can also reproduce asexually, the negative impact of infertility on fitness might be diminished. From an animal breeder perspective this is an evolutionary dead-end since they are mainly obligately sexual while a plant breeder can ignore the sexual deficit due to the ability of plants to propagate asexually thus hybrids are often selected for their novel, and sometimes, improved traits. The question remains if it is a good idea for a species to be hybridised, the answer is- it depends; it depends on the F1 hybrid's characteristics such as viability, asexual reproduction or growth, and other novel features.

1.2. Heterosis

The evolutionary outcomes of hybridisation will depend on differences between the parental genotypes. Hybridisation can bring alleles together in combinations that have not previously been tested by natural selection thus the expectation is for intermediate or worse performance of the F1 hybrids in comparison to the parental average in the ancestral conditions (Barton, 2001). However, in some cases, the F1 hybrid shows performs better than one or both parents in the ancestral or in novel environmental conditions, a phenomenon known as heterosis or hybrid vigour (Shull, 1908). Heterosis relies on crossing two divergent parental populations leading to a superior fitness of the F1 hybrid or other desirable trait. Heterosis might be related to increased heterozygosity of the F1 hybrid which might buffer the effect of deleterious alleles or generate positive interactions between parental alleles, thus providing the F1 hybrid with a genetic background better equipped to deal with ancestral and new environmental conditions (Dobzhansky, 1950).

Hybridization is not only important for speciation and adaptation of the parental populations but has also an intrinsic economic value since its discovery in the beginning of the 20th century (Shull, 1908). Heterosis is widely observed in our food chain from crop plants production to cattle breeding. Seeds from divergent and highly

inbred parents (F1 hybrid seeds) are still the most common type of seeds used in agriculture worldwide for crop plants such as maize, sugar beets, spinach, and sunflower crops (Edgerton, 2009). These crop seeds are characterized by giving rise to offspring with greater biomass and speed of development (Shull, 1908; East, 1936). From an economic standpoint, F1 hybrid seeds are also utilitarian because they produce homogeneous higher yield crops that are frequently infertile; an asset for business trades because it forces farmers to buy a new set of F1 hybrid seeds every season. Over the last century plant breeders have used classical techniques to improve heterosis by artificially selecting genetic and phenotypic divergent inbred parents and crossing them to create increasingly adapted F1 hybrids; F1 hybrids produced from low-yield inbred parents can have yield advantages up to 120% over their parents (Moll *et al*., 1965). Crop yield advantages are normally phenotypic traits such as faster growth, bigger fruit size, and higher seed number; these phenotypic traits are mainly related to plant growth, and not to cell size or other characteristics (Birchler *et al*., 2010). Thus plant breeders tend to care about phenotypic traits that affect crop yield, which is not necessarily related to fitness, in the sense that F1 hybrids may not be able to outcompete their parents without human intervention. However, it is possible for F1 hybrids, in particular for species that reproduce asexually, to exhibit a competitive advantage over their parent in the field. As such, heterosis can be measure as yield and growth in plants or it can also be measured as asexual growth or mitosis in microorganisms.

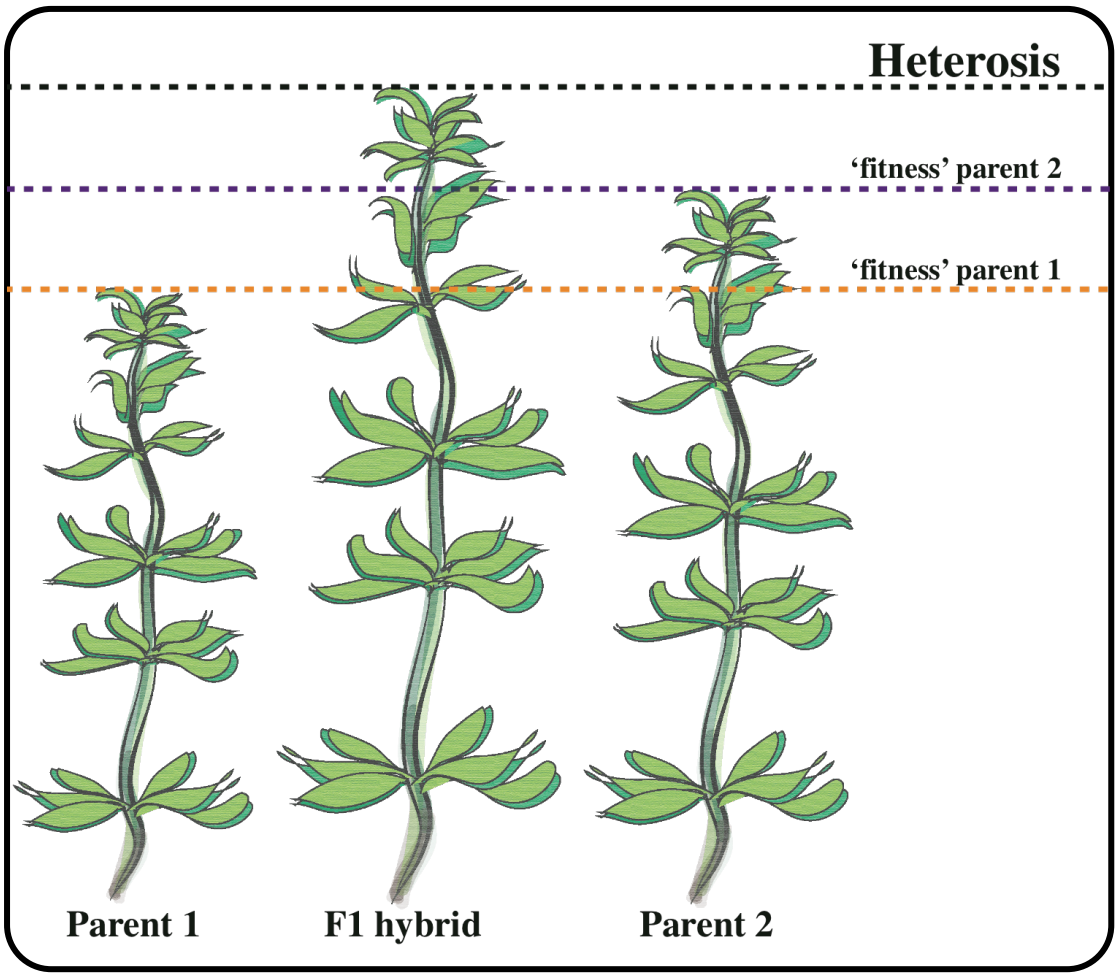


Figure 1: F1 hybrid displays higher fitness than both parents a phenomenon known as heterosis. Fitness here is depicted as size of the plants. F1 hybrid made from a cross between parent 1 and parent 2. Fitness of parent 1 (orange line) is lower than fitness of parent 2 (purple line), which is lower than fitness of the F1 hybrid (black line). Heterosis can be defined as higher fitness in the F1 hybrid relative to both parents.

2. Genetic mechanisms for heterosis

Heterosis refers to the F1 hybrid increased size, yield or vigor in relation to one or both parents (Shull, 1948), which researchers translated to asexual fitness in microorganisms. Heterosis can be analysed as a quantitative trait and thus explained by similar genetic mechanisms (East, 1936). Three principal genetic mechanisms have been proposed to explain heterosis: Dominance, Overdominance and Epistasis (Figure 1). These mechanisms may act independently of each other, but can also occur simultaneously in the F1 hybrid. However, their relative effects on heterosis remain unknown (for review see Lippman & Zamir, 2007; Birchler *et al*., 2010).

2.1. Dominance

The dominance mechanism of heterosis refers to simple or reciprocal complementation of superior dominant alleles over inferior recessive alleles at multiple loci which contribute to F1 hybrid superior fitness (Jones, 1917). Evidence of the dominance mechanism has been found in rice (Li *et al*., 2008; Shen *et al*., 2014), maize (Feher *et al*., 2014) and yeast (Zörgö *et al*., 2012; Plech *et al*., 2014; Shapira *et al*., 2014).

This mechanism relies on individuals from different populations carrying recessive deleterious alleles at different loci. These alleles occur by mutation and can be maintained in the population due to their recessive nature: they are hidden from selection in heterozygotes because they are complementated by the wild-type dominant allele. Breeding experiments use inbred parental lineages; in this case, inbreeding eliminates heterozygosity causing recessive alleles to become homozygous causing the phenotypic consequences. When recessive alleles are deleterious, parental lineages display reduced fitness, the 'inbreeding depression' can be so severe that inbred lineages my go extinct. Crossing two different inbred parents, introduces a gain in heterozygosity in the F1 hybrid and the recessive deleterious alleles of one parental genotype can be complemented by the more favourable dominant alleles of the other parent at multiple loci. The higher number of complemented deleterious alleles renders the F1 hybrid higher fitness in relation to one or both parents. This complementation can be simple when refers to one parent over the other parent at multiple loci or reciprocal on both parents complement each other at multiple loci. If

this mechanism was exclusively responsible for heterosis, genetic divergence between parents should have a positive effect on heterosis because two closely related parents are more prone to share similar recessive deleterious alleles (Moll *et al*., 1965).

2.2. Overdominance

The overdominance mechanism refers to positive allelic interactions that give an advantage to heterozygotes at one or more loci which contribute to the superior fitness of F1 hybrids (Shull, 1948). Evidence for the overdominance mechanism has been found in rice (Luo *et al*., 2001; Li *et al*., 2008) maize (Hollick & Chandler, 1998) and yeast (Shapira *et al*., 2014).

This mechanism relies on the number of heterozygous over homozygous sites of the individuals, and assumes the heterozygous locus have an advantage over either homozygous locus. When inbred parents are selected their heterozygosity is removed through self-fertilization of other types of inbreeding. By crossing two different highly homozygous parents the resulting F1 hybrid has a gain in heterozygosity. The higher number of heterozygous locus renders the F1 hybrid higher fitness in relation to one or both parents (Shull, 1948).

2.3. Epistasis

The epistasis mechanism refers to the interactions between non-allelic genes at two or more loci, which contribute to the phenotype exceeding each individual gene, in this specific case the phenotype would be the F1 hybrid superior fitness (Powers, 1944). Evidences for the epistasis mechanism has been found in rice (Luo *et al*., 2001; Li *et al*., 2008; Shen *et al*., 2014) and yeast (Shapira *et al*., 2014).

This mechanism relies on one locus expression having a positive effect on other loci, and few or even one interaction might have a great impact on heterosis. Under epistasis mechanism, heterosis over one or both parents is the result of positive interactions between alleles at different loci (Powers, 1944). Detection of epistasis can be difficult, one locus effect can be altered or masked by the effect of another locus, but if more than one locus are involved, detection of epistasis can become increasingly complicated. If the contribution of each parental allele with the overall contribution of both parental alleles in the F1 hybrid is outside the expected of additive, dominant or overdominant effects we assume an epistatic effect.

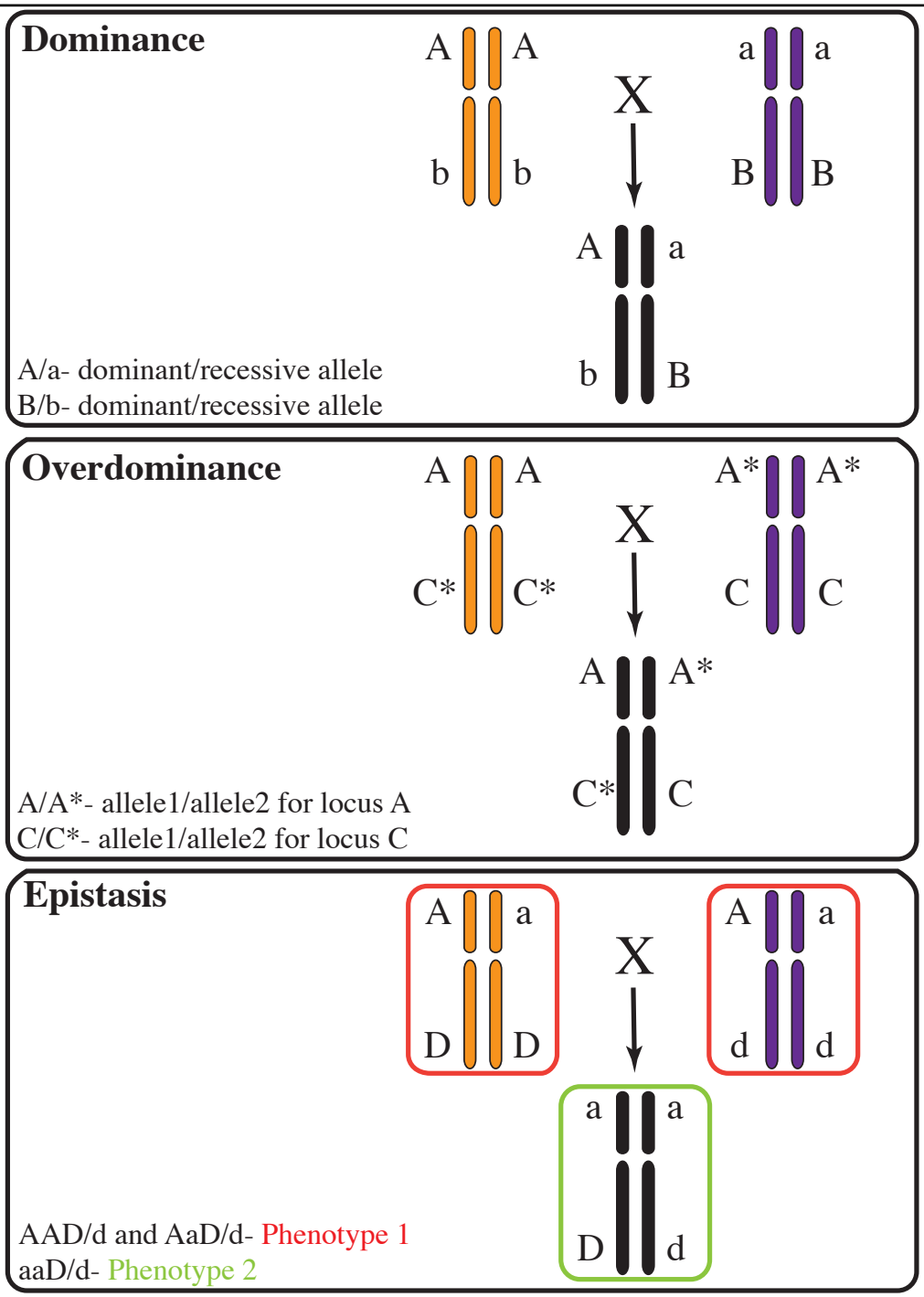


Figure 2: Genetic mechanisms for heterosis. Orange chromosomes refer to parent 1, purple chromosomes refer to parent 2, and black chromosomes refer to the F1 hybrid. Dominance mechanism refers to the complementation of the recessive deleterious alleles (a allele by A allele and b allele by B allele) between parents that renders the F1 hybrid higher fitness. Overdominance mechanism refers to allelic interactions in heterozygous locus (A allele with A* allele and C allele with C* allele) that renders the F1 hybrid higher fitness. Epistasis mechanism refers to the interaction between alleles of two or more locus (D/d locus will produce phenotype 1 in red, if locus is AA or Aa whereas if locus is aa will produce phenotype 2 in green independent of D/d locus) that renders the F1 hybrid higher fitness.

2.4. Direct observations of heterosis

Direct observations of heterosis and other resulting events have supported the above genetic mechanisms for heterosis; inbred individuals display a decrease in growth, or inbreeding depression (Waser, 1993), genetic divergence of the parents has a positive relationship with heterosis (Moll *et al*., 1965), and inbred or domesticated parents have a greater positive effect of heterosis (Plech *et al*., 2014). However there are some observations that have been neglected and are not directly explained by any of the above genetic mechanisms:

For example, artificially purging the deleterious alleles from parental inbreds does not impact the amount of heterosis of their F1 hybrid (Duvick, 1999). Duvick (1999) compared heterosis of crosses between improved inbred lines of several crop plants purged of recessive deleterious alleles, with crosses between inbred lines with accumulation of recessive deleterious alleles, and found no decrease on heterosis for the improved inbred lines. These results go against the dominance mechanism for heterosis because the accumulation of superior alleles in the inbred parent should lead to less recessive deleterious alleles being complemented in the F1 hybrid. Also, increasing hybrid ploidy has a positive effect on heterosis (East, 1936). East (1936) cross two allotetraploid hybrids (individual with four times more chromosomes than the haploid individual) with known positive heterosis, and identified increasingly higher levels of heterosis, higher than either of the allotetraploid parents. Thus polyploidy was positive correlated with the advantages in the F1 hybrid fitness or heterosis an event it is not explained by any of the above genetic mechanisms. And finally, aneuploid hybrids such as haploids carrying an extra pair of chromosomes show lower heterosis than similar diploids (Birchler *et al*., 2007). These results cannot be explained by either dominance or overdominance genetic mechanisms because hybrids would be expected to have an optimum fitness based on the parents, and this fitness should not increase just because there is an extra copy of the beneficial allele(s).

There is a need for a new model possibly a molecular that explains heterosis for the above observations of heterosis and works for a variety of organisms (Birchler *et al*., 2010).

3. Molecular mechanisms for heterosis

A molecular model for heterosis needs to describe the increase fitness of the F1 hybrid in comparison to its inbred parents (Shull, 1908) and has to encompass the above observations (Birchler *et al*., 2007). The successful model has to be species-independent (Arnold & Hodges, 1995), cell-based (Birchler *et al*., 2010), and has to be evolutionary conserved (Goff, 2011).

3.1. Multigenic molecular model for heterosis

A molecular model has been proposed to describe multigenic heterosis; this model is based on the ability of individual hybrid cells to distinguish between parental alleles based on the relative stability of the encoded proteins, and take advantage of allele specific expression to conserve energy and promote growth (Goff, 2011).

Goff (2011) molecular model relies on the follow mechanisms: First, the F1 hybrid transcribes and translates two different parental alleles into proteins with different stabilities. Second, the F1 hybrid receives a signal that identifies the favourable or superior allele over the inferior allele based on the protein stabilities. Third, the F1 hybrid has to be able to distinguish between parental alleles. And finally, the F1 hybrid preferential transcribes the superior parental allele over the inferior parental allele.

Goff (2011) molecular model for multigenic heterosis requires the F1 hybrid to have the ability to differentially transcribe the parental alleles, or allelic specific expression. The F1 hybrid fitness advantage or heterosis is due to its ability of the F1 hybrid to preferentially transcribe the superior alleles of both parents instead of the inferior alleles that are also present. The model is assumes transcription and translation of proteins is an energetically demanding process, therefore transcribing and translating only the superior alleles from both parents makes for more efficient energy use in the F1 hybrid and consequently promoting higher growth for the F1 hybrid (Goff, 2011).

3.2. Transcription expectations for the molecular model

Goff's (2011) molecular model for multigenic heterosis can be tested by F1 hybrid transcriptome analysis According to this model, in the F1 hybrid the superior alleles

would be up-regulated and the inferior alleles down-regulated, creating the optimal combination of superior alleles being transcribed, which renders the F1 hybrid advantageous (Figure 2).

New technological advances such as RNA-Seq can accurately measure the presence and quantity of mRNA for all genes of a biological sample at a given moment of time, also known as the transcription profile. Orthologous genes normally retain the same or similar functions between parental genomes but have slightly different genetic sequences (Jensen, 2001). When mapping the reads belonging to a given orthologous gene in the F1 hybrid, the reads from one parental allele would map to a specific parental genome, and the reads from the other parental allele would map to another parental genome (Wittkopp *et al*., 2004). The presence and quantity of reads that differentially map to alleles from each parental genome is a measure of differential allelic specific expression (Wittkopp *et al*., 2004). Thus, under Goff's model, superior alleles would have more reads mapping to the orthologous genes than the inferior alleles if they were translated into more stable proteins. The evaluation of superior or inferior alleles depends on the environment the individuals are in; a parental allele could be favourable in one environment but deleterious in another environment (Gasch *et al*., 2000). The F1 hybrid has the possibility to have multiple transcription profiles adapted to different environments, more than its individual parents (Clowers *et al*., 2015). Thus allele specific expression might also depend on the environmental conditions experienced by the F1 hybrid. The best combinations of pattern and quantity of parental alleles transcribed and the ability to change this transcription upon environmental change gives the F1 hybrid an advantage over its parents (Goff, 2011).

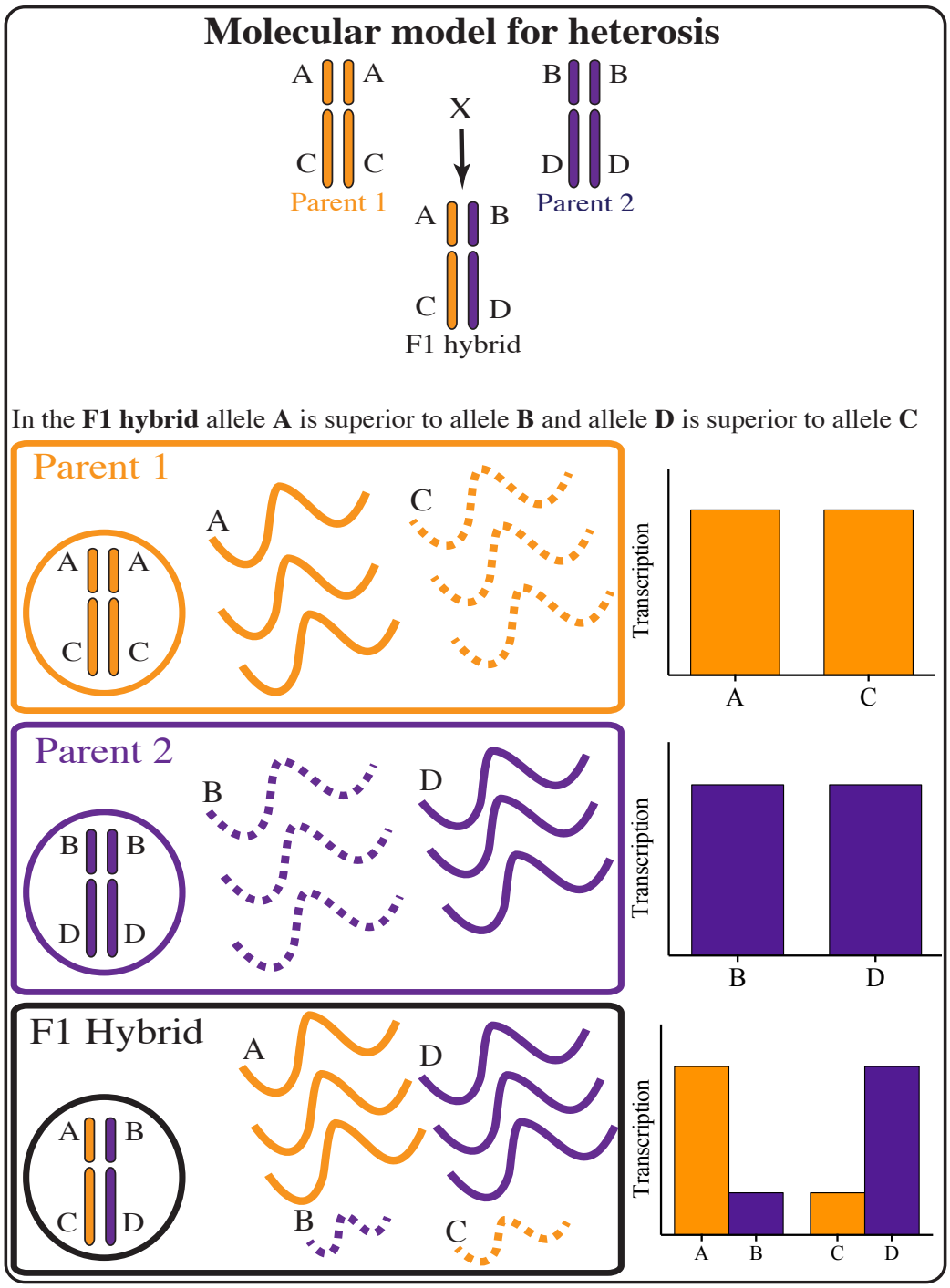


Figure 3: Molecular mechanism for heterosis. Parent 1 in orange, parent 2 in purple and F1 hybrid in orange and purple. Solid lines represent superior alleles that translate into stable proteins and dashed lines represent inferior alleles that translate into unstable proteins. Allele A from parent 1 is superior to allele B from parent 2 and allele D from parent 2 is superior to allele C from parent 1. Graphs refer to the allele transcription. F1 hybrid preferentially transcribes superior allele A from parent 1 over inferior allele B from parent 2, and preferentially transcribes superior allele D from parent 2 over inferior allele C from parent 1.

Mechanisms for heterosis

Type	Mechanism	Description
Genetic	**Dominance**	Complementation of recessive deleterious alleles in F1 hybrid
	Overdominance	Positive allelic interactions at heterozygous locus in the F1 hybrid
	Epistasis	Positive non-allelic interactions between two or more loci in the F1 hybrid
Molecular	**Multigenic heterosis** (Goff, 2011)	Preferentially transcription of superior parental alleles over deleterious parental alleles at one or more loci in the F1 hybrid

Summary Table 1: Different mechanisms for heterosis or fitness advantage of the F1 hybrid.

4. *Saccharomyces* yeasts

4.1. Saccharomyces yeasts as a model system

Saccharomyces yeasts are a group of eukaryotic single-celled microorganisms from the *Ascomycota* phylum. *Saccharomyces* yeasts have a close relationship with humans and for thousands of years have been used for baking and production of alcoholic beverages (McGovern *et al*., 2004). Their relative small genome led to *Saccharomyces cerevisiae* being the first fully-sequenced eukaryote genome (Whole Genome Sequence or WGS) (Goffeau *et al*., 1996). Since then, population genomics studies have provided robust and extensive data on *Saccharomyces* yeast species isolated from different habitats around the world (Liti *et al*., 2009).

S. cerevisiae and six other yeast species belong to *Saccharomyces* yeasts complex, they show high levels of genetic divergence, but can however readily hybridise (Naumov *et al*., 2000). The ability of highly diverged species to hybridise makes *Saccharomyces* yeasts a very interesting group to study hybridisation and consequently heterosis. Also, *Saccharomyces* yeasts are easy to work under laboratory conditions; they are well characterized both genetically and phenotypically, their genome can be easily manipulated, and they have short generation times. These features allow for easy and accurate measures of fitness due to the ability to grow them in large populations under controlled and repeatable conditions (Scannell *et al*., 2011).

4.2. Saccharomyces yeasts reproduction

Saccharomyces yeasts have facultative sexual reproduction; they can reproduce asexually by budding of a diploid or haploid cell (mitosis) or reproduce sexually when sporulation is induced of a diploid individual leading to the production of four haploid spores inside a tetrad (meiosis) (Figure 4, Herskowitz, 1988). Sexual cycles are normally induced when conditions are stressful, when conditions improve, the haploid spores might germinate and recover their diploid state by crossing with a spore from another ascus (outcrossing), crossing with a spore from the same ascus (inbreeding), or by germinating, inducing mate-type switching and crossing within the same haploid colony (self-fertilization or autodiplodization) (Herskowitz, 1988).

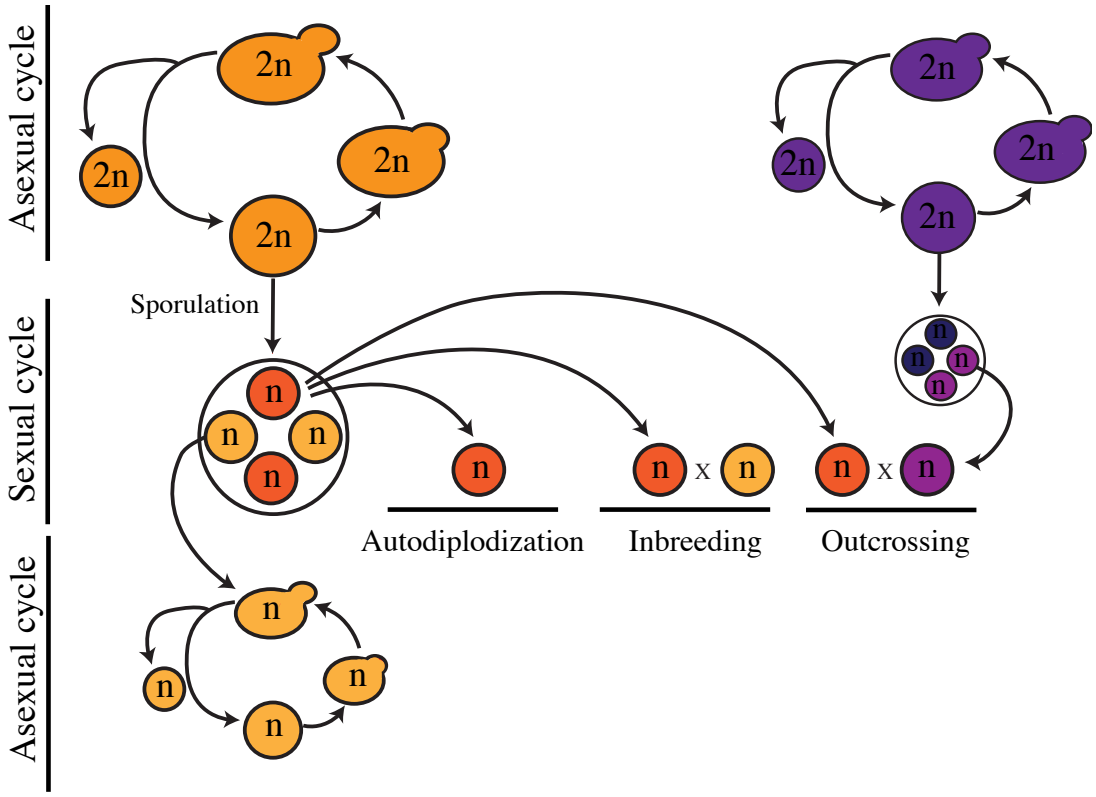


Figure 4: Asexual and sexual reproduction in *Saccharomyces* yeasts. Orange colour refers parent 1 and purple colour refers to parent 2. Parents 1 and 2 from divergent populations or different species. Asexual cycle in diploid (orange and purple (2n)), and haploid yeasts (light orange (n)) by vegetative growth resulting in the formation of a copy of the mother cell. Sexual cycle in diploid yeasts (2n) by sporulation and formation of a tetrad with four haploid spores (n). The haploid spores can Autodiplodized, Inbreed with a sister spore from the same tetrad or Outcross with a spore from another tetrad.

4.3. Saccharomyces yeasts ecology

Saccharomyces yeasts are used in human activities such as baking and in the making of alcoholic beverages. However they can also be found in other unlikely habitats and become a potential pathogen for humans. *Saccharomyces* yeasts like *S. cerevisiae* have been readily isolated from domesticated habitats for centuries but also from wild habitats such as tree exudates (Naumov *et al.*, 1998), while wild *S. paradoxus* isolates, a exclusively wild *Saccharomyces* species was only identified in the last century from tree exudates in Russia (Batshinskaya, 1914), since then, wild yeasts have been isolated from around the world from hardwood bark, soil and leaf liter around mainly oak trees (Kowallik & Greig, 2016). Thus the most likely natural

habitat of wild *Saccharomyces* yeasts would be related to bark or leaf litter of trees, supporting this idea, an extensive genomic analysis revealed that most wild *Saccharomyces* isolates have not been affected by domestication (Fay & Benavides, 2005).

4.3.1 Domesticated Saccharomyces yeasts

Saccharomyces yeasts from domesticated habitats have a close relationship with human activities, such as wineries or breweries, or a close relationship with humans, such as pathogen yeasts infections of immune-compromised patients (Muller & McCusker, 2009). Human activity has lead to a possible admixture or outcrossing between domesticated yeast strains, which produce mosaic or mixed genetic backgrounds in domesticated strains (Yue *et al*., 2017). These habitats are considered simple and sugar-rich, and do not require maintenance of all cellular functions, they are ideal for the emergence and maintenance of recessive deleterious mutations, because they promote genetic drift of recessive alleles that are not under the selective pressure (Figure 5) (Zörgö *et al*., 2012). Under strong directional selection, specific phenotypes might be acquired to cope with high ethanol levels in breweries or wineries (Casey & Ingledew, 1986) or faster growth rate in clinics like pseudohyphal growth (Muller *et al*., 2011). Due to the recessive nature of these mutations and most yeast genes being haplosufficent (Delneri *et al*., 2008), they are able accumulate throughout the genome. These deleterious alleles would be difficult to purge from domesticated populations because of their reduced effective population size (Gu *et al*., 2005) and the low rates of sexual reproduction with only one in every 50 000 divisions being meiotic (Ruderfer *et al*., 2006). Evidence of admixture between domesticated yeasts (Schacherer *et al*., 2009), and outcrossing thought t be more common in domesticated environments (Magwene, 2014)might also account for the high levels of heterozygosity measured domesticated natural isolates (Magwene *et al*., 2011).

4.3.2. Wild Saccharomyces yeasts

Saccharomyces yeasts from wild habitats have been found around the globe mainly associated with oak tree bark and leaf litter (Kowallik & Greig, 2016). These habitats are considered more stressful, and possibly with more selective pressures like season

fluctuations on temperature, food resources and others (Goncalves *et al*., 2011; Kowallik & Greig, 2016). The *Saccharomyces* yeasts from wild habitats would easily purge any potential deleterious mutations; due to their strong selection pressures and higher rates of outcrossing. Outcrossing in wild isolates would lead to the formation of the haploid spores and haploidization would expose recessive deleterious alleles to selection, which would eliminate haploids with deleterious alleles even if they were recessive (Delneri *et al*., 2008). These deleterious alleles would be purge from wild populations due to their 'higher' rates of outcrossing (higher than domesticated isolates); population genomic studies of *S. paradoxus* indicated outcrossing to occur once every 1 000 generations for European populations or once every 3 000 generations for east Asia populations (Tsai *et al*., 2008).The lack of recessive deleterious alleles and the higher rates of outcrossing might also account for the low heterozygosity measured in wild natural isolates (Magwene *et al*., 2011).

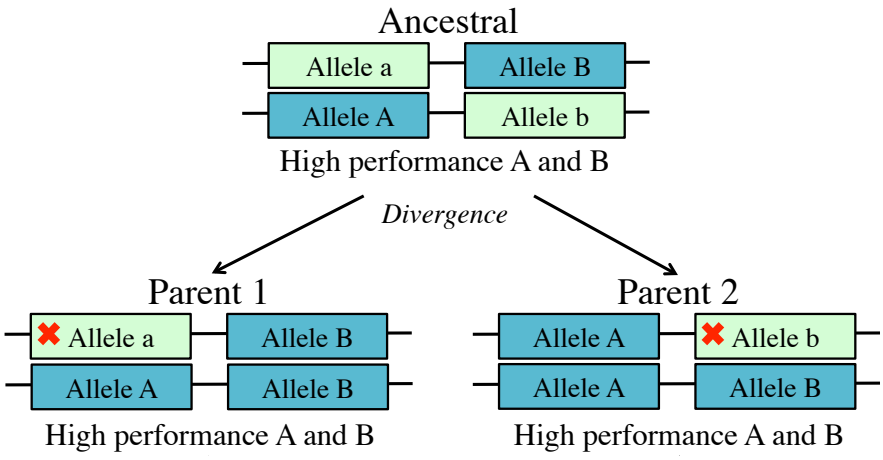


Figure 5: Accumulation of recessive deleterious alleles. Ancestral population with heterozygous locus 1 with recessive allele a (light blue) and dominant allele A (blue), and locus 2 with dominant allele B (blue) and recessive allele b (light blue), ancestral with high performance allele A and B. Ancestral population diverges in to populations Parent 1 or Parent 2; Parent 1 with deleterious mutation in recessive allele a (red cross), Parent 2 with deleterious mutation in recessive allele b. Both Parent 1 and Parent 2 with high performance allele A and B and no effect of the deleterious allele.

4.3.3. Saccharomyces yeasts isolation

A caveat of ecology studies of *Saccharomyces* is the way we isolate yeast from their natural habitats. Standard practice uses enrichment cultures where a physical isolate is placed in rich growth medium and incubated. Enrichment cultures select for strains

that grow fast and can outcompete other microbes present (Goddard & Greig, 2015), they do not provide any information about the quantity or quality (e.g. ploidy) of yeasts in a sample (Goddard *et al*., 2010). Therefore enrichment cultures might not represent *Saccharomyces* natural samples. If a sample is composed of haploid homothallic spores, enrichment culture will induce spores to autodiplodize, resulting in colonies consisting of completely homozygous diploids. Instead, if the sample is composed of vegetative mitotic diploids, enrichment cultures will only select for fast growing strains, maintaining the heterozygosity of the primary isolates (Magwene *et al*., 2011). The different numbers of heterozygous loci could therefore be skewed due to the ploidy of the of primary isolates (Goddard *et al*., 2010); if the domesticated isolates are mainly in the diploid form enrichment cultures would isolate high heterozygosity yeast strains while if wild isolates are in haploid form enrichment cultures would isolate low heterozygosity strains, as in Magwene *et al*. (2011). Nevertheless the primary isolates are repeatedly induced to sporulate and autodiplodize when they are brought into the laboratory so to form an highly homozygous and stable workable yeast strain which is easy to cross (Lindegren, 1945) and sequence (Liti *et al*., 2009)

4.4. Saccharomyces yeasts species

Saccharomyces yeasts form a complex of seven closely related, genetically tractable yeast species with similar morphologies (Figure 6) (Vaughan & Martini, 1987). These yeast come from a variety of habitats around the world (for review see Boynton & Greig, 2014): *S. cerevisiae* has been isolated from around the world and is mostly associated with human fermentations, or as potential pathogen in medical clinics but also occurs in wild habitats such as bark, soil leaf litter and even in insects, which are thought to be dispersion vectors (Fay & Benavides, 2005; Stefanini *et al*., 2012). In contrast *S. paradoxus,* has been isolated in Europe, Asia and North America but is exclusively to wild habitats such as bark, soil and leaf litter (Kowallik & Greig, 2016). Other *Saccharomyces* species like *S. mikatae, S. kudriavzevii* and *S. arboricola* have been isolated in East Asia and are found in bark, decaying leaves and soil in east Asia (Naumov *et al*., 2000; Wang & Bai, 2008). *S. eubayanus* is associated with beech trees in South America (Libkind *et al*., 2011) and similar trees in east Asia (Bing *et al*., 2014). And finally *S. uvarum,* found around the world associated mostly

with human fermentations such as wineries and breweries but also found in hardwood bark soil and insects (Almeida *et al*., 2014).

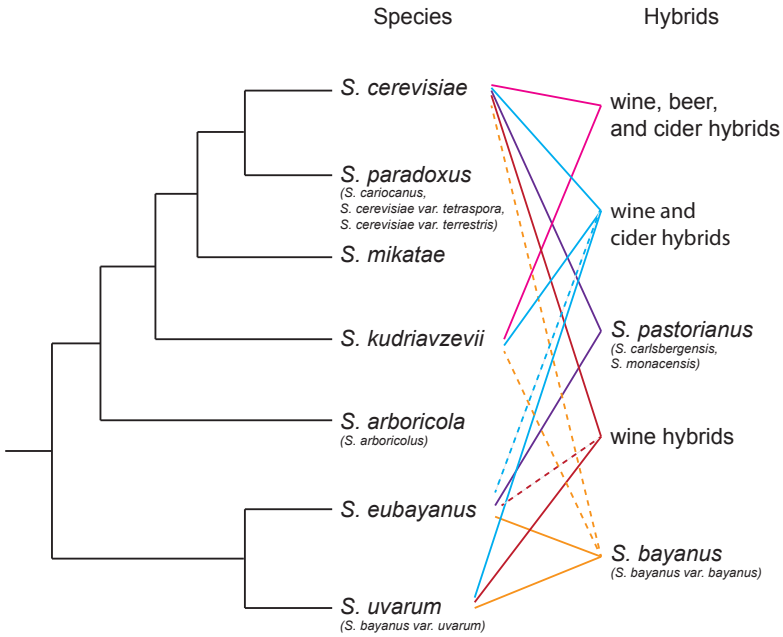


Figure 6: *Saccharomyces* species phylogenetic relationships and frequent isolated hybrids. Most recent cladogram for *Saccharomyces* yeasts on the left, and hybridization event on the right as lines. Dashed lines represent introgressions from more than two species into the hybrid. Synonyms are given in parentheses below species names (Figure reproduced without modification from Boynton & Greig, 2014 licensed under CC BY 4.0).

