Dysfunctional Gene Regulation of the Meiotic Program in an Interspecific Yeast Hybrid

Devjanee Swain Lenz
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds
Part of the Biology Commons, and the Genetics and Genomics Commons

Recommended Citation
Swain Lenz, Devjanee, "Dysfunctional Gene Regulation of the Meiotic Program in an Interspecific Yeast Hybrid" (2013). Arts & Sciences Electronic Theses and Dissertations. 1039.
https://openscholarship.wustl.edu/art_sci_etds/1039

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Dysfunctional Gene Regulation of the Meiotic Program in an Interspecific Yeast Hybrid

by

Devjanee Swain Lenz

A dissertation presented to the Graduate School of Arts and Sciences of Washington University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

August 2013

St. Louis, Missouri
# Table of Contents

LIST OF FIGURES ........................................................................................................ iv
LIST OF TABLES ........................................................................................................... v
LIST OF SUPPLEMENTAL FIGURES ......................................................................... vi
LIST OF SUPPLEMENTAL TABLES ............................................................................. vii
Acknowledgements ..................................................................................................... viii
Dedication ..................................................................................................................... x
ABSTRACT OF THE DISSERTATION .................................................................... xi

Chapter 1: Introduction .............................................................................................. 1
  Reproductive Isolation ............................................................................................ 2
  Chromosomal divergence and reproductive isolation ........................................... 3
  The genetic basis of post-zygotic reproductive isolation ....................................... 3
  What types of genes contribute to RI? ................................................................... 4
  Regulatory evolution and reproductive isolation .................................................... 5
  Yeast as an evolutionary model ................................................................................ 7
  The yeast life cycle .................................................................................................. 7
  Reproductive isolation in yeast .............................................................................. 8
    Pre-zygotic isolation in yeast .............................................................................. 8
    Post-zygotic isolation in yeast ............................................................................ 9
  Focus of dissertation work ..................................................................................... 12

Chapter 2: Methods to identify and verify incompatible genes ............................... 15
  Abstract .................................................................................................................. 16
  Part 1 – Pre-zygotic isolation in yeast ................................................................. 16
    Introduction ......................................................................................................... 16
    Results and Discussion ...................................................................................... 17
  Part 2 – Post-zygotic isolation in yeast ................................................................. 18
    Introduction ......................................................................................................... 18
    Results and Discussion ...................................................................................... 19
  Part 3 – Conclusions ............................................................................................. 22
  Part 4 – Materials and Methods .......................................................................... 24
    Strains.................................................................................................................. 24
    Measurement of Agglutination Index ................................................................ 25
    Recessive screen media and growth .................................................................. 25
    Tetrad dissections ................................................................................................. 26

Figures and Tables ...................................................................................................... 27

Chapter 3: Heterochronic meiotic misexpression in an interspecific yeast hybrid ......................................................... 31
  Abstract .................................................................................................................. 32
  Introduction ............................................................................................................ 32
  Results ..................................................................................................................... 34
  Hybrid sporulation is similar to S. paradoxus ....................................................... 34
  The hybrid has a longer transition between meiosis I and II than its parents........ 34
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA-seq profiling during meiosis</td>
<td>35</td>
</tr>
<tr>
<td>Heterochronic changes in the hybrid</td>
<td>36</td>
</tr>
<tr>
<td>Trans-acting factors dominate heterochronic differences between the hybrid and its parents</td>
<td>37</td>
</tr>
<tr>
<td>Cis- and trans-acting changes in gene expression</td>
<td>37</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>38</td>
</tr>
<tr>
<td>Meiotic differences between the hybrid and its parents</td>
<td>38</td>
</tr>
<tr>
<td>The anti-recombination model of hybrid sterility</td>
<td>40</td>
</tr>
<tr>
<td>Cis-only and trans-only changes between S. cerevisiae and S. paradoxus</td>
<td>40</td>
</tr>
<tr>
<td>Cis-trans opposing and compensatory changes between the parents and their relationship to reproductive isolation</td>
<td>41</td>
</tr>
<tr>
<td><strong>Conclusions</strong></td>
<td>42</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>42</td>
</tr>
<tr>
<td>Strains</td>
<td>42</td>
</tr>
<tr>
<td>Growth and sporulation conditions</td>
<td>43</td>
</tr>
<tr>
<td>Sampling</td>
<td>43</td>
</tr>
<tr>
<td>Illumina indexing library, alignments and mapping</td>
<td>43</td>
</tr>
<tr>
<td>Differential gene expression measurements</td>
<td>44</td>
</tr>
<tr>
<td>Other statistical analyses</td>
<td>44</td>
</tr>
<tr>
<td>Data access</td>
<td>46</td>
</tr>
<tr>
<td><strong>Acknowledgements</strong></td>
<td>46</td>
</tr>
<tr>
<td><strong>Figures and Tables</strong></td>
<td>47</td>
</tr>
<tr>
<td><strong>Supplementary Figures and Tables</strong></td>
<td>52</td>
</tr>
<tr>
<td><strong>Chapter 4: Conclusion</strong></td>
<td>57</td>
</tr>
<tr>
<td>The relationship between dysfunctional hybrid expression and reproductive isolation</td>
<td>58</td>
</tr>
<tr>
<td><strong>Dysfunctional pathway regulation in hybrids and genetic incompatibilities</strong></td>
<td>60</td>
</tr>
<tr>
<td><strong>Future directions for yeast speciation</strong></td>
<td>62</td>
</tr>
<tr>
<td>Traditional and modern approaches to speciation genomics studies</td>
<td>62</td>
</tr>
<tr>
<td>Reassessment of the anti-recombination model</td>
<td>63</td>
</tr>
<tr>
<td>Why do null mutants in some laboratory backgrounds increase spore viability?</td>
<td>63</td>
</tr>
<tr>
<td>What are the evolutionary forces driving dysfunctional regulation in hybrids?</td>
<td>64</td>
</tr>
<tr>
<td><strong>Final Thoughts</strong></td>
<td>66</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>67</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