All the *Saccharomyces* yeasts have the ability to hybridise (Figure 6), whether artificially in the laboratory or 'naturally occurring' in wineries and breweries. So why do we consider them separate *Saccharomyces* species? Earlier taxonomy of *Saccharomyces* yeasts was based on carbon and nitrogen assimilation tests or homo-heterothallism of the yeast populations, however researchers soon realized that this type of classification was not suitable (Naumov, 1996). Based on the modern concept of biological species:

'species are groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups' isolated.'- (Mayr, 1942)

So even though *Saccharomyces* yeasts species can hybridise, their offspring have very low fertility which creates an intrinsic post-zygotic barrier between *Saccharomyces* species (Naumov, 1996).

4.5. Saccharomyces yeasts hybridisation

There is evidence of naturally occurring hybrids between yeast species, especially between domesticated populations of yeasts used for the production of alcoholic beverages (Figure 6) (de Barros Lopes *et al*., 2002; González *et al*., 2006; Lopandic *et al*., 2007). The detection of naturally occurring hybridisation events in domesticated habitats, such as wine and beer fermentations, indicates that the hybrids might have been selected due to their beneficial and convenient characteristics. For hybrid species to occur, two different yeast species must cross and create an inter-specific F1 hybrid that is viable and can eventually produce viable spores. Normally inter-specific F1 hybrids created under laboratory conditions are viable and can grow asexually but are also infertile with only 1% viable spores (Greig, 2009).

4.5.1. Hybridisation in domesticated habitats

New genomic tools enabled the identification of several hybridisation events between *Saccharomyces* species, and all of them have been traced back to domesticated habitats such as wineries and breweries (Boynton & Greig, 2014). New techniques, like Whole Genome Sequencing (WGS) were used to detect hybrids between *S. cerevisiae, S. kudriavzevii, S. eubayanus,* and *S. uvarum* (de Barros Lopes *et al*., 2002; González *et al*., 2006; Lopandic *et al*., 2007). Lager beer yeast *S. pastorianus* is one of the best case studies of hybridization, however the circumstances of this hybridisation event are relatively unknown. *S. pastorianus,* a cross between *S. cerevisiae* and *S. eubayanus*, was first used in Bavaria to brew lager beer around the sixteen-century. But *S. eubayanus* was only recently described and isolated from southern beech trees in South America (Libkind *et al*., 2011) or similar trees in east Asia (Bing *et al*., 2014). So there are two possible explanations, either *S. eubayanus* from South American population was brought to Europe by transatlantic traders and then crossed with *S. cerevisiae* (Libkind *et al*., 2011), or *S. eubayanus* from east Asia populations was brought to Europe by traders using the Silk Road (Bing *et al*., 2014). Next-Generation Sequencing (NGS) data has favoured the Silk Road hypothesis by having *S. eubayanus* isolated from east Asia with sequence similarities of 99.8% to *S. pastorianus* over 99.4% of *S. eubayanus* isolated from south America; this depicts the power of the new tools we have to identify hybridisation events (Bing *et al*., 2014).

4.5.2. Hybridisation in wild habitats

No hybridisation events between *Saccharomyces* species have been detected in wild habitats or between exclusively wild *Saccharomyces* species such as *S. paradoxus*. Even though *S. cerevisiae* and *S. paradoxus*, *S. kudriavzevii* can be sympatric, hybrids between these species have not been found in wild habitats (Sampaio & Goncalves, 2008). This is surprising, because *S. cerevisiae* and *S. paradoxus* have little to no phenotypic differences and easily hybridise under laboratory conditions (Greig, 2009). However *S. cerevisiae* and *S. paradoxus* are 13% genetically diverged, according to Single Nucleotide Polymorphism (SNP) data (Liti *et al*., 2009). Hybridisation between *S. cerevisiae* and *S. paradoxus* would require (1) the parental yeast species to co-occur, which they do but are not necessarily sharing the same ecological niche (Sampaio & Goncalves, 2008), (2) to reproduce sexually, which can be a rare event (Ruderfer *et al*., 2006; Tsai *et al*., 2008), and (3) their F1 hybrid should be viable and eventually also fertile, which might be difficult but not impossible because 99% of the spores produced by the F1 hybrid are unviable (Greig, 2009).

Saccharomyces yeasts

Characteristic	Domesticated isolates	Wild isolates
Habitat	Wineries, Breweries and Clinics	Bark, Soil and Leaf litter
Selection	Strong directional	Strong
Sex rate	1: 50 000	1: 1 000 – 1: 3 000
Heterozygosity	High	Low
Genome	Accumulation of many recessive deleterious alleles	No accumulation of few recessive deleterious alleles

Summary Table 2: Comparison between domesticated and wild isolates of *Saccharomyces* yeasts.

5. Heterosis studies in *Saccharomyces* yeasts

Heterosis in *Saccharomyces* yeasts was first reported by Lindegren (1953) when the F1 hybrid had an advantage in maltose fermentation due to complementation of one dominant enzyme over a less advantageous enzyme. Nowadays we report heterosis in a more general view as the higher asexual growth or fitness advantage of the F1 hybrid in comparison to one or both parents (Zörgö *et al*., 2012). The fitness advantage can be a measure of competitive growth, when there is a direct competition between the F1 hybrid and one of its parents, or a measure of maximum growth rate in isolation, when the asexual growth of the F1 hybrid and its parents are measured independently (Zörgö *et al*., 2012; Plech *et al*., 2014; Shapira *et al*., 2014). The F1 hybrid fitness can be compared to the mid-parent (an estimation of the parent's average fitness- MP), to the best-parent (or the parent with the highest fitness- BP), or to the worst-parent (or the parent with the lowest fitness- WP). Heterosis will be a measure of the difference between the F1 hybrid and the mid-parent or best-parent fitnesses. Heterosis can be positive if it is above mid-parent or best-parent measures or negative when it is below the mid-parent or worst-parent measures (Figure 7). Most of the studies on heterosis in yeast have focused on intra-specific F1 hybrids using crosses of *S. cerevisiae* populations (Zörgö *et al*., 2012; Plech *et al*., 2014; Shapira *et al*., 2014). Only one study (Blein-Nicolas *et al*., 2015), and now the studies of my thesis, have tried to detect heterosis in yeast inter-specific F1 hybrids.

5.1. Heterosis in intra-specific S. cerevisiae F1 hybrids

Zörgö *et al*. (2012) crossed nine homozygous *S. cerevisiae* strains and compared the fitness of the F1 hybrids to that of the autodiplodized parents in 56 different environments. This study found positive mid-parent heterosis (MPH) to be common and a fraction of the F1 hybrids had positive best-parent heterosis (BPH). Heterosis was not correlated with genetic divergence of parental populations. These results were in agreement with dominance mechanism for heterosis where complementation of recessive deleterious alleles of one parent are complemented by the high performance alleles of the other parent, and this relation is reciprocal leading to a higher fitness of the F1 hybrid over its parents (Zörgö *et al*., 2012).

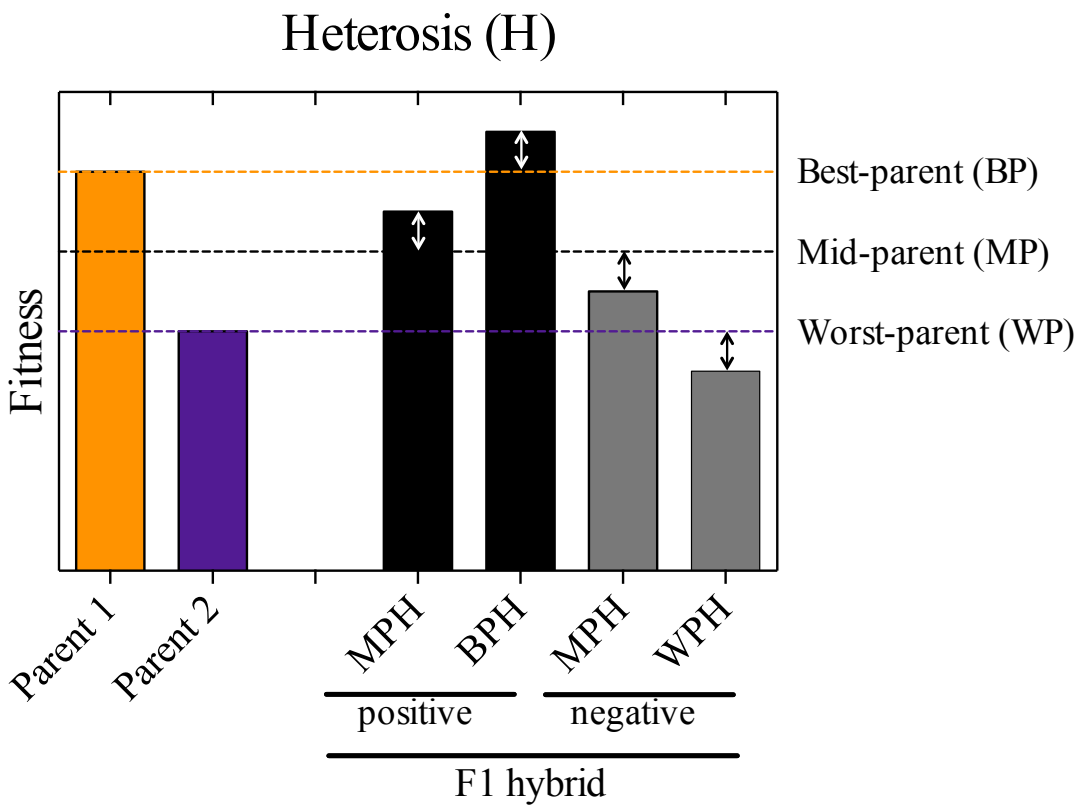


Figure 7: Measures of heterosis by comparing F1 hybrid to parental fitnesses. Parent 1 in red, Parent 2 in yellow, F1 hybrid in orange and light orange. Best-Parent (BP), Mid-parent (MP) and Worst-parent (WP) fitness in dashed lines. F1 hybrid with positive heterosis in orange, and F1 hybrid with negative heterosis in light orange. Advantage or disadvantage in relation to BP, MP or WP described as arrows. By order: parent 1, parent 2, positive MPH; positive BPH; negative MPH; negative WPH.

5.2. Heterosis in domesticated intra-specific F1 hybrids

To test the effect of domestication on heterosis Plech *et al*. (2014) measured the fitness of intra-specific F1 hybrids made from crosses between domesticated populations or crosses between wild populations of *S. cerevisiae*. They crossed 22 homozygous parents that have been previously autodiplodized and measured the fitness of the F1 hybrids in eleven different environments. This study showed crosses between domesticated populations were characterized by mid-parent heterosis while crosses between wild populations exhibit no heterosis. Heterosis was correlated with the genetic divergence of domesticated parental populations. As in the previous study, heterosis was mainly due to dominance. These results also suggested that the autodiplodized forms (parents) of domesticated natural isolates had lower fitness than

the wild natural isolates. Heterosis was exclusive to domesticated crosses and this might be attributed to the accumulation of recessive deleterious alleles in domesticated habitats (Plech *et al*., 2014).

5.3. Heterosis in wild intra-specific F1 hybrids

Shapira *et al*. (2014) tested the fitness of intra-specific F1 hybrids between *S. cerevisiae* wild populations. They crossed sixteen homozygous parents and measured the individual fitness of the resulting F1 hybrids in five different environments. This study showed mid- but not best-parent heterosis for the F1 hybrids of wild *S. cerevisiae* populations. As in previous studies they identified as a genetic mechanism behind heterosis, moreover they invoked overdominance and epistasis mechanisms to explain instances where dominance could not solely explain the heterosis identified (Shapira *et al*., 2014).

5.4. Heterosis in inter-specific F1 hybrids

Inter-specific crosses between different species of *Saccharomyces* yeasts have not been widely described in the literature. Thus direct measures of heterosis have not been reported as in the previous cited studies, which is remarkable because different yeast species can readily hybridise and have viable offspring. In addition several hybridisation events have been observed between domesticated yeast species (de Barros Lopes *et al*., 2002; González *et al*., 2006; Lopandic *et al*., 2007).

Blein-Nicolas *et al*. (2015) used inter-specific F1 hybrids between *S. cerevisiae* and *S. uvarum*, a previously reported hybridisation for wine and cider fermentations (Figure 6). They measured heterosis by comparing the protein inheritance of the inter-specific F1 hybrid with intra-specific F1 hybrids between *S. cerevisiae* and *S. uvarum* populations in two different temperatures. The inter-specific F1 hybrids showed several genes with positive mid-parent heterosis for protein abundance data collected by mass-spectometry, and a smaller set of genes with best-parent heterosis, while intra-specific F1 hybrids showed a balance between positive and negative heterosis. Blein-Nicolas *et al*. (2015). These results were not explained by any of the genetic mechanism previously put forward but advanced that heterosis for proteins occurred (Blein-Nicolas *et al*., 2015).

5.5. Summary of heterosis studies

Heterosis was first described in *Saccharomyces* yeasts over half a century ago (Lindegren, 1953), the current heterosis studies focus on the the genetic mechanisms behind the growth advantages of the F1 hybrids by crossing domesticated or wild parental strains of divergent population of the same species (Zörgö *et al*., 2012; Plech *et al*., 2014; Shapira *et al*., 2014), or focus on the protein inheritance of F1 hybrids from different yeast species (Blein-Nicolas *et al*., 2015). However questions around heterosis in *Saccharomyces* yeasts remain unanswered; I would like to test if heterosis is present in *Saccharomyces* F1 hybrids made by crossing different *Saccharomyces* species or from crossing two divergent populations of a wild *Saccharomyces* yeast in different environments, I would also like to examin how heterosis manifests at the transcriptome level in different environments, and finally I would like to test if the heterosis is persistent due to the parental background of *Saccharomyces* yeasts or if it is highly dependent on the laboratory manipulation of the parental strains. I hope this thesis will elucidate some of these questions and provide a new perspective of heterosis in *Saccharomyces* yeast.

Heterosis in *Saccharomyces* yeasts

F1 hybrids	Authors	Heterosis	Mechanism
Intra-specific	*Zörgö* (2012)	Mid and Best-parent heterosis for *S. cerevisiae* cross	Dominance
	Plech (2014)	Mid-parent heterosis for *S. cerevisiae* domesticated cross	Dominance
	Shapira (2014)	Mid-parent heterosis for *S. cerevisiae* wild cross	Dominance, Overdominance and Epistasis
Inter-specific	*Blein-Nicolas* (2015)	Mid and Best-parent heterosis for protein abundance in inter-specific *S. cerevisiae* and *S. uvarum* cross	?

Summary Table 3: Description and comparison between heterosis studies of intra-specific F1 hybrids and inter-specific F1 hybrids.

Thesis Outline & Authors Contributions

My thesis explores heterosis and the mechanisms contributing to heterosis in a set of yeast F1 hybrids by comparing hybrid's asexual fitness with those of its parents. I used *Saccharomyces* yeasts because of several attractive features like their the ability to readily produce viable F1 hybrids from crosses between diverged yeast species, and their straightforward measures of asexual fitness. Due to the lack of empirical data, I wanted to find evidence of positive heterosis for yeast F1 hybrids originated from a variety of different crosses. Then I wanted to determine if there were any specific patterns of heterosis at a transcriptome level. And finally, I review of my work and other similar heterosis studies, and tested the importance of the ecological background of F1 hybrids on heterosis.

This thesis is organized in three independent but related chapters, described below. I was responsible for conducting the experiments, data analysis and writing of all three chapters, therefore I am the first author of all three chapters of my thesis. My PhD supervisor Dr. Duncan Greig was the mind behind the theory and supervised both experimental and analytic work; therefore Dr. Duncan Greig is the last author of all the three chapters of this thesis. Dr. Rike Stelkens and Dr. David Rogers had a significant intellectual input in the analytic work of Chapter I and Chapter II respectively; therefore they are mention as authors in both chapters.

In **Chapter I**, I measured heterosis in a set of F1 hybrids by direct competitions between F1 hybrid and one of its parents in pairwise comparisons. Inter-specific F1 hybrids were made by crossing *S. cerevisiae* and *S. paradoxus,* and intra-specific F1 hybrids by crossing *S. paradoxus* wild populations. I found the competitive growth of the F1 hybrids relative to the average of its parents, a measure of mid-parent heterosis, was correlated with the difference in parental growth relative to their F1 hybrid, a measure of phenotypic divergence. Inter-specific F1 hybrids showed stronger heterosis than intra-specific hybrids. In order to manipulate parental phenotypic divergence independently of genotype, I also measured the competitive growth of a single inter-specific F1 hybrid relative to both its parents in twelve different environments. I not only identified a strong relationship between parental

phenotypic divergence and mid-parent heterosis as before, but also, a weak relationship between phenotypic divergence and best-parent heterosis. These results suggested reciprocal complementation of deleterious alleles to be one of the main mechanisms behind heterosis. My work proposes that wild isolates of *Saccharomyces* yeasts have a lower load of recessive deleterious alleles than domesticated isolates, which explains the absence of heterosis for crosses between wild isolates (intra-specific F1 hybrids) and the presence of heterosis for crosses between wild and domesticated isolates (inter-specific F1 hybrids).

In **Chapter II**, I used a previously tested inter-specific F1 hybrid that showed high levels of best-parent heterosis, and compare its transcription profile with the transcription profile of its parents in two different environments. The environments chosen gave the F1 hybrid an advantage in relation to both parents and each environment favoured one or the other parent. F1 hybrid transcription profile was more similar to fitter parent for the specific environment the hybrid was in, moreover allelic specific expression was prevelant in the F1 hybrid, which unrevals the importance of *cis*-regulated transcription. These results were consistent with complementation of low fitness alleles of one parent by high fitness alleles of the other parent, as a mechanism for heterosis at the transcriptome level. Interestingly, the F1 hybrid not only regulated its transcription by inducing fitter parental alleles for a specific environment, but also had the ability to modify its transcription upon environmental change. My work proposes that the ability to regulate both parental alleles and preferentially transcribe the more advantageous parental alleles for multiple loci rendered the F1 hybrid an intrinsic advantage that might explain the occurrence of heterosis.

In **Chapter III**, I used *S. cerevisiae* primary isolates from domesticated and wild habitats with known heterozygosity and measured their spore viability and asexual fitness in comparison to corresponding derived forms. High heterozygosity was characteristic of domesticated primary isolates and low heterozygosity of wild primary isolates. From the primary isolates I formed monosporic clones by self-fertilization or autodiplodization of one derived haploid spore, and I also made inbred forms by crossing two monosporic clones from the same primary isolate. I identified a

strong negative relationship between heterozygosity of the primary isolates and spore viability, and a negative relationship between heterozygosity of the primary isolates and the growth of their monosporic clones. For inbred forms fitness was similar to their respective primary isolates for both high and low heterozygosity. The results supported the claim that highly heterozygous domesticated strains of *S. cerevisiae* carry multiple recessive deleterious alleles, which upon haploidization or autodiplodization become exposed and might render their haploid and the monosporic clones unviable or less well adapted. However by crossing monosporic clones fitness was restored in the inbred forms by complementation of the recessive deleterious alleles. My work reveals that most studies that identified heterosis used domesticated monosporic clones as parents of the F1 hybrid, as in our study crosses among divergent monosporic clones increase fitness by complementation of dominant high-fitness alleles, giving the illusion of heterosis.

In the next pages, I went through a detail description of the three projects in three separated chapters. The chapters are divided into Introduction, Material and Methods, Results, Discussion and Supplementary Material, and the corresponded figures are inserted throughout the chapters. In the end of the chapters I prepared an overall Conclusion & Perspectives advanced by my PhD work. And finally, I provided a Glossary & List of Abbreviations with the goal of facilitating reading and also a overall Figures & Tables index.

Chapter I

Heterosis in hybrids within and between yeast species.

Bernardes, J. P.[1], Stelkens, R[1]. B. & Greig, D.[1,2]

1. Max Planck Institute for Evolutionary Biology, Experimental Evolution Group, Plön, Germany

2. The Galton Laboratory, Department of Genetics, Evolution, and Environment, University College London, London, UK

Published in Journal of Evolutionary Biology

December, 2016

ISSN: 1420-9101

DOI: 10.1111/jeb.13023

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13023

Heterosis in hybrids within and between yeast species

J. P. BERNARDES*, R. B. STELKENS* & D. GREIG*†

*Max Planck Institute for Evolutionary Biology, Plön, Germany
†The Galton Laboratory, Department of Genetics, Evolution, and Environment, University College London, London, UK

Keywords:

F1 hybrid;
heterosis;
hybrid vigour;
Saccharomyces cerevisiae;
Saccharomyces paradoxus;
speciation.

Abstract

The performance of hybrids relative to their parents is an important factor in speciation research. We measured the growth of 46 *Saccharomyces* yeast F1 interspecific and intraspecific hybrids, relative to the growth of each of their parents, in pairwise competition assays. We found that the growth of a hybrid relative to the average of its parents, a measure of mid-parent heterosis, correlated with the difference in parental growth relative to their hybrid, a measure of phenotypic divergence, which is consistent with simple complementation of low fitness alleles in one parent by high fitness alleles in the other. Interspecific hybrids showed stronger heterosis than intraspecific hybrids. To manipulate parental phenotypic divergence independently of genotype, we also measured the competitive growth of a single interspecific hybrid relative to its parents in 12 different environments. In these assays, we not only identified a strong relationship between parental phenotypic divergence and mid-parent heterosis as before, but, more tentatively, a weak relationship between phenotypic divergence and best-parent heterosis, suggesting that complementation of deleterious mutations was not the sole cause of interspecific heterosis. Our results show that mating between different species can be beneficial, and demonstrate that competition assays between parents and offspring are a useful way to study the evolutionary consequences of hybridization.

Introduction

When individuals from different species or from genetically distinct populations mate, they may produce hybrid offspring (Barton & Hewitt, 1985).

Hybridization can bring alleles together in combinations that have never before been exposed to natural selection, often with unpredictable results. Genetic incompatibilities between independently diverged alleles at different loci might reduce hybrid fertility or viability, restricting gene flow between diverging populations (Orr & Turelli, 2001) through Bateson–Dobzhansky–Muller (hereafter BDM) incompatibilities. But interactions among novel combinations of alleles from different populations or species can also increase aspects of hybrid fitness (Shull, 1948). There is evidence from a variety of taxa including plants (Rieseberg *et al.*, 2003), fish (Nolte & Sheets, 2005), insects (Schwarz *et al.*, 2005) and yeast (Stelkens *et al.*, 2014) that hybrids can colonize new environments which are inaccessible to their parents. Thus, hybridization can increase or decrease fitness, and both promote or prevent speciation (Barton & Hewitt, 1985).

It is difficult to determine experimentally the factors that can enable hybrids to outcompete their parents. Various traits contribute to the single trait called fitness, including traits that affect viability (e.g. vigour, survival, growth rate) and those that affect sexual reproduction (e.g. mating success, fertility, fecundity). Hybridization can simultaneously improve some fitness-determining traits, such as vigour, while diminishing others, such as fertility. Different generations of hybrids may also be affected differently; for example, 'hybrid breakdown' describes a reduction in fitness affecting later, but not earlier, generations of hybrids, due to homozygous recessive allelic incompatibilities (Edmands, 2002; Stelkens *et al.*, 2015). And because

Correspondence: Joana P. Bernardes, Max Planck Institute for Evolutionary Biology, August-Thienemann Straße 2, 24306 Plön, Germany.
Tel.: +49 4522 763 234; fax: +49 4522 763 260;
e-mail: bernardes@evolbio.mpg.de

hybridization can greatly increase phenotypic variance, it is possible for some hybrid individuals to be much fitter than their parents, even when most are much less fit or even inviable. Hybrid effects on fitness may also depend on the local environment; for example, BDM incompatibilities often depend on environmental conditions (Nosil, 2012). Thus, to evaluate the evolutionary potential of a hybrid, it is helpful to sample hybrid fitnesses in multiple environments (Lexer *et al.*, 2003; Rieseberg *et al.*, 2003; Stelkens *et al.*, 2014).

Many of these complexities can be avoided using a simple experimental model system. The facultatively sexual yeasts of the *Saccharomyces sensu stricto* species complex are ideal for experimental studies of hybridization. They are well characterized genetically and phenotypically, they have short generation times, and they are easy to cultivate in large populations under controlled and repeatable conditions (Scannell *et al.*, 2011). All members of the *sensu stricto* complex can mate with each other, forming diploid F1 hybrids (Naumov, 1996). Diploids do not have sexes or mating types, but they can undergo meiosis to produce haploid gametes of two mating types, 'a' and 'alpha', which can fuse to restore diploidy and complete the sexual cycle. Because both haploids and diploids can undergo mitosis, individuals can be isolated and propagated as clones, allowing the effects of hybridization to be studied at all life stages and across many generations. Different genotypes can be genetically marked so that they and their offspring can be distinguished, allowing competitive growth assays in a common environment. These advantages allow different methods that are not possible in traditional plant or animal model systems, and although the results from yeast may not be directly applicable to obligate outcrossing species, they are likely to be relevant to a large number of other sexual microbial eukaryotes.

The most striking and best studied characteristic of diploid F1 hybrids between different *Saccharomyces sensu stricto* species is their greatly reduced sexual fertility: <1% of the gametes they produce are viable (Hunter *et al.*, 1996). BDM incompatibilities contributing to this interspecific F1 hybrid gamete inviability have not been found (Kao *et al.*, 2010). Instead, antirecombination has been shown to be the major cause of yeast F1 hybrid sterility (Hunter *et al.*, 1996). When chromosomes from different parents are sufficiently diverged, they cannot crossover during meiosis and so fail to segregate accurately. The genomes of *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* differ at about 14% of nucleotides, which impairs chromosome crossing-over and meiotic segregation so much that most gametes produced by F1 hybrids lack essential chromosomes and are inviable. Despite the low number of gametes that survive F1 hybrid meiosis, those that do, and that are capable of mating, can form F2 hybrids. Some of these F2 hybrids are both viable and sexually fertile, capable of producing viable gametes themselves, yet reproductively isolated from their parents by their new chromosome compositions, thus demonstrating a potential mechanism of hybrid speciation (Greig *et al.*, 2002).

Although much work has concentrated on the reduced sexual fertility of interspecies *Saccharomyces* F1 hybrids, there has been relatively little work on the competitive ability of the F1 hybrids themselves. This is surprising, because the first challenge a new F1 yeast hybrid faces is not its sexual fertility, but its viability and ability to compete under asexual growth. Interspecies F1 hybrid vigour has not been systematically quantified by competitive asexual growth assays against parent species, as far as we know. As most yeast reproduction is by asexual diploid mitosis, the competitive growth of F1 hybrids is likely to determine the success of further generations when hybrids compete for the same resource as their parents: in principle, high F1 asexual competitiveness could completely compensate for their low sexual fertility, or conversely, low F1 asexual competitiveness might greatly strengthen the barrier already established between species. Thus, the ability of F1 hybrids to compete against their parents is evolutionarily important. Furthermore, F1 hybrids are ideal for studying the net contribution of all genetic effects of hybridization at all loci: a single diploid genotype captures the entire range of genetic differences between its two parents. This contrasts to F2 hybrids in which parental differences between loci and within loci are reduced by recombination and segregation respectively, as a result of the preceding sexual cycle. F2 hybrids derived from the same two parent species can vary genetically, containing any combination or proportion of parental alleles, and therefore being more or less affected by hybridization. This presents a sampling problem for researchers studying speciation, particularly with interspecific yeast crosses, where many F2 hybrid individuals have zero viability. For these practical and evolutionary reasons, we set out to measure the factors that affect the competitive ability of F1 hybrids relative to their parents.

When genetically diverged parents mate, their F1 hybrid offspring inherit a complete set of alleles from both parents and might therefore be expected to be phenotypically intermediate. However, parental phenotypes often interact nonadditively, producing hybrid trait values that are different from the average of the parental trait values, and which can even fall outside the range of parental values. For many crosses, these nonadditive genetic interactions may reduce viability enough that the F1 hybrids are rendered completely inviable, preventing traits from being quantified. But fitness-determining traits can also be enhanced by the high heterozygosity of hybrids relative to their parents: this is known as hybrid vigour or heterosis (Shull, 1948). In this article, we will use the term positive heterosis (or sometimes just heterosis) to refer to an

increase in fitness of a F1 hybrid due to heterozygosity and the term negative heterosis when F1 hybrid fitness is diminished due to heterozygosity.

It should be possible to predict the strength and sign of heterosis in F1 hybrids from the characteristics of their parents, so many experimental studies have measured the effect of evolutionary divergence between parents on F1 hybrid traits (for review see Edmands, 2002). One would expect that there should be an optimum level of divergence, between the occurrences of inbreeding depression and outbreeding depression, at which positive heterosis for fitness is maximized, so natural selection should act on mating systems to achieve this intermediate level of outcrossing (Waser, 1993). However, although some researchers do find such a humped-shaped relationship between parental evolutionary divergence and hybrid traits (Moll *et al.*, 1965), others find only positive relationships (Xiao *et al.*, 1996; Gonzáles *et al.*, 2007), negative relationships (McClelland & Naish, 2007; Pekkala *et al.*, 2012) or no relationship at all (Hung *et al.*, 2012). One problem with such experiments is determining which traits to measure. Yeast has an advantage over other study systems in that experimental strains can be made homozygous and propagated as pure clones, so parents and hybrids can be grown simultaneously in a common environment to determine their direct competitive ability in standardized and repeatable – albeit artificial and highly simplified – conditions. This method is commonly used in experimental evolution studies with asexual microbes to determine relative fitness (Lenski, 1991). Relative fitness is the evolutionary important measure: this is what natural selection acts to improve and in competition in batch cultures allows several fitness-associated traits – such as faster maximal growth rate, shorter lag phase, higher carrying capacity or better survival in stationary phase (Vasi *et al.*, 1994) – to be incorporated in a single evolutionary-relevant measurement, albeit one that excludes sexual parts of the life cycle. Studying heterosis in yeast can also help address another practical problem, in that the best measure of evolutionary distance between parents is not obvious: geographic distance, difference in local environments, general phenotypic divergence in multiple traits and general genetic divergence in DNA sequences or markers have all been used (for review see Edmands, 2002). The ability of yeast to grow clonally allows the same genotypes to be tested and retested in different ways, potentially allowing one measure of parental divergence to be manipulated independently of another. For example, genetic distance can be fixed and phenotypic distance varied by retesting the same genotypes in different environments in which their phenotypic differences vary.

Here, we used F1 hybrids between wild *S. paradoxus* parents differing by up to 4% in nucleotide divergence and between *S. paradoxus* and *S. cerevisiae* parents differing by up to 14%. These crosses represent much greater genetic divergence than the intraspecific *S. cerevisiae* hybrids used in previous yeast studies on heterosis, which were <1% divergent according to SNP data (Zörgö *et al.*, 2012; Plech *et al.*, 2014; Shapira *et al.*, 2014). Rather than measuring growth rates in isolation, we determined the growth of these hybrids relative to their parents in direct competition. We determined the relationship between heterosis and both genetic divergence (genome sequence divergence) and phenotypic divergence (the difference in competitive growth) of the parents. Then, to determine the relationship between heterosis and phenotypic divergence independently from genetic divergence, we retested the competitive growth of a single interspecific hybrid relative to its parents under different environmental conditions, to manipulate parental phenotypic divergence.

Materials and methods

Strains and hybrid crosses

We used 32 homozygous strains of *S. paradoxus* and *S. cerevisiae* from the National Collection of Yeast Cultures (NCYC, http://www.ncyc.co.uk/) to produce 46 F1 hybrids: 28 intraspecific hybrids between *S. paradoxus* and *S. paradoxus* and 18 interspecific hybrids between *S. paradoxus* and *S. cerevisiae* (Stelkens *et al.*, 2014). Strains and crosses were selected to maximize the ranges of genetic and phenotypic divergence within a manageable set of hybrids (Stelkens *et al.*, 2014). Strains are available on request (see Table S1). Parental strains came from around the world. Most of the *S. paradoxus* strains were collected from oak trees, but *S. cerevisiae* strains came from diverse habitats with high ecological diversity such as soil, trees, diseased human tissue, faeces, insects, fruit, beer and wine (Liti *et al.*, 2009; see Table S1). F1 hybrid strains were made by mixing equal volumes of the haploid parental strains of opposite mating types, mating overnight on YEPD agar (1% yeast extract, 2% peptone, 2% glucose, 2% agar), streaking onto new YEPD plates and replica-plating the resulting single colonies onto KAC agar (1% potassium acetate, 0.1% yeast extract, 0.05% glucose, 2% agar) to induce sporulation. After 48 h incubation, we identified the colonies that had sporulated (and were therefore founded by mated diploids) using a microscope and selecting the corresponding colony from YEPD plate. These pure diploid F1 hybrids were stored frozen at −80 °C 20% glycerol stock for later use. The parental haploid strains used to make F1 hybrids strains were genetically marked with one of the two dominant homozygous alleles conferring resistance to the antibiotics G418 and hygromycin: the resulting F1 hybrid was resistant to both antibiotics (*ho::HYGMX/ho::HYGMX, ura3::KANMX/ura3::KANMX*), whereas the

parental diploid strains were homozygous for the wild-type alleles (*HO/HO, URA3/URA3*) and thus sensitive to the antibiotics. Gene transformation was carried out by following methods in Gietz & Woods (2002).

Measuring heterosis using competitive growth assays

We measured the competitive growth of every diploid hybrid relative to both of its diploid parents using replicated assays in 5 mL liquid YEPD (1% yeast extract, 2% peptone, 2% glucose) shaken cultures at 30 °C. Each assay tested a hybrid strain against one of its parents. The hybrid and both its parental strains were grown in isolation for 24 h before mixing the hybrid with each parent separately in equal volumes. A 50 μL sample of the mixture was used to inoculate 5 mL of fresh sterile medium, and the initial (t0) cell number of the hybrid and parental strains was estimated by taking a 100 μL sample, serially diluting it and plating it to solid YEPD agar before incubating it for 2 days to yield ~200 colonies. The proportion of hybrid colonies was determined by replica-plating to YEPD agar, supplemented with 400 mg of the antibiotic G418 in every litre of medium (0.04% final concentration of G418). Multiplying by the dilution factor allowed the initial number of the hybrid and parent cells in the culture to be determined. Meanwhile, the freshly inoculated medium was incubated for 1 day before a second 100 μL sample was removed, and the final (t1) number of each cell type was determined by serial dilution and replica-plating as before. The competitive growth of the hybrid relative to its parent was determined by the ratio of their Malthusian growth parameters (Lenski, 1991). Each assay was replicated independently three times using the same strains but different primary cultures, and then, the mean of these three competitive growth measurements was taken and log-transformed. Every hybrid was tested against both of its parents, producing two log-transformed hybrid competitive growth values, one relative to each parent. The higher value is a measure of the performance of a hybrid relative to its less competitive parent, whereas the lower values represent its performance against the more competitive parent, which we therefore use as our measure of best-parent heterosis. Thus, the average of the two hybrid competitive growth values is our measure of mid-parent heterosis. Heterosis values below zero mean the parent(s) outperform the hybrid, whereas heterosis values higher than zero mean the hybrid outperforms its parent(s). The absolute difference between the two values represents the difference between the competitive growths of the two parents relative to their hybrid and is therefore a measure of phenotypic divergence for competitive growth against a common competitor (the hybrid). Genetic divergence between the parents in each cross was calculated using SNP data (personal communication with Gianni Liti), by dividing the number of bases that differed between species by the total number of aligned bases.

To quantify any systematic effect on competitive growth due to the genetic markers used to distinguish hybrids (*ho::HYGMX/ho::HYGMX, ura3::KANMX/ura3::KANMX*) from their parents (*HO/HO, URA3/URA3*), we competed each parental diploid against a marked (*ho::HYGMX/ho::HYGMX, ura3::KANMX/ura3::KANMX*) version of the same parental strain, under the same conditions used for the competitive growth assays between parents and offspring described above.

Effect of environment on heterosis

To determine the effect of phenotypic divergence independently from the genetic divergence of the parent strains, we measured heterosis in a single interspecific hybrid under different environmental conditions. To facilitate future investigation into the molecular mechanisms of heterosis, we chose two genetically tractable laboratory strains as parents: s288c (*S. cerevisiae*) and N17 (*S. paradoxus*). We again used genetic markers to identify competing strains in our growth assays. The parents were marked with dominant drug resistance cassettes conferring resistance to G418 and to hygromycin as a heterozygote in the same locus, *ura3* (i.e. *ura3::KANMX/ura3::HYGMX*). The hybrid was simply homozygous for a *ura3* deletion (thus *ura3/ura3*) and sensitive to the two drugs. Gene transformation was carried out by following methods in Gietz & Woods (2002). Mid-parent heterosis and best-parent heterosis were measured as before using competitive growth assays replicated three times, except that instead of conducting the assays in YEPD medium at 30 °C, we used 12 different media. Assays were all conducted in shaken liquid minimum medium with added uracil (MIN+URA: 0.67% yeast nitrogen base without amino acids, 2% glucose, 0.003% uracil) with the following supplements: caffeine (10%, 30 °C), zinc sulphate (10%, 30 °C), citric acid (10%, 30 °C), acetylsalicylic acid or aspirin (10%, 30 °C), sodium chloride (10%, 30 °C), peroxide (10%, 30 °C), nipagin (10%, 30 °C), ethanol (1%, 30 °C), lithium acetate (1%, 30 °C), dimethyl sulphoxide or DMSO (1%, 30 °C) as well as at 15 and 30 °C with no supplement.

To test for any systematic effect on growth of the genetic markers used to identify competing strains, we ran control assays in which each drug-resistant diploid (*ura3::KANMX/ura3::HYGMX*) s288c and N17 parent was competed against an isogenic drug-sensitive diploid containing only a homozygous *ura3* deletion (*ura3/ura3*) diploids. Assays were conducted as described above in the different supplemented media (not including temperature this time) and replicated three times independently, using the same strains but different primary cultures.

Statistical analysis

All statistical analyses were performed in R (R version 3.0.2, packages: 'lawstat' version 2.4.1, 'lme4' version 1.17 and 'nlme' version 3.1-120). Individual statistical tests are listed in the Results. All hybrid relative competitive growth measures were log-transformed to produce measures of positive or negative heterosis for competitive growth. We tested the intraspecies and interspecies competitive growth for normality (Shapiro–Wilk test: mid-parent heterosis: $W = 0.953$, $P = 0.063$, Best-parent heterosis: $W = 0.984$, $P = 0.781$) and homogeneity of variances (Levene's test: mid-parent heterosis: $F_{1,44} = 0.041$, $P = 0.840$, best-parent heterosis: $F_{1,44} = 0.001$, $P = 0.971$) to ensure the correct use of parametric tests.

Results

Colony counts from all competitive growth assays are provided as Tables S1 and S2. Note that although some authors use the word heterosis only to refer to cases of hybrid outperforming parents, our measure of heterosis can be negative (see Methods), which thus describes parent outcompeting their hybrids.

Heterosis and genetic divergence in different crosses

Genetic divergence ranged from 0.06% to 14% (Table S1), but because of the global population structure of *S. paradoxus*, divergence clustered into four categories (Fig. 1): hybrids between *S. paradoxus* parents from within the same continent (i.e. within Europe, Asia or America, resulting in <1% sequence divergence) with similar competitive growth than the parental average (group mean = −0.001%), hybrids between *S. paradoxus* parents from adjacent continents (i.e. between Europe and Asia, resulting in 1–2% sequence divergence) with similar competitive growth to the parental average (group mean = 0.012%), hybrids between *S. paradoxus* parents from continents isolated by oceans (i.e. crosses between America and Europe and between America and Asia, resulting in 3–4% sequence divergence) also with similar competitive growth to the parental average (group mean = −0.001%) and finally interspecific hybrids between *S. paradoxus* and *S. cerevisiae* (13–14% sequence divergence) with higher competitive growth than the parental average (group mean = 4.3%).

Overall, there was a significant increase in mid-parent heterosis for relative competitive growth with increasing genetic divergence ($F_{44} = 2127$, $P < 0.001$, Fig. 1), but the relationship was driven entirely by the interspecific hybrids, which as a group showed strong and significant positive mid-parent heterosis with hybrids on average growing 4.3% better than the average parent (one-sample *t*-test: $t_{18} = 7.142$, $P < 0.001$). The intraspecific hybrids grew on average 0.2% better than their parents, but not significantly (one-sample *t*-test: $t_{26} = 0.628$, $P = 0.536$). Interspecific hybrids had significantly higher mid-parent heterosis for competitive growth than intraspecific hybrids (two-sample *t*-test: $t_{44} = 4.547$, $P < 0.001$). There was no significant relationship between genetic divergence and heterosis within intraspecific hybrids as a group ($F_{1,25} = 0.108$, $P = 0.746$), nor within interspecific hybrids as a group ($F_{1,17} = 2.883$, $P = 0.108$).

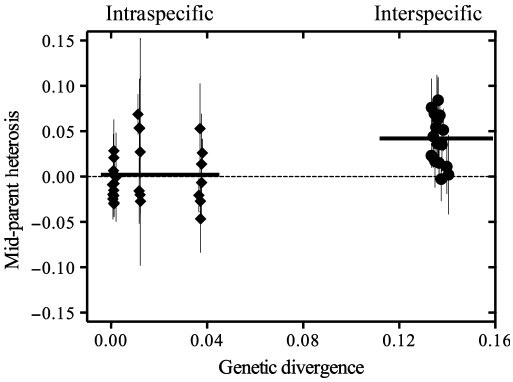


Fig. 1 Mid-parent heterosis in intraspecific and interspecific hybrids. Horizontal lines indicate the average mid-parent heterosis for intraspecific (mean = 0.004) and interspecific hybrids (mean = 0.045). Points with error bars indicate the means and standard deviations, respectively, of the replicates measures of mid-parent heterosis (see Methods). Diamonds indicate intraspecific hybrids, which are crosses between *Saccharomyces paradoxus* strains, and circles indicate interspecific hybrids, which are crosses between *Saccharomyces cerevisiae* and *S. paradoxus*.

Best-parent heterosis for competitive growth also increased significantly with genetic distance ($F_{1,44} = 10.49$, $P = 0.002$), but, as for mid-parent heterosis, the relation was driven by the higher best-parent heterosis of the interspecific hybrid group compared to the intraspecific group (Fig. S1). Interspecific hybrids had significantly higher best-parent heterosis for competitive growth than intraspecific hybrids (two-sample *t*-test: $t_{44} = 3.307$, $P = 0.002$), but there was no significant relationship between genetic divergence and best-parent heterosis within either of the two subgroups (Fig. S1: intraspecific hybrids: $F_{1,26} = 0.003$, $P = 0.954$; interspecific hybrids: $F_{1,18} = 0.397$, $P = 0.535$). Interspecific hybrids grew on average 0.5% better than their best parent but not significantly (one-sample *t*-test: $t_{18} = 0.812$, $P = 0.427$). Intraspecific hybrids grew on average 2% worse than their best parents, a significant difference (one-sample *t*-test: $t_{26} = 4$, $P < 0.001$).

Heterosis and phenotypic divergence in different crosses

Mid-parent heterosis significantly increased with phenotypic divergence of the parents (i.e. the absolute difference between the competitive growth of the two parents relative to their hybrid $F_{1,44} = 25.73$, $P < 0.001$ – Fig. 2). Unlike the general relationship between genetic divergence and heterosis discussed above, this relationship did not appear to be driven by any outlying group of strains; however, we note that phenotypic divergence was positively correlated with genetic divergence (Fig. S2: $F_{1,44} = 8.535$, $P = 0.006$). There was no significant relationship between best-parent heterosis and phenotypic divergence ($F_{1,44} = 0.235$, $P = 0.630$).

Effect of genetic marker in different crosses

The genetic markers (*ho::HYGMX/ho::HYGMX, ura3::KANMX/ura3::KANMX*) used to identify hybrids from their competing parent strains had a significant cost on competitive growth when tested in the 32 parent genetic backgrounds (one-sample *t*-test: $t_{30} = 2.065$, $P = 0.047$). On average, unmarked parents grew 2.07% (SD = ±0.026%) better than the marked versions of the same strains. In the competitions between hybrids and their parents, the hybrids were marked, so the cost of the marker might cause a systematic underestimation of the strength of positive heterosis. To account for this, we adjusted all log-transformed hybrid relative competitive growth rate values by adding the average log-transformed growth advantage of unmarked parents relative to unmarked parents (Fig. S3, Table S3). This adjustment made some of our results more significant. As before the adjustment, the interspecific hybrid had significant mid-parent heterosis (one-sample *t*-test: $t_{18} = 10.08$, $P < 0.001$), but their best-parent heterosis was now also significant after the adjustment (one-sample *t*-test: $t_{18} = 4.246$, $P < 0.001$). As before, intraspecific hybrids show no best-parent heterosis (one-sample *t*-test: $t_{26} = 0.135$, $P = 0.894$), but they now show significant mid-parent heterosis (one-sample *t*-test: $t_{26} = 3.574$, $P = 0.001$). As before, mid-parent heterosis significantly increased with phenotypic divergence across the entire set of crosses ($F_{1,44} = 25.73$, $P < 0.001$), and best-parent heterosis remained unrelated to phenotypic divergence ($F_{1,44} = 0.235$, $P = 0.630$). As before, interspecific hybrids had significantly higher heterosis for competitive growth than intraspecific hybrids, both for mid-parent heterosis (two-sample *t*-test: $t_{44} = 4.681$, $P < 0.001$) and for best-parent heterosis (two-sample *t*-test: $t_{44} = 3.307$, $P = 0.002$). Thus, although some differences became significant that were previous not significant, the adjustment did not change the pattern of the effect or our interpretation. We therefore present and discuss the more conservative, unadjusted heterosis values in the main body of the manuscript, but provide the adjusted values as Table S3.

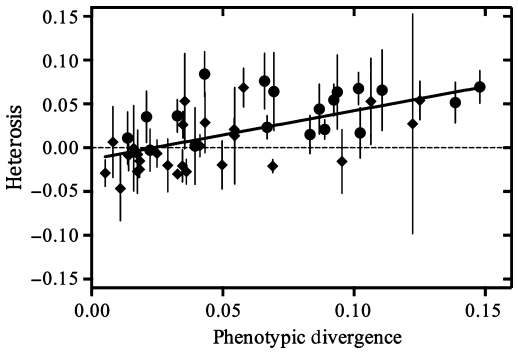


Fig. 2 The relationship between mid-parent heterosis for competitive growth and phenotypic divergence. Solid line indicates a significant positive correlation ($r_{44} = 0.607$, $P < 0.001$). Points and error bars as for Fig. 1.

Heterosis in different environments

To investigate the effect of phenotypic divergence independently of genetic divergence, we tested the competitive growth of an interspecific hybrid relative to its parents in different environments. The interspecific hybrid (s288c x N17) we tested grew on average 13% better than the average of its parents across 12 different environments (Fig. 3, Tables S2 and S4), and it grew significantly better than at least one of its parent in all environments (one-sample *t*-test corrected for multiple comparisons using the Holm–Sidak method: see Table S4 for statistics). In ten of the twelve environments (all except for aspirin and zinc sulphate), the competitive growth of the interspecific hybrid was higher against the *S. paradoxus* parent than against *S. cerevisiae* parent. Phenotypic distance correlated with both mid-parent heterosis ($F_{1,10} = 150.4$, $P < 0.001$, Fig. 4a) and best-parent heterosis ($F_{1,10} = 5.684$, $P = 0.038$, Fig. 4b) across all environments.

Effect of genetic marker in different environments

The marker (*ura3::KANMX/ura3::HYGMX*) used to identify the parent strains in the experiment in different environments increased competitive growth by an average of 1.21% relative to the marker carried by the hybrids (*ura3/ura3*), when both markers were tested in the parental genetic back grounds in all environments (Table S3). Thus, the benefit of the parental marker might cause an underestimation of heterosis. To adjust for this, we added the log-transformed measured growth advantage of the parental marker, for each

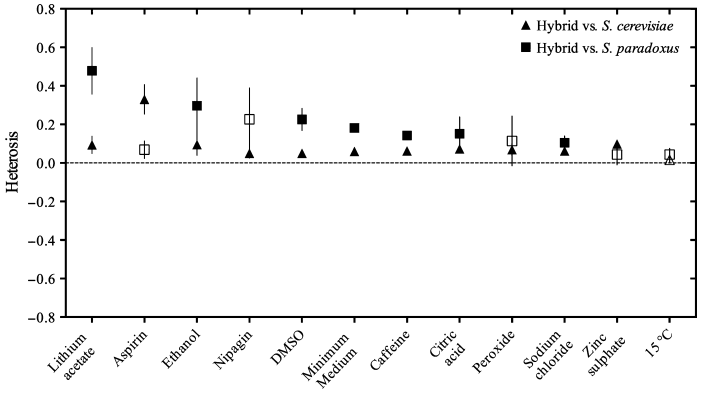


Fig. 3 Heterosis for competitive growth of a single interspecific cross in twelve different environments. Triangles show average heterosis relative to the *Saccharomyces cerevisiae* parent; squares show average heterosis relative to *Saccharomyces paradoxus* parent. Open shapes indicate heterosis not significant after correction for multiple testing (Table S4; see Results). Error bars indicate standard deviation of the mean of the replicate measurements.

parent in each environment except low temperature, to the log-transformed competitive growth of the hybrids relative to each parent in each environment except low temperature (Table S3). The adjustment generally increased our estimates of heterosis, but did not change our interpretation of the results. As before, the interspecific hybrid grew significantly better than at least one of its parents for all the environments tested (one-sample *t*-test corrected for multiple comparisons using the Holm–Sidak method: see Table S4 for statistics, Fig. S4). Adjusting for the measured marker effect did not affect the relationship between phenotypic distance and mid-parent heterosis, which stayed significantly positive ($F_{1,9} = 75.50$, $P < 0.001$), but it made the correlation between phenotypic distance and best-parent heterosis not significant ($F_{1,9}<0.001$, $P = 0.976$). Because the unadjusted results provide a more conservative measure of heterosis, we present and discuss the unadjusted results here, but we provide the adjusted results as supporting data (Table S3).

Discussion

Here, we find that hybrids between *S. cerevisiae* and the wild species *S. paradoxus* can grow on average 4.3% better than their parents in direct competition. In contrast, crosses between genetically diverged *S. paradoxus* strains had much less strong heterosis. We show that the strength of heterosis is best predicted by the difference in the competitive growth rates of parents relative to their common hybrid, both when different strains are tested in the same environment, and when the same strains are tested in different environments.

Recent studies of intraspecific *S. cerevisiae* x *S. cerevisiae* crosses have attributed positive heterosis to complementation of deleterious alleles that have accumulated in this species as the result of its domestication by humans (Zörgö *et al.*, 2012; Plech *et al.*, 2014). Cellular functions that are maintained in the wild may be lost in simplified winery or brewery habitats. Two features of yeast domestication might exacerbate this process: drift due to reduced effective population size and disruptive selection in different environments allowing fixation of loss-of-function mutations in different metabolic pathways. Zörgö *et al.* (2012) crossed nine genetically diverged *S. cerevisiae* strains in all pairwise combinations and grew the F1 hybrids asexually under various environmental conditions. Mid-parent heterosis was prevalent and was correlated with poor parental growth, consistent with the simple complementation of loss-of-function mutations that reduce growth in the experimental environment. A follow-up study with larger sample of parental strains confirmed that heterosis was indeed much more likely when parents originated from domesticated, rather than natural environments (Plech *et al.*, 2014).

Could the presence of deleterious mutations in *S. cerevisiae* due to domestication explain the general heterosis we observe when it is crossed to wild *S. paradoxus* strains that lack such mutations? In our experiments, *S. paradoxus* x *S. paradoxus* crosses have much lower heterosis than our *S. cerevisiae* x *S. paradoxus* crosses (Fig. 1), consistent with the wild species having fewer deleterious mutations (or less deleterious mutations). We also found that the larger the difference in parental competitive growth, the stronger the mid-parent heterosis was, both in the full set of crosses tested in a single environment (Fig. 2) and in a single *S. cerevisiae* x *S. paradoxus* cross tested in multiple environments (Fig. 4a). Simple complementation of recessive deleterious mutations in one parent, such as a domesticated *S. cerevisiae* strain, by functional alleles in another, such as a wild *S. paradoxus* strain, would be expected to give exactly this pattern of autocorrelation. To visualize this, imagine that *S. cerevisiae* strains carrying recessive deleterious mutations with different effect sizes (and therefore with different low fitnesses) are crossed to *S. paradoxus* strains lacking such deleterious

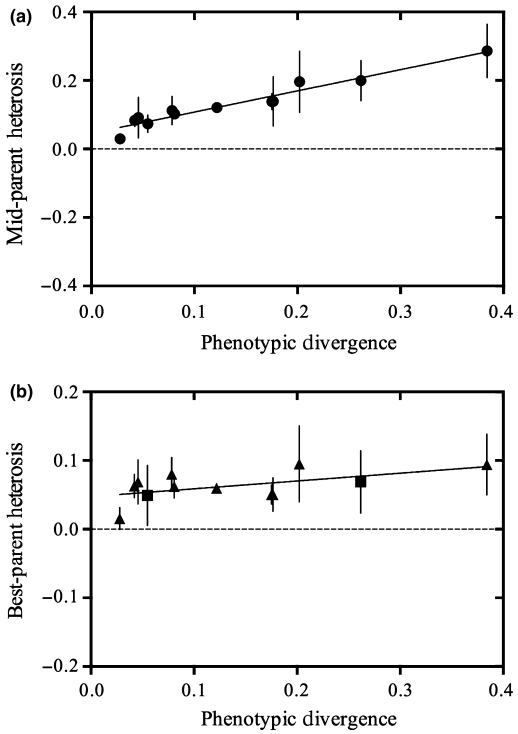


Fig. 4 The relationship between parental phenotypic divergence and heterosis of a single interspecific hybrid in twelve different environments. (a) Mid-parent heterosis. Circles with error bars indicate means and standard deviations, respectively, of the replicate measures of mid-parent heterosis (see Methods). Solid line indicates a significant positive correlation ($r_{10} = 0.968$, $P < 0.001$). (b) Best-parent heterosis. Points with error bars indicate means and standard deviations, respectively, of the replicates measures of best-parent heterosis (see Methods). Triangles indicate that the best parent was *Saccharomyces cerevisiae* parent, and squares indicate that the best parent was *Saccharomyces paradoxus*. Solid line indicates a significant positive correlation ($r_{10} = 0.566$, $P = 0.038$).

mutations (and therefore of approximately equal, high fitness). Under this simple complementation model, all recessive defects will be complemented so all hybrids will have approximately equal high fitness, but those hybrids showing the strongest mid-parent heterosis will be those whose parents have the largest fitness difference and thus the lowest mid-parent fitness. Zörgö *et al.* (2012) found such a relationship between the difference in growth between *S. cerevisiae* parents and the mid-parent heterosis of their resulting intraspecific hybrids and also interpreted it as simple complementation of domestication defects in one parent by wild-type alleles in another.

However, several aspects of our data make this model of simple complementation of defective *S. cerevisiae* alleles by functional *S. paradoxus* alleles questionable. Plech *et al.* (2014) found that intraspecific heterosis was more prevalent when *S. cerevisiae* parents had been isolated from human-made habitats rather than wild habitats. But we did not find significant higher heterosis in the 13 interspecific crosses made with *S. cerevisiae* strains from human habitats, than in the six interspecific crosses made with *S. cerevisiae* isolated from natural habitats (crosses with *S. cerevisiae* strains from human habitats grew only 1.2% better, two-sample *t*-test: $t_{17} = 0.935$, $P = 0.363$), but we note that this test has little power, especially given that the domestication history of a strain cannot reliably be inferred from the habitat it was isolated from. A clearer prediction of the simple complementation model is that if interspecific heterosis was due to growth defects in *S. cerevisiae*, then *S. cerevisiae* parents should grow less well than *S. paradoxus* parents in competition with their shared hybrids. But in general, the opposite was true: for 13 out of 19 hybrids (not a significant majority, one-way two-tailed chi-squared test: $\chi^2_1 = 1.746$, $P = 0.186$) and for 10 out of 12 environments (a significant majority, one-way two-tailed chi-square test: $\chi^2_1 = 3$, $P = 0.042$), the *S. cerevisiae* parent actually grew better than the *S. paradoxus* parent, relative to their common hybrid. Finally, and perhaps, most importantly, simple complementation of defective *S. cerevisiae* alleles by functional *S. paradoxus* alleles is expected to produce only mid-parent heterosis, in which the hybrid grows at best as well as the functional *S. paradoxus* parent, not best-parent heterosis in which it grows better. Best-parent heterosis can occur when two parents carrying defects at different loci are crossed (Zörgö *et al.*, 2012; Plech *et al.*, 2014). For example when a strain with loss-of-function mutation in one of the genes in the galactose utilization pathway was crossed to a strain with a loss-of-function mutation in another gene of the same pathway, function was restored because the defects were recessive and the intraspecies cross grew better on galactose than either of its parents (Zörgö *et al.*, 2012). However, we see evidence for best-parent heterosis in our interspecies hybrids, both in multiple crosses after the marker effect is corrected for (Fig. S3 and Table S3) and in the single hybrid we studied, which could outcompete both parents in many different environments (Fig. 3). Although recessive deleterious mutations might be fixed in *S. cerevisiae* strains because of relaxed selection due to domestication, we would not expect such mutations in *S. paradoxus*, which is undomesticated, so we would not expect best-parent heterosis, nor would we expect it to correlate with phenotypic divergence (Fig. 4b). Thus, our results suggest that mechanisms in addition to complementation of recessive deleterious alleles, such as overdominance, might also contribute to best-parent heterosis of interspecies yeast hybrids,

although we note that our sample size is too small to be conclusive.

A mechanism that can explain the presence of fixed recessive deleterious mutations in both species also presents a caveat that applies to all yeast heterosis studies to date, as far as we know. The parental diploids we used were monosporic isolates, which were originally derived from single haploids that were allowed to divide mitotically, switch mating type, and mate with their identical haploid clone mates to produce perfectly homozygous diploids (Liti *et al.*, 2009). This is a standard practice to produce pure genetic backgrounds that can be sequenced and studied without the complications of segregating genetic variation (Liti *et al.*, 2009). However, there is evidence that natural strains can be highly heterozygous (Magwene *et al.*, 2011), so deriving monosporic isolates would homozygose any recessive deleterious mutations that were previously masked reducing the monosporic strains' fitness relative to their heterozygous parents. Crosses among different monosporic strains would then restore fitness by complementation, giving the illusion of heterosis, even though the resulting F1 hybrids would not necessarily be any fitter than their heterozygous grandparents. It is not easy to eliminate this potential artefact, because most strains available in collections have been treated in this way. Further, natural strains of *Saccharomyces* are usually isolated by enrichment culture, in which an environmental sample (typically a piece of oak bark) is placed into rich liquid growth medium and incubated, before cells from the resulting mixed culture are isolated and their species identified. If oak bark samples usually contain *Saccharomyces* haploid spores rather than vegetative diploid cells, then the rapid germination and growth conditions provided by enrichment culture might promote mating type switching and homozygosis of recessive deleterious mutations, rather than mating with other spores to produce heterozygotes, as might occur under natural conditions. A challenge for yeast biologists studying evolution is therefore to identify a natural source of vegetatively growing *Saccharomyces* from which samples could be taken directly, which without enrichment culturing.

The positive relationship between parental phenotypic divergence and the strength of heterosis, as well as the general heterosis we find in interspecies hybrids, suggests that mating between species might be advantageous. However, any benefit of interspecies hybrids have under mitosis would have to outweigh the cost they suffer under meiosis: 99% of the gametes produced by F1 hybrids are inviable (Hunter *et al.*, 1996), so only if mitotic divisions greatly outnumber meiotic divisions could their increased vigour compensate for their decreased fertility. This might be possible: an estimated based on population genetic suggests that 1000 mitotic divisions occur for every meiosis in wild oak-associated *S. paradoxus* (Tsai *et al.*, 2008), and a F1 hybrid cell with a growth advantage of 4.3% over a cell of its parent species would need only 175 mitotic generations before its population was over 100 times larger (i.e. large enough to compensate for the ~99% spores that die from F1 hybrid meiosis). Indeed, yeast hybrids are well known, especially in wine and beer industry, environments, where, perhaps, meiosis is not required. Best known is *S. pastorianus* the hybrid used to produce low temperature fermented larger beer, which benefits from a combination of the ethanol resistance of its *S. cerevisiae* parent and the cold tolerance of its *S. eubayanus* parent (Vaughan & Martini, 1987; Libkind *et al.*, 2011), but many other hybrids of *S. cerevisiae*, *S. kudriavzevii*, *S. uvarum* and *S. eubayanus* have been found in wine and cider too (Lopandic *et al.*, 2007; Sipiczki, 2008). Genomic methods are now identifying an increasing number of hybrids between *S. cerevisiae* and *S. paradoxus* outside fermentation environments and examples of introgression of *S. cerevisiae* genes into majority wild *S. paradoxus* genomes (Liti *et al.*, 2006) and *vice versa* (Muller & McCusker, 2009), indicating that many sexual cycles occurred since hybridization and suggesting that be benefits of yeast hybridization can indeed sometimes outweigh their fertility costs.

There is increasing awareness in the role that hybridization has played in the evolution of a wide range of species (see the special issue of *Journal of Evolutionary Biology*, 26(2) 2013; Seehausen, 2004; Mallet, 2007; Schumer *et al.*, 2014), not least on our own (Sankararaman *et al.*, 2014). The importance of that role depends very much on the ability of the hybrid to compete against nonhybrids, and yeast offers a useful way to assess the factors contributing to the relative fitness of hybrids.

Acknowledgments

We thank the members of the Greig laboratory for their help in the laboratory and Arne Nolte and Dietrich Ober for discussion. The very thorough comments of the anonymous reviewers greatly improved the manuscript. This research was supported by the Max Planck Society.

Conflict of interests

The authors declare no conflict of interests.

References

Barton, N.H. & Hewitt, G.M. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* **16**: 113–148.

Edmands, S. 2002. Does parental divergence predict reproductive compatibility? *Trends Ecol. Evol.* **17**: 520–527.

Gietz, R.D. & Woods, R.A. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350**: 87–96.

Gonzáles, W.L., Suárez, L.H. & Medel, R. 2007. Outcrossing increases infection success in the holoparasitic mistletoe *Tristerix aphyllus* (Loranthaceae). *Evol. Ecol.* **21**: 173–183.

Greig, D., Borts, R.H., Louis, E.J. & Travisano, M. 2002. Epistasis and hybrid sterility in *Saccharomyces*. *Proc. R. Soc. Biol. Sci.* **269**: 1167–1171.

Hung, H.Y., Browne, C., Guill, K., Coles, N., Eller, M., Garcia, A. *et al.* 2012. The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. *Heredity* **108**: 490–499.

Hunter, N., Chambers, S.R., Louis, E.J. & Borts, R.H. 1996. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* **15**: 1726–1733.

Kao, K.C., Schwartz, K. & 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 Genet.* **6**: 10010–10038.

Lenski, R.E. 1991. Quantifying fitness and gene stability in microorganisms. *Biotechnology* **15**: 173–192.

Lexer, C., Welch, M.E., Raymond, O. & Rieseberg, L.H. 2003. The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): selection on transgressive characters in a novel hybrid habitat. *Evolution* **57**: 1989–2000.

Libkind, D., Hittinger, C.T., Valério, E., Gonçalves, C., Dover, J., Johnston, M. *et al.* 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc. Natl. Acad. Sci. USA* **108**: 14539–14544.

Liti, G., Barton, D.B.H. & Louis, E.J. 2006. Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. *Genetics* **174**: 839–850.

Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A. *et al.* 2009. Population genomics of domestic and wild yeasts. *Nature* **458**: 337–341.

Lopandic, K., Gangl, H., Wallner, E., Tscheik, G., Leitner, G., Querol, A. *et al.* 2007. Genetically different wine yeasts isolated from Austrian vine-growing regions influence wine aroma differently and contain putative hybrids between *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii*. *FEMS Yeast Res.* **7**: 953–965.

Magwene, P.M., Kayıkçı, Ö., Granek, J.A., Reininga, J.M., Scholl, Z. & Murray, D. 2011. Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **108**: 1987–1992.

Mallet, J. 2007. Hybrid speciation. *Nature* **446**: 279–283.

McClelland, E.K. & Naish, K.A. 2007. What is the fitness outcome of crossing unrelated fish populations? A meta-analysis and an evaluation of future research directions. *Conserv. Genet.* **8**: 397–416.

Moll, R.H., Lonnquist, J.H., Fortuno, J.V. & Johnson, E.C. 1965. The relationship of heterosis and genetic divergence in maize. *Genetics* **52**: 139–144.

Muller, L.A. & McCusker, J.H. 2009. A multispecies-based taxonomic microarray reveals interspecies hybridization and introgression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **9**: 143–152.

Naumov, G.L. 1996. Genetic identification of biological species in the *Saccharomyces sensu stricto* complex. *J. Ind. Microbiol.* **17**: 295–302.

Nolte, A.W. & Sheets, H.D. 2005. Shape based assignment tests suggest transgressive phenotypes in natural sculpin hybrids (Teleostei, Scorpaeniformes, Cottidae). *Front. Zool.* **2**: 11.

Nosil, P. 2012. *Ecological Speciation*, 1 edn. Oxford Series in Ecology and Evolution, Oxford, UK.

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

Pekkala, N., Knott, K.E., Kotiaho, J.S., Nissinen, K. & Puurtinen, M. 2012. The benefits of interpopulation hybridization diminish with increasing divergence of small populations. *J. Evol. Biol.* **25**: 2181–2183.

Plech, M., de Visser, J.A. & Korona, R. 2014. Heterosis is prevalent among domesticated but not wild strains of *Saccharomyces cerevisiae*. *G3* **4**: 315–323.

Rieseberg, L.H., Raymond, O., Rosenthal, D.M., Lai, Z., Livingstone, K., Nakazato, T. *et al.* 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**: 1211–1216.

Sankararaman, S., Mallick, S., Dannemann, M., Prufer, K., Kelso, J., Paabo, S. *et al.* 2014. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**: 354–357.

Scannell, D.R., Zill, O.A., Rokas, A., Payen, C., Dunham, M.J., Eisen, M.B. *et al.* 2011. The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* Genus. *G3* **1**: 11–25.

Schumer, M., Rosenthal, G.G. & Andolfatto, P. 2014. How common in homoploid speciation? *Evolution* **68**: 1553–1560.

Schwarz, D., Matta, B., Shakir-Botteri, N.L. & McPheron, B.A. 2005. Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* **436**: 546–549.

Seehausen, O. 2004. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**: 198–207.

Shapira, R., Levy, T., Shaked, S., Fridman, E. & David, L. 2014. Extensive heterosis in growth of yeast hybrids is explained by a combination of genetic models. *Heredity* **113**: 316–326.

Shull, G.H. 1948. What Is "Heterosis"? *Genetics* **33**: 439–446.

Sipiczki, M. 2008. Interspecies hybridization and recombination in Saccharomyces wine yeasts. *FEMS Yeast Res.* **8**: 996–1007.

Stelkens, R.B., Brockhurst, M.A., Hurst, G.D., Miller, E.L. & Greig, D. 2014. The effect of hybrid transgression on environmental tolerance in experimental yeast crosses. *J. Evol. Biol.* **27**: 2507–2519.

Stelkens, R.B., Schmid, C. & Seehausen, O. 2015. Hybrid breakdown in cichlid fish. *PLoS ONE* **10**: e0127207.

Tsai, I.J., Bensasson, D., Burt, A. & Koufopanou, V. 2008. Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle. *Proc. Natl. Acad. Sci. USA* **105**: 4957–4962.

Vasi, F., Travisano, M. & Lenski, R.E. 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. *Am. Nat.* **144**: 432–456.

Vaughan, M.A. & Martini, A. 1987. Three newly delimited species of *Saccharomyces sensu stricto*. *Antonie Van Leeuwenhoek* **53**: 77–84.

Waser, N.M. 1993. Population structure, optimal outbreeding, and assortative mating in angiosperms. In: *The Natural*

History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives (N.W. Thornhill, ed.), pp. 173–199. University of Chicago Press, Chicago, USA.

Xiao, J., Li, J., Yuan, L., McCouch, S.R. & Tanksley, S.D. 1996. Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theor. Appl. Genet.* **92**: 637–643.

Zörgö, E., Gjuvsland, A., Cubillos, F.A., Louis, E.J., Liti, G., Blomberg, A. *et al.* 2012. Life history shapes trait heredity by accumulation of loss-of-function alleles in Yeast. *Mol. Biol. Evol.* **29**: 1781–1789.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1 Best-parent heterosis in intra-specific and interspecific hybrids.

Figure S2 Relationship between genetic divergence and phenotypic divergence.

Figure S3 Heterosis in intra-specific and inter-specific hybrids adjusted for marker effect.

Figure S4 Heterosis for competitive growth of a single inter-specific cross in twelve different environments, adjusted for marker effect.

Table S1 Description of the strains and crosses, raw data for the competition between hybrids and its parents, phenotypic divergence, mid-parent heterosis and best-parent heterosis data.

Table S2 Description of the environments, raw data for the competition in different environments between inter-specific hybrid and its parents, phenotypic divergence, mid-parent heterosis and best-parent heterosis data.

Table S3 Raw data for the competition between marked and unmarked strains for different crosses and different environment competition, adjusted phenotypic divergence, mid-parent heterosis and best-parent heterosis for markers' effect.

Table S4 Multiple comparisons table for the competition between inter-specific hybrid and its parents in different environments, and adjusted multiple comparisons tables for markers' effect.

Received 27 July 2016; revised 14 November 2016; accepted 28 November 2016

5. Supplementary material

Table S1: Description of the parental strains by species, habitat, and genetic and phenotypic divergence. Competitive growth of F1 hybrids: mid-parent heterosis (MPH) and best-parent heterosis (BPH) log-transformed.

Hybrid	Parent 1	Parent 1 habitat	Parent 2	Parent 2 habitat	Genetic divergence	Phenotypic divergence	MPH (LOG)	BPH (LOG)
1	*S. par*	quercus	*S. par*	quercus	0,001	0,008	0,007	0,002
2	*S. par*	drosophila	*S. par*	soil beneath quercus	0,002	0,016	-0,001	-0,009
3	*S. par*	quercus	*S. par*	soil beneath quercus	0,038	0,035	0,026	0,009
4	*S. par*	quercus	*S. par*	drosophila	0,038	0,025	-0,007	-0,019
5	*S. par*	quercus	*S. par*	quercus	0,001	0,041	0,002	-0,018
6	*S. par*	quercus	*S. par*	quercus	0,011	0,058	0,069	0,040
7	*S. par*	quercus	*S. cer*	vaginatis	0,135	0,092	0,054	0,008
8	*S. par*	quercus	*S. cer*	bee	0,136	0,033	0,036	0,020
9	*S. par*	quercus	*S. par*	quercus	0,001	0,050	-0,019	-0,044
10	*S. par*	drosophila	*S. par*	quercus	0,037	0,011	-0,047	-0,052
11	*S. par*	quercus	*S. par*	quercus	0,012	0,122	0,027	-0,034
12	*S. par*	quercus	*S. par*	quercus	0,001	0,033	-0,030	-0,046
13	*S. par*	mor soil	*S. par*	quercus	0,001	0,018	-0,015	-0,024
14	*S. par*	quercus	*S. cer*	wine	0,134	0,089	0,021	-0,023
15	*S. par*	quercus	*S. cer*	feaces	0,133	0,066	0,076	0,043
16	*S. par*	quercus	*S. cer*	soil beneath quercus	0,135	0,102	0,017	-0,034
17	*S. cer*	feaces	*S. par*	soil beneath quercus	0,1384	0,139	0,051	-0,018
18	*S. cer*	barrel	*S. par*	quercus	0,136	0,094	0,064	0,017
19	*S. par*	quercus	*S. cer*	vaginites	0,137	0,102	0,067	0,017
20	*S. par*	mor soil	*S. cer*	bee	0,137	0,083	0,015	-0,026
21	*S. par*	quercus	*S. cer*	grapes	0,133	0,067	0,023	-0,010
22	*S. cer*	soil	*S. par*	quercus	0,134	0,087	0,044	0,001
23	*S. par*	mor soil	*S. cer*	wine	0,137	0,022	-0,003	-0,014
24	*S. cer*	beer	*S. par*	drosophila	0,140	0,039	0,002	-0,018
25	*S. cer*	wine	*S. par*	quercus	0,135	0,148	0,069	-0,005
26	*S. par*	quercus	*S. cer*	beer	0,136	0,069	0,064	0,029
27	*S. cer*	fermentation must	*S. par*	soil beneath quercus	0,138	0,021	0,035	0,025
28	*S. cer*	soil	*S. par*	quercus	0,136	0,111	0,066	0,010
29	*S. cer*	beer	*S. par*	soil beneath quercus	0,140	0,014	0,011	0,004
30	*S. par*	mor soil	*S. par*	quercus	0,012	0,125	0,054	-0,009
31	*S. par*	mor soil	*S. par*	quercus	0,012	0,036	-0,027	-0,045
32	*S. par*	quercus	*S. par*	quercus	0,012	0,095	-0,016	-0,063
33	*S. par*	quercus	*S. par*	quercus	0,012	0,035	0,053	0,035
34	*S. par*	quercus	*S. par*	quercus	0,012	0,029	-0,020	-0,034
35	*S. par*	quercus	*S. par*	quercus	0,001	0,043	0,029	0,007
36	*S. par*	quercus	*S. par*	quercus	0,001	0,069	-0,021	-0,055
37	*S. par*	mor soil	*S. par*	quercus	0,001	0,054	0,021	-0,006
38	*S. par*	mor soil	*S. par*	quercus	0,001	0,018	-0,008	-0,016
39	*S. par*	quercus	*S. par*	quercus	0,001	0,014	-0,009	-0,016
40	*S. par*	quercus	*S. par*	quercus	0,001	0,005	-0,029	-0,031

Table S2: Description of the environments, raw data for the competition between inter-specific F1 hybrid and its parents in different environments. Mid-parent (MPH) and best-parent heterosis (BPH) values were log-transformed from the competitions, and phenotypic divergence was calculated through the difference between competitions.

Environments	*S. cerevisiae* *vs.* F1 hybrid	*S. paradoxus* vs. F1 hybrid	MPH (LOG)	BPH (LOG)	Phenotypic divergence
Lithium acetate	0,094	0,478	0,286	0,094	0,384
Aspirin	0,330	0,069	0,200	0,069	0,262
Ethanol	0,095	0,297	0,196	0,095	0,202
Nipagin	0,050	0,227	0,139	0,050	0,176
DMSO	0,050	0,226	0,138	0,050	0,175
Minimum medium	0,060	0,182	0,121	0,060	0,122
Caffeine	0,062	0,143	0,102	0,062	0,081
Citric acid	0,073	0,151	0,112	0,073	0,078
Peroxide	0,069	0,114	0,092	0,069	0,046
Sodium Chloride	0,063	0,105	0,084	0,063	0,042
Zinc Sulphate	0,098	0,043	0,071	0,043	0,055
15°C	0,016	0,043	0,029	0,016	0,028

Figure S1: Best-parent heterosis of intra-specific and inter-specific F1 hybrids. Horizontal lines indicate the average best-parent heterosis for intra-specific (mean=-0.02) and inter-specific hybrids (mean=0.005). Points with error bars indicate the means and standard deviations, respectively, of the replicates measures of best-parent heterosis (see Methods Chapter I). Diamonds indicate intra-specific hybrids, which are crosses between *S. paradoxus* populations and circles indicate inter-specific hybrids, which are crosses between *S. cerevisiae* and *S. paradoxus*.

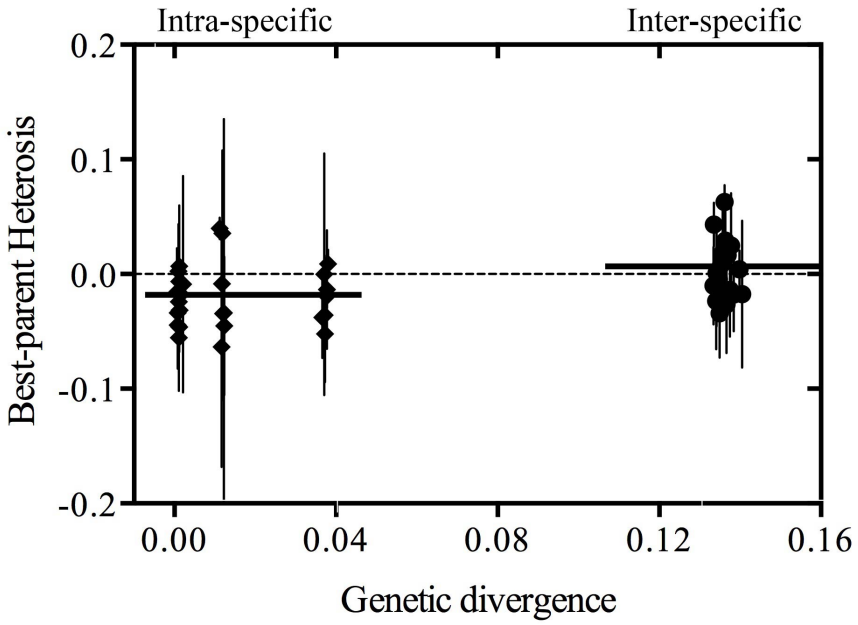


Figure S2: Relationship between genetic divergence and phenotypic divergence. Line indicates a significant positive correlation ($F_{1,44}$=8.535, P=0.006). Points and error bars indicate means and standard deviations of replicate measures of phenotypic divergence.

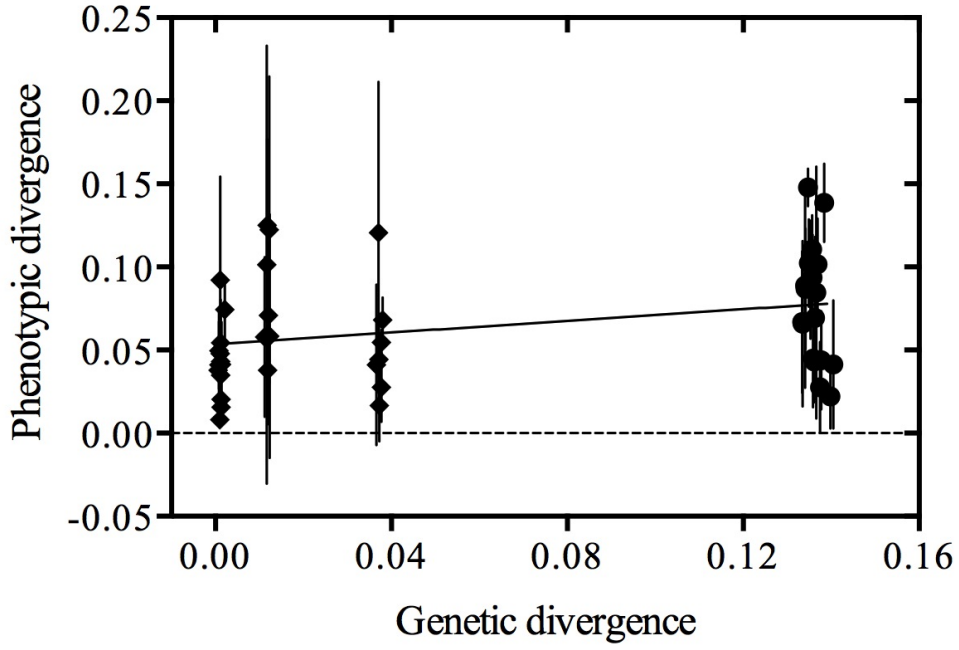


Figure S3: Mid-parent and best-parent heterosis of intra-specific and inter-specific hybrids adjusted for marker effect. Mid-parent heterosis: horizontal lines indicate the average best-parent heterosis for intra-specific (mean=0.002) and inter-specific hybrids (mean=0.051). Best-parent heterosis: horizontal lines indicate the average best-parent heterosis for intra-specific (mean=0.000) and inter-specific hybrids (mean=0.02). Points with error bars indicate the means and standard deviations, respectively, of the replicates measures of mid-parent and best-parent heterosis adjusted for marker effect (see Chapter I Methods). Diamonds indicate intra-specific hybrids, which are crosses between *S. paradoxus* populations and circles indicate inter-specific hybrids, which are crosses between *S. cerevisiae* and *S. paradoxus*.