FIGURE 2- 1. PRE-ZYGOTIC ISOLATION: AGGLUTINATION INDEXES OF MATING PAIRS ........................................ 27
FIGURE 2- 2. POST-ZYGOTIC ISOLATION: DEFICIENCY SCREEN ............................................................................. 28

FIGURE 3- 1. PROFILES OF MEIOTIC DIVISIONS DURING SPORULATION .......................................................... 47
FIGURE 3- 2. PRINCIPAL COORDINATE ANALYSIS OF 352 DIFFERENTIALLY EXPRESSED GENES ........... 48
FIGURE 3- 3. TEMPORAL CHANGES IN MISEXRESSED GENES DURING SPORULATION ............................. 49
FIGURE 3- 4. GENES CLASSIFIED INTO DIFFERENT CATEGORIES OF CIS- OR TRANS-ACTING EXPRESSION DIFFERENCES ............................................................................................................. 50
LIST OF TABLES

TABLE 1- 1. D-M INCOMPATIBLE GENES THAT EFFECT HYBRID LETHALITY AND STERILITY BETWEEN SPECIES ........................................................................................................................................................................ 14

TABLE 2- 1. SPORE VIABILITY FOR ZIP2_{SACE}::KANMX4 HYBRIDS ................................................................................................................................. 29
TABLE 2- 2. STRAINS USED IN THIS CHAPTER .............................................................................................................................................. 30

TABLE 3- 1. NUMBER OF HYBRID GENES SIGNIFICANTLY DIFFERENT FROM BOTH PARENTS ......................... 51
TABLE 3- 2. GENES WITH OPPOSING AND COMPENSATORY cis-trans EXPRESSION DIVERGENCE........ 51
LIST OF SUPPLEMENTAL FIGURES

Table 3S-1. Strains used in this study .......................................................... 56
Table 3S-2. Classification of genes into different categories of expression divergence. 56
LIST OF SUPPLEMENTAL TABLES

Table 3S-1. Strains Used in This Study

Table 3S-2. Classification of Genes into Different Categories of Expression Divergence
Acknowledgements

My dissertation would not have been possible without the following professional, personal and financial support. My research has been funded by the National Institute of Health Training Grant and the Division of Biology & Biomedical Sciences (Molecular Genetics and Genomics Program).

I appreciate the opportunity to perform my thesis work in Justin Fay’s laboratory. I thank all former and current members of the lab for their help. Specifically I would like to acknowledge Justin Fay for his help with my statistical analyses; Linda Riles for her help with taming *S. paradoxus* and her friendship; Katie Hyma and Kara Powder for great scientific conversations, constructive criticism, continued support and friendship; and especially Betsy Engle. Not everyone is lucky enough to work (or travel well!) with someone who is not only a great scientist, but also one of their best friends.

I’d like to thank the members of my committee, especially Barak Cohen for his mentorship, scientific advice, and listening skills. I thank Jim Skeath and Tim Schedl for running the Genetics program – it’s obvious they care very much for the students; Ursula Goodenough for being a voice of reason; Steve Johnson and Jim Cheverud for scientific conversations and encouraging me to step out of my comfort zone.