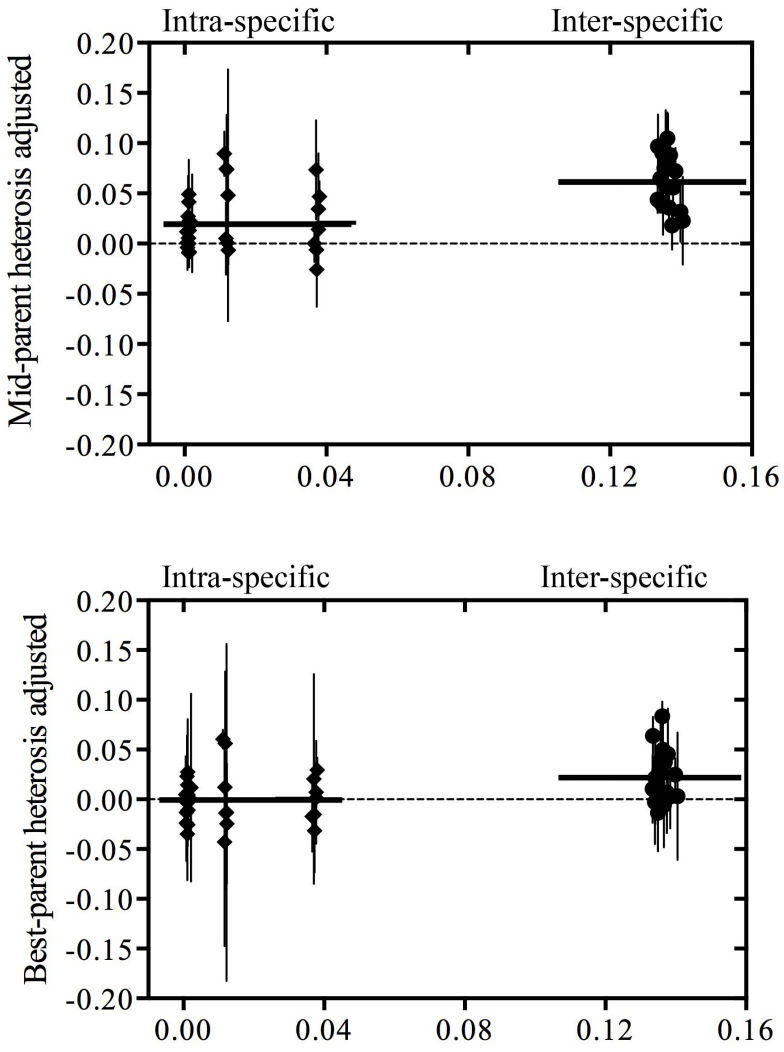
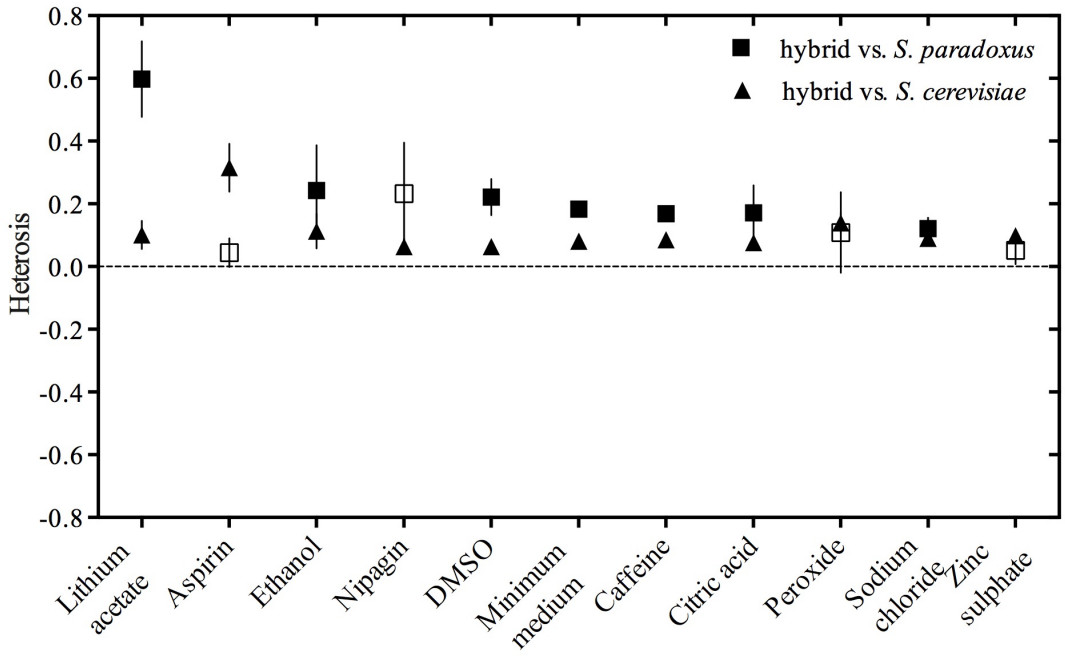


Figure S4: Heterosis for competitive growth of a single inter-specific cross in twelve different environments, adjusted for marker effect. Triangles show average heterosis relative to the *S. cerevisiae* parent, squares show average heterosis relative to *S. paradoxus* parent. Open shapes indicate heterosis not significant after correction for multiple testing. Error bars indicate standard deviation of the mean of the replicate measurements.



Chapter II

Hybrid transcription is similar to the fitter parent for a specific environment.

Bernardes, J. P.[1], Rogers, D.[1] & Greig, D.[1,2]

1. Max Planck Institute for Evolutionary Biology, Experimental Evolution Group, Plön, Germany

2. The Galton Laboratory, Department of Genetics, Evolution, and Environment, University College London, London, UK

1. Introduction

When two closely related species hybridise, the hybrid might display transcription features that are different from those of its parents: this deregulated transcription can take the form of over- or under-transcription of genes in comparison with one or both parental species (Gibson *et al*., 2004), and has been described in a variety of taxa including Drosophila (Ranz *et al*., 2004), yeast (Tirosh *et al*., 2006), and sea urchin (Nielsen *et al*., 2000). The causes of the hybrid deregulated transcription have been linked to divergence of regulatory pathways that control transcription in the parental species (for review see Wittkopp and Kalay (2012)), such as variations in affinity between transcription factors and their binding sites (Johnson & Porter, 2000), or variations in transcription initiation due to promoter elements like TATA boxes (Tirosh *et al*., 2006). If we were to cross two closely related species, the first hybrid generation (F1 hybrid) would share half of its genome with each parental species, and the two parental alleles would be exposed to the same *trans*-regulatory environment, while *cis*-regulatory effects would control only the parental alleles they were linked to (Wittkopp *et al*., 2004). Thus any difference in the F1 hybrid transcription between alleles of a specific gene is due to *cis*-regulation, known as allele specific expression, while a difference in F1 hybrid transcription of both parental alleles in comparison to its parents is due to *trans*-regulation. There can also be cases when both *cis*- and *trans*-regulation affect transcription, then we consider gene transcription to be *cis-trans* regulated or *cis-trans* antagonistic, when *cis*- and *trans*-regulations have opposite directionalities (Figure S1).

The F1 hybrid gene transcription can be inherited from its parents' in a conserved manner or it can vary from one or both parental species; (1) when the F1 hybrid gene transcription is similar to the average of both parental species we classify it as additive; (2) when the F1 hybrid gene transcription is similar to one of the parents but not the other we classify it as dominant; (3) when the F1 hybrid gene transcription is an extreme form of dominant and goes beyond a parental species we classify it as overdominant, also known as misexpression (Figure S2).

When genetically diverged parents mate, their F1 hybrid can, in some cases, display a fitness advantage in relation to both parental species or the average of them, a

phenomenon known as heterosis or hybrid vigour (Shull, 1908). What causes the F1 hybrid higher fitness in comparison to its parents has been a question many have tried to answer; hybrids are characterized by being highly heterozygous, this could lead to a fitness increase due to simple or reciprocal complementation of low-fitness parental alleles by high-fitness parental alleles, or the heterozygous loci might also have an intrinsic advantage.

F1 hybrid transcription depends on the regulation of parental alleles and how adapted they are to the environment. Goff (2011) molecular model for multigenic heterosis, assumes the F1 hybrid cells are able to distinguish between parental alleles and preferentially transcribed the fitter allele for each heterozygous pair of parental alleles (Goff, 2011). This model would encompass a mechanism where the fitter parental allele for a particular environment is preferentially transcribed over the less fit parental allele, based on the stability of the resulting proteins, and also, a mechanism where both parental alleles for a particular environment are up-regulated, based on the stability of the resulting proteins (Goff, 2011). We can test the model by comparing the F1 hybrid transcription with its parental transcription. We expect the F1 hybrid to be enriched for allele specific expression (*cis*-regulated transcription- Figure S1) of the favourable alleles of both parents so the F1 hybrid transcription resembles or is dominated by the fittest parent alleles for a particular environment. Also, we expect the F1 hybrid transcription to have the plasticity to be modified upon environmental change, because the relative fitness alleles would vary according to the environment the individuals are in.

One way to infer differences in transcription between F1 hybrids and their parents, and consequently F1 hybrid parental alleles, is to use environments that have an impact on their fitnesses. In our previous project we explored the effect of different environments in the competitive growth of a inter-specific F1 hybrid when in direct competition against its parental species; we crossed a *Saccharomyces cerevisiae* strain with a *S. paradoxus* strain and we tested the resulting F1 hybrid in direct competition with its parents in different environments (Chapter I and Bernardes *et al*., 2016). In some environments *S. cerevisiae* grew better than *S. paradoxus*, in others vice versa, but the F1 hybrid grew better than both its parents in all cases (i.e. there was best-

parent heterosis). Using the differences in fitness between the F1 hybrid and the parental species in a specific environment could help unveil the importance of a specific transcription response of the F1 hybrid. When yeast is exposed to different environments, its transcription profile is specific to the environment the yeast is in and has adapted to (Gasch *et al*., 2000). Upon change of environmental conditions, yeast cells adjust their transcription to the new conditions, this response can be specific to the environment, by acting on genes from pathways which are involved in dealing with the new conditions (Alexandre *et al*., 2001), or the response can be general, affecting several genes regardless of the type of environmental change, know as environmental stress response (ESR) (Gasch *et al*., 2000). A specific yeast species can have great variations in its transcription profile between different environments; *S. cerevisiae* displays both specific and general transcription responses and has evidence considerable transcription variations under several different environments such as heat shock, osmotic shock and depletion of essential compounds (Gasch *et al*., 2000). Different *Saccharomyces* yeasts also have specific transcription under the same environment stress depending on the lifestyle of the individual species (Brion *et al*., 2016). Environmental stress responses in fermentative yeasts like *S. cerevisiae* show greater variations when compared to transitional lifestyle yeast *Lachancea kluyveri* (Brion et al., 2016) than to the more divergent but fermentative fission yeast *Schizosaccharomyces pombe* (Gasch, 2007). Inter-specific F1 hybrids have the potential to use the common features they have with their parental species, and induce a transcription profile that is similar to the fitter parent for a particular environment. Thus the F1 hybrid would be better adapted to deal with a variety of different environments, more than its individual parents on they own.

Most studies that have analysed the transcription profiles of F1 hybrids aimed to unveil how gene transcription is regulated either by *cis*-regulation, *trans*-regulation or both (Tirosh *et al*., 2009; Emerson *et al*., 2010; Schaefke *et al*., 2013; Artieri & Fraser, 2014; McManus *et al*., 2014; Wang *et al*., 2015). Surprisingly there seems to be several contradictions between different studies. This could be due to variations in methodology, such as, environments used, the type of F1 hybrid (intra-specific vs. inter-specific), or variations in techniques- earlier studies used microarrays while contemporary studies use Next-Generation Sequencing tools such as RNA-seq. For

example, previous studies on a hybrid cross between *S. paradoxus* and *S. cerevisiae* species, showed either similar levels of *cis*- and *trans*-regulation (Artieri & Fraser, 2014), or much higher levels of *trans*-regulation over *cis*-regulated transcription (McManus *et al*., 2014), while a hybrid cross between *S. cerevisiae* populations also showed *trans*-regulation was more common than *cis*-regulation (Emerson *et al*., 2010; Schaefke *et al*., 2013). The McManus *et al*. (2014) study on an inter-specific F1 hybrid (*S. cerevisiae* crossed with *S. paradoxus*) revealed low levels of misexpression and high level of *S. paradoxus* dominated inheritance for the F1 hybrid transcription, while Wang *et al*. (2015) using an inter-specific F1 hybrid between *S. cerevisiae* and *S. bayanus* showed low levels of additive inheritance and most of the variation in the F1 hybrids transcription to be *S. cerevisiae* dominated. Nonetheless, there were some coherent results between studies; the responses to different environments indicated *trans*-regulated F1 hybrid transcription, while the divergence between parental species indicated *cis*-regulated F1 hybrid transcription (Tirosh *et al*., 2009; Wang *et al*., 2015). Also, different studies found that genes containing a TATA box element(s) in their promoter region were more likely to be differentially transcribed between the F1 hybrid and its parental species (Tirosh *et al*., 2006; Schaefke *et al*., 2013; Wang *et al*., 2015), and evidenced *trans* but not *cis*-regulated genes to be enriched for TATA box elements (Tirosh *et al*., 2009), or both *cis* and *trans*-regulated genes to be enriched for TATA box elements (Wang *et al*., 2015). The TATA box is a conserved element that bounds to a TATA-biding protein and affects the initiation of gene transcription. Only 20% of the yeast genes contain TATA box in their promoter region and those genes have been associated with stress-related response (Tirosh *et al*., 2006). Environmental-stress response related genes (ESR) have evidence a significant differential transcription in a variety of environmental conditions for the same yeast strain (Gasch *et al*., 2000), and this response seems to be common for a variety of yeast species (Gasch, 2007). These genes can also have a significant effect in the F1 hybrid transcription because they constitute around 16% of yeast genes. Whilst a lot of work on F1 hybrids has focused on regulation in transcription in the parental strain, there has been little work on the F1 hybrid transcription itself.

We set out to compare an inter-specific F1 hybrid transcription to its parents under different environmental conditions. For that, we grew two yeast species, *S. cerevisia*e,

S. paradoxus, and its F1 hybrid in two environments in which a different species was superior but in which the F1 hybrid was better than both (Figure 1). We hypothesized that the best-parent heterosis shown by the F1 hybrid was the result of it preferentially transcribing the superior allele at each heterozygous locus (dominance mechanism). According to Goff (2011) model for multigenic heterosis in F1 hybrid, the higher fitness allele is preferentially expressed over the lower-fitness alleles at multiple heterozygous loci. This model explains the advantage of the F1 hybrid over its parents by (1) using allelic specific expression to preferentially transcribe higher and (2) promoting environmental dependent transcription so to transcribe the better equipped parental alleles for a specific environment. We therefore aimed to test if high levels of allele specific expression (*cis*-regulated transcription) in the F1 hybrid could indicate why the hybrid is better fit than both parents for the environments used, and, if so, if the F1 hybrid transcription profile resembles the fitter parent for a specific environment and has the plasticity to change its transcription profile upon environmental change.

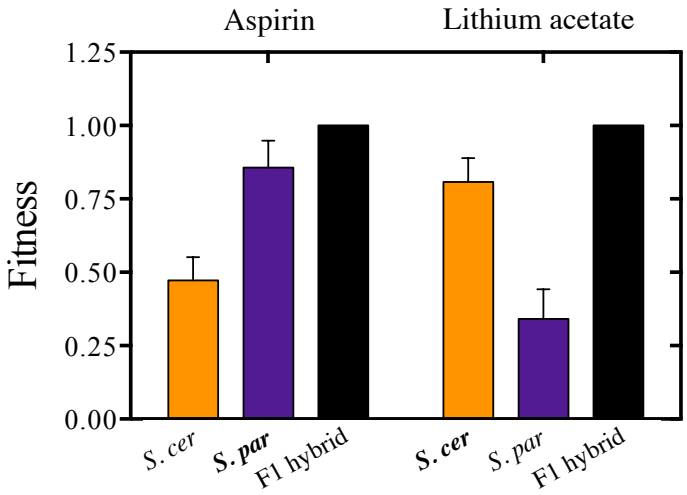


Figure 1: Best-parent heterosis for an inter-specific F1 hybrid in aspirin and lithium acetate environments. Fitness of *S. cerevisiae* parent (*S. cer* in orange), *S. paradoxus* parent (*S. par* in purple), and their F1 hybrid (black). Error bars indicate standard deviation of the mean of three replicate measurements. Parental species with a fitness advantage in the specific environment in **bold.**

In our study, we not only identified the F1 hybrid transcription profile to be significantly correlated to the fitter parent for a specific environment, but also to be enriched for genes whose transcription was dominated by the fittest parent. In addition, our analysis revealed *cis*-regulated transcription was more pervasive than *trans*-regulated transcription, thus indicating the importance of allelic specific

expression in the F1 hybrid under these particular environments. Our results also suggested that the F1 hybrid had the ability to display different transcription profiles under different environments. Overall these results suggest that the F1 hybrid plasticity in transcription might yield its advantage over the parents in different environments, and are in agreement with Goff (2011) model for multigenic heterosis that is based on the ability of hybrid cells to distinguish between fitter parental alleles.

2. Materials and Methods

Yeast strains

We used one *S. cerevisiae* strain (s288c- Cer), one *S. paradoxus* strain (N17- Par) and their inter-specific F1 hybrid (s288c x N17- Hyb or Fcer and Fpar). We chose these strains due to their genetically tractability and well-annotated genome (Bergström *et al*., 2014). All the strains were *ura* auxotrophs. i.e. homozygous for a *ura3* deletion, therefore if there is a fitness effect of this marker, it would be similar for all the strains.

Transcriptome extraction

We grew the yeast strains in minimum media plus uracil (MIN+URA: 0.67% yeast nitrogen base without amino acids, 2%glucose, 0.003% uracil) for 24h with the addition of aspirin (10%) or lithium acetate (1%). We chose these two environments because of the significant levels of heterosis displayed by the F1 hybrid in competition with the *S. paradoxus* parent in lithium acetate, and with the *S. cerevisiae* parent in aspirin (Chapter I and Bernardes *et al*., 2016). We grew the two parental strains and the F1 hybrid separately in aspirin and in lithium acetate in three independent replicates. We extracted the transcriptome of every sample with Master Pure™ Yeast purification Kit from Epicenter®. We sequenced the transcriptome using Illumina HiSeq technology (1x75bp) with rRNA removal. In total 18 samples were sequenced.

Sequence alignment

We mapped the high quality reads identified by FastQC (Andrews, 2010), to a mock hybrid genome with hisat2 software (Kim *et al*., 2015), on average 4.3 million reads were mapped per sample. The mock genome was made by concatenating *S. cerevisiae* (s288c) and *S. paradoxus* (CBS432) genomes as in McManus *et al*. (2014), the genomes used were based on Bergström *et al*. (2014) reference genomes. We identified high accuracy in the mapped reads (Table 1), with *S. cerevisiae* sample reads mapping >99.9% to the s288c genome, and *S. paradoxus* sample reads mapping >99.8% to the CBS432 genome, in addition the hybrid sample reads mapped slightly better to CBS432 genome (51% in aspirin and 50.5% in lithium acetate) than to the s288c genome in both environments (48.9% in aspirin and 49.5% in lithium acetate).

Statistical analysis- correlation and MDS

Our gene list was comprised of 5357 orthologous genes, based on common genes between *S. cerevisisae* (s288c) and *S. paradoxus* (CBS432). We used HTSeq (Anders *et al*., 2015) to count the reads per gene, and DESeq2 (Love *et al*., 2014) to normalize the read counts of every sample by estimation of size factors and estimation of dispersion, which normalized the depth and dispersion parameters of every sample so to make the samples comparable. We used the normalized read counts to verify the correlations between our strain-environment combinations, and to create a Multi-Dimensional Scaling (MDS) analysis based on Euclidean distances. We used Spearman correlation method because it better preserves the relative rank relationships between samples, and is less influenced by skewness and outliers (Reeb *et al*., 2015). We compared the correlations of the strain-environment combinations with a one-tail correlation test. We compared strain-environment combinations transcription by estimating the log fold change of every orthologous gene and the corresponding P-value adjusted for multiple comparisons. For the F1 hybrid samples we used two types of data: we either sum the reads that mapped to *S. cerevisiae* genome and *S. paradoxus* genome for every gene as a overall hybrid transcription (Hyb), or we separated the hybrid reads into two groups; reads that map to the *S. cerevisiae* genome (Fcer) and reads that map to the *S. paradoxus* genome (Fpar), as an estimation of allelic specific expression; when the number of reads mapped to the *S. cerevisiae* parent (Fcer) was different from the number of reads mapped to the *S.*

paradoxus parent (Fpar), or vice-versa, we classified these genes as having differential allelic specific expression in the F1 hybrid, meaning one parental allele was differentially transcribed from the other parental allele in the F1 hybrid.

Environment	Samples	Total reads	Mapped to s288c	Mapped to CBS432	Mapped to s288c (%)	Mapped to CBS432 (%)
Aspirin	Cer_1	3 330 022	3 328 527	1 495	99.96	0.05
	Cer_2	4 984 190	4 981 355	2 835	99.94	0.06
	Cer_3	4 444 183	4 441 873	2 210	99.95	0.05
	Par_1	3 012 208	5 009	3 007 199	0.17	99.83
	Par_2	4 349 332	7 001	4 342 331	0.16	99.84
	Par_3	4 544 130	7 608	4 536 522	0.17	99.83
	Hyb_1	4 081 434	1 999 179	2 082 255	48.98	51.02
	Hyb_2	4 290 424	2 098 242	2 192 182	48.90	51.09
	Hyb_3	4 450 982	2 187 693	2 263 289	49.15	50.85
Lithium acetate	Cer_1	2 915 001	2 913 340	1 661	99.94	0.06
	Cer_2	3 387 356	3 385 587	1 769	99.95	0.05
	Cer_3	2 521 850	2 520 481	1 369	99.95	0.05
	Par_1	4 422 380	6 544	4 415 836	0.15	99.85
	Par_2	6 093 155	9 656	6 083 499	0.16	99.84
	Par_3	5 697 444	5 743	5 691 701	0.10	99.9
	Hyb_1	4 945 944	2 450 013	2 495 931	49.54	50.46
	Hyb_2	5 273 607	2 607 322	2 666 285	49.44	50.56
	Hyb_3	4 703 861	2 329 583	2 374 278	49.52	50.47

Table 1: RNA-seq reads for samples mapped to *S. cerevisiae* and *S. paradoxus* genomes. Total reads of every sample that mapped to *S. cerevisiae* (s288c) and *S. paradoxus* (CBS432) genomes. Samples were comprised of *S. cerevisiae* (Cer), *S. paradoxus* (Par) and their hybrid (Hyb) in two different environments, aspirin and lithium acetate. The numbers in samples indicate the three replicates.

TATA box and ESR

We used the list of environmental stress response genes (ESR), and a list of genes that contain a TATA box element in their promoter region. We obtained a list of ESR genes based on Gasch *et al*. (2000), comprised of 813 genes previously identified for *S. cerevisiae* in different environments, and a list of genes containing a TATA box elements in their promoter region for various yeast species from Basehoar *et al*. (2004) and selected for the 1004 genes in common between *S. cerevisiae* and *S. paradoxus*. We used these lists to detect any enrichment of these particular elements.

F1 Hybrid transcription analysis

We compared the changes in gene transcription between *S. cerevisiae* parent (Cer) and the F1 hybrid (Hyb) with changes in gene transcription between *S. paradoxus*

parent (Par) and the F1 hybrid (Hyb) in aspirin and lithium acetate environments. We analysed the log fold change values of the transcribed genes between Cer-Hyb and Par-Hyb comparisons in both environments; log fold changes within [-1.25; 1.25] indicated conserved transcription (based on McManus *et al*., 2014)), while changes outside the conserved transcription limits were classified as additive, overdominant, and dominant transcription. F1 hybrid transcription was additive when it was an intermediate value between the parental transcriptions (Cer>Hyb>Par or Cer<Hyb<Par), F1 hybrid transcription was *S. paradoxus* overdominant (or *S. cerevisiae* underdominant) when is higher or lower than both Par and Cer (Hyb>Par>Cer and Cer>Par>Hyb), F1 hybrid transcription was *S. cerevisiae* overdominant (or *S. paradoxus* underdominant) when is higher than both Cer and Par (Hyb>Cer>Par and Par>Cer>Hyb), F1 hybrid transcription was *S. cerevisiae* dominant when the F1 hybrid transcription was similar Cer but not to Par (Hyb=Cer ≠ Par), and F1 hybrid transcription was *S. paradoxus* dominant when the F1 hybrid transcription was similar to Par and but not to Cer (Hyb=Par ≠ Cer).

Allelic specific transcription analysis

We compared the changes in gene transcription between *S. cerevisiae* parent (Cer) and *S. paradoxus* parent (Par), and between allelic specific expression in the F1 hybrid (Fcer compared to Fpar), for aspirin and lithium acetate environments. We applied a model with DESeq2 based on (Lovell *et al*., 2016), which takes in account the genotype (Cer or Par) and the generation (parent or F1 hybrid) to assess genes that were *cis*-regulated, and trans-regulated:

$$log2q_{ij} = \beta_0 + \beta_A A_i + \beta_F F_i + \beta_{AF} A_i * F_i$$

For a sample *i* and gene *j*, A_i=1 if sample has a Cer genotype and A_i=0 has a Par genotype and, F_i=1 if sample is from a parent (Cer or Par) and F_i=0 if sample is from the F1 hybrid (Fcer or Fpar). We used a Likelihood-ratio test (LRT) to compare our model with and without interaction, and took into account estimation of size factor and estimation of dispersion. *Cis*-regulated genes were selected by determine β_A or the difference between genotype within the F1 hybrid (Fcer *vs*. Fpar, P<0.05), and *trans*-regulated genes were selected by determine the interaction β_{AF} or the difference between genotype between generations ((Fcer *vs*. Fpar)/(Cer *vs*. Par), P<0.05). Genes with both *cis* and *trans*-regulation characteristics were identified, and distinguished

into two categories: *cis-trans* regulated (*cis-trans* (+)) when the F1 hybrid difference (log fold change) between alleles was higher or lower than the parental difference, or vice-versa, and *cis-trans* antagonistic (*cis-trans* (-)) when the F1 hybrid difference between alleles was inverse to the parental difference. In addition, we plotted the log fold change of comparison between parental transcriptions (Cer-Par) against the log fold change of allelic specific expression of the parental alleles in the F1 hybrid (Fcer-Fpar). We expected a regression with slope 0 if there was only *trans*-regulation, and a regression with slope 1 if there was only *cis*-regulation of the F1 hybrid parental alleles.

Gene regulation by environment and genotype

We calculated how many genes were differentially transcribed between environments within species (compared Cer, Par and F between environments), and how many genes were differentially transcribed between parents and F1 hybrid in the same environment (Cer-F and Par-F within environments). Moreover we applied a model with DESeq2 based on (Lovell *et al*., 2016), which takes in account the genotype (Cer or Par and Fcer or Fpar) and the environment (aspirin or lithium acetate), to assess genes whose differentially transcription was due to environment or to genotype, or to the interaction between environment and genotype:

$$log2q_{ij} = \beta_0 + \beta_A A_i + \beta_E E_i + \beta_{AE} A_i * E_i$$

For a sample *i* and gene *j*, A_i=1 if sample has a Cer genotype and A_i=0 has a Par genotype and, E_i=1 if sample is from aspirin and E_i=0 if sample is from lithium acetate. We used a Likelihood-ratio test (LRT) to compare our model with and without interaction, and took into account the estimation of size factor and estimation of dispersion. Changes in gene transcription due to the genotype (Cer/Fcer *vs*. Par/Fpar) were selected by determine β_A, and changes in gene transcription due to environment (aspirin *vs*. lithium acetate) were calculated by determine β_E, finally the interaction between genotype and environment was identified by the coefficient β_{AE} or the difference between genotype between environments (aspirin[Cer *vs*. Par]/lithium acetate[Cer *vs*. Par], P<0.05).

3. Results

The read counts for the orthologous genes of the same strain-environment combinations were organized in three independent replicates, and the strain-environment combinations were identified as their transcription profiles or the pattern and quantity of genes transcribed. The results were divided into two groups; first we analysed the transcription profiles of *S. cerevisiae* parent (Cer), *S. paradoxus* parent (Par) and their F1 hybrid (Hyb), and second we analysed the transcription profiles of Cer parent, Par parent and their F1 hybrid allelic specific expression when mapped to *S. cerevisiae* genome (Fcer) and to *S. paradoxus* genome (Fpar). Finally we compared both groups to combine transcription classification and transcription regulation.

F1 hybrid comparisons with parental strains

Correlations

We normalized the raw read counts for the parents and the F1 hybrid by taking in account the estimation of size factors and estimation of dispersion for negative binomial distributed data. We used the normalized data to perform a correlation test using the Spearman method. All comparisons between replicates of the same strain-environment group were highly correlated (ro>0.95- Figure 2). Correlations between parents (Cer-Par) in the different environments evidenced the lowest correlation between the strain-environment combinations. The F1 hybrid was more correlated between aspirin and lithium acetate (average=0.871) than the Cer (average=0.851) and Par (average=0.844). Further analysis revealed the F1 hybrid was equally correlated to Par and Cer parents in aspirin ($z=34.13$, $P>0.710$), and to be equally correlated to Cer and Par parents in lithium acetate ($z=14.76$, $P>0.119$).

MDS

We compared the similarities between the normalized transcription profiles of the parents and the F1 hybrid in both environments with the use of Multi-Dimensional Scaling (MDS). Samples were separated into six groups of strain-environment combinations (Figure 3); transcription profiles of samples grown in aspirin showed fewer similarities than in lithium acetate. In aspirin, Hyb transcription profile showed more similarities to with Par parent than to Cer parent, and clustered between the two

parents. In lithium acetate, parents showed great similarities, thus clustering together, while Hyb showed fewer similarities with both Cer and Par parents. The Hyb show higher similarities between aspirin and lithium acetate than both its parents.

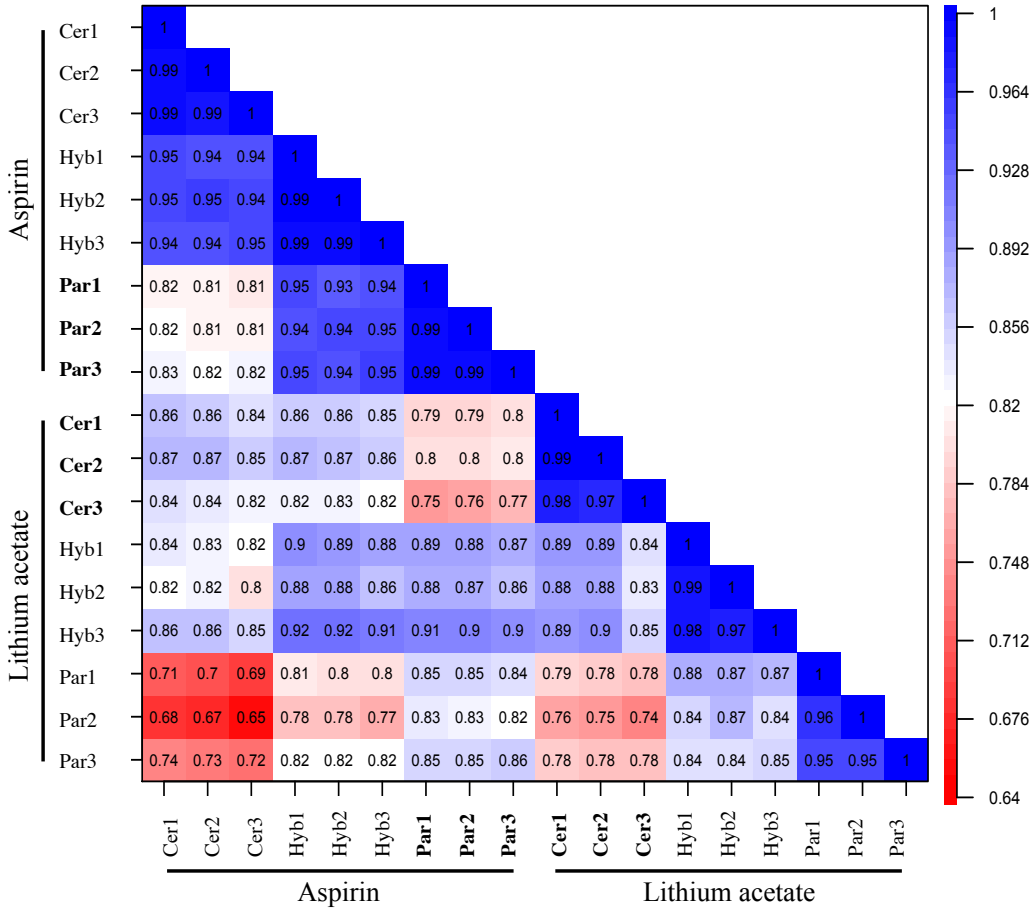


Figure 2: Correlation plot between replicate parental and F1 hybrid samples. Cer (*S. cerevisiae*), F1 hybrid (Hyb), and Par (*S. paradoxus*) in aspirin and lithium acetate environments. Blue represented higher correlations and red lower correlations. Parental species with a fitness advantage in the specific environment with a **bold** legend.

F1 Hybrid transcription analysis

We calculated the differential transcription of all orthologous genes in pairwise comparisons between Cer parent and F1 hybrid (Cer-Hyb), and between Par parent and the F1 hybrid (Par-Hyb) for both environments. We analysed log fold change values for all orthologous genes between the two comparisons in both environments (Figure S1). We selected for genes that were differentially transcribed for one or both comparisons (P-value<0.05, adjusted for multiple comparisons) and analysed their log fold change values in the respective environments (Figure 4).

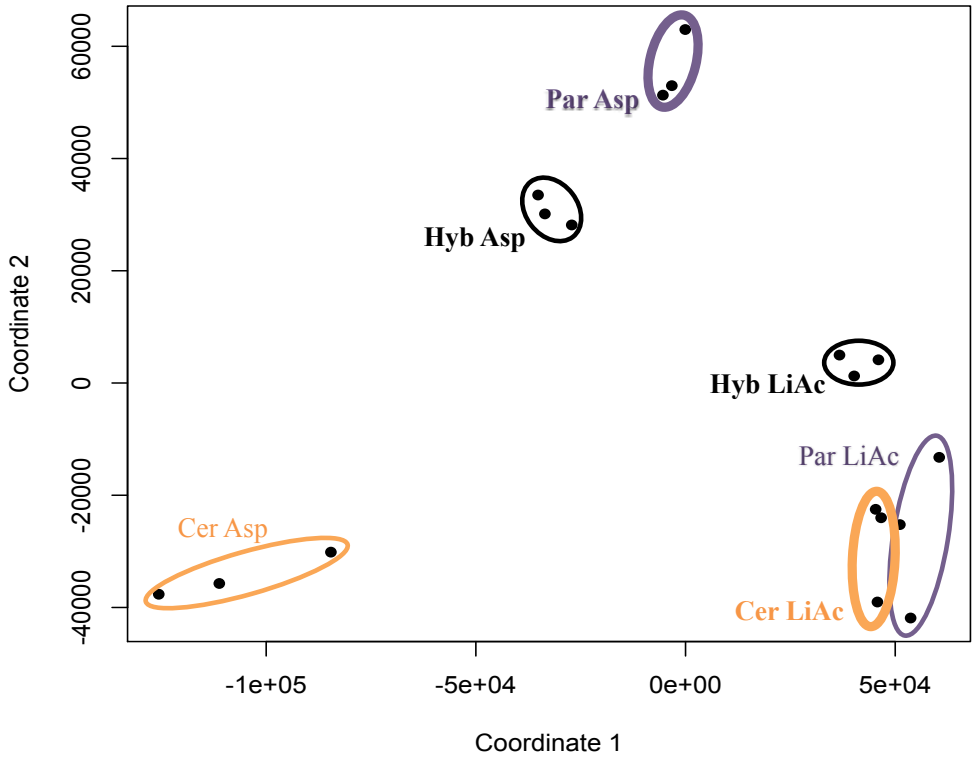


Figure 3: MDS analysis of parental and F1 hybrid transcription profiles. Cer (*S. cerevisiae*), Par (*S. paradoxus*), and their F1 hybrid (Hyb) in in aspirin and lithium acetate environments. Cer in orange, Par in purple, and Hyb in black. Aspirin environment (Asp) and lithium acetate (LiAc). Parental species with a fitness advantage in the specific environment within a thicker circle and with **bold** labels.

In aspirin, 3964 genes were differentially transcribed in one or both comparisons; 3034 genes were differentially transcribed between Cer-Hyb, 2587 genes were differentially transcribed between Par-Hyb, and 1657 genes in common between the two comparisons. We identified 223 genes that were uniquely differentially transcribed between Par-Hyb, while 360 genes, significantly more, were uniquely differentially transcribed between Cer-Hyb ($z=4.038$, $df=1$, $P<0.001$). The uniquely differentially transcribed genes between Cer-Hyb were enriched for mitochondrial translation and single organism process (GO terms- http://www.geneontology.org/). In lithium acetate, 3995 were differentially transcribed in one or both comparisons; 2991 genes were differentially transcribed between Cer-Hyb, while 2324 genes were differentially transcribed between Par-Hyb, and 1320 genes in common between both comparisons. We detected 428 genes that were uniquely differentially transcribed

between Cer-Hyb, while 261 genes, a significant lower number, were uniquely differentially transcribed between Par-Hyb ($z=4.531$, $P<0.001$). The genes that were uniquely differentially transcribed between Par-Hyb were enriched for catalytic activity. Further analysis indicated that 549 genes were differentially transcribed between Cer-Hyb and between Par-Hyb in both environments. These genes were enriched for processes such as carboxylic acid metabolic, oxidation-reduction and nitrogen compound. Moreover, from the 549 genes 32 were genes with a TATA box element in their promoter significantly less than expected by random ($\chi^2=6.75$, $P=0.009$), while 105 genes were in common with the ESR gene list significantly more than expected by random ($\chi^2=3.838$, $P=0.050$).

In aspirin, the majority of the differentially transcribed F1 hybrid genes (3361 genes or 84.8%) had a log fold change within the limits of parental transcription ([-1.25; 1.5]) and was classified has conserved. The remaining genes transcription was classified as: additive for 47 genes when Hyb was an intermediate of both parents (7.8%), *S. paradoxus* overdominant for 17 genes (2.8%), *S. cerevisiae* overdominant for 6 genes (1%), *S. paradoxus* dominant for 390 genes when Hyb was similar to *S. paradoxus* but not to *S. cerevisiae* (64.7%) and only 143 genes were *S. cerevisiae* dominant when Hyb was similar to *S. cerevisiae* but not to *S. paradoxus* (23.7%) (Figure 4A). In lithium acetate, most of the differentially transcribed hybrid genes (3141 genes or 78.6%) had a log fold change within the limits of parental transcription ([-1.25; 1.5]) and were classified has conserved. The remaining genes transcription was classified as: additive for 47 genes (5.5 %), *S. paradoxus* overdominant for 27 genes (3.2%), *S. cerevisiae* overdominant for 39 genes (4.6%), *S. cerevisiae* dominated for 412 genes (48.2%), and only 326 genes were *S. paradoxus* dominated (38.2%) (Figure 4B).

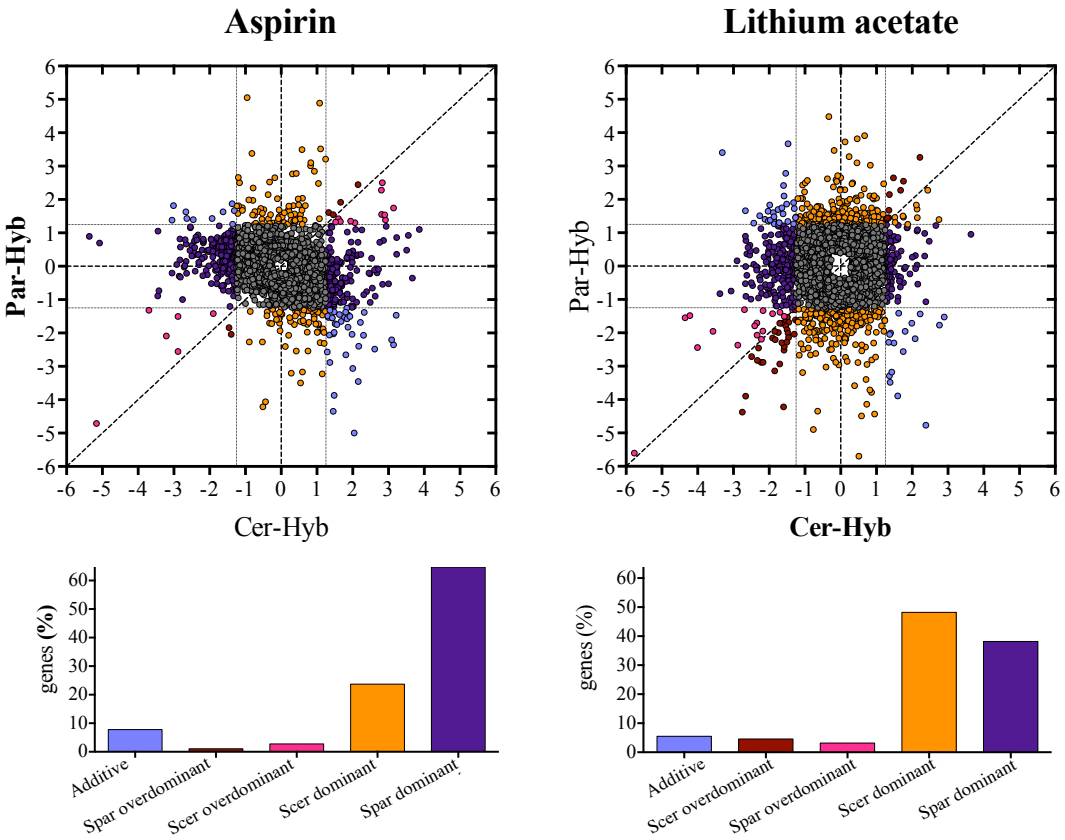


Figure 4: Changes in F1 hybrid transcription in comparison with Cer and Par parents in aspirin and lithium acetate environments. Axis refers to log fold change values for Cer-Hyb comparison against log fold change for Par-Hyb comparison. A and B graphs refer to percentage of non-conserved genes that were differentially transcribed in aspirin and lithium acetate. Points, Classification (%), and Bars refer to F1 hybrid transcription as conserved (grey), Additive (blue), *S. paradoxus* overdominant (Spar overdominant-crissom), *S. cerevisiae* overdominant (Scer overdominant-pink), *S. cerevisiae* dominant (Scer dominant-orange), and *S. paradoxus* dominant (Spar dominant- purple). Parental species with a fitness advantage in the specific environment with a **bold** legend.

F1 hybrid allelic specific expression

Correlations

We normalized the raw read counts for the parents and the allelic specific expression of the F1 hybrid by taking in account the estimation of size factors and estimation of dispersion for negative binomial distributed data. We used the normalized data to perform a correlation test using the spearman method. All comparisons between replicates of the same strain-environment combinations were highly correlated

(ro>0.95- Figure 5). Correlations between parents in different environments evidenced the lowest correlation between strain-environment combinations, while the F1 hybrid allelic specific expression showed better correlations to their respective parents in the two environments. We compared the parents with the F1 hybrid allele specific expression in aspirin and identified a significant higher correlation between Par-Fpar than between Cer-Fcer (z=5.31, P<0.001), while in lithium acetate there was a higher correlation between Cer-Fcer in comparison to Par-Fpar but was not significant (z=2.54, P<0.001).

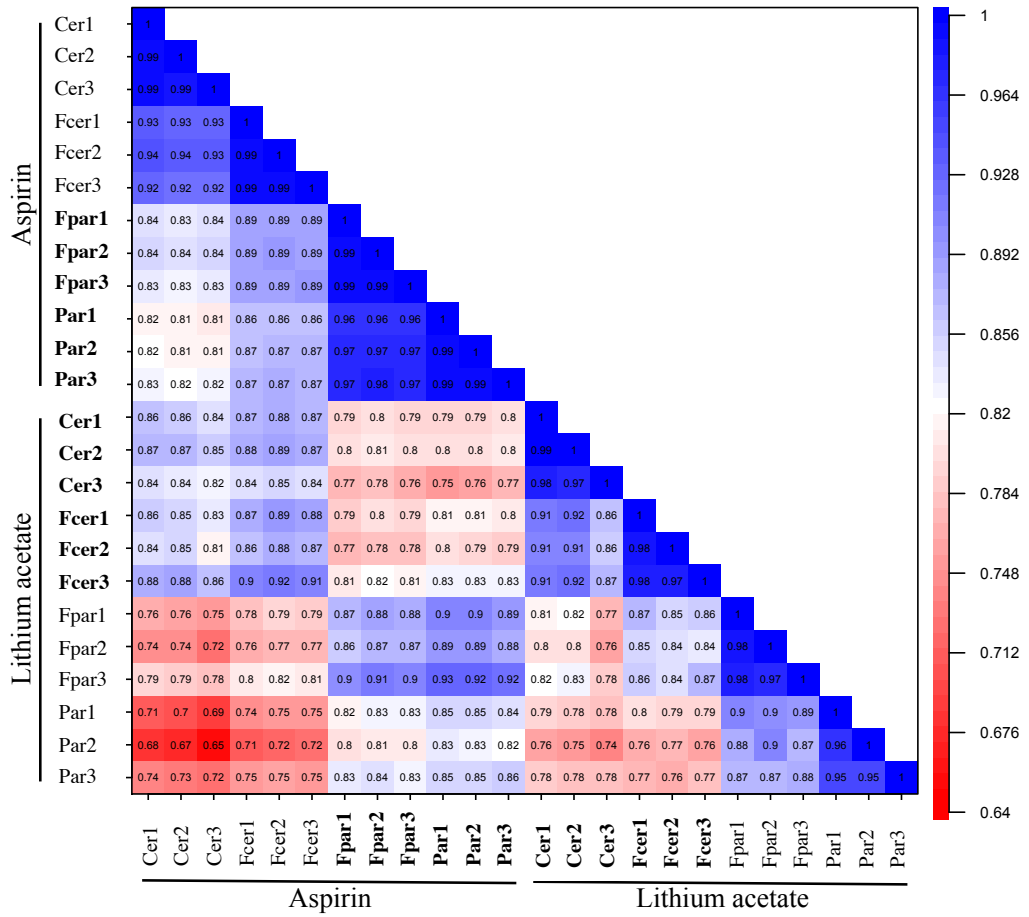


Figure 5: Correlation plot between replicate parental and F1 hybrid allelic specific samples. Cer (*S. cerevisiae*), F1 hybrid (Fcer and Fpar), and Par (*S. paradoxus*) in aspirin and lithium acetate environments. Blue represents highly correlations and red low correlations. Parental species with a fitness advantage in the specific environment with a **bold** legend.

MDS

We used the normalized transcription profiles to compare the similarities between parents and the F1 hybrid's allelic specific expression in both environments by the means of a MDS. The samples were separated into eight groups of strain-environment combinations (Figure 6); the transcription profiles of the parents grown in aspirin showed fewer similarities than the parents grown in lithium acetate. In aspirin, the transcription profile of the F1 hybrid allelic specific expression Fpar showed greater similarities to the Par parent than Fcer showed to the Cer parent. In lithium acetate the opposite happened; the transcription profile of the F1 hybrid allelic specific expression Fcer showed greater similarities to both Cer and Par parents than Fpar to the Par parent.

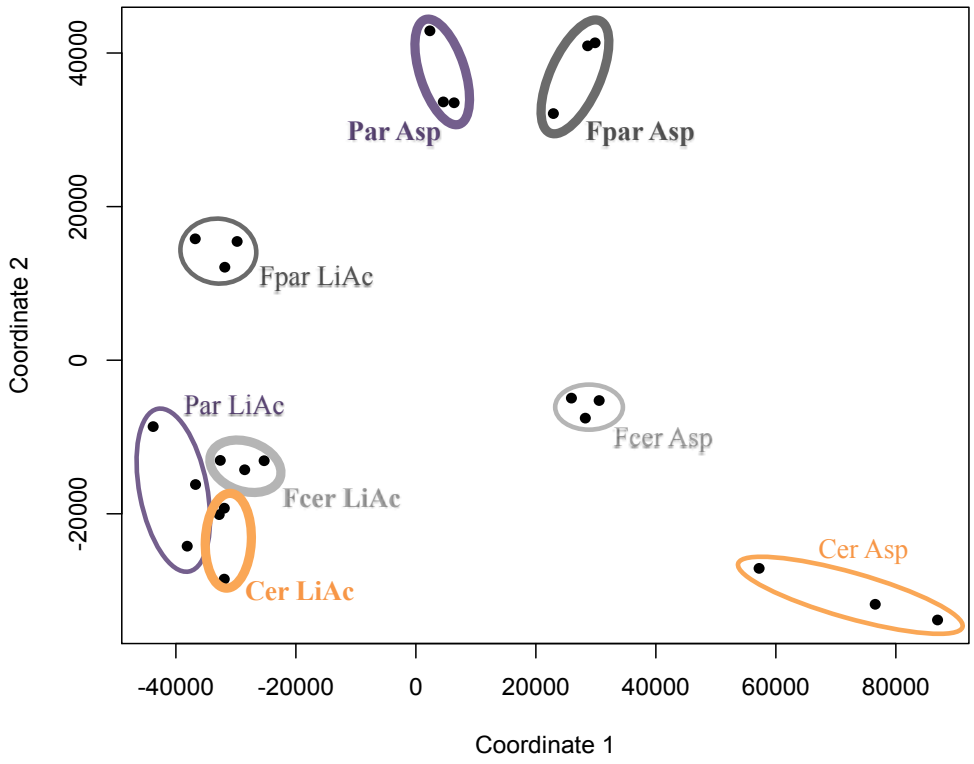


Figure 6: MDS analysis of parental and F1 hybrid allelelic specific transcription profiles. Cer (*S. cerevisiae*), Par (*S. paradoxus*), and their F1 hybrid (Fcer and Fpar) in aspirin and lithium acetate environments. Cer in orange, Par in purple, Fcer in light grey, and Fpar in dark grey. Aspirin environment (Asp) and lithium acetate (LiAc). Parental species with a fitness advantage in the specific environment within a thicker circle and with **bold** labels.

Allelic specific expression analysis

We calculated the differential transcription of all orthologous genes in pairwise comparisons between parents (Cer-Par) and between allelic specific expressions of the F1 hybrid (Fcer-Fpar). We analysed log fold change values for every orthologous gene between the two comparisons. In aspirin, there were 3886 genes differentially transcribed; 2867 genes between parents (Cer-Par), 3179 genes between parental alleles in the F1 hybrid (Fcer-Fpar), and 2160 genes in common between both comparison. In lithium acetate, there were 4607 genes differentially transcribed; 3741 genes between parents, 3968 between parental alleles in the F1 hybrid, and 3102 genes in common between both comparisons.

We used a model to calculate the coefficients for all orthologous genes that reflect *cis*-regulation and *trans*-regulation of transcription. For *cis*-regulation we selected for genes with differential transcription between parental alleles in the F1 hybrid (Fcer-Fpar) in aspirin and lithium acetate, while for *trans*-regulation we determined genes with differential transcription between genotypes (Cer-Par or Fcer-Fpar) and between generations (parent-F1 hybrid) (Figure 7). Genes that showed evidence of both *cis*- and *trans*-regulated transcription were classified into *cis-trans* regulated or antagonistic. We identified in aspirin, 1795 *cis*-regulated genes (33.5%), 767 *trans*-regulated genes (14.3%), and 1416 genes with both *cis*- and *trans*-regulation characteristics; 903 genes were *cis-trans* regulated (*cis-trans* (+) 17%) and 508 genes were *cis-trans* antagonistic (*cis-trans* (-) 9.5%). We identified in lithium acetate, 1848 *cis*-regulated genes (34.5%), 776 *trans*-regulated genes (14.5%), and 938 genes with both *cis*- and *trans*-regulation characteristics; 527 genes were *cis-trans* (*cis-trans* (+) 9.8%) and 411 genes were *cis-trans* antagonistic (*cis-trans* (-) 7.7%). Further analysis identified a significant enrichment for TATA box elements in the promoter region for *trans*-regulated genes in lithium acetate (χ^2=4.762, P=0.029) but not in aspirin.

We compared the log fold change of the parents against the log fold change of the allelic specific expression in the F1hybrid for both environments (Figure 7). There was a close positive relationship between parental differences and F1 hybrid allelic specific differences. The difference in gene transcription between parents (Cer-Par) tended to be higher than the difference between allelic specific expressions of F1 hybrid alleles (Fcer-Fpar) in both environments; aspirin with a slope of 0.566 ($F_{1,5356}$=5066, P<0.001) and lithium acetate with a slope of 0.500 for lithium acetate

($F_{1,5356}$=3480, P<0.001). Thus *cis*- and *trans*-regulated transcriptions were present. We compared for each orthologous genes its classification (additive, overdominant, or dominant) with its regulation of transcription and all classes in both environments evidenced a higher number of *cis*-regulated genes than trans-regulated genes.

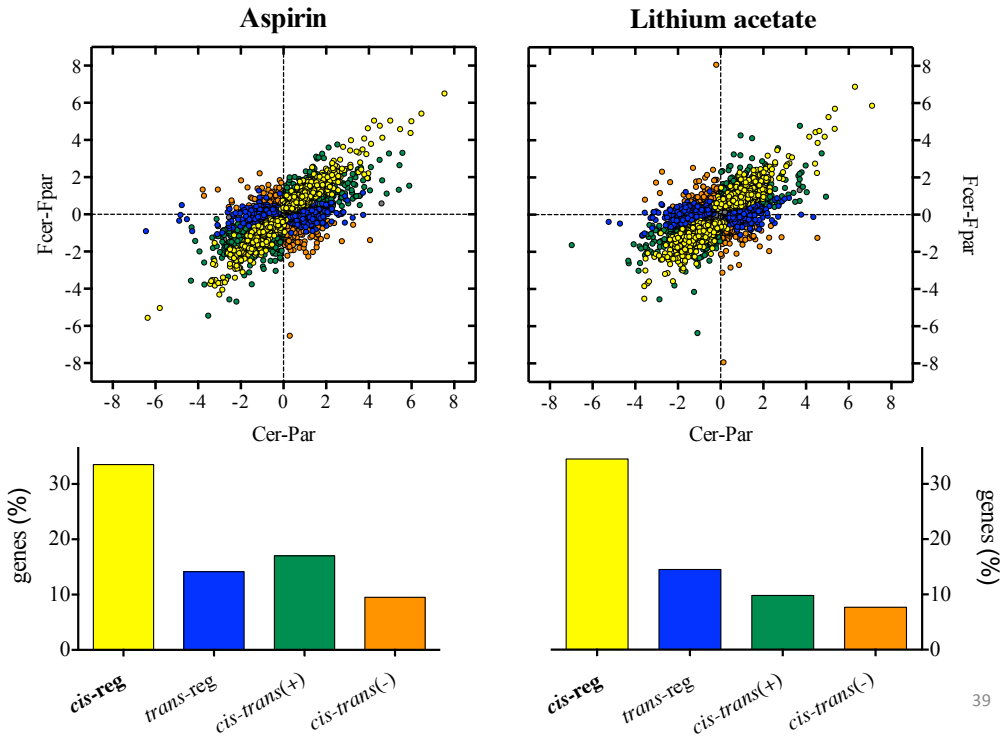


Figure 7: Changes in parental transcription and allelic specific expression of the F1 hybrid in aspirin and lithium acetate environments. Axis refers to log fold change values for Cer-Par comparison against log fold change for Fcer-Fpar comparison. A and B graphs refer to percentage of genes that were *cis*-regulated genes (yellow), *trans*-regulated genes (blue), *cis-trans* (+) regulated genes (green) and *cis-trans* (-) regulated genes (orange).

Gene regulation by environment and species

Based on the raw read counts, normalized as previously described, we calculated the differential transcription of all orthologous genes in comparisons between environments (aspirin and lithium acetate) for Cer parent, Par parent, F1 hybrid, and the allelic specific expression of the F1 hybrid (Fcer and Fpar). We analysed log fold change values for every gene between environments. For Cer parent there were 3171 genes differentially transcribed between environments, 2722 for Par parent, 3001 genes for the F1 hybrid, 2744 genes for the Fcer allele, and 2767 genes for Fpar allele. There were 1237 genes in common between the differentially transcribed genes of

Cer, Par and F, from those a significant number of 230 genes were in common with the ESR gene list (χ^2=4.452, P=0.035). We used a model to calculate the coefficients for all orthologous genes that indicated if changes in transcription were due to genotype, to environment or to the interaction between genotype and environment. Our results showed 1165 genes (22%) with only a genotype effect, 1112 genes (21%) were affected by only environment, and a similar number of 1070 genes (21%) were affected by the interaction between genotype and environment. The other 36% of genes did not have a specific effect in transcription. Further analysis revealed that genes that had changes in transcription due to genotype were enriched for *cis*-regulation in transcription (51%), while genes that had changes in transcription due to environment were enriched for *trans*-regulation (17%). The genes with changes in transcription due to the interaction between genotype and environment were enriched for *cis*-regulated genes as well (36%) (Figure 8).

4. Discussion

Here we found how the transcription of two yeast species, *S. cerevisiae* (Cer) and *S. paradoxus* (Par), relates to the transcription of their F1 hybrid. We analysed the transcription profile of the F1 hybrid and its parents in two different environments, aspirin and lithium acetate. The environments used in this study were chosen because previous work that had identified an advantage of *S. paradoxus* over *S. cerevisiae* in aspirin, and an advantage of *S. cerevisiae* over *S. paradoxus* in lithium acetate, while their F1 hybrid outcompeted both parents in both environments (Chapter I and Bernardes *et al*., 2016). We were intrigued by the F1 hybrid's ability to outcompete both parents, a phenomenon know as heterosis (Shull, 1908) , and if variations in the F1 hybrid's transcription could explain its fitness advantage. We wanted to show that instead of a deregulation in transcription, there could be induction and a directionality of the variations in the F1 hybrid transcription so it resembles the fitter parent for a specific environment (Goff, 2011).

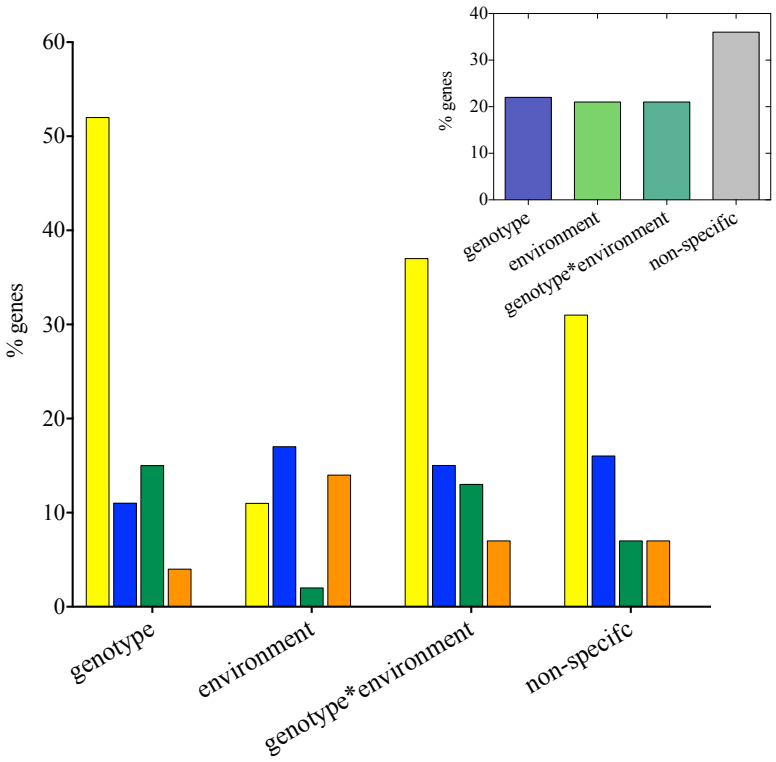


Figure 8: Classification of gene regulation by genotype and/or environment. Gene regulation identified as *cis*-regulation (yellow), *trans*-regulation (blue), *cis-trans* regulation (green) and *cis-trans* antagonistic (orange) by changes in gene transcription due to genotype (blue), environment (green), genotype-environment interaction (blue-green), and non-specific (grey).

We identified a high similarity between the F1 hybrid and the fitter parent for a specific environment when we analysed the F1hybrid transcription as a whole. The MDS analysis (Figure 3) indicated that the F1 hybrid was capable of coordinating its transcription to resemble the fitter *S. paradoxus* parent in aspirin environment. In addition, the majority of differentially transcribed genes of the F1 hybrid with a non-conserved inheritance (Figure 4) were *S. paradoxus* dominant in aspirin (64.7% over 23.7% for *S. cerevisiae* dominant) and *S. cerevisiae* dominant in lithium acetate (48.2% over 38.2% for *S. paradoxus* dominant) (Figure 4). A recent study by McManus *et al*. (2014), also classified the transcription of a similar inter-specific yeast F1 hybrid and identified most of the non-conserved F1 hybrid genes to be *S. paradoxus* dominated over *S. cerevisiae*, unfortunately this study did not produce any fitness assays for the environments used, so the reason for the abundance of *S. paradoxus* dominant transcription of F1 hybrid remains unknown. It is difficult to distinguish if the F1 hybrid transcription resembles *S. paradoxus* transcription because it was advantageous or if it was due to stochasticity, although randomness seems unlikely because the F1 hybrid *S. paradoxus* dominated transcription was two thirds higher than the *S. cerevisiae* dominated transcription (McManus *et al*., 2014).

Also, a similar study used a inter-specific F1 hybrid between *S. cerevisiae* and *S. bayanus* (Wang *et al*., 2015) which showed greater levels of *S. cerevisiae* dominated transcription over *S. bayanus*, again the question remains if this was due to pure chance or if it was dependent on the environment the researchers grew the yeast in? Our study might answer this question; because the environments were carefully chosen, we knew *a priori* that *S. paradoxus* would perform better in aspirin and *S. cerevisiae* in lithium acetate. When the F1 transcription was dominated by *S. paradoxus*-like transcription in aspirin and then shifts to be dominated by *S. cerevisiae*-like transcription in lithium acetate indicated a coordination transcription by F1 hybrid. This might render the F1 hybrid the fitness advantage we had previously detected when in direct competition with its parents (Chapter I and Bernardes *et al*., 2016). Thus the prevalence of the dominant transcription of the fitter parent for a specific environment indicates a directionality of F1 hybrid transcription, and the plasticity to modify its transcription environmental change (Gasch *et al*., 2000).

S. cerevisiae and *S. paradoxus* are closely related yeast species that can hybridise, even though their populations can genetically diverged up to 14%, according to SNP data. The genetic divergence between parents of the F1 hybrid gives us the ability to distinguish in hybrids between reads that mapped to *S. cerevisiae* from reads that mapped to *S. paradoxus*. The distinction between parental alleles in the F1 hybrid enables the detection of allelic specific expression (Figure S1); if the parental alleles are differentially transcribed we identify them as being *cis*-regulated, while differences in the transcription of both parental alleles in comparison to the parents is due to *trans*-regulation. We wanted to test if the ability of the F1 hybrid to resemble the fitter parent previously detected was based on the capacity of allelic specific expression?

We identified a high similarity between the F1 hybrid allelic specific expression and the fitter parent for a specific environment. Both correlation plot and MDS analysis (Figure 5 and Figure 6) indicated that the F1 hybrid transcription of *S. paradoxus* alleles was more similar to the *S. paradoxus* parent than the transcription of *S. cerevisiae* alleles to the *S. cerevisiae* parent in aspirin, the environment that favours *S.*

paradoxus (z=5.31, P<0.001). While the inverse occurred in lithium acetate, the environment that favours *S. cerevisiae*, where F1 hybrid transcription of *S. cerevisiae* alleles was more similar to the *S. cerevisiae* parent than the transcription of *S. paradoxus* alleles to the *S. paradoxus* parent (z=2.54, P<0.001).

The F1 hybrid distinguishs between parental alleles and preferentially transcribes the fitter parental allele for a given locus using allelic specific expression. Not only the ability to coordinate the allelic specific expression to resemble the fitter parent occurres, but the F1 hybrid also has the plasticity to modify its transcription upon environmental change. Both these occurrences are in agreement with Goff (2011) model for multigenic heterosis; where the F1 hybrid itself can induce transcription of the most advantageous parental alleles for a specific environment. Even though this model is based in the stability of the resulting proteins and we did not collect this type of data, the transcription response that the model refers to was undeniable. However this model relies on the ability of an individual cell to produce a transcriptional response upon changes in protein stability, and such mechanism has never been detected. We can also think of a different model, where different cells acquire divergent transcription profiles and, depending on the environment, one transcription profile will be favoured over the others. The cell(s) with the fitter transcription profile, i. e. the transcription profile that resembles the fitter parent for a specific environment acquires a fitness advantage that spreads the advantageous transcription profile throughout the cell population. In our study we cannot detect if different cells display different transcription profiles because our data only summarizes the average transcription profile of a cell population, however the response to this model would be similar to what we detected. Neither models for molecular heterosis can be tested with this study but further experiments, using techniques such as single cell sequence could provide the advances needed to test the basis for heterosis at a molecular level. Nevertheless, as far as we know, this was the first time heterosis at the transcriptome level was directly tested and described in literature.

In addition, we also identified *cis*-regulation to be more pervasive than *trans*-regulation under both environments (Figure 7), the striking higher levels of *cis*-regulation in the F1 hybrid are also supported Goff (2011) model for multigenic

heterosis at molecular level. These results were not in agreement with previous studies, which identified *trans*-regulation to be a more common regulation of transcription between intra-specific F1 hybrids and their parents (Emerson *et al*., 2010; Schaefke *et al*., 2013). These contradicting results could be due to a decrease in power of identifying allelic specific expression in intra-specific F1 hybrids. Our results also differ from Artieri and Fraser (2014) where they found similar levels of *cis*- and *trans*-regulation using an inter-specific F1 hybrid between *S. paradoxus* and *S. cerevisiae*. Once more, different methodologies might be the reason behind such contradicting results; Artieri and Fraser (2014) compare the F1 hybrid to haploid *S. cerevisiae* parent and diploid *S. paradoxus* parent, so the increased variation of *trans*-regulation might be completely biases because of the different states of ploidy of the parents. This would also explain the enrichment for mating and fertilization genes they identified, because the haploid *S. cerevisiae* parent would be transcribing mate-specific genes but not diploid specific genes like the diploid *S. paradoxus* parent would. Thus the literature seems to disagree in which type of regulation is the major force affecting the divergence between parental transcriptions, the differences could be due to the variety of environments or the type of F1 hybrids we and other studies used, so we will abstain to speculate or give definitive proof on this subject. Nevertheless our data supports the claim that *cis*-regulated transcription is more pervasive for species divergence and *trans*-regulated transcription has a bigger effect in response to environment (Figure 8) as in Tirosh *et al*. (2009).

We detected low levels of misexpression in the F1 hybrid, identified as *S. paradoxus* and *S. cerevisiae* overdominant transcription (Figure 4 and Figure S3); overall there were 0.45% genes misexpressed in aspirin, and 1.2% in lithium acetate, much lower than previously detected by McManus *et al*. (2014) where 7% of genes were misexpressed, and more similar but still lower than Tirosh *et al*. (2009) where 1-6% of the genes were misexpressed. This evidences how different methodologies might affect the results of different studies; the three studies used an inter-specific F1 hybrid between *S. paradoxus* and *S. cerevisiae*, however Tirosh *et al*. (2009) used microarrays to measure variations between the F1 hybrid and its parents, while McManus *et al*. (2014) used RNA-seq but different statistical methods from the ones used in this study.

Interestingly differentially transcribed genes between environment for *S. cerevisiae*, *S. paradoxus* and their F1 hybrid, revealed enrichment for Environmental-Stress Response genes (ESR) (χ^2=4.452, P=0.035). Also the differentially transcribed genes between parents and the F1 hybrid in common between aspirin and lithium acetate environments showed enrichment for ESR genes (χ^2=3.838, P=0.050). These genes are commonly differentially transcribed when yeast copes with environmental change (Gasch *et al*., 2000), and up- or down-regulation of these genes seems to be common for a variety of yeast species (Gasch, 2007). Both aspirin and lithium acetate environments were affected by ESR general response, indicating these environments were stressful enough to induce a general response to stress in both parents and the F1 hybrid. Our results supported and added to the claim that ESR genes are crucial in transcription response upon environmental change. On the contrary, the presence of TATA box elements did not affect differentially transcribed genes between the parents and the F1 hybrid for either environments, and we could only account for an enrichment of differentially transcribed *trans*-regulated genes in lithium acetate (χ^2=4.762, P=0.029). TATA box presence was previous related with divergence of evolutionary lineages such as two different yeast species, and the genes with this element are normally common in the promoters of stress-related genes (Tirosh *et al*., 2006). However our results did not indicate any effect of the TATA box elements in the F1 hybrid differentially transcribed or stress-related genes thus its presence was not a generalized factor as in Tirosh *et al*. (2009) and Wang *et al*. (2015).

When we analysed the unique differentially transcribed genes between *S. cerevisiae* and F1 hybrid in aspirin environment, there was an enriched for mitochondrial translation and single organism process, meaning the genes responsible for the translation of proteins by the mitochondrion were up or down-related in the F1 hybrid in comparison to *S. cerevisiae*. Interestingly, the first metabolite of aspirin (salicylate) is known to be an uncoupler and an inhibitor of mitochondrial electron transportation (Norman *et al*., 2004). It seems to be favourable for F1 hybrid to have mitochondrial translation genes similar to *S. paradoxus*, which might be better equipped to deal with high levels of aspirin and give the F1 hybrid and the *S. paradoxus* an intrinsic fitness advantage in this environment. Why would *S. paradoxus* be more suited to withstand the effects of aspirin? This might be due to its ecology; *S. paradoxus* has been

isolated from oak bark around the world (Liti *et al*., 2009) and bark of white willows contains salicin, a chemical similar to salicylate, so it might be possible *S. paradoxus* populations to have been in contact with similar compounds, and therefor be better adapted to deal with aspirin than the domesticated *S. cerevisiae*.

Our study shed light on the regulation of transcription of an inter-specific F1 hybrid made from crossing two closely related species. We proposed that the transcription regulation of F1 hybrid does not only reflect variation due to parental divergence but is also dependent on the environment the F1 hybrid is in. The environment influenced the F1 hybrid transcription of parental alleles in a way that resembled the fitter parent for a specific environment, and the F1 hybrid had the plasticity to modify its transcription upon environmental change. Finally high levels of heterosis previously described for the F1 hybrid (Chapter I and Bernardes *et al*., 2016) might result from a molecular mechanism for heterosis described by Goff (2011) where parental alleles were differentially transcribed in order to resemble the transcription parent for a specific environment. This phenomenon might be one of main contributors for variations in fitness between divergent parental population and their hybrid offspring.

5. Supplementary Material

Figure S1: Changes in gene transcription according to gene regulation of parents and their F1 hybrid. Changes in gene transcription between *S. cerevisiae* parent (Cer-orange), *S. paradoxus* parent (Par-purple) and allelic specific expression of the F1 hybrid mapped to *S. cerevisiae* (Fcer-light grey) and mapped to *S. paradoxus* (Fpar- dark grey). Gene transcription is *cis*-regulated when parental alleles in the hybrid are differentially transcribed. Gene transcription is *trans*-regulated when the parental alleles in the hybrid are not differentially transcribed. Gene transcription is *cis-trans* (+) regulated when the parental alleles in the hybrid are differentially transcribed and are not differential transcribed between parents or when the parental alleles in the hybrid are not differentially transcribed and are differential transcribed between parents. Gene transcription is *cis-trans* (-) regulated or antagonistic when the parental alleles in the hybrid are differentially transcribed and are differential transcribed between parents on the opposite direction.

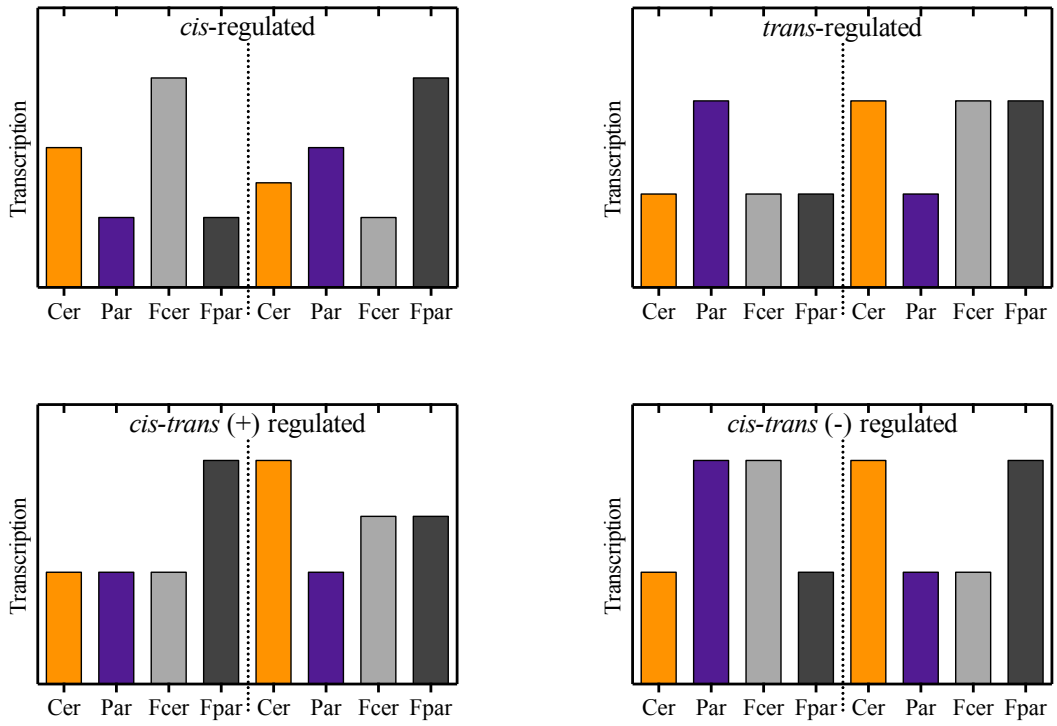


Figure S2: Changes in gene transcription according to the F1 hybrid transcription in comparison with both parents. Changes in gene transcription between *S. cerevisiae* parent (Cer-orange), *S. paradoxus* parent (Par-purple) and the F1 hybrid (Hyb-black). F1 hybrid transcription is conserved when its similar to both parents (Hyb=Par=Cer), F1 hybrid transcription is additive when it is an intermediate between the parental transcriptions (Cer>Hyb>Par or Cer<Hyb<Par), F1 hybrid transcription is *S. paradoxus* overdominant (Spar overdominant or Scer underdominant) when is higher than both Par and Cer or lower than both Cer and Par (Hyb>Par>Cer and Cer>Par>Hyb), F1 hybrid transcription *S. cerevisiae* overdominant (Scer overdomiant or Spar underdominant) when is higher than both Cer and Par or lower than both Par and Cer (Hyb>Cer>Par and Par>Cer>Hyb), F1 hybrid transcription is *S. cerevisiae* dominant (Scer dominant) when is similar Cer but not to Par (Hyb=Cer ≠ Par), F1 hybrid transcription is *S. paradoxus* dominant (Spar dominant) when is similar to Par and but not to Cer (Hyb=Par ≠ Cer).

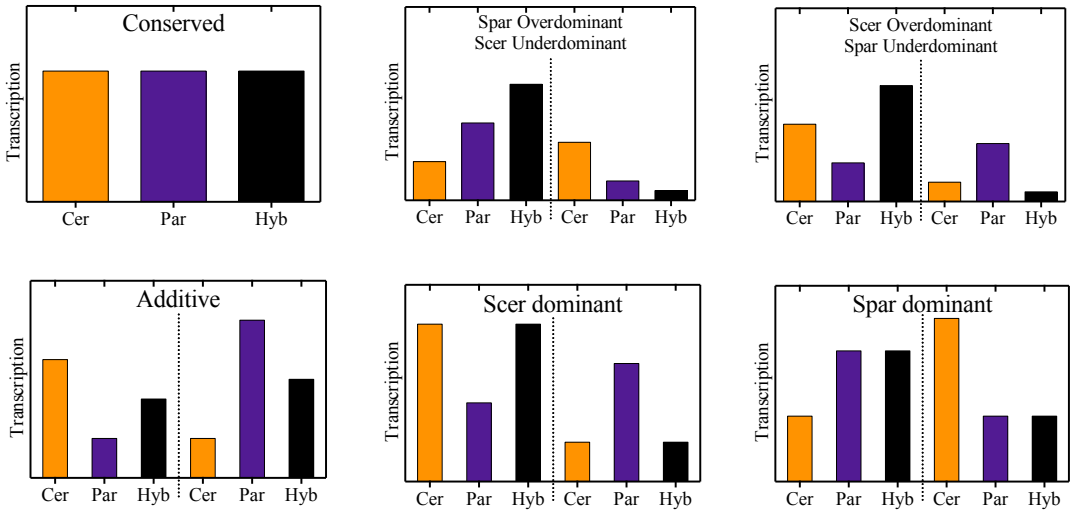
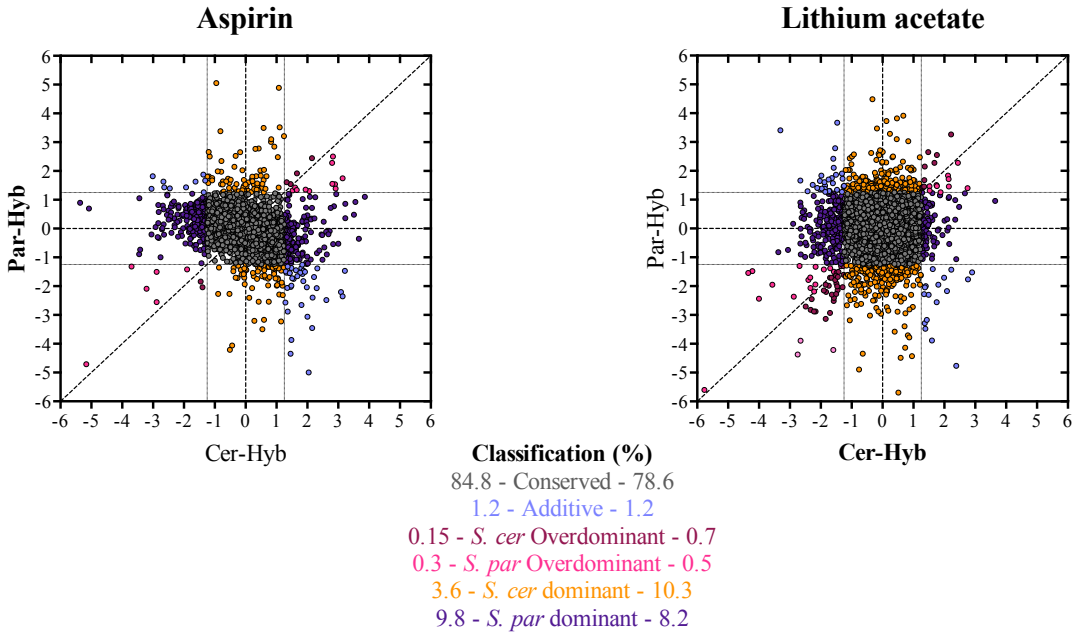


Figure S3: Changes in F1 hybrid transcription (Hyb) in comparison to *S. cerevisiae* (Cer) and *S. paradoxus* (Par) parents in aspirin and lithium acetate environments. Axis refers to log fold change values for Cer-Hyb comparison against log fold change for Par-Hyb comparison. Points refer to F1 hybrid transcription as Conserved (grey), Additive (blue), *S. cerevisiae* overdominant (Scer overdominant or Spar underdominant- crissom), *S. paradoxus* overdominant (Spar overdominant or Scer underdominant- pink), *S. cerevisiae* dominant (Scer dominant- orange), and *S. paradoxus* dominant (Spar dominant- purple). At center percentages for Conserved, Additive, *S. paradoxus* overdominant (Spar overdominant), *S. cerevisiae* overdominant (Scer overdominant), *S. paradoxus* dominant (Spar dominant), and *S. cerevisiae* dominant (Scer dominant). Percentages of inheritance of F1 hybrid gene transcription in the center. Parental species with a fitness advantage in the specific environment in **bold.**



Chapter III

Monosporic cloning of laboratory strains might explain F1 hybrid heterosis

Bernardes, J. P.[1] & Greig, D.[1,2]

1. Max Planck Institute for Evolutionary Biology, Experimental Evolution Group, Plön, Germany

2. The Galton Laboratory, Department of Genetics, Evolution, and Environment, University College London, London, UK

1. Introduction

When two divergent populations have the ability and the opportunity to mate and do so, their first hybrid generation (F1 hybrid) inherits a complete set of alleles from both parents, thus the characteristics of parental genomes and their interaction will influence the overall hybrid fitness. When the F1 hybrid displays enhanced viability or fitness in comparison to one or both parents, we refer to this phenomenon as heterosis or hybrid vigour (Shull, 1948). *Saccharomyces* yeasts have recently become an important model system for investigating the genetic basic of heterosis. The facultative sexual reproduction of *Saccharomyces* yeast offers two important features for the study of heterosis. First, their promiscuous sexual stage allows crosses between highly genetically diverged strains or even different species (Naumov, 1996). Second, haploid vegetative growth and autodiploidization allows the propagation of perfectly homozygous strains and repeated crosses can be made using genotypically identical isolates.