I’ve had the fortune to be a part of the scientific community at the Center for Genome Sciences and Wash U. Thank you to Kim Lorenz, Drew Michaels, Priya Sudarsanam and Bin Wang for advice on experimental design, execution and analysis. Thank you to GTAC, especially Paul Cliften, Mark Johnson, and Nicole Rockweiler for help with sequencing. Thank you Shirley McKinney for saving some of my experiments with her super fast dish washing skills. I thank Sung Chun, Aaron Spivak, Ryan Christianson and Robert Ziegler for computational help. I also thank Scott Doniger, Jim Dover, Justin Gerke, Jay Gertz, and Josh Witten for introducing me into yeast genetics and being great distractions during my earlier years. I would also like to acknowledge YSP for always giving me a greater reason to stay committed to science.

I have a special thank you for my “team.” Thank you to Doctor Nancy Friesen, Dr. Karen Winters and Student Health; Dr. Rahul Rastogi, Dr. Beverly Field and BJC’s Pain Management; the Rehabilitation Institute and STEPP, especially Jean, Theresa and my support group; Wash U’s Physical Therapy, especially Sylvia Czuppon and Tracy Spitznaegle; and Marty Ritter, Patty Zorn and Amy Smoot. You have taught me to relive my life, and I honestly could not have finished my dissertation without your medical care.

I’d like to thank my friends and family for their constant support and love. Thank you to Betsy, Kara, Abby Buchwalter, Kim Delaney, Sarah Naylor, Geoff Ravillious, Erin Pickar and Han Tan for being wonderful colleagues and friends.

My siblings, Tooshar and Joy, are my best friends, and have always been the most supportive people in my life. I love them and their spouses, Christy and John, very much. I also have very supportive in-laws from my husband’s side of our family, Oma, Opa, Mutti, Papa and Norman. My dog-baby Dawn always snuggles with me when I’m down. But of course I love my niece Ellora more. I’d especially like to thank my parents for being my role models. I thank them for giving me what they could not have, and for showing me that no obstacles are too big. I am proud to be their daughter, and I love them with all my heart.

To Fabian,
What a Wonderful World
ABSTRACT OF THE DISSERTATION

Dysfunctional gene regulation of the meiotic program in an interspecific yeast hybrid

by

Devjanee Swain Lenz

Doctor of Philosophy in Molecular Genetics and Genomics

Washington University in St. Louis, 2013

Professor Justin Fay, Chair

Speciation, the evolutionary process by which species arise, is a fundamental biological concept. One of the major goals of evolutionary genetics is to understand the genetic basis of reproductive isolation (RI), a collection of barriers that prevents two species from forming viable or fertile hybrids. Given the lack of viable or fertile hybrid progeny, identifying genes that impact RI as well as their functions has been difficult. Thus we lack a view of the total genetic contribution to RI. Recently the budding yeast 
Saccharomyces
has served as an evolutionary model, as abundant sequence, expression, and phenotypic data exist for the model organism 
S. cerevisiae
and its closest relatives. Additionally we can manipulate the yeast genome and control its environment arguably more than any other organism. Hence I developed assays to catalog all genes contributing to RI between 
S. cerevisiae
and its closest known relative 
S. paradoxus,
which can form sterile hybrids under laboratory conditions. Chapter 2 details my utilization of accessible genetic tools for yeast to understand the total contribution of genes to RI. Though I unveiled multiple problems with studying speciation genetics using standard methods in yeast, I acquired valuable information about the biology of hybrids. For instance, I determine that yeast hybrids are highly sensitive to background mutations, commonly generated in yeast transgenesis, resulting in experimental artifacts. Using this knowledge, I took advantage of the emergence of next-generation sequencing in Chapter 3 to analyze wild type hybrid and parental genome expression to understand the relationship between gene expression and RI.

My main objectives in my dissertation are to understand dysfunctional hybrid gene regulation in the context of RI in yeast and to ascertain subsets of genes whose expression is disrupted. Thus I measured genome-wide changes in gene expression over the course of meiosis for 
S. cerevisiae,
S.
paradoxus and their sterile hybrid. I show that misexpressed genes in a yeast hybrid result from earlier activation of the meiotic program relative to its parents. This heterochrony is expected under the anti-recombination model of RI in yeast. I also find an increase in dysfunctional regulation in genes that are involved with sporulation, mitochondrial function, rRNA processing and translation. Genes in these pathways could contribute to RI. My dissertation adds to the field of speciation genetics, as it lends an example of a time-dependent relationship between dysfunctional hybrid regulation and RI for yeast species, as well as identifies candidate genes that could contribute to RI.
Chapter 1: Introduction
Speciation, the evolutionary process by which new species arise, is a fundamental biological concept. Few theories elegantly weave together so many areas of biology, including but not limited to genetics, population genetics, ecology, behavior, developmental biology, biochemistry, biogeography and biodiversity. As such the study of speciation touches most if not all aspects of biology. The process of speciation is also intimately linked to most biologists’ core interests: we seek to understand the diversity of the life we observe, be it variation within human traits that lead to disease or what makes humans different than the rest of primate family. Often, this diversity is due the same genetic, behavioral, or morphological factors that can influence speciation itself.