We have previously shown that the F1 hybrid of inter-specific crosses between *S. cerevisiae* and *S. paradoxus* strains generally exhibit more heterosis than do intra-specific crosses within *S. paradoxus* species suggesting that higher genetic divergence between parents might increase the likelihood of heterosis (Bernardes *et al*., 2016). However, while heterosis is rarely detected in F1 hybrids of intra-specific crosses between *S. paradoxus* strains, intra-specific crosses between *S. cerevisiae* strains often result in heterosis in the F1 hybrids (Zörgö *et al*., 2012), a surprising result given the much greater genetic divergence between *S. paradoxus* populations relative to *S. cerevisiae* populations (Liti et al., 2009). Further studies have found that heterosis is much more prevalent in F1 hybrids from crosses involving domesticated *S. cerevisiae* strains than it is in those involving wild primary isolates (Plech *et al*., 2014). Thus, it appears that domesticated *S. cerevisiae* strains are unusual, and crossing them to wild strains of either *S. cerevisiae* or *S. paradoxus* tends to produce offspring with higher fitness than the domesticated parents.

One important factor that can be a caveat in heterosis studies is the way we manipulate primary isolates in the laboratory. When preparing a yeast strain coming from a domesticated of a wild habitat for laboratory work, it is invariably made

homozygous by monosporic cloning. Monosporic cloning or autodiplodization encompasses inducing sporulation of the primary isolate and allow the haploid spores to grow vegetatively, mate-type switch followed by mating within the same haploid colony to produce homozygous diploids. Any recessive deleterious alleles that were heterozygous in the primary isolate will be exposed by this artificial step, suppressing fitness of the derived monosporic clones. When the monosporic clones from different primary isolates are crossed together to form an F1 hybrid their fitness will be restored, giving the impression of heterosis if we compare to the parental monosporic clone, when the fair comparison should be with the primary heterozygous isolate.

One clear difference between domesticated and wild primary isolates is the much higher level of heterozygosity observed in domesticated primary isolates. Magwene *et al*. (2011) surveyed *S. cerevisiae* primary isolates from domesticated and wild habitats and measured their heterozygosity; they identified that domesticated isolates from clinics and vineyards had high heterozygosity while wild isolates from woodlands had much lower heterozygosity. Since wild primary isolates are largely homozygous, any recessive deleterious alleles will be exposed to selection and purged from the population. Because yeast primarily grows as vegetative diploids, sexual cycles are rare specially in the continuous culture conditions of domesticated habitats (Ruderfer *et al*., 2006), where simplified environments like in wineries or breweries do not require the maintenance of all cellular functions as in the wild habitat (Zörgö *et al*., 2012). The maintenance of high levels of heterozygosity in the domesticated isolates suggests that recessive deleterious alleles are not exposed to selection and may accumulate within domesticated populations. Not exposed to selection, that is, until they are brought into the laboratory and passed through a single-spore stage.

The highly heterozygous domesticated primary isolates have not been used in heterosis studies, instead most studies used derived monosporic clones as parental strains for the F1 hybrids. The monosporic clones are used as parents because it is easier to cross them with other genetically divergent monosporic clones in repeated crosses, allowing the F1 hybrids to be similar and comparable.

Domesticated primary isolates of *S. cerevisiae* display high levels of heterozygosity (Magwene *et al*., 2011). Thus, we proposed the observed heterosis of the F1 hybrid of domesticated strains (Zörgö *et al*., 2012; Plech *et al*., 2014) is in fact an artefact of comparing the fitness of high heterozygous F1 hybrid to highly homozygous monosporic clones parents. Wild primary isolates of *S. cerevisiae* and *S. paradoxus* have low levels of heterozygosity (Johnson *et al*., 2004; Magwene *et al*., 2011) and therefore are likely to carry fewer recessive deleterious alleles. If these wild primary isolates are made homozygous into monosporic clones, there will be few recessive deleterious alleles exposed, consequently, if we were to cross two of these monosporic clones, there would be no particular increase of heterozygosity or overall fitness. Thus the F1 hybrid fitness would be similar to the original wild primary isolate and to its monosporic clones parents. Thus, we proposed that the lack of heterosis in F1 hybrids of *S. paradoxus* populations (Chapter I and Bernardes *et al*., 2016) is due to the lack of heterozygosity of the wild populations.

We expect a monosporic clone to immediately fix a proportion of deleterious recessive alleles present in the primary heterozygous isolate. When two monosporic clones from divergent primary isolates are crossed, much of the heterozygosity will be restored and once again recessive mutations will be masked, restoring fitness and giving the illusion of heterosis. We predict that monosporic clones derived from highly heterozygous domesticated isolates will exhibit a decrease in fitness more severe than the monosporic clones derived from less heterozygous wild primary isolates, and consequently, a greater fitness advantage of the F1 hybrid made from a cross of monosporic clones derived from highly heterozygous domesticated isolates. We argue that the F1 hybrid increased fitness will not be higher than the fitness of the original heterozygous domesticated isolate, and higher than the monosporic clone parent, if the primary isolate has multiple masked deleterious alleles (Figure 1).

We used a set of *S. cerevisiae* primary isolates with known heterozygosity, isolated from both domesticated and wild habitats (Magwene *et al*., 2011). We set out to replicate the normal workflow in the laboratory by deriving monosporic clones from the primary isolates and then crossing back the monosporic clones to make a inbred form of the primary isolates (Figure S1). We measured spore viability and calculated

the asexual growth or fitness of the primary isolates, their derived monosporic clones and inbred crosses between different monosporic diploids derived from the same primary isolate, by calculating their maximum growth rate. We expected high heterozygosity domesticated isolates to have higher growth than the monosporic clones, and inbred crosses to have growth intermediate between them. We also expected the differences in gorwths between these three types to be less in low heterozygosity wild isolates than in high heterozygosity domesticated isolates (Figure S1). We found that monosporic clones growth is negative related with their primary isolates heterozygosity. Our results suggested that indeed high heterozygosity primary isolates accumulate recessive deleterious alleles that are exposed in the monosporic clones, and their use in heterosis studies and can lead to inflated levels of heterosis in the F1 hybrids.

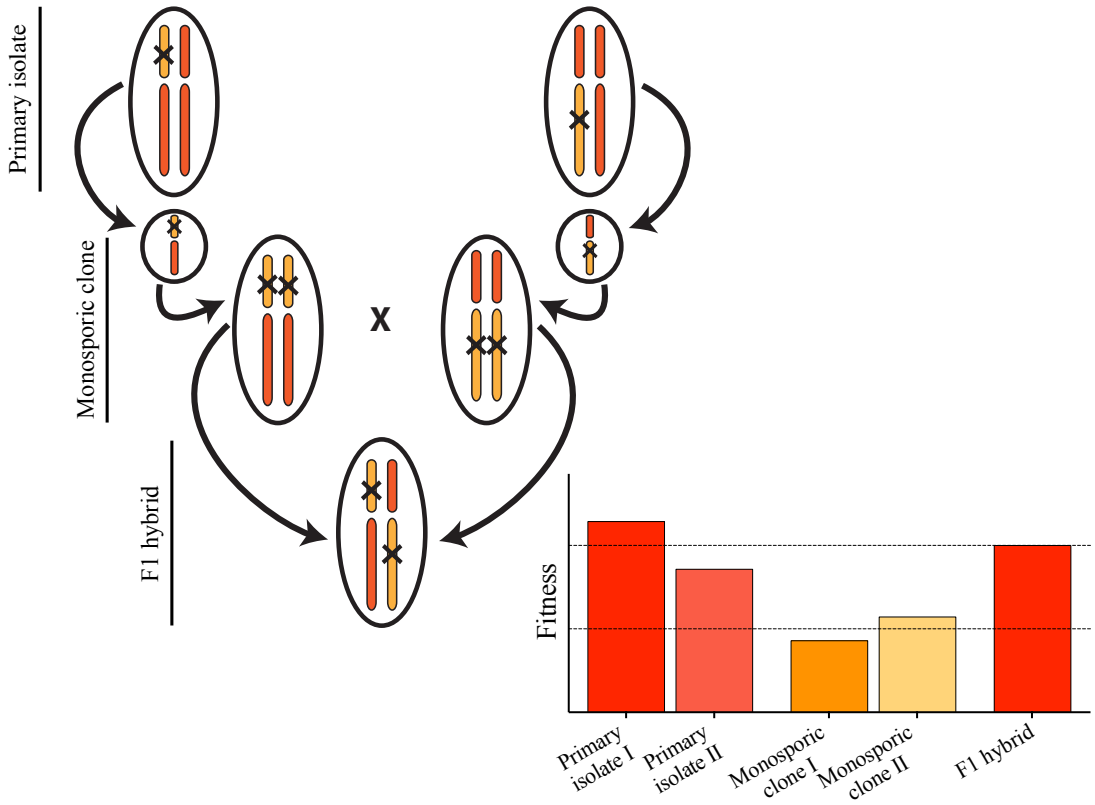


Figure 1: Diagram of the workflow of heterosis studies and the expected asexual growth or fitness of derived *Saccharomyces* yeasts. Primary isolate is sporulated, and from one single spore a monosporic clone is derived, a cross between two monosporic clones from different primary isolates produces a F1 hybrid. Recessive deleterious alleles are represented as crosses. Recessive deleterious alleles are complemented in the primary isolate and the F1 hybrid, but are in the homozygous in the monosporic clone. Primary isolates fitness in red, derived monosporic clones fitness in orange, and F1 hybrid fitness in red. F1 hybrids have

higher fitness or heterosis when compared to the monosporic clones but not higher fitness than the primary isolates their monosporic parents were derived from.

2. Methods and Materials

Primary isolates

We used twelve *S. cerevisiae* strains with known heterozygosity as primary isolates, provided by Magwene *et al*. (2011), the strains were isolated from diverse range of habitats that we classified as either domesticated or wild. We also classified these twelve primary isolates into three heterozygosity groups: low heterozygosity with less than 600 heterozygous sites, intermediate heterozygosity with less than 8 000 and more than 6 000 heterozygous sites, and high heterozygosity with more than 20 000 heterozygous sites (Table 1). The two low heterozygosity strains were isolated one from oak and one was a laboratory strain, the four intermediate heterozygosity strains were isolated from vineyards and muscadine grapes, and one from diseased human tissue in a clinic, and the six high heterozygosity strains were mostly isolated from sick tissue in clinics, and one of them from a rotting fig (precursor of a widely used laboratory strain s288c).

Sporulation

We took every primary isolate and induced sporulation using Elrod *et al*. (2009) method: we grew the strains in YEPD (1% yeast extract, 2% peptone, 2% glucose) for 24h and inoculate 200 cells from the overnight culture into KAC (1%) for another 48h at 30°C. We dissected over 50 tetrads per strain with a dissecting microscope (MSM System 400 by Singer®), and we separated the four spores of each tetrad in a YEPD dissection plates. We incubated the YEPD plates with the dissected tetrads at 30°C for 48h and counted the number of visible colonies formed by each tetrad for each strain, this gave us a number of viable spores or rate of spore viability which we the compared with heterozygosity of the primary isolates.

Strain code	Habitat	Habitat class	Heterozygosity	Heterozygosity class
PMY068	Lab	Laboratory	337	Low
PMY017	Oak	Wild	551	Low
PMY093	Mus grape	Wild/Domesticated	4086	Intermediate
PMY110	Vineyard	Wild/Domesticated	6045	Intermediate
PMY112	Vineyard	Wild/Domesticated	6480	Intermediate
PMY131	Clinical	Domesticated	7248	Intermediate
PMY141	Clinical	Domesticated	22229	High
PMY142	Clinical	Domesticated	22987	High
PMY144	Clinical	Domesticated	23852	High
PMY070	Fig	?	24420	High
PMY127	Clinical	Domesticated	33457	High
PMY132	Clinical	Domesticated	37148	High

Table 1: Characteristics of the twelve primary isolates. Strains code as in Magwene *et al*. (2011), original habitat of isolation and the classification of the habitat (wild or domesticated). Measures of heterozygosity or number of heterozygous sites of the primary isolates, and heterozygosity classification into groups (low, intermediate or high heterozygosity).

Monosporic clones

We used the YEPD dissection plates with visible colonies and replica-plate them to fresh KAC plates and score the ability of the derived colonies to produce spores. If the derived monosporic clones produced spores it meant they had the ability to autodiplodized by germinating into haploid cells that underwent mate-type switching and mated, forming completely diploid monosporic clone, which are highly homozygous. If the derived monosporic clones did not have the ability to sporulate it meant they remained a stable haploid, and were not use for further analysis because of the difficulty to discern the effect of ploidy in the overall fitness (Gerstein *et al*., 2011). For every primary isolate three random tetrads were selected and stored in the -80°C freezer for future analysis. From nine primary isolates we selected three monosporic clones from three different tetrads.

Inbred crosses

We used a subset of the primary isolates; one low heterozygosity strain (PMY017), two intermediate heterozygosity strains (PMY110 and PMY112) and two high heterozygosity strains (PMY142 and PMY127), and marked their three monosporic clones with two antibiotic markers conferring resistance to either hygromycin or G418 (*ura3::HYGMX*/*URA3* or *ura3::KANMX*/*URA3*). Gene transformations were done by following methods in Gietz and Woods (2002) with modifications. We then proceed to cross the transformed monosporic clones derived from the same primary isolate, by crossing the G418 transformed monosporic clones with the hygromycin transformed monosporic clone. We replica-plated the cross YEPD plates to KAC plates, and incubate them for 5 days, followed by a replica-plating to fresh YEPD plates marked with both antibiotics (G418 and hygromycin). We selected for colonies with both antibiotic resistances (*ura3::HYGMX* /*ura3::KANMX*) that were the result of a successfully cross and called them inbred forms. From five primary isolates we selected three inbreds from three possible crosses between monosporic clones (Figure S2).

Growth rate measures

We calculated the growth curve of twelve primary isolates, 27 monosporic clones and 45 inbred crosses using three replicated assays. We prepared a 96-well plate with 200µL of YEPD in each well and add 2µL of the overnight culture (1:100 dilution). We sealed the plate with Breathe-Easy® membrane and measure their growth in a plate reader at 30°C with orbital shaking every 10 minutes. We took measures of OD_{660} every 30 minutes for 24 hours, and use the measurements to calculate growth rates (dx/dt) to create a growth curve using three replicates of every strain of interest. We calculated the maximum growth rate of every single growth curve by selecting for the highest dx/dt, or the lowest doubling time.

$$\frac{dx}{dt} = (log(OD_{t+1}/OD_t) \times 100)$$

3. Results

We tested twelve *S. cerevisiae* primary isolates with known heterozygosity provided by Magwene *et al*. (2011) and we classified each strain into low, intermediate or high heterozygosity. There were two low heterozygosity strains, one isolated from oak and the other a laboratory strain, four intermediate heterozygosity strains isolated from clinical habitats, vineyards and muscadine grape, and six high heterozygosity strains, five isolated from clinical habitats and one isolated from a fig which is a s288c precursor, a widely used and first ever sequenced *S. cerevisiae* strain.

Sporulation

We induced sporulation of the primary isolates using Elrod *et al*. (2009) method (see methods). For every primary isolate we dissected over 50 tetrads by physically separating the four haploid spores on a YEPD plate. We measured germination time and the number of spores that formed visible colonies. We measured spore viability by calculating the ratio of spores that produced visible colonies over the total number of spores dissected (Figure S3 and Table S1). We then compared the spore viability with the heterozygosity of primary isolates and there was a significant negative relationship between heterozygosity and spore viability rate ($F_{1,10}$=79.07, P<0.001, Figure 2). High heterozygosity strains produced one or two viable spores per four spores tetrad, intermediate heterozygosity strains produced three or four viable spores per four spores tetrad, and low heterozygosity strains produced four viable spores per tetrad (Figure S3). Finally, the strain PMY132 with the highest heterozygosity was able to produced only one viable spore from the 50 tetrads dissected, therefor we did not used PMY132 for any further analysis.

Monosporic clones

We induced sporulation of the dissected colonies formed by sporulation of the primary isolates, and scored the ability of the cells in these colonies to themselves form spores (Table S1). The strains that were able to form spores were declared autodiplodized diploids, meaning the haploid spores had the ability to germinate into haploid cells that underwent mate-type switching and mated, forming homozygous monosporic clones. The colonies derived from one low heterozygosity primary isolate

(PMY068) and one high heterozygosity primary isolate (PMY070), remained haploid and thus failed to sporulate while PMY141 displayed a majority of haploid colonies but we were able to recovered diploid colonies (Table S1).

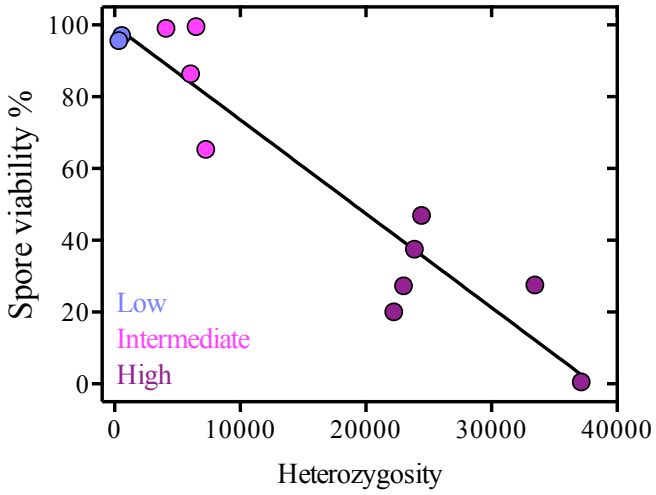


Figure 2: Negative relationship between heterozygosity and spore viability rate. Low heterozygosity primary isolates in blue, intermediate in pink and high in purple. Spore viability rate was a measure of spores that form visible colonies over the total of spores dissected.

For every primary isolate able to produce spores that autodiplodized, we selected three monosporic clones (monosporic clone A, B and C). We calculated the maximum growth rate of three monosporic clones and the primary isolates in three replicates (Table S2). There was no significant difference between the overall maximum growth rate of monosporic clones and the primary isolates (Wilcoxon signed-rank test: W=-25, P=0.164). We also individually calculated if there was a significant difference between the maximum growth rates of the three monosporic clones and their correspondent primary isolate (multiple t-tests: Figure 3 and Table S3). For the primary isolate with low heterozygosity (PMY017) there was no significant disadvantage between the maximum growth rates of the monosporic clones in relation to the primary isolate. For three of the intermediate heterozygosity primary isolates (PMY093, PMY110 and PMY131) there was no significant disadvantage of the monosporic clones in relation to the primary isolates, while one intermediate heterozygosity (PMY112) showed at least two monosporic clones with significant lower maximum growth rates in comparison to the primary isolate. For high heterozygosity primary isolates, only one (PMY142) had no significant disadvantage of the monosporic clones in relation to the primary isolates, while the other three high heterozygosity primary isolates (PMY141, PMY144 and PMY127) displayed a significant lower maximum growth rate for at least one monosporic clone in

comparison to their primary isolates (Figure 3 and Table S3). We also identified a significant negative relationship between heterozygosity of the primary isolate and the maximum growth rate of the monosporic clones relative to their correspondent primary isolates ($F_{1,79}$=10.03, P=0.002, Figure 6a).

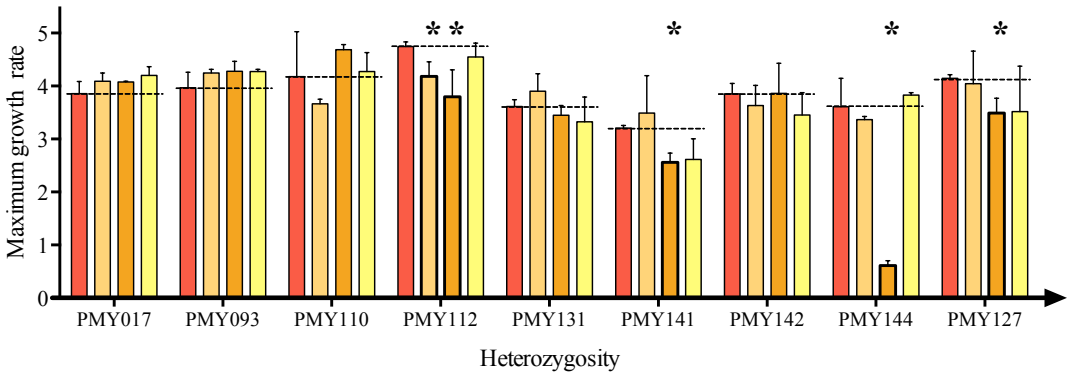


Figure 3: Maximum growth rate of primary isolates and the derived monosporic clones. Primary isolates in red, monosporic clone A in light orange, monosporic clone B in orange and monosporic clone C in yellow. Bars with error indicate the standard deviation, and the dashed lines indicate the average maximum growth rate of the primary isolates. Black asterisk and thicker bars indicate significantly lower maximum growth rate of the monosporic clone in comparison to the correspondent primary isolate.

Inbred crosses

For a subset of primary isolates representing each heterozygosity category, we made every pairwise cross, which we refer to as an "inbred cross", between all three monosporic clones and measured their maximum growth rate (Figure S2 and Table S4). There was no significant difference between the overall maximum growth of the inbreds and the primary isolates (Wilcoxon signed-rank test: W=-3, P=0.812). However we found in several cases that the inbred crosses had higher growth than their primary isolates (multiple t-tests: Figure 4 and Table S5). For the primary isolate with low heterozygosity (PMY017), inbred cross 2 and inbred cross 3 showed a significant advantage in maximum growth rate in relation to the primary isolate. For intermediate heterozygosity primary isolates, PMY110 inbred cross 3 maximum growth rate was significantly higher than its primary isolate, while for PMY112 inbred cross 1 had a significant decrease in maximum growth rates when compared to the primary isolate. For high heterozygosity primary isolates; PMY142 inbred cross 3 showed an advantage in maximum growth rates in comparison to its primary isolate,

while PMY127 displayed inbred cross 1 with a significant lower maximum growth rate in comparison to its primary isolate (Figure 4, Table S5). We also identified a slightly significant negative relationship between heterozygosity of the primary isolate and the maximum growth rate of the inbred crosses relative to their correspondent primary isolates ($F_{1,43}$=4.294, P=0.044, Figure 6b).

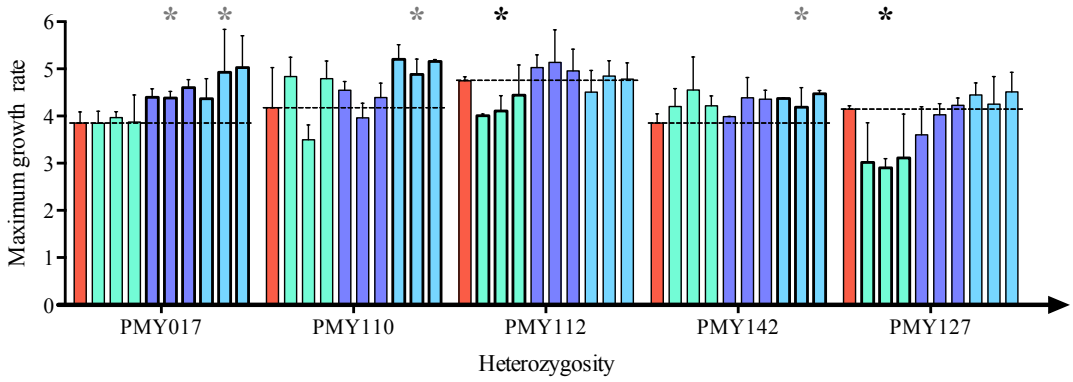


Figure 4: Maximum growth rate of primary isolates and the derived inbred crosses. Primary isolates in red, inbred cross 1 in aqua blue, inbred cross 2 in dark blue and inbred cross 3 in blue, same bar colour indicate biological replicates of the same cross (Figure S2). Bars with error indicate the standard deviation, and the lines indicate the average maximum growth rate of primary isolate. Black asterisk and thicker bars indicate significantly lower maximum growth rate of the inbreds in comparison to their primary isolate. Grey asterisk and thicker bars indicate significantly higher maximum growth rate of the inbreds in comparison to their primary isolate.

We identified no difference between maximum growth rate of the inbreds crosses and the monosporic clones used in the inbred crosses (Wilcoxon signed-rank tests: W=62, P=0.083). In addition we calculated if there was a significant difference between the maximum growth rates of each inbred cross in comparison to the correspondent monosporic clones used as parents (multiple t-tests: Figure 5 and Table S6). Nine inbred crosses of the 15 inbreds produced had a significant higher maximum growth rate than the average of their monosporic parents and only one inbred cross (PMY127) had a significant higher maximum growth rate in comparison with the average monosporic clones (AxB) (Figure 5 and Table S6). For the primary isolate with low heterozygosity (PMY017), inbred 2 and 3 showed a significant advantage in maximum growth rate in relation to monosporic clones (AxC and BxC). For intermediate heterozygosity primary isolates, PMY110 inbred cross 3 had a significant higher maximum growth rate than the average monosporic clone (BxC),

while for PMY112 two inbreds crosses 2 and 3 had a significant higher maximum growth rates in comparison to the average monosporic clonese. For high heterozygosity strains; all PMY142 inbred crosses (1, 2 and 3) showed an advantage of maximum growth rates in comparison to their corresponding monosporic clones, while PMY127 displayed the only inbred cross 1 with a significant decrease in maximum growth rate in comparison to the average monosporic clone (AxC) but still a inbred cross 3 with a significant higher maximum growth rate than the correspondent monosporic clones (Figure 5, Table S6).

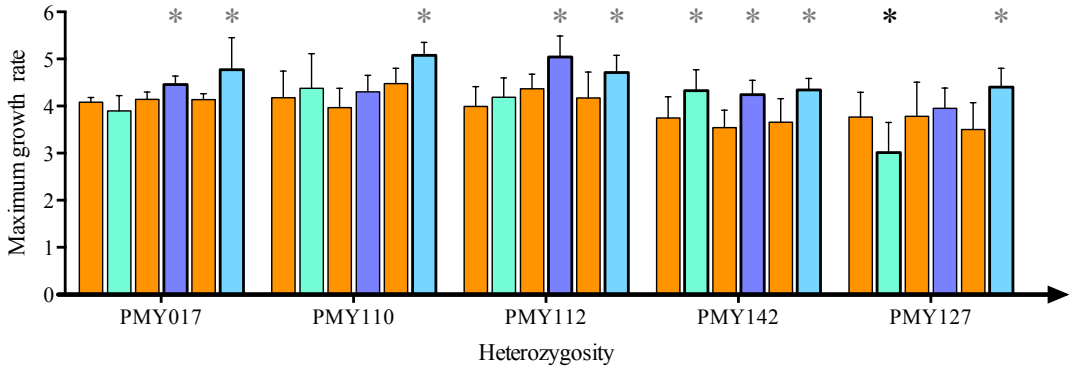


Figure 5: Maximum growth rate of monosporic clones and the correspondent inbred crosses. Monosporic clones in orange and inbred cross 1 in aqua blue, inbred cross 2 in dark blue and inbred cross 3 in blue. Bars with error bars indicate the standard deviations. Black asterisk and thicker bars indicate significantly lower maximum growth of the inbred cross in comparison to the average of monosporic clones used as parents. Grey asterisk and thicker bars indicate significantly higher maximum growth of the inbred cross in comparison to the average of monosporic clones.

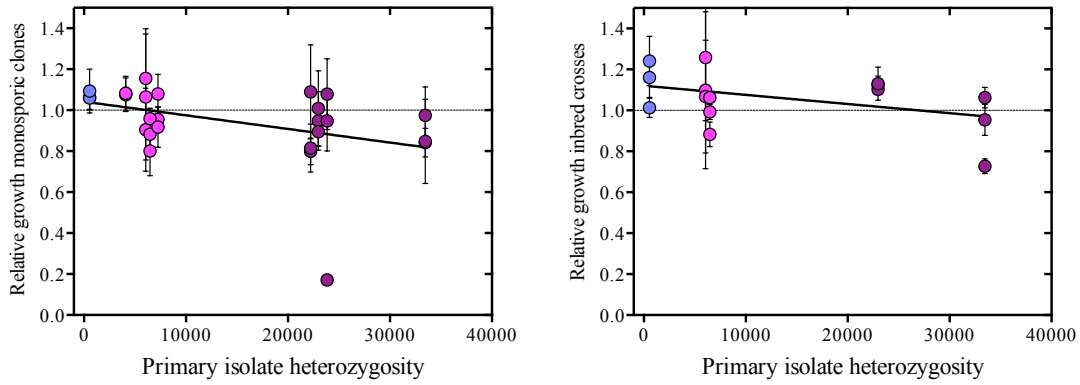


Figure 6: Negative relationships between relative growth of the monosporic clones and heterozygosity of the primary isolates (a), and between inbred crosses and heterozygosity of the primary isolates (b). Relative growth measured as a ratio of maximum growth rate of monosporic clones or inbred crosses over primary isolates. Low heterozygosity strains in blue, intermediate heterozygosity strains in pink and high heterozygosty strains in the purple.

4. Discussion

Here we used primary isolates of *S. cerevisiae* with know heterozygosity (Magwene *et al*., 2011) to replicate the workflow of heterosis studies with *Saccharomyces* yeasts; we induced sporulation of the primary isolates and let the haploid spores autodiplodized into monosporic clones which we then crossed to form inbred crosses. We showed that spore viability of the primary isolates had a strong negative relationship with primary isolate heterozygosity. In addition we identified a decrease in asexual growth or fitness of monosporic clones in comparison to their primary isolates, while inbred crosses displayed a mix of increase and decrease growth when compared to the primary isolates but a clear growth advantage when compared to the monosporic clones used to as their parents. Finally we identified an effect of primary isolates heterozygosity in the subsequent growth of the derived monosporic clones and inbred crosses.

For primary isolates we used six strains isolated from disease tissue in human patients from clinics around California (McCusker *et al*., 1994), one strain isolated from a rotting fig which is a precursor of the widely used s288c strain (Mortimer & Johnston, 1986), two strains from a vineyard in Australia and one from muscadine grape from North America, one strain isolated from a oak tree in North America, and one laboratory strain used as control in Magwene *et al*. (2011). We induced sporulation of the primary isolates; the primary isolate with the highest heterozygosity (PMY132), a clinical strain, did not produce any viable spores. Under the microscope we identified a low tetrad production and the small size spores within the tetrads. Thus the high number of heterozygous sites could be masking possible recessive deleterious alleles that prevent adequate sexual reproduction or sporulation, but have no affect in asexual reproduction or vegetative growth of the highly heterozygous primary isolate. From the primary isolates that were able to produce spores and developed into visible colonies, there were two strains (PMY068 and PMY070) that were not able to produce diploid colonies and remained in a haploid form. These two primary isolates were the low heterozygosity laboratory strain PMY068 (MLY61), where *HO* deletions are a common characteristic, and the high heterozygosity strain isolated from a rotting fig PMY070 (EM93). As stated before, we knew *a priori* that EM93 was the primary contributor of the reference strain s288c (Mortimer & Johnston,

1986), and its *HO* deletion might be the reason why; having a stable haploid form that grows vegetatively makes the strain attractive for a variety of studies. This might have been the reasoning behind the Lindegren's mailing of the haploid form of the EM93 strain to several colleagues around the world in late 1940's (Mortimer & Johnston, 1986), which culminated in s288c being the most well study strain in yeast biology.

From the beginning of our study we identified differences between primary isolates with high or low heterozygosity; the number of viable spores per tetrad was lower for primary isolates with high heterozygosity and the mode was either one or two viable spores per tetrad, while for low heterozygosity primary isolates the mode was four viable spores per four spore tetrad (Figure S3). More importantly the rate of viable spores was significantly negative related with heterozygosity of the primary isolate ($F_{1,10}$=79.07, P<0.001, Figure 2). These results indicated the possibility of an accumulation of recessive deleterious alleles masked in high heterozygosity primary isolates, which, upon sporulation, form tetrads with haploid spores with no visible physical problems (except PMY132) but that lack the ability to properly germinize and form a visible colony. If primary isolates recessive deleterious alleles were masked, upon sporulation or haploidization these alleles would be immediately fixed, which might render the haploid spores unviable. Other explanations seem unlikely; high heterozygosity primary isolates come mostly from clinical samples (McCusker *et al*., 1994) and no relationship has been identified between this habitat and low sporulation viability (Muller & McCusker, 2009). However our results showed similarities with an older study by Mortimer *et al*. (1994); where they identified eight heterozygous sites in 43 *Saccharomyces* strains isolated from grape must in Italy, they also identified a significant negative relationship between the number of heterozygous sites and spore viability. However comparisons between the two studies can be difficult due to the different scale number of heterozygous sites (Mortimer *et al*., 1994; Magwene *et al*., 2011), nonetheless both our studies suggested that low spore viability is related with the accumulation of deleterious alleles in domesticated *Saccharomyces* strains high heterozygosity.

For all significant differences in growth between monosporic clones and their primary isolates, the monosporic clones were always less fit than their primary isolates

(multiple t-tests in Table S3 and Figure 3), even though this effect was not significant (Wilcoxon signed-rank test: W=-25, P=0.148). In addition, monosporic clones derived from low heterozygosity primary isolate showed no decrease in growth, while most monosporic clones derived from high heterozygosity primary isolates displayed at least one monosporic clone with a significant decrease in growth, and there was a the negative relationship between monosporic clone growth and heterozygosity of the primary isolate ($F_{1,79}$=10.03, P=0.002, Figure 6a). Thus supporting the claim that highly heterozygous domesticated strains of *S. cerevisiae* might carry multiple recessive deleterious alleles, which upon autodiplodization are exposed and make the monosporic clones less fit, while low heterozygous wild strains of *S. cerevisiae* do not accumulate this type of alleles. We also identified that the effect of recessive deleterious alleles exposed in the monosporic clones might be meatigated by crossing two monosporic clones: We tested inbred crosses made from crossing two monosporic clones derived from the same primary isolate in order to restore heterozygosity. The inbred crosses had an increase in growth when compared to the parental monosporic clones (multiple t-tests Table S6 and Figure 5), however this trend was not significant (Wilcoxon signed-rank tests: W=62, P=0.083). Surprisingly there was still a slight negative relationship between inbred growth and primary isolate heterozygosity ($F_{1,43}$=4.294, P=0.044, Figure 6b), however inbred crosses derived from high heterozygosity primary isolates had similar growth to their primary isolates (multiple t-tests Table S5 and Figure 4). This indicated that most inbred crosses derived from high heterozygosity primary isolates were fitter than the parental monosporic clones but had similar fitness than the primary isolates they were derived from, while inbred crosses derived from low heterozygosity primary isolates were slighty fitter than the parental monosporic clones which were as fit as the primary isolates they were derived from.

This outcome might explain the disparity in results between heterosis studies where monosporic clones have been systematically used as parental strains of F1 hybrids (Zörgö *et al*., 2012; Plech *et al*., 2014; Bernardes *et al*., 2016). The studies that used domesticated *S. cerevisiae* strains with high heterozygosity identified high levels of heterosis (Zörgö *et al*., 2012; Plech *et al*., 2014), while studies that used wild *S. cerevisiae* strains with low heterozygosity do not identified heterosis (Plech *et al*.,

2014; Bernardes *et al*., 2016). By using domesticated *S. cerevisiae* monosporic clones as parents overall parental fitness is compromised from the beginning because in possible of recessive deleterious alleles will have been exposed by autodiplodization, when crossing two monosporic clones from divergent (F1 hybrids) or even similar (inbred crosses) populations the fitness would be increased by simple complementation of those recessive deleterious alleles, giving the illusion of heterosis (Zörgö *et al*., 2012; Plech *et al*., 2014). However the resulting F1 hybrids or inbred crosses are not necessarily fitter than the primary isolates their monosporic parents were derived from. In contrast, wild *S. cerevisiae* or *S. paradoxus* monosporic clones would be prone to have less deleterious alleles, thus crossing monosporic parents from different populations leads to no significant fitness increase in either or inbred crosses or F1 hybrids (Plech *et al*., 2014; Bernardes *et al*., 2016). It would be interesting to study heterosis of a group of F1 hybrids produce by crossing primary isolates instead of their monosporic clones. This study would have to be extensive because crosses between primary isolates create a variety of F1 hybrids with diverse genotypes and consequently phenotypes when the primary isolates are highly heterozygous. With this study we could compare the primary isolates with the F1 hybrid identify if heterosis occurs in primary isolates of *Saccharomyces* yeasts and it is not only a response to manipulation in the laboratory environment.

High heterozygosity has been observed in domesticated but not wild primary isolates (Magwene *et al*., 2011). The maintenance of high levels of heterozygosity suggests recessive deleterious alleles are not exposed to selection in domesticated habitats and will accumulate, moreover human activity is know to facilitate outcrossing which also promotes high heterozygosity (Magwene, 2014). In domesticated habitats simplified environments do not require the maintenance of all cellular functions and promote specific phenotypes, such as, pseudohyphal growth and robust cell wall for clinical primary isolates (Muller *et al*., 2011). In primary isolates, recessive deleterious alleles have small to no effect in growth because they are masked by functional wild-type alleles (heterozygous locus), autodiplodization exposes the recessive deleterious alleles (homozygous locus), which might cause a decrease in fitness in monosporic clones derived from high heterozygous primary isolates. By crossing two monosporic clones, the previously exposed recessive deleterious alleles become complemented

creating a positive or neutral effect on growth depending on which monosporic clones were crossed. However additional studies would be great to directly and systematically measure the presence of such recessive deleterious alleles both *S. cerevisiae* and *S. paradoxus*, by measuring fitness of haplodiplodized and autodiplodized strains derived from high heterozygosity primary isolates, followed by whole genome sequence we might potentially identify recessive deleterious alleles present in the primary isolate.

Heterosis studies are not only important due to their applications in agriculture crops and cattle breeding, but are also a useful tool to discern characteristics acquired during evolutionary history of the parental populations. Thus it is important to provide a critical analysis in the way we conduct heterosis studies by keeping the evolutionary history of the parental population in mind. Yeast studies due to their simplicity and reproducibility, together with their ability of being a facultative sexual model system offers useful tools to assess the factors contributing heterosis and sheds light on the reasons behind heterosis.

5. Supplementary Material

Figure S1: Aim of our study using three forms of *S. cerevisiae* strains with known heterozygosity and measuring their asexual growth or fitness. Primary isolate with know heterozygosity was sporulated, we selected one spore that formed one visible diploid colony as monosporic clones, and a cross between monosporic clones from the same primary isolate produced an inbred. Primary isolate with high heterozygosity depicted as red-orange genomes, and primary isolate with low heterozygosity depicted as purple genomes. Recessive deleterious alleles are represented as crosses in high heterozygosity strains. For high heterozygosity primary isolates we expected the primary isolate to have a higher fitness than both monosporic clones and similar or higher than the inbred. For low heterozygosity primary isolates we expect similar fitness between the primary isolate, the monosporic clone and F1 inbred.

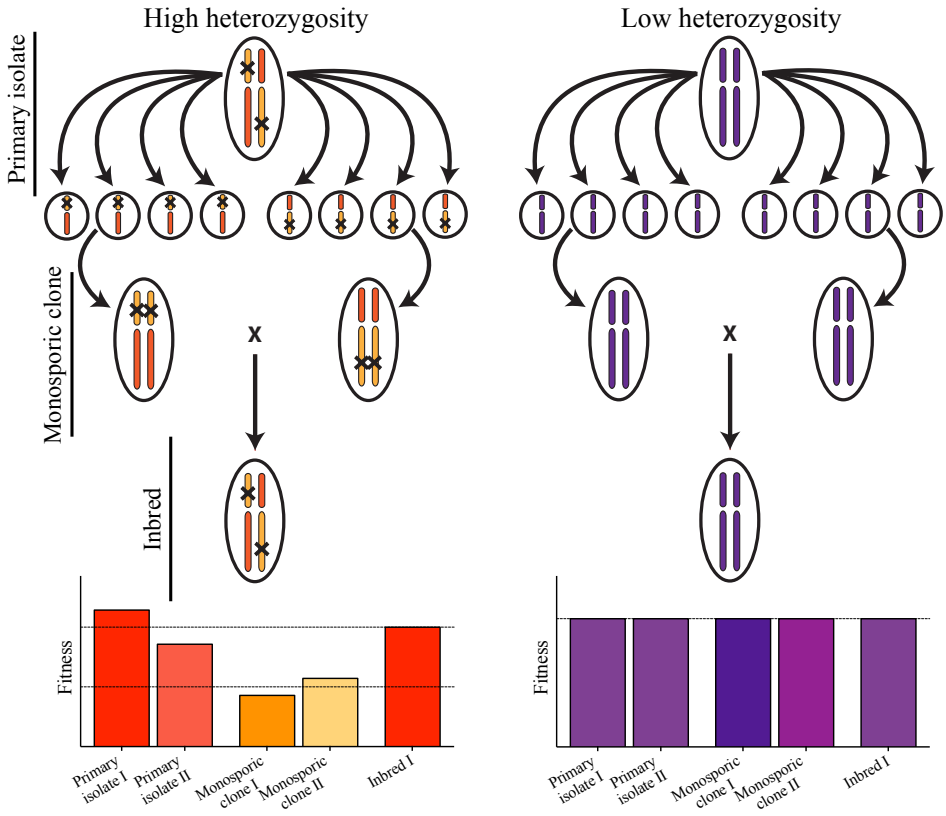


Figure S2: Diagram of the derived strains (monosporic clones and inbreds) from five primary isolates. The primary isolates are depicted on top, for each primary isolate three monosporic clones were isolated (A, B and C), these were transformed with two antibiotic markers (KANMX in red and HYGMX in green), which were crossed between monosporic clones to form the inbreds (depicted in different blue tones), for each primary isolate three inbreds were formed (1, 2 and 3). We tested the maximum growth rate of the primary isolates, their three monosporic clones and their three inbreds independently in three replicates.

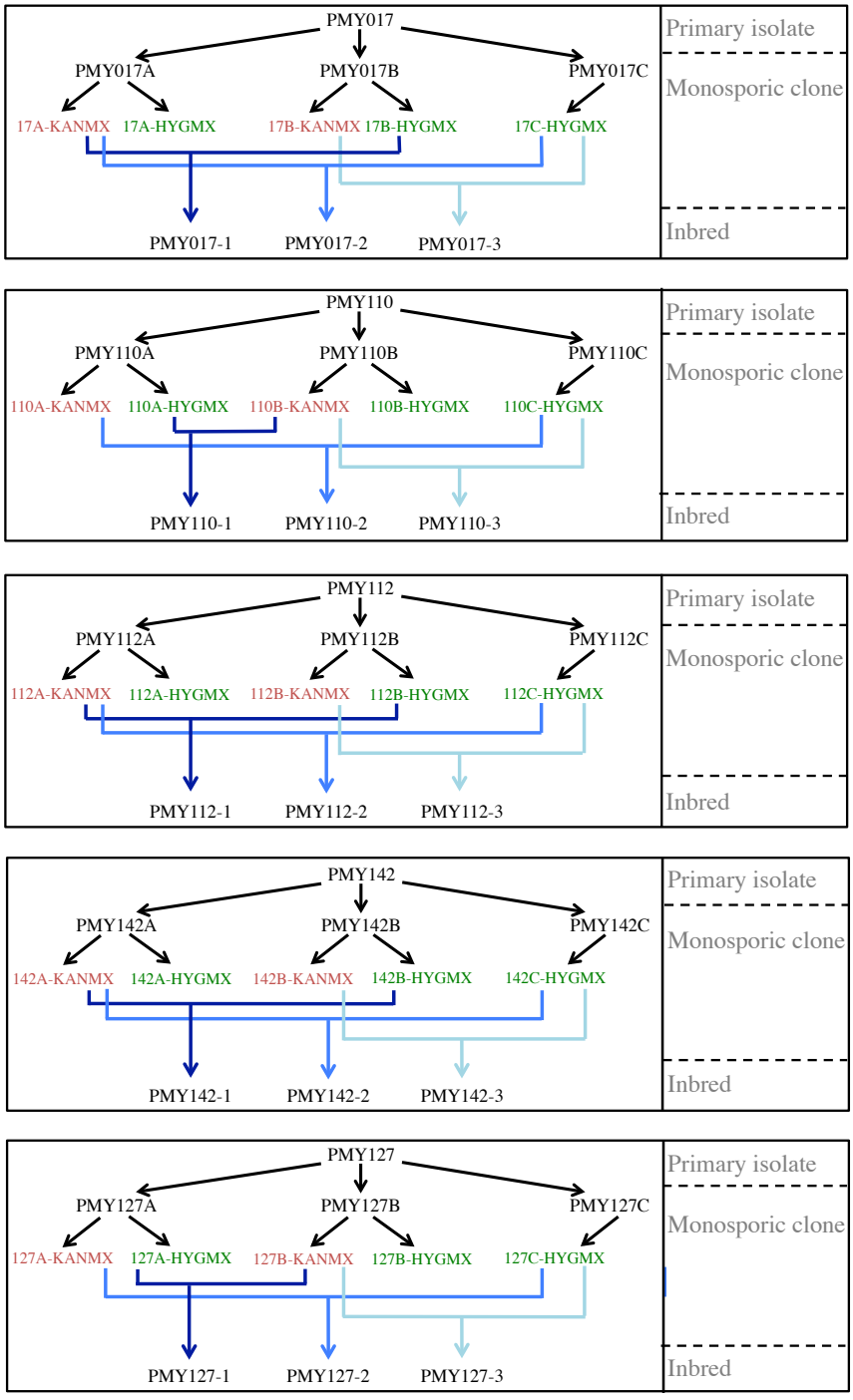


Figure S3: Percentage of viable spores for each tetrad produced by the primary isolate. Viable spores by score by the ability to form visible colonies. Percentage of 4 viable spores tetrads (black), 3 viable spores tetrads (dark grey), 2 viable spore tetrads (grey), 1 viable spore tetrads (light grey) and 0 viable spores (white). Two low heterozygosity strains surrounded by blue, three intermediate heterozygosity strains surrounded by pink and six high heterozygosty strains surrounded by purple.

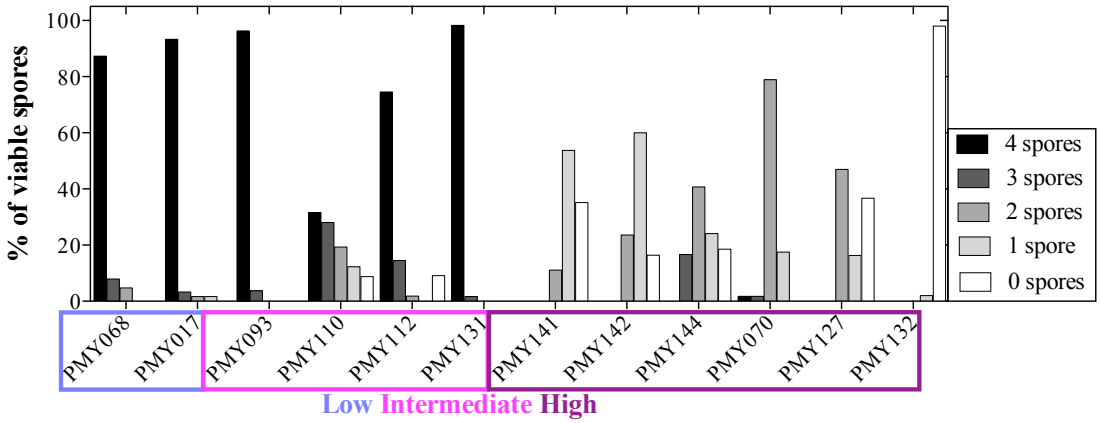


Table S1: Primary isolates sporulation details. Strain code for each primary isolate, number of days needed for the tetrads to form, number of tetrads dissected, total number of spores dissected and quantity of viable spores, i.e. spores that formed visible colonies. Spore viability (%) as a ratio of viable spores over the total number of spores dissected. Heterozygosity measured by calculating the heterozygous sites for a array of SNPs in Magwene *et al*. (2011). Monosporic clones formed were classified as either haploid or diploid depending on the ability to sporulate.

Strain code	Sporulation time (days)	Tetrads dissected	Total spores	Viable spores	Spore viability	Heterozygosity	Heterozygosity classification	Monosporic clones
PMY068	4	63	252	241	95.635	337	Low	haploid
PMY017	2	60	240	233	97.083	551	Low	diploid
PMY093	2	54	216	214	99.074	4086	Intermediate	diploid
PMY110	5	55	220	190	86.364	6045	Intermediate	diploid
PMY112	5	59	236	235	99.576	6480	Intermediate	diploid
PMY131	2	57	228	149	65.351	7248	Intermediate	diploid
PMY141	5	54	216	41	20.098	22229	High	diploid
PMY142	2	55	220	59	27.315	22987	High	diploid
PMY144	4	54	216	84	37.5	23852	High	diploid
PMY070	5	57	228	107	46.930	24420	High	haploid
PMY127	5	49	196	54	27.551	33457	High	diploid
PMY132	5	50	200	1	0.5	37148	High	-

Table S2: Primary isolates and their derived monosporic clones maximum growth rates. Maximum growth rate of the primary isolate and monosporic clones in three replicates (1, 2 and 3). Monosporic clones were selected by the ability of derived spores from the primary isolates to autodiplodized (become diploid). Monosporic clones A, B and C refer to autodiplodized colonies randomly selected from different tetrads of the same primary isolate.

Strain code	Primary isolate			Monosporic clone A			Monosporic clone B			Monosporic clone C		
	1	2	3	1	2	3	1	2	3	1	2	3
PMY017	3.814	4.103	3.647	3.920	4.227	4.121	4.080	4.087	4.060	4.177	4.046	4.374
PMY093	3.828	4.301	3.765	4.184	4.238	4.319	4.170	4.498	4.167	4.321	4.241	4.260
PMY110	4.501	4.815	3.212	3.760	3.592	3.641	4.658	4.792	4.607	4.451	3.862	4.506
PMY112	4.735	4.838	4.667	3.964	4.099	4.491	4.012	3.212	4.163	4.814	4.539	4.292
PMY131	3.493	3.742	3.613	4.059	4.127	3.521	3.550	3.562	3.238	2.997	3.862	3.122
PMY141	3.255	3.156	3.212	3.572	4.153	2.750	2.365	2.687	2.630	2.572	2.250	3.025
PMY142	3.627	4.005	3.921	3.943	3.212	3.742	4.188	3.212	4.191	3.069	3.902	3.387
PMY144	3.012	4.044	3.772	3.343	3.324	3.432	0.593	0.710	0.529	3.829	3.784	3.874
PMY127	4.105	4.227	4.106	3.343	4.498	4.292	3.772	3.484	3.212	2.572	3.742	4.238

Table S3: Multiple unpaired t-tests between maximum growth rates of the primary isolate and the derived monosporic clones (A, B and C). P-value corrected for multiple comparisons, mean of the primary isolates, mean of the monosporic clones, difference between primary isolates and their monosporic clones (Difference), and standard error of the difference of means (SE of difference). T-ratio for every comparison with 4 degrees of freedom for t-tests. Highlight in green for significant lower growth of the monosporic clones in comparison to the primary isolates.

Monosporic clone A Strain code	P-value	Primary isolate mean	Monosporic clone mean	Difference	SE of difference	t ratio
PMY017	0.218	3.855	4.089	-0.235	0.161	1.460
PMY093	0.180	3.965	4.247	-0.282	0.174	1.623
PMY110	0.358	4.176	3.664	0.512	0.493	1.038
PMY112	0.028	4.747	4.184	0.562	0.166	3.391
PMY131	0.234	3.616	3.903	-0.286	0.205	1.399
PMY141	0.524	3.208	3.492	-0.284	0.408	0.697
PMY142	0.425	3.851	3.632	0.219	0.246	0.888
PMY144	0.478	3.609	3.367	0.243	0.310	0.782
PMY127	0.791	4.146	4.045	0.101	0.358	0.283

Monosporic clone B Strain code	P-value	Primary isolate mean	Monosporic clone mean	Difference	SE of difference	t ratio
PMY017	0.173	3.855	4.076	-0.221	0.133	1.656
PMY093	0.195	3.965	4.278	-0.314	0.202	1.555
PMY110	0.360	4.176	4.686	-0.510	0.494	1.032
PMY112	0.034	4.747	3.796	0.951	0.300	3.175
PMY131	0.265	3.616	3.450	0.166	0.128	1.296
PMY141	0.003	3.208	2.561	0.647	0.103	6.257
PMY142	0.973	3.851	3.864	-0.012	0.346	0.036
PMY144	0.001	3.609	0.611	2.998	0.313	9.577
PMY127	0.017	4.146	3.489	0.657	0.167	3.939

Monosporic clone C Strain code	P-value	Primary isolate mean	Monosporic clone mean	Difference	SE of difference	t ratio
PMY017	0.103	3.855	4.199	-0.344	0.164	2.103
PMY093	0.145	3.965	4.274	-0.309	0.171	1.810
PMY110	0.864	4.176	4.273	-0.097	0.532	0.182
PMY112	0.279	4.747	4.548	0.198	0.159	1.252
PMY131	0.359	3.616	3.327	0.289	0.279	1.035
PMY141	0.059	3.208	2.615	0.592	0.227	2.613
PMY142	0.212	3.851	3.453	0.399	0.268	1.485
PMY144	0.517	3.609	3.829	-0.220	0.310	0.710
PMY127	0.273	4.146	3.517	0.628	0.495	1.268

Table S4: Inbreds maximum growth rates. Maximum growth rate of the inbreds in three replicates (1, 2 and 3). Inbreds were made by crossing two different monosporic clones derived from the same primary isolates and marked with antibiotic markers. Inbreds 1, 2 and 3 refer to three different crosses (indicated in the first column of the table and in Figure S2), and A, B and C refer the three independent crosses selected from the same monosporic clone parental cross.

Inbred 1 (cross)	**Inbred 1A**			**Inbred 1B**			**Inbred 1C**		
	1	2	3	1	2	3	1	2	3
17A x 17B	3.863	3.584	4.095	4.067	3.829	4.007	4.052	3.222	4.337
110A x 110B	4.591	4.616	5.312	3.236	3.847	3.415	4.377	5.104	4.902
112A x 112B	3.976	3.996	4.049	4.466	3.832	4.018	5.004	3.746	4.581
142A x 142B	3.980	4.638	3.997	4.275	4.035	5.349	3.998	4.240	4.416
127A x 127B	2.400	2.681	3.975	3.058	2.687	2.967	3.687	2.038	3.609

Inbred 2 (cross)	**Inbred 2A**			**Inbred 2B**			**Inbred 2C**		
	1	2	3	1	2	3	1	2	3
17A x 17C	4.186	4.513	4.487	4.399	4.235	4.511	4.515	4.490	4.799
110A x 110C	4.632	4.336	4.676	3.642	3.995	4.253	4.308	4.138	4.728
112A x 112C	4.724	5.136	5.228	5.491	4.343	5.582	4.491	5.413	4.967
142A x 142C	3.990	3.990	3.989	4.013	4.288	4.859	4.185	4.562	4.324
127A x 127C	3.001	3.623	4.185	3.959	3.839	4.287	4.406	4.150	4.128

Inbred 3 (cross)	**Inbred 3A**			**Inbred 3B**			**Inbred 3C**		
	1	2	3	1	2	3	1	2	3
17B x 17C	4.332	3.964	4.810	5.966	4.532	4.287	5.240	5.570	4.266
110B x 110C	5.437	5.320	4.843	4.590	4.830	5.230	5.202	5.143	5.123
112B x 112C	4.784	4.765	3.974	4.492	4.928	5.125	4.933	4.388	5.027
142B x 142C	4.385	4.374	4.358	4.146	3.791	4.619	4.547	4.414	4.461
127B x 127C	4.227	4.727	4.384	3.639	4.309	4.809	4.906	4.085	4.551

Table S5: Multiple unpaired t-tests between maximum growth rates of the primary isolate and the derived inbreds crosses (inbreds as 1, 2 and 3). P-value corrected for multiple comparisons, mean of the primary isolates, mean of the inbred crosses, difference between primary isolates and inbreds (Difference), and standard error of the difference of means (SE of difference). T-ratio for every comparison with 4 degrees of freedom for t-tests. Highlight in green for lower growth of the inbreds in comparison to the primary isolates, and highlighted in red for higher growth of the inbreds in comparison to the primary isolates.

Strain	P-value	Primary isolate	Inbred 1	Difference	SE of difference	t ratio
PMY017	0.849	3.855	3.895	-0.041	0.207	0.196
PMY110	0.698	4.176	4.378	-0.202	0.505	0.400
PMY112	0.046	4.747	4.185	0.562	0.246	2.280
PMY142	0.112	3.851	4.325	-0.474	0.272	1.745
PMY127	0.014	4.146	3.011	1.135	0.382	2.967

Strain	P-value	Primary isolate	Inbred 2	Difference	SE of difference	t ratio
PMY017	0.001	3.855	4.459	-0.605	0.127	4.776
PMY110	0.712	4.176	4.301	-0.125	0.329	0.380
PMY112	0.293	4.747	5.042	-0.295	0.266	1.109
PMY142	0.066	3.851	4.244	-0.393	0.191	2.063
PMY127	0.469	4.146	3.953	0.193	0.256	0.752

Strain	P-value	Primary isolate	Inbred 3	Difference	SE of difference	t ratio
PMY017	0.049	3.855	4.774	-0.920	0.410	2.242
PMY110	0.013	4.176	5.080	-0.904	0.300	3.008
PMY112	0.880	4.747	4.713	0.034	0.220	0.155
PMY142	0.011	3.851	4.344	-0.493	0.158	3.117
PMY127	0.304	4.146	4.404	-0.258	0.238	1.083

Table S6: Multiple unpaired t-tests between maximum growth rates of the inbred cross and the average maximum growth rates of the monosporic clones used to make the inbred crosses. P-value corrected for multiple comparisons, mean of the monosporic clones used to make the correspondent inbred, mean of the inbreds (inbred 1 2 and 3), difference between monosporic clones and inbreds (Difference), and standard error of the difference of means (SE of difference). T-ratio for every comparison with 4 degrees of freedom for t-tests. Highlight in green for lower growth of the inbreds in comparison to the monosporic clones, and highlighted in red for higher growth of the inbreds in comparison to the monosporic clones.

Strain	P-value	Monosporic (AxB)	Inbred 1	Difference	SE of difference	t ratio
PMY017	0.201	4.083	3.895	0.187	0.139	1.346
PMY110	0.577	4.175	4.378	-0.203	0.355	0.572
PMY112	0.389	3.990	4.185	-0.195	0.219	0.890
PMY142	0.029	3.748	4.325	-0.577	0.235	2.457
PMY127	0.032	3.767	3.011	0.756	0.316	2.395

Strain	P-value	Monosporic (AxC)	Inbred 2	Difference	SE of difference	t ratio
PMY017	0.004	4.144	4.459	-0.315	0.090	3.519
PMY110	0.116	3.969	4.301	-0.332	0.197	1.686
PMY112	0.007	4.367	5.042	-0.675	0.210	3.220
PMY142	0.001	3.543	4.244	-0.702	0.175	4.022
PMY127	0.571	3.781	3.953	-0.172	0.296	0.582

Strain	P-value	Monosporic (BxC)	Inbred 3	Difference	SE of difference	t ratio
PMY017	0.043	4.137	4.774	-0.637	0.283	2.248
PMY110	0.002	4.479	5.080	-0.600	0.154	3.889
PMY112	0.038	4.172	4.713	-0.541	0.235	2.304
PMY142	0.003	3.658	4.344	-0.686	0.192	3.570
PMY127	0.003	3.503	4.404	-0.901	0.249	3.624

Conclusion & Perspectives

The aim of my thesis was to identify the mechanisms underlying heterosis. I used *Saccharomyces* yeast as a model system because of its practicality in evolution studies, where fast and repeatable growth measures are useful to directly estimate asexual growth or fitness, and small genome size with high quality reference genomes allow for an thorough analysis of transcription. *Saccharomyces* yeasts have the ability to form hybrids between highly genetically diverged populations or species. I was interested in identifying heterosis in F1 hybrids from crosses that had not been tested before, in discerning genetic and molecular mechanisms behind heterosis, and in detecting heterosis predictors than can be applied on a broader scale.

1. Heterosis and ecological background of F1 hybrids

My thesis supported the idea that the ecological background of the parental strains had a significant effect on the strenght of heterosis. A previous study by Plech *et al*. (2014) identified heterosis in crosses between domesticated but not wild *S. cerevisiae* parents. In the first Chapter, I identified both mid- and best-parent heterosis for crosses between wild *S. paradoxus* and domesticated *S. cerevisiae* parents, but no heterosis for crosses between divergent wild *S. paradoxus* parents (Chapter I and Bernardes *et al*., 2016).

Heterosis studies in yeast, as described in Chapter I and Bernardes *et al*. (2016), or by Zörgö *et al*. (2012) and Plech *et al*. (2014) use monosporic clones as parental strains, as is conventional in yeast genetics studies. Monosporic clones are derived from single haploid spores by mating-type switching and autodiplodization, and are thus perfectly homozygous, except at the mating type locus. Domesticated yeast populations tend to accumulate recessive deleterious alleles (Magwene, 2014) evidenced by the high heterozygosity levels displayed in clinical and other domesticated *S. cerevisiae* primary isolates (Magwene *et al*., 2011). Domesticated habitats have optimized growth conditions which often select only for specific phenotypes, such as high ethanol tolerance in wineries and breweries (Casey & Ingledew, 1986), or pseudo hyphae growth in medical clinics (Muller *et al*., 2011). Thus strong directional selection in domesticated habitats relaxes selection for other cellular functions and allows for accumulation of recessive deleterious alleles in these

yeast populations. Wild yeast populations are not expected to accumulate as many recessive deleterious alleles evidenced by the low heterozygosity levels displayed by *S. cerevisiae* primary isolates from oak (Magwene *et al*., 2011). Wild habitats are characterized by strong selective pressures, such as season fluctuations on temperature and food resources, which can maintain balancing selection on multiple traits (Goncalves *et al*., 2011; Kowallik & Greig, 2016). Also, such stressful conditions might lead to starvation, which is known to induce sexual reproduction in yeast; the formation of haploid yeasts exposes recessive deleterious alleles to selection, purging yeast with such alleles out of the population. Monosporic clones are formed by autodiploidization of a haploid spore, exposing previously masked deleterious alleles as homozygous. If domesticated primary isolates tend to accumulate more recessive deleterious alleles than wild primary isolates, then more deleterious alleles become exposed in monosporic clones with a domesticated background. Thus monosporic clones with domesticated background have a fitness decrease in comparison to their primary isolate, while monosporic clones with a wild background have a similar fitness to their primary isolate because less deleterious alleles become homozygous (Chapter III). When crossing two monosporic clones from diverged yeast populations recessive deleterious alleles might be complemented, giving the F1 hybrid an advantage in relation to its homozygous monosporic parents. Thus by crossing a domesticated monosporic parent with another diverged domesticated or wild monosporic parent we are increasing the chances of complementing deleterious alleles and having a fitter individual than its parents or heterosis. While by crossing two divergent wild monosporic parents, there will be a smaller number of deleterious alleles to complement, so small the resulting individual shows no advantage or no heterosis. Consequently, heterosis is highly dependent on the ecology of *Saccharomyces* yeast natural populations (domesticated or wild) because of its impact on parental strains fitness. Moreover, by using only monosporic clones or perfectly homozygous strains as parents we might be failing to capture the diversity of hybridisation outcomes.

My thesis work does not clarify if the F1 hybrids are fitter than the domesticated or wild primary isolates the monosporic parents were derived from. To answer this question I would need to gather more direct data; I suggest repeating the F1 hybrid crosses using *Saccharomyces* yeast primary isolates collected directly from

domesticated and wild habitats as parental strains. By re-testing the F1 hybrid asexual fitness and competitive growth against their heterozygous parents (primary isolates), I would expect greater fitness variability than the one detected in Chapter I, due to the diversity in parental genotypes and phenotypes. Heterosis would depend on the amount of recessive deleterious alleles complemented in the F1 hybrid cross, and if these number is lower or higher than the primary isolates. Such study would determine whether heterosis depends on advantages of the F1 hybrid or on disadvantages of the monosporic parents, and would clarify the importance of certain genetic mechanisms behind heterosis of *Saccharomyces* yeasts (see *2. Heterosis genetic mechanisms*).

2. Genetic mechanisms for heterosis

Throughout my thesis I intended to discern genetic mechanisms that contribute to heterosis in *Saccharomyces* yeasts. To simplify the analysis I focused solely on dominance and overdominance mechanisms for heterosis, due to the difficulties of identifying epistasis events in even simple model organisms such as *Saccharomyces* yeast. However, a recently published gene-interaction map of ~6000 *S. cerevisiae* genes might be a useful tool to disentangle how epistatic interactions might be affecting F1 hybrids fitness (Costanzo *et al.*, 2016). Heterosis studies in *Saccharomyces* yeasts mainly support dominance genetic mechanism for heterosis (Zörgö *et al.*, 2012; Plech *et al.*, 2014; Shapira *et al.*, 2014) but cannot completely discard overdominance mechanism for heterosis (Shapira *et al.*, 2014). In Chapter I, F1 hybrids with a domesticated background displayed significant competitive advantages, even though, wild *S. paradoxus* parents showed lower competitive growth than domesticated *S. cerevisiae* parents for almost every environment tested. Moreover, best-parent heterosis was prevalent in crosses between different species, however crosses between highly divergent wild yeast populations displayed no heterosis (Chapter I and Bernardes *et al.*, 2016).

Heterosis can be attributed to reciprocal complementation of the recessive deleterious alleles in the F1 hybrid genome, or dominance mechanism for heterosis, where one less advantageous or deleterious parental allele is complemented by a more advantageous or wild-type parental allele in a dominant interaction at multiple loci (Shull, 1948; Zörgö *et al.*, 2012). By crossing individuals from divergent populations

with accumulation of recessive deleterious alleles, these alleles are more likely to be complemented by functional wild-type alleles, giving the F1 hybrid a fitness advantage over its parents. While crossing individuals from divergent populations with less recessive deleterious alleles, will lead to a smaller advantage in F1 hybrid in comparison to its parents. As such, we expect F1 hybrids with a domesticated background to have higher heterosis because, as discussed before, they are more likely to complement recessive deleterious alleles than a F1 hybrid with an exclusively wild background (as discussed in *1. Heterosis and ecological background of F1 hybrids*). However dominance mechanism cannot explain why wild *S. paradoxus* parents, with less recessive deleterious alleles, had a lower competitive growth than the domesticated *S. cerevisiae* parent for the majority of environments. One explanation for this occurrence refers to the characteristics of the environments chosen in Chapter I; where all environments were highly artificial and the competitions were done under laboratory conditions. Thus even if *S. paradoxus* parents had a competitive advantage in their wild natural habitat, this advantage would be hard to capture under artificial laboratory conditions, in contrast, *S. cerevisiae* parents which have been used in the laboratory for decades (Mortimer & Johnston, 1986) should easily outcompete less adapted wild yeast.

Prevalent best-parent heterosis in Chapter I might be better explain by overdominance mechanism for heterosis, which attributes the F1 hybrid advantage to numerous heterozygous loci (Shull, 1948; Shapira *et al*., 2014). In this case, having one *S. cerevisiae* and one *S. paradoxus* parental allele for a certain locus produces a positive interaction that contributes to a superior fitness of the F1 hybrid in comparison to their correspondent homozygous *S. cerevisiae* or *S. paradoxus* parental locus. Thus, significant best-parent heterosis for almost every environment might be due to the positive interaction of highly divergent alleles, instead of complementation of deleterious alleles, because one would expect at least for one environment the domesticated parent to be better adapted than the F1 hybrid cross. According to overdominance mechanism, crosses between highly divergent wild *S. paradoxus* strains (1-5% genetic divergence) should have higher heterosis than crosses between less divergent domesticated *S. cerevisiae* strains (>1% genetic divergence). However, our results showed the opposite; the genetic divergence of wild parental strains crosses had no relationship with the F1 hybrid fitness, and F1 hybrids with a wild

background showed on average no heterosis for competitive growth, as expected by the overdominance mechanism for heterosis.

Thus the genetic mechanism(s) behind heterosis remain a debatable subject. Using the study suggested previously (*1. Heterosis and ecological background of F1 hybrids)*, where primary isolates are used as parental strains of F1 hybrids instead of monosporic clones might be useful to directly measure overdominance. By crossing two divergent domesticated primary isolates, we control for 'inbreeding depression' of monosporic clones normally used as parents. If these F1 hybrid crosses display overall heterosis, its likely caused by overdominant interactions, because deleterious alleles would be complemented in both F1 hybrids and domesticated primary isolates used as parents, thus heterosis would not have been due to complementation of deleterious alleles. Genetic mechanisms behind heterosis remain a controversial topic even after a century of heterosis studies in different model systems. Advances in molecular techniques might clarify such mechanisms, and lead to a new mechanism for heterosis.

3. Molecular mechanisms for heterosis

The results of my thesis supported a molecular mechanism for heterosis where the F1 hybrid has the ability to regulate its transcription according to the surrounding environment. Previous studies that used *S. cerevisiae* hybrids mainly focused on regulation of transcription, and they identified widespread *cis*-regulated transcription however they did not provide any complementary fitness studies (Tirosh *et al*., 2009; McManus *et al*., 2014). Thus, as far as I know, heterosis has never been tested at the transcriptome level; Chapter II presents a thorough analysis of transcriptome of the F1 hybrid and its parents in different environments. The environments were chosen *a priori* so each environment favoured one of the parental strains, while the F1 hybrid had a competitive advantage under both environments (Chapter I and Bernardes *et al*., 2016). The F1 hybrid transcription profiles resembled the fitter parent for a specific environment, and the majority of differentially transcribed genes were *cis*-regulated (Chapter II).

Transcriptome analysis showed that the F1 hybrid transcription resembled *S. paradoxus* parent for the environment where *S. paradoxus* had an advantage, and F1 hybrid transcription resembled *S. cerevisiae* parent for the environment where *S.*

cerevisiae had an advantage. These results hint for a molecular mechanism behind heterosis where the F1 hybrid has the ability to change its transcription profile depending on the environment. The ability to modify *Saccharomyces* yeast transcription profile was first described in Gasch *et al*. (2000), where they identified several genes with differential transcription upon environmental change. If there is variation in transcription of *S. cerevisiae* upon environmental change, the variation should be greater in a F1 hybrid cross between two divergent species, due to their high heterozygosity. The environment surrounding the F1 hybrid might determine advantageous parental alleles at multiple loci, and these might be preferentially transcribe accordingly. The ability to distinguish advantageous parental alleles was first proposed by Goff (2011) molecular model for multigenic heterosis, where the ability of the hybrid cell to detect the advantageous parental alleles depends on the stability of encoded proteins. Thus less advantageous parental alleles produce less stable proteins, which signal or feedback to preferential transcribe the advantageous parental alleles. However stability of encoded proteins is a difficult characteristic to measure because of protein's transient state, and, although Goff's (2011) model seems convincing, it relies on a feedback process that remains elusive to modern techniques. Another simpler model might also explain these results; F1 hybrid displays transcription profile diversity, i.e. different transcription profiles with one or the other parental allele differentially transcribed at multiple loci. If we grow the F1 hybrid under a specific environment where one transcription profile is more advantageous, this transcription profile would have a higher growth and it would spread in the F1 hybrid cell population. When we sequence the transcriptome of *Saccharomyces* yeasts growing in a specific environment, we sequence not one cell but a part of the cell population, thus it is impossible to accurately determine the diversity of transcription profiles.

Both molecular models assume that heterosis is due to the ability of the F1 hybrid to resemble the fitter parent in a specific environment, using preferential transcription of advantageous alleles. The difference between the models is how the F1 hybrid discerns the advantageous transcription profile, whether by protein stability or by an advantage in growth of a specific transcription profile, further studies are necessary to clarify which model holds true. It would be interesting to use single-cell sequencing technology to study transcription profile diversity within the F1 hybrid cell population

at different time periods. If the majority of cells display one heterotic transcription profile consistently throughout time, would suggest heterosis depends on signal or feedback process, which can be based on protein stability, as hypothesize by Goff (2011). While if transcription profiles are diverse and this diversity diminishes throughout time so a specific heterotic trancription profile dominates the cell population, would suggest F1 hybrid transcription initially displays stochasticity but will be overcome by one heterotic transcription profile. Nonetheless, more data is needed to support any molecular model for heterosis, and as shown in my thesis, *Saccharomyces* yeast is an excellent model system to dissect allele specific expression of a hybrid genome due to the availability good quality reference genomes and high divergent *Saccharomyces* populations (Liti *et al*., 2009). However I do recommend uniformity in transcriptome analysis methods in future studies, and even some previous measures of competitive growth or fitness, to give researchers a sense of how transcription might be influencing the fitness of an individual.

4. Heterosis predictions

My thesis proposed certain parental characteristics which impact heterosis in *Saccharomyces* yeast. Heterosis has been thoroughly studied for over century, thus common parental characteristics that produce a higher yield hybrid in crop plants have been systematically identified (Shull, 1908; Fehr & Hadley, 1980). Throughout my thesis I observed similar patterns; parental phenotypic divergence was positive related to F1 hybrid competitive growth (Chapter I and Bernardes *et al*., 2016), and indirectly, highly inbred monosporic clones derived from high heterozygosity domesticated primary isolates displayed lower asexual growth than monosporic clones derived from low heterozygosity wild primary isolates (Chapter III).

Heterosis was primarily identified and studied in crop plants (Shull, 1908; East, 1936), where crosses between divergent and highly inbred parents resulted in offspring with greater biomass than both inbred parents. Parental divergence could be a measure of genetic or phenotypic divergence; I did not identify any relationship between heterosis and genetic divergence (Chapter I and Bernardes *et al*., 2016) neither did Zörgö *et al*. (2012) or Plech *et al*. (2014) studies, I did identify a positive relationship between phenotypic divergence and heterosis. Phenotypic divergence, or difference between divergent *Saccharomyces* parents competitive growth was positive

related to heterosis in a wide variety of crosses, between wild populations of *S. paradoxus* and between domesticated and wild *Saccharomyces* species, and in different environments (Chapter I and Bernardes *et al*., 2016). Thus suggesting phenotypic divergence can be a general predictor for heterosis in *Saccharomyces* yeast as it already is in crop plants. Another good predictor for heterosis relates heterozygosity of the primary isolates from which we normally derived the F1 hybrid parental strains. Heterozygosity of the primary isolates can be indicative of the number of recessive deleterious alleles present yeast populations. Autodiplodization or inbreeding of highly heterozygous primary isolates (domesticated) results in inbred parents with lower fitness due to exposure of deleterious alleles, while autodiplodization of low heterozygous primary isolates (wild) exposes less deleterious alleles producing inbred parents with higher fitness (Chapter III). Thus if we were to cross two highly inbred parents derived from domesticated populations with high heterozygosity we would expect several deleterious alleles to be complemented in the F1 hybrid, giving it a fitness advantage over its inbred parents. While if we were to cross two monosporic clones derived from divergent wild populations only a smaller number of alleles would be complemented in the F1 hybrid.

Inbreeding and phenotypic divergence of parental strains are readily measured in *Saccharomyces* yeast, but further applications of these patterns should be applied in useful areas such as crop plant breeding and cattle breeding. Studies in *Saccharomyces* yeasts can be easily applied to crop plant breeding, due to similarities in domestication and reproduction of yeast and most crop plants (Birchler *et al*., 2010). Both *Saccharomyces* yeasts and crop plants can be selected for highly inbred strains, by autodiplodization in yeast and rounds of self-fertilization in crop plants, specific and desirable phenotypic characteristics can be favoured. Both plants and yeast can be easily hybridise and tolerate crosses between different species (Fehr & Hadley, 1980), producing viable but infertile hybrids, however infertility can be overcome by their fast asexual reproduction that leads to a higher yield in size and number of (infertile) seeds, used for human consumption in crop plants. Thus similar parental characteristics that contribute to heterosis in *Saccharomyces* yeasts such as parental divergence, inbreeding and heterozygosity may also contribute to heterosis in crop plants. However these characteristics are challenging to apply in other economically important goods such as animal cattle; selection of favourable and

desirable characters through inbreeding can be much more time consuming, due to big generation times in mammals, and also hybridisation can be prevented by pre- and pos-zygotic barriers which are much more strict than for crop plants or and *Saccharomyces* yeast. But the parallelisms between crop plants and *Saccharomyces* yeast outcomes of hybridisation are astonishing, and further studies are desirable to apply what we have learned from *Saccharomyces* yeasts heterosis studies and predictions to crop plant breeding.

5. Isolation of Saccharomyces yeasts

In my thesis I only used *Saccharomyces* yeasts that had been previously isolated by enrichment cultures (Chapter I, Bernardes *et al*. (2016), Chapter II and Chapter III), this method is a widespread technique used to isolate *Saccharomyces* yeast from natural environments. However, the use of enrichment cultures might be depleting a great proportion of *Saccharomyces* yeast diversity from our sampling.

A caveat in heterosis studies is the way we isolate *Saccharomyces* from their natural habitats or how we get primary isolates. Standard practice uses enrichment cultures where environmental samples are placed in rich growth medium and then incubated. This method distorts the quantity and diversity of yeast populations present in the sample because it selects for fast growing inbred diploids present (Goddard & Greig, 2015). For example, if the environmental samples are composed of haploid homothallic spores, rich growth medium and incubation will induce spores to autodiplodize depleting the population of heterozygosity by forming completely homozygous diploid strains (Mortimer *et al*., 1994), and diminishing diversity by selecting for a fast growing primary isolates, culminating in the isolation of fast growing primary isolates with low heterozygosity. While when the environmental samples are composed of vegetative mitotic diploids, enrichment cultures will select for genotypes that grow fast in artificial media, not capturing slow growing, and potentially heterozygous, yeasts that might be fit in their natural environment (Goddard *et al*., 2010). This might explain why we identify low heterozygosity levels for *Saccharomyces* isolated from wild habitats, and high heterozygosity levels for *Saccharomyces* isolated from wine ferments (Knight & Goddard, 2015) or other domesticated environments (Magwene *et al*., 2011), because wild yeast populations are more likely to go through a haploid form more often due to strong selection that

induces higher rates of sexual reproduction (Tsai *et al*., 2008), while domesticated strains are more likely to remain diploids (Cubillos *et al*., 2009) maintaining the levels of heterozygosity of the environmental sample but diminishing any diversity that might be present.

It would be interesting to develop other method for *Saccharomyces* yeasts isolation that does not constrict yeast diversity by selecting for autodiplodization or fast growth. By not selecting for fast growth in artificial media we could probably capture a higher diversity of *Saccharomyces*, also it would be useful to identify different autodiplodized haploid spores that were collected in the environmental sample. If there was such method we could compare the *Saccharomyces* populations isolated with the enrichment cultures and have a critical discussion in whether enrichment culture is truly capturing the diversity of *Saccharomyces* yeasts present in the natural habitats.

6. General conclusion

Heterosis is one of the most interesting phenomenons in biology and its importance to world economy is undeniable. Agriculture and crop breeding have been the focus of heterosis studies for over a century due to their direct applications and relevance to human consumption. The rise of *Saccharomyces* yeasts as a model system for the study of heterosis comes with a promise of simplifying the study of genetic and molecular mechanisms behind heterosis. Thus a molecular mechanism that explains multigenic heterosis at a cell level for multiple species under different environmental conditions can be paramount in future advances and applications of heterosis. Heterosis applications in crop and cattle breeding can benefit human society, highlighting the impact of evolutionary biology studies on our lives.

Glossary & List of Abbreviations

Additive transcription
When the F1 hybrid gene transcription is similar to the average of the parental strains.

Allele specific expression
Specific transcription of the parental alleles in the F1 hybrid, when the F1 hybrid has the ability to differentiate parental alleles. Distinguish parental alleles transcription by mapping the RNA-seq reads of the F1 hybrid to two divergent parental genomes. If a read maps to both genomes in the same amount, there is no allele specific expression.

Asexual reproduction
Diploid or haploid yeasts grow vegetatively by budding. The offspring is identical to the parent in terms of genome and ploidy.

Autodiplodization
Germination of a haploid spore, followed by mate-type switching of the haploid cell and cross with a sister haploid cell from the same spore-colony. Complete duplication of a haploid genome to a diploid genome. Complete homozygosity upon autodiplodization.

Best-parent heterosis (BPH)
When the F1 hybrid fitness is higher than the parent with the highest fitness.

Competitive growth
Measure of asexual growth of a test strain against a control. In my thesis, asexual growth of F1 hybrid when in direct competition with the parental strains in pairwise comparisons.

***Cis*-regulated transcription**
Transcription is regulated by elements closely linked to the gene(s) they affect (i. e. enhancers or TATA box). *Cis*-regulatory effects control only the parental alleles they are linked to in the F1 hybrid. *Cis*-regulation can be identified by significant differences in allelic specific expression in the F1 hybrid.

Conserved transcription
When the F1 hybrid gene transcription is similar to both parental strains.

Diploid
Yeast cell contains a pair of homologous chromosomes, 32 chromosomes in total.

Domestic habitat
Man made habitats where yeast is isolated from, normally wineries, breweries or clinics. Relative simplified habitat with strong directional selection, characterized by very low rates of sex (lower than wild habitats). Yeast populations isolated from this habitat display high heterozygosity, and consequently accumulation of recessive deleterious alleles.

Dominance genetic mechanism
Heterosis mechanisms by which the F1 hybrid has a higher fitness due to complementation of recessive deleterious alleles of the parental strains.

Dominant Transcription
When the F1 hybrid gene transcription is similar to one of the parental strains but not the other parental strain.

Enrichment culture
Method of isolating yeast from primary natural habitats, which uses rich culture media selecting for rapid growth yeasts. If the yeast isolates are haploid there are high chances for autodiplodization, and if the yeast isolates are diploids they remain diploids.

Environmental-stress related genes (ESR)
Several genes with specific transcription response regardless of the environmental change. ESR genes comprise around 16% of the yeast genome.

Epistasis genetic mechanism
Heterosis mechanism by which the F1 hybrid has a higher fitness due to interaction between alleles from unlinked loci.

F1 hybrid
First generation of an intra-specific or an inter-specific cross between divergent parental populations.

Facultative sexual reproduction
Ability to reproduce sexually or asexually.

Fitness
In this thesis, fitness refers to a measure of asexual growth such as competitive growth (Chapter I and II) or maximum growth rate (Chapter III).

Haploid
Yeast cell contains a homologous chromosome, 16 chromosomes in total.

Heterosis
Advantage of the F1 hybrid in comparison to one or both parental species. Also known as hybrid vigour.

Heterozygosity
Number or heterozygous sites in a given genome.

Heterozygous
When a gene has two different alleles for a certain locus.

Homozygous
When a gene has the same two alleles for a certain locus.

Hybridisation
Cross between two individuals of divergent populations (intra-specific) or between two individuals of different species (inter-specific). Hybridisation is detected by direct observation or by analysis of population genomic data.

Inbred
Cross between two spores of the same tetrad. Complementation of haploid genomes into a diploid similar to the original parent (Chapter III).

Inbreeding depression
Low fitness of a hybrid individual due to inbreeding. Normally cause by decrease in heterozygosity, which might expose recessive deleterious alleles.

Inter-specific hybridisation
Result of a cross between two individuals of different species.

Intra-specific hybridisation
Result of a cross between two individuals of divergent populations.

Maximum growth rate
Measure of asexual growth taken from the steepest slope of a certain growth curve when a strain grows in isolation ($MG=(LOG(OD_{t+1}/OD_t))$ x100).

Mid-parent heterosis (MPH)
When the F1 hybrid fitness is higher than the parental average.

Misexpression
When the F1 hybrid gene transcription is significantly higher or lower than both parental strains, also refer to as overdominant for a particular strain.

Monosporic clone
Result of autodiplodization. Complete homozygous strains normally used in the laboratory for experiments or sequence due to their stability.

Negative heterosis
When a F1 hybrid has lower fitness than one or both parents. Disagreeing with Shull (1941) definition for heterosis.

Next Generation Sequencing (NGS)
Or High-throughput sequencing refers to modern sequencing techniques like Illumina sequencing. These sequencing techniques are normally used in whole-genome sequencing (WGS) and RNA-seq.

Outcrossing
Cross between two spores from tetrads or ascus belonging to divergent yeast populations. Result in a strain with higher heterozygosity than the parental strains.

Overdominance genetic mechanism
Mechanism by which the F1 hybrid has a higher fitness due to the intrinsic advantage of heterozygous loci.

Overdominant transcription
Extreme form of dominant transcription when the transcription value is not similar to any of the parents but goes beyond one specific parental strains.

Positive heterosis
When a F1 hybrid has a similar or higher fitness than one or both parents.

Recessive deleterious alleles
The result of recessive deleterious mutations. When a locus is dominant and heterozygous with one recessive deleterious allele and one wild-type allele, there will be no negative effect on the individual (masked deleterious allele). When a locus is homozygous with both alleles recessive deleterious alleles or is haploid with recessive deleterious alleles, there will be a negative effect on the individual (exposed deleterious alleles).

RNA-Seq
Whole transcriptome sequencing by Next-generation sequencing (NGS). Account of the presence and quantity of mRNA in a biological sample at a given moment in time (transcription profile).

***Saccharomyces* yeasts**
Saccharomyces yeasts form a complex of seven closely related, genetically tractable yeast species with similar morphologies. Species are *S. cerevisiae*, *S. paradoxus*, *S. mikatae*, *S. kudriavzevii*, *S. arboricola*, *S. eubayanus*, and *S. uvarum*.

Sexual reproduction
Diploid yeast sporulates and forms a tetrad with four haploid spores within. Each spore is able to autodiplodized, to inbreed or to outcross. The offspring will be different from the parental strains if the haploid spore outcross. The offspring will be similar to the parent if the haploid spore inbreeds or autodiplodizes.

Single Nucleotide Polymorphism (SNP)
DNA sequence variation occurring when a single nucleotide differs between individuals of different species or divergent populations.

Spore viability
Number of viable spores over the total number of spores. In my thesis, viable spores form visible colonies (Chapter III).

TATA box
Conserved element in the promoter region of a gene that bounds to a TATA-biding protein and affects on the initiation of gene transcription. Only 20% of the yeast genes contain TATA box in their promoter region and have been associated with stress-related response.

Tetrad
Or ascus is the sac where the four haploid spores are stored that is formed upon sexual reproduction.

***Trans*-regulated transcription**
Transcription is regulated by elements not linked to the gene(s) they affect (i. e. transcription factors). *Trans*-regulatory effects affect both parental alleles of a particular gene in the F1 hybrid.

Transcription profile
Presence and quantity of genes transcribed by an.

Whole Genome Sequence (WGS)
Modern sequence technique that determine the complete DNA sequence of the genome of an individual or biological sample.

Wild habitat
Natural habitats where yeast is isolated normally bark or leaf litter of trees. Relative stressful and complex habitats with strong selection, characterized by low rates of sex (higher than domesticated habitats). Yeast populations isolated from this habitat display low heterozygosity, and consequently no accumulation of recessive deleterious alleles.

Worst-parent heterosis
F1 hybrid fitness lower than the parent with the lowest fitness. Disagreeing with Shull (1941) definition for heterosis.

Acknowledgements

I would like to thank my supervisor Dr. Duncan Greig for all the support and for opportunity of doing my PhD in the Experimental Evolution Group in such a prestigious institute as Max Planck Institute for Evolutionary Biology. This work enabled me to grow both professionally and as a scientist and would not have been possible without the encouragement and enthusiasm of my PhD supervisor, who motivated me to do better on every single step.

I would also like to thank the referees who have kindly accepted to judge my PhD thesis. Dr. Arne Nolte for his enthusiasm and advice in how to write and discuss my different projects, and of course the great food talks. And Dr. Eva Stukenbrock for accepting my invitation and once more judging the work of a *Saccharomyces* yeast PhD student, I am sure you are already a specialist in *Saccharomyces* yeasts.

As for my colleagues in the Experimental Evolution group, I would like to thank my 'classmates' Ozan Bozdag and Vienna Kowallik, my 'officemate' Primrose Boynton, and my 'labmate' Ellen McConnell for showing me the right path(s) through my PhD studies and for always being there helping me to see the 'light in the end of the tunnel', weather with encouraging scientific words or with laughs at 'feierabend' beers. I would like also to thanks Gunda Dechow-Seligmann for standing by me, even when my workbench seemed like a lost cause.

A special thanks to my co-authors in two of my projects, the input to my work was truly invaluable. To Dr. Rike Stelkens for sharing part of her discoveries with me, supporting me as a young scientist and being a great example of what women in science should be. And to Dr. David Rogers for always being positive even when negative, for pushing me to do more and to do better, I hope you continue to inspire students to pray to the PCR Gods, and thank you for making the time in Plön great!

I would also like to thank my co-workers in MPI for Evolutionary Biology, especially Dr. Lutz Becks for the nice conversations and on point opinions on our train rides. To Emre who will always deny his help, and Jens for our great late night science talks.

To the theory group for adopting me and allowing me to be on their group activities- you have the best group activities. And finally to the greater than great Noemi Woltermann and Svenja Hoier for all the help, at work or at my broken old house, and for 'integrating' me in the best possible way!

I also have to thank my family. My parent for giving me the opportunity to study and pursue my dreams; these last years were tough but I was only able to finish the PhD because you were always there for me. To my brother Carlos Bernardes and his wife Daniela Dias, my fellow scientists and tireless supporters in reviewing my thesis or giving me baby cats to play with every time it was needed.

To my partner Jan Marxsen for making me happy in the most stressful times. To my roommates, Bocki und Willer, for always giving me a purpose even when the purpose was doing the dishes! To my grrrlz Flicke, Nathalia, Lisa, Alexa, Julia, Svetlana etc. for the good times, the bad times, and the dancing times, I love you all. To Herb for allowing me the pleasure of having a hobby and making horrible shows with him. To Moritz, Tim and Hinnerk for fests, the screams, and the tears, and for making my world better.

And of course, in a more serious note I would like to thank Dr. Christine Otto-Morris, my therapist in Kiel. Doing a PhD was hard on my mental health, much harder than I was willing to admit. Only with professional help I was able to persevered and finish my work. I will forever be grateful for your readiness in taking me as your patient, even when I thought I was a lost cause.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Most, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Vainola, R., Wolf, J.B. & Zinner, D. 2013. Hybridization and speciation. *J Evol Biol* **26:** 229-246.

Alexandre, H., Ansanay-Galeote, V., Dequin, S. & Blondin, B. 2001. Global gene expression during short-term ethanol stress in Saccharomyces cerevisiae. *FEBS Letters* **498:** 98-103.

Almeida, P., Gonçalves, C., Teixeira, S., Libkind, D., Bontrager, M., Masneuf-Pomarède, I., Albertin, W., Durrens, P., Sherman, D.J., Marullo, P., Todd Hittinger, C., Gonçalves, P. & Sampaio, J.P. 2014. A Gondwanan imprint on global diversity and domestication of wine and cider yeast Saccharomyces uvarum. **5:** 4044.

Anders, S., Pyl, P.T. & Huber, W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31:** 166-169.

Andrews, S. 2010. FastQC A Quality Control tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Artieri, C.G. & Fraser, H.B. 2014. Evolution at two levels of gene expression in yeast. *Genome Res* **24:** 411-421.

Alexandre, H., Ansanay-Galeote, V., Dequin, S. & Blondin, B. 2001. Global gene expression during short-term ethanol stress in Saccharomyces cerevisiae. *FEBS Letters* **498:** 98-103.

Almeida, P., Gonçalves, C., Teixeira, S., Libkind, D., Bontrager, M., Masneuf-Pomarède, I., Albertin, W., Durrens, P., Sherman, D.J., Marullo, P., Todd Hittinger, C., Gonçalves, P. & Sampaio, J.P. 2014. A Gondwanan imprint on global diversity and domestication of wine and cider yeast Saccharomyces uvarum. **5:** 4044.

Anders, S., Pyl, P.T. & Huber, W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31:** 166-169.

Anderson, W.S. 1939. Fertile mare mules. *Journal of Heredity* **30:** 549-551.

Andrews, S. 2010. FastQC A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Arnold, M.L. & Hodges, S.A. 1995. Are natural hybrids fit or unfit relative to their parents? *TREE* **10:** 67-71.

Artieri, C.G. & Fraser, H.B. 2014. Evolution at two levels of gene expression in yeast. *Genome Res* **24:** 411-421.

Barton, N.H. & Hewitt, G.M. 1985. Analysis of hybrid zones. *Annual Reviews of Ecology and Systematics* **16:** 113-148.

Barton, N.H. 2001. The role of hybridization in evolution. *Molecular Ecology* **10:** 551-568.

Basehoar, A.D., Zanton, S.J. & Pugh, B.F. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116:** 699-709.

Batshinskaya, A.A. 1914. En twicklung sges chichte und Kultur desneuen Hefepilzes Saccharomyces paradoxus. *Journal of Microbiology Epidemiology Immunobiology* **1:** 231-232.

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., Warringer, J., Durbin, R. & Liti, G. 2014. A high-definition view of functional genetic variation from natural yeast genomes. *Molecular Biology and Evolution*.

Bernardes, J.P., Stelkens, R.B. & Greig, D. 2016. Heterosis in hybrids within and between yeast species. *Journal of Evolutionary Biology*.

Bing, J., Han, P., Liu, W., Wang, Q. & Bai, F. 2014. Evidence for a far east Asian origin of lager beer yeast. *Current Biology* **24:** R380-R381.

Birchler, J.A., Yao, H. & Chudalayandi, S. 2007. Biological consequences of dosage dependent gene regulatory systems. *Biochim Biophys Acta* **1769:** 422-428.

Birchler, J.A., Yao, H., Chudalayandi, S., Vaiman, D. & Veitia, R.A. 2010. Heterosis. *Plant Cell* **22:** 2105-2112.

Blein-Nicolas, M., Albertin, W., da Silva, T., Valot, B., Balliau, T., Masneuf-Pomarède, I., Bely, M., Marullo, P., Sicard, D., Dillmann, C., de Vienne, D. & Zivy, M. 2015. A systems approach to elucidate heterosis of protein abundances in yeast. *Molecular & Cellular Proteomics : MCP* **14:** 2056-2071.

Boynton, P.J. & Greig, D. 2014. The ecology and evolution of non-domesticated Saccharomyces species. *Yeast* **31:** 449-462.

Brion, C., Pflieger, D., Souali-Crespo, S., Friedrich, A. & Schacherer, J. 2016. Differences in environmental stress response among yeasts is consistent with species-specific lifestyles. *Mol Biol Cell* **27:** 1694-1705.

Casey, G.P. & Ingledew, W.M. 1986. Ethanol tolerance in yeasts. *Crit Rev Microbiol* **13:** 219-280.

Clowers, K.J., Will, J.L. & Gasch, A.P. 2015. A unique ecological niche fosters hybridization of oak-tree and vineyard isolates of Saccharomyces cerevisiae. *Molecular Ecology* **24:** 5886-5898.

Costanzo, M.C., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D., Pelechano, V., Styles, E.B., Billmann, M., van Leeuwen, J., van Dyk, N., Lin, Z., Kuzmin, E., Nelson, J., Piotrowski, J.S., Srikumar, T., Bahr, S., Chen, Y., Deshpande, R., Kurat, C.F., Li, S.C., Li, Z., Usaj, M.M., Okada, H., Pascoe, N., San Luis, B., Sharifpoor, S., Shuteriqi, E., Simpkins, S.W., Snider, J., Suresh, H.G., Tan, Y., Zhu, H., Malod-Dognin, N., Janjic, V., Przulj, N., Troyanskaya, O.G., Stagljar, I., Xia, T., Ohya, Y., Gingras, A., Raught, B., Boutros, M., Steinmetz, L.M., Moore, C.L., Rosebrock, A.P., Caudy, A.A., Myers, C.L., Andrews, B.J. & Boone, C. 2016. A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**.

Cubillos, F.A., Vasquez, C., Faugeron, S., Ganga, A. & Martinez, C. 2009. Self-fertilization is the main sexual reproduction mechanism in native wine yeast populations. *FEMS Microbiol Ecol* **67:** 162-170.

de Barros Lopes, M., Bellon, J.R., Shirley, N.J. & Ganter, P.F. 2002. Evidence for multiple interspecific hybridization in Saccharomyces sensu stricto species. *FEMS Yeast Research* **1:** 323-331.

Delneri, D., Hoyle, D.C., Gkargkas, K., Cross, E.J.M., Rash, B., Zeef, L., Leong, H.S., Davey, H.M., Hayes, A., Kell, D.B., Griffith, G.W. & Oliver, S.G. 2008. Identification and characterization of high-flux-control genes of yeast through competition analyses in continuous cultures. *Nat Genet* **40:** 113-117.

Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *The American Naturalist* **74:** 312-321.

Dobzhansky, T. 1950. Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of Drosophila pseudoobscura. *Genetics* **35:** 288-302.

Duvick, D.N. 1999. Heterosis: Feeding people and protecting natural resources. In: *The genetic and exploitation of heterosis in crops*. American Society of Agronomy, Crop Science Society of America.

East, E.M. 1936. Heterosis. *Genetics* **21:** 375-397.

Edgerton, M.D. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiology* **149:** 7-13.

Elrod, S.L., Chen, S.M., Schwartz, K. & Shuster, E.O. 2009. Optimizing sporulation conditions for different Saccharomyces cerevisiae strain backgrounds. *Methods Mol Biol* **557:** 21-26.

Emerson, J.J., Hsieh, L.C., Sung, H.M., Wang, T.Y., Huang, C.J., Lu, H.H., Lu, M.Y., Wu, S.H. & Li, W.H. 2010. Natural selection on cis and trans regulation in yeasts. *Genome Res* **20:** 826-836.

Fay, J.C. & Benavides, J.A. 2005. Evidence for domesticated and wild populations of Saccharomyces cerevisiae. *PLOS Genetics* **1:** e5.

Feher, K., Lisec, J., Romisch-Margl, L., Selbig, J., Gierl, A., Piepho, H.P., Nikoloski, Z. & Willmitzer, L. 2014. Deducing hybrid performance from parental metabolic profiles of young primary roots of maize by using a multivariate diallel approach. *PLoS One* **9:** e85435.

Fehr, W.R. & Hadley, H.H. 1980. *Hybridization of Crop Plants*. American Society of Agronomy, Crop Science Society of America, Madison, WI.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D. & Brown, P.O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell* **11:** 4241-4257.

Gasch, A.P. 2007. Comparative genomics of the environmental stress response in ascomycete fungi. *Yeast* **24:** 961-976.

Gerstein, A.C., Cleathero, L.A., Mandegar, M.A. & Otto, S.P. 2011. Haploids adapt faster than diploids across a range of environments. *Journal of Evolutionary Biology* **24:** 531-540.

Gibson, G., Riley-Berger, R., Harshman, L., Kopp, A., Vacha, S., Nuzhdin, S. & Wayne, M. 2004. Extensive sex-specific nonadditivity of gene expression in Drosophila melanogaster. *Genetics* **167:** 1791-1799.

Gietz, R.D. & Woods, R.A. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymology* **350:** 87-96.

Goddard, M.R., Anfang, N., Tang, R., Gardner, R.C. & Jun, C. 2010. A distinct population of Saccharomyces cerevisiae in New Zealand: evidence for local dispersal by insects and human-aided global dispersal in oak barrels. *Environ Microbiol* **12:** 63-73.

Goddard, M.R. & Greig, D. 2015. Saccharomyces cerevisiae: a nomadic yeast with no niche? *FEMS Yeast Res* **15**.

Goff, S.A. 2011. A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding. *New Phytol* **189:** 923-937.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S.G. 1996. Life with 6000 genes. *Science* **274:** 546-567.

Goncalves, P., Valerio, E., Correia, C., de Almeida, J.M. & Sampaio, J.P. 2011. Evidence for divergent evolution of growth temperature preference in sympatric Saccharomyces species. *PLoS One* **6:** e20739.

González, S.S., Barrio, E., Gafner, J. & Querol, A. 2006. Natural hybrids from Saccharomyces cerevisiae, Saccharomyces bayanus and Saccharomyces kudriavzevii in wine fermentations. *FEMS Yeast Research* **6:** 1221-1234.

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

Gu, Z., David, L., Petrov, D., Jones, T., Davis, R.W. & Steinmetz, L.M. 2005. Elevated evolutionary rates in the laboratory strain of Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 1092-1097.

Herskowitz, I. 1988. Life cycle of the budding yeast Saccharomyces cerevisiae. *Microbiological Reviews* **52:** 536-553.

Hollick, J.B. & Chandler, V.L. 1998. Epigenetic allelic states of a maize transcriptional regulatory locus exhibit overdominant gene action. *Genetics* **150:** 891-897.

Jensen, R.A. 2001. Orthologs and paralogs - we need to get it right. *Genome Biology* **2:** interactions1002.1001-interactions1002.1003.

Johnson, L.J., Koufopanou, V., Goddard, M.R., Hetherington, R., Schäfer, S.M. & Burt, A. 2004. Population genetics of the wild yeast Saccharomyces paradoxus. *Genetics* **166:** 43-52.

Johnson, N.A. & Porter, A.H. 2000. Rapid speciation via parallel, directional selection on regulatory genetic pathways. *J Theor Biol* **205:** 527-542.

Jones, D.F. 1917. *Dominance of linked factors as a means of accounting for heterosis*. . NY USA: John Wiley & Sons.

Kim, D., Langmead, B. & Salzberg, S.L. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Meth* **12:** 357-360.

Knight, S. & Goddard, M.R. 2015. Quantifying separation and similarity in a Saccharomyces cerevisiae metapopulation. *ISME J* **9:** 361-370.

Kowallik, V. & Greig, D. 2016. A systematic forest survey showing an association of Saccharomyces paradoxus with oak leaf litter. *Environ Microbiol Rep*.

Li, L., Lu, K.P., Chen, Z., Mu, T., Hu, Z. & Li, X. 2008. Dominance, overdominance and epistasis condition the heterosis in two heterotic rice hybrids. *Genetics* **180:** 1725-1742.

Libkind, D., Hittinger, C.T., Valério, E., Gonçalves, C., Dover, J., Johnston, M., Gonçalves, P. & Sampaio, J.P. 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proceedings of the National Academy of Sciences* **108:** 14539-14544.

Lindegren, C.C. 1945. Yeast Genetics : Life Cycles, Cytology, Hybridization, Vitamin Synthesis, and Adaptive Enzymes. *Bacteriological Reviews* **9:** 111-170.

Lindegren, C.C.B., J. Edgar; Calle, Juan De Dios 1953. Heterosis in Saccharomyces. *Nature* **172:** 800-802.

Lippman, Z.B. & Zamir, D. 2007. Heterosis: revisiting the magic. *Trends in Genetics* **23:** 60-66.

Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., Tsai, I.J., Bergman, C.M., Bensasson, D., O'Kelly, M.J., van Oudenaarden, A., Barton, D.B., Bailes, E., Nguyen, A.N., Jones, M., Quail, M.A., Goodhead, I., Sims, S., Smith, F., Blomberg, A., Durbin, R. & Louis, E.J. 2009. Population genomics of domestic and wild yeasts. *Nature* **458:** 337-341.

Lopandic, K., Gangl, H., Wallner, E., Tscheik, G., Leitner, G., Querol, A., Borth, N., Breitenbach, M., Prillinger, H. & Tiefenbrunner, W. 2007. Genetically different wine yeasts isolated from Austrian vine-growing regions influence wine aroma differently and contain putative hybrids between *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii*. *FEMS Yeast Research* **7:** 953-965.

Love, M.I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15:** 550.

Lovell, J.T., Schwartz, S., Lowry, D.B., Shakirov, E.V., Bonnette, J.E., Weng, X., Wang, M., Johnson, J., Sreedasyam, A., Plott, C., Jenkins, J., Schmutz, J. & Juenger, T.E. 2016. Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C4 grass. *Genome Res* **26:** 510-518.

Luo, L.J., Li, Z.K., Mei, H.W., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S. & Paterson, A.H. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* **158:** 1755-1771.

Magwene, P.M., Kayıkçı, Ö., Granek, J.A., Reininga, J.M., Scholl, Z. & Murray, D. 2011. Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* **108:** 1987-1992.

Magwene, P.M. 2014. Revisiting Mortimer's Genome Renewal Hypothesis: Heterozygosity, homothallism, and the potential for adaptation in yeast. In: *Ecological Genomics: Ecology and the Evolution of Genes and Genomes* (C. R. Landry & N. Aubin-Horth, eds), pp. 37-48. Springer Netherlands, Dordrecht.

Mayr, E. 1942. *Systematics and origin of species, from a view point of a zoologist*. Columbia University Press, New York.

McCusker, J.H., Clemons, K.V., Stevens, D.A. & Davis, R.W. 1994. Genetic characterization of pathogenic Saccharomyces cerevisiae isolates. *Genetics* **136:** 1261-1269.

McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R., Moreau, R.A., Nuñez, A., Butrym, E.D., Richards, M.P., Wang, C., Cheng, G., Zhao, Z. & Wang, C. 2004. Fermented beverages of pre- and proto-historic China. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 17593-17598.

McManus, C.J., May, G.E., Spealman, P. & Shteyman, A. 2014. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome Res* **24:** 422-430.

Moll, R.H., Lonnquist, J.H., Fortuno, J.V. & Johnson, E.C. 1965. The relationship of heterosis and genetic divergence in maize. *Genetics* **52:** 139-144.

Mortimer, R.K. & Johnston, J.R. 1986. Genealogy of the principal strains of the yeast genetic stock center. *Genetics* **113:** 35.

Mortimer, R.K., Romano, P., Suzzi, G. & Polsinelli, M. 1994. Genome renewal: a new phenomenon revealed from a genetic study of 43 strains of Saccharomyces cerevisiae derived from natural fermentation of grape musts. *Yeast* **10:** 1543-1552.

Muller, L.A., Lucas, J.E., Georgianna, D.R. & McCusker, J.H. 2011. Genome-wide association analysis of clinical vs. nonclinical origin provides insights into Saccharomyces cerevisiae pathogenesis. *Mol Ecol* **20:** 4085-4097.

Muller, L.A.H. & McCusker, J.H. 2009. Microsatellite analysis of genetic diversity among clinical and nonclinical Saccharomyces cerevisiae isolates suggests heterozygote advantage in clinical environments. *Molecular Ecology* **18:** 2779-2786.

Naumov, G.I., Naumova, E.S. & Sniegowski, P.D. 1998. Saccharomyces paradoxus and Saccharomyces cerevisiae are associated with exudates of North American oaks. *Canadian Journal of Microbiology* **44:** 1045-1050.

Naumov, G.I., James, S.A., Naumova, E.S., Louis, E.J. & Roberts, I.N. 2000. Three new species in the Saccharomyces sensu stricto complex: Saccharomyces cariocanus, Saccharomyces kudriavzevii and Saccharomyces mikatae. *Int J Syst Evol Microbiol* **50**.

Naumov, G.l. 1996. Genetic identification of biological species in the *Saccharomyces* sensu stricto complex. *Journal of Industrial Microbiology* **17:** 295-302.

Nielsen, M.G., Wilson, K.A., Raff, E.C. & Raff, R.A. 2000. Novel gene expression patterns in hybrid embryos between species with different modes of development. *Evol Dev* **2:** 133-144.

Norman, C., Howell, K.A., Millar, A.H., Whelan, J.M. & Day, D.A. 2004. Salicylic acid is an uncoupler and inhibitor of mitochondrial electron transport. *Plant Physiology* **134:** 492-501.

Plech, M., de Visser, J.A. & Korona, R. 2014. Heterosis is prevalent among domesticated but not wild strains of *Saccharomyces cerevisiae*. *G3* **4:** 315-323.

Powers, L. 1944. An expansion of Jones's theory for the explanation of heterosis. *The American Naturalist* **78:** 275-280.

Ranz, J.M., Namgyal, K., Gibson, G. & Hartl, D.L. 2004. Anomalies in the expression profile of Interspecific hybrids of Drosophila melanogaster and Drosophila simulans. *Genome Research* **14:** 373-379.

Reeb, P.D., Bramardi, S.J. & Steibel, J.P. 2015. Assessing dissimilarity measures for sample-based hierarchical clustering of RNA sequencing data using plasmode datasets. *PLOS ONE* **10:** e0132310.

Rieseberg, L.H., Sinervo, B., Linder, C.R., Ungerer, M.C. & Arias, D.M. 1996. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* **272:** 741-745.

Rong, R., Chandley, A.C., Song, J., McBeath, S., Tan, P.P., Bai, Q. & Speed, R.M. 1988. A fertile mule and hinny in China. *Cytogenet Cell Genet* **47:** 134-139.

Ruderfer, D.M., Pratt, S.C., Seidel, H.S. & Kruglyak, L. 2006. Population genomic analysis of outcrossing and recombination in yeast. *Nat Genet* **38:** 1077-1081.

Sampaio, J.P. & Goncalves, P. 2008. Natural populations of Saccharomyces kudriavzevii in Portugal are associated with oak bark and are sympatric with S. cerevisiae and S. paradoxus. *Appl Environ Microbiol* **74**.

Scannell, D.R., Zill, O.A., Rokas, A., Payen, C., Dunham, M.J., Eisen, M.B., Rine, J., Johnston, M. & Hittinger, C.T. 2011. The awesome power of yeast

evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces* sensu stricto Genus. *G3* **1:** 11-25.

Schacherer, J., Shapiro, J.A., Ruderfer, D.M. & Kruglyak, L. 2009. Comprehensive polymorphism survey elucidates population structure of S. cerevisiae. *Nature* **458:** 342-345.

Schaefke, B., Emerson, J.J., Wang, T.Y., Lu, M.Y., Hsieh, L.C. & Li, W.H. 2013. Inheritance of gene expression level and selective constraints on trans- and cis-regulatory changes in yeast. *Mol Biol Evol* **30:** 2121-2133.

Shapira, R., Levy, T., Shaked, S., Fridman, E. & David, L. 2014. Extensive heterosis in growth of yeast hybrids is explained by a combination of genetic models. *Heredity* **113:** 316-326.

Shen, G., Zhan, W., Chen, H. & Xing, Y. 2014. Dominance and epistasis are the main contributors to heterosis for plant height in rice. *Plant Science* **215–216:** 11-18.

Shull, G.H. 1908. The composition of a field of Maize. *Journal of Heredity* **os-4:** 296-301.

Shull, G.H. 1948. What Is "Heterosis"? *Genetics* **33:** 439-446.

Stefanini, I., Dapporto, L., Legras, J.L., Calabretta, A., Di Paola, M., De Filippo, C., Viola, R., Capretti, P., Polsinelli, M., Turillazzi, S. & Cavalieri, D. 2012. Role of social wasps in Saccharomyces cerevisiae ecology and evolution. *Proceedings of the National Academy of Sciences* **109:** 13398-13403.

Tirosh, I., Weinberger, A., Carmi, M. & Barkai, N. 2006. A genetic signature of interspecies variations in gene expression. *Nat Genet* **38:** 830-834.

Tirosh, I., Reikhav, S., Levy, A.A. & Barkai, N. 2009. A yeast hybrid provides insight into the evolution of gene expression regulation. *Science* **324:** 659-662.

Tsai, I.J., Bensasson, D., Burt, A. & Koufopanou, V. 2008. Population genomics of the wild yeast *Saccharomyces paradoxus*: Quantifying the life cycle. *Proceedings of the National Academy of Sciences* **105:** 4957-4962.

Vaughan, M.A. & Martini, A. 1987. Three newly delimited species of *Saccharomyces* sensu stricto. *Antonie van Leeuwenhoek* **53:** 77-84.

Wang, S. & Bai, F. 2008. Saccharomyces arboricolus sp. nov., a yeast species from tree bark. *International Journal of Systematic and Evolutionary Microbiology* **58:** 510-514.

Wang, Z., Sun, X., Zhao, Y., Guo, X., Jiang, H., Li, H. & Gu, Z. 2015. Evolution of gene regulation during transcription and translation. *Genome Biology and Evolution* **7:** 1155-1167.

Waser, N.M. 1993. Population structure, optimal outbreeding, and assortative mating in angiosperms. In: *The natural history of inbreeding and outbreeding: theoretical and empirical perspectives*. (N. W. Thornhill, ed., pp. pp. 173-199. University of Chicago Press, Chicago, USA.

Wittkopp, P.J., Haerum, B.K. & Clark, A.G. 2004. Evolutionary changes in cis and trans gene regulation. *Nature* **430:** 85-88.

Wittkopp, P.J. & Kalay, G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet* **13:** 59-69.

Yue, J.-X., Li, J., Aigrain, L., Hallin, J., Persson, K., Oliver, K., Bergstrom, A., Coupland, P., Warringer, J., Lagomarsino, M.C., Fischer, G., Durbin, R. & Liti, G. 2017. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat Genet* **49:** 913-924.

Zörgö, E., Gjuvsland, A., Cubillos, F.A., Louis, E.J., Liti, G., Blomberg, A., Omholt, S.W. & Warringer, J. 2012. Life history shapes trait heredity by accumulation

of loss-of-function alleles in Yeast. *Molecular Biology and Evolution* **29:** 1781-1789.

Figures & Tables Index

Introduction

Chapter I

Chapter II

Chapter III

List of Publications

Bernardes, J. P.[1], Stelkens, R[1]. B. & Greig, D.[1,2], 2016, Heterosis in hybrids within and between yeast species., Journal of Evolutionary Biology, ISSN: 1420-9101, DOI: 10.1111/jeb.13023

1. Max Planck Institute for Evolutionary Biology, Experimental Evolution Group, Plön, Germany

2. The Galton Laboratory, Department of Genetics, Evolution, and Environment, University College London, London, UK

Curriculum Vitae

JOANA P. BERNARDES

PhD Candidate EVOLUTIONARY BIOLOGY

PROFILE

Recently PhD in Evolutionary Biology, using Yeast Hybrids Studying their Fitness Advantages or Heterosis. Background in Evolution of Yeast Populations and Evolution and Development of Insects Immune System.

INTERESTS

Microbiology and Evolution
Hybridization and Evolution
Agriculture applications
Yeast Evolution Studies

PUBLICATIONS

Bernardes, J. P., Stelkens, R.B. and Greig, D. (2016), "Heterosis in hybrids within and between yeast species." Journal of Evolutionary Biology, DOI: 10.1111/jeb.13023

EDUCATION & WORK EXPERIENCE

PhD Evolutionary Biology 2014-present
Max-Planck Institute for Evolutionary Biology (MPI)
"Heterosis in Yeast Hybrids"
Supervision under Dr. Duncan Greig

Teacher of Sciences 2012-2013
School Group of Monte-Abrãao
High-school and middle-school levels

MSc thesis Evolutionary Biology and Development 2011-2012
Institute of Biology Leiden
"Roles of the Serosa in the Immune Defense of the beetle '*Tribolium castaneum*'"
Supervision under Prof. Dr. Maurijn van der Zee
Grade 18 in 20 possible

MSc Evolutionary Biology and Development 2010-2012
Faculty of Science of the University of Lisbon
Grade 17 in 20 possible

BSc Evolutionary Biology and Development 2006-2010
Faculty of Science of the University of Lisbon
Grade 14 in 20 possible

SKILLS

MICROBIOLOGY ●●●●●
Cultivation and fitness assays of microorganisms .
Genetic transformations in microorganisms and PCR.

STATISTICAL ANALYSIS ●●●○○
WGS and RNA-Seq transcriptome analysis. Uni and Multivariate analysis.

COMPUTATION ●●●●○
Python, R, Microsoft Office and Adobe Illustrator.

CONTACT

jpimentabernardes@gmail.com | bernardes@evolbio.mpg.de
+4915205993516 | linkedin.com/in/joana-pimenta-bernardes

JOANA P. BERNARDES

PhD Candidate EVOLUTIONARY BIOLOGY

PROFILE

Recently PhD in Evolutionary Biology, using Yeast Hybrids Studying their Fitness Advantages or Heterosis. Background in Evolution of Yeast Populations and Evolution and Development of Insects Immune System.

CHOSEN ACADEMIC ACTIVITIES

Invited speaker at "ESEB 2017" in in Groningen	-
Invited speaker at AQUAVIT in MPI	2017
Invited speaker at Evolutionary Theory Department in MPI	2017
Invited speaker at Community Dynamics Group in MPI	2017
Invited speaker at "Fungal Symposium' in MPI	2016
Poster presentation at "Evolution 2016" in Austin	2016
Poster presentation at "Life Science Nord"	2015
Poster presentation at "ESEB 2015" in Lausanne	2015
Poster presentation at "SMBE 2015" in Vienna	2015
Invited speaker at "Science Club" at the Department of Evolutionary Ecology in CAU in Kiel	2015
Invited speaker at "10th APBE Meeting' in Lisbon	2014
Invited presentation at "NGS analysis" at POPGEN in Vienna	2014
Invited presentation at "Quantitative Evolutionary Biology" in Şirince	2014
Student Representative of IMPRS students	2014-2016
Poster presentation at "EVO-DEVO 2012" in Lisbon	2012

WORKSHOP & QUALIFICATIONS

Multivariate statistic analysis in R course at MPI	2015
Workshop in NGS at POPGEN in Vienna	2014
Workshop in Quantitative Evolutionary Biology in Şirince	2014
Evolutionary Genetics course at MPI	2014
Evolutionary Medicine course at CAU in Kiel	2014
Biostatistics course at CAU in Kiel	2014

AWARDS

PhD Stipend from Max-Planck Research School	2014-present
Best poster in "Life Science Nord"	2015
MSc Scholarship from Erasmus Program	2011

REFFERENCES

Dr. Duncan Greig
d.greig@evolbio.mpg.de
d.greig@ucl.ac.uk
+49 4522 763 234
+44 20 7679 5106

Prof. Dr. Maurjin van der Zee
m.van.der.zee@biology.leidenuniv.nl
+31 71 527 4885

CONTACT
jpimentabernardes@gmail.com | bernardes@evolbio.mpg.de
+4915205993516 | linkedin.com/in/joana-pimenta-bernardes

Declaration

Hereby I declare that:

i) Apart from my supervisor's guidance, the content and design of this dissertation is the product of my own work. The co-author's contributions to specific chapters are listed in the thesis outline section.

ii) This thesis has not already been submitted either partially or wholly as part of a doctoral degree to another examination body, and no other materials are published or submitted for publication than indicated in the thesis.

iii) The preparation of the thesis has been subjected to the Rules of Good Scientific Practice of the German Research Foundation.

Kiel, November 2017

_____

Joana Pimenta Bernardes