One of the major goals of evolutionary genetics is to understand the mechanistic basis of speciation. As defined by the biological species concept (BSC), species are “groups of actually or potentially interbreeding natural populations, which are reproductively isolated [RI] from other such groups.” RI is a collection of mechanistic barriers that prevents two distinct species from forming either viable or fertile hybrids. The amount of RI between two species can be quantitatively measured and thus RI is as a complex, quantitative trait. Therefore, it is possible to deduce the genetic bases and mechanisms that contribute to the phenotypes of RI and speciation.

Reproductive Isolation

RI barriers include both pre- and post-zygotic isolation. Pre-zygotic isolation consists of barriers that prevent two gametes from fusing and forming a zygote. These barriers include temporal and behavioral isolation, in which mating times or behavior between two species are not recognized by each other. Pre-zygotic isolation also includes mechanistic isolation, in which two species physically cannot mate with one another, for instance, if their genitalia are not compatible. If two species can mate, gametic isolation can occur, in which the gamete cannot fuse to form a zygote. However, if a zygote is formed between two species, post-zygotic barriers exist to prevent the hybrid zygote from either developing or producing its own offspring. Post-zygotic isolation includes hybrid inviability, in which the zygote cannot survive; hybrid sterility, in which a viable hybrid cannot produce offspring; and finally hybrid breakdown, in which first generation (F1) hybrids are fertile, but subsequent generations (F2, F3, etc.) of hybrids become less and less fit.
Post-zygotic isolation can be described as either extrinsic or intrinsic. Extrinsic post-zygotic isolation depends on environmental factors and results from hybrid phenotypes that perform poorly in a given environment\(^2\). Conversely, intrinsic post-zygotic isolation, for instance hybrid inviability and sterility, results from developmental problems in hybrids regardless of environmental influences\(^2\). These problems stem from divergence in parental genomes, such as chromosomal rearrangements or even single nucleotide mutations that are genetically incompatible in a single hybrid genome. Intrinsic isolation contributes to either viable or infertile hybrids, which are phenotypes that can be measured in a quantitative fashion, and thus genotypes underlying those phenotypes can be uncovered. The remainder of my introduction as well as dissertation research will focus on intrinsic factors contributing to RI.

**Chromosomal divergence and reproductive isolation**

Large-scale chromosomal mutations are known to cause post-zygotic RI due to improper synapsis during meiosis\(^3\). These include differences in chromosome number, in which hybrids contain a number of chromosomes that cannot produce four fertile gametes. This also includes gross chromosome inversions and translocations between parents that prevent proper synapsis, lead to improper chromosomal segregation, and ultimately result in gamete death.

**The genetic basis of post-zygotic reproductive isolation**

The Dobzhansky-Muller (D-M) model posits how genes contribute to intrinsic RI\(^3,4\). Genetic divergence between two species produces speciation genes that cannot properly function together in a single hybrid genome. Speciation genes, or preferably incompatible genes, always maintain reproductive isolation, but may or may not cause speciation. Under the simplest model, a gene from one species cannot interact with another gene from another species. However, RI is a collection of multiple barriers, and most likely multiple genes contribute to RI\(^2\).

Much of our understanding of the genetic contribution to RI comes from pre-zygotic isolation in which gametes cannot fuse. There is extensive information concerning the evolutionary forces that drive change in the genes that create a “lock and key” mechanism between sperm and egg, most notably in sea urchins\(^5\). Recently work has focused on understanding behavioral isolation in insects, and we now know genes that encode pheromones play a role in isolation\(^6-8\).
Over the past 30 years more attention has been given to understanding the genetic contribution to post-zygotic isolation. Abundant genomic evidence shows that genes contribute to post-zygotic RI\(^{9-16}\), and a handful of studies have performed the appropriate genetic studies to demonstrate that divergent genes contribute to RI. The genes listed in Table 1-1 list mapped genes (not genetic regions), which contribute to post-zygotic RI between species. Most work in hybrid incompatibilities has been performed in *Drosophila*, but speciation genetics has also been successful in finding genes that contribute to RI in fungi and mammals\(^{17-19}\). Traditional mapping by chromosome introgression and genetic markers was utilized during the 1980s to uncover the first incompatible gene, *OdsH*, which contributes to hybrid male sterility in *Drosophila*\(^{20,21}\). However the gene that is incompatible with *OdsH* is still unknown. In the 1990s, molecular tools for *Drosophila* became more advanced, and more efficient methods of finding genes that contribute to RI via screens utilizing P element and deficiency collections\(^{10}\).

Genetic regions that rescued hybrid phenotypes were mapped and in 2006, the first study uncovering both genes of a D-M incompatibility between species of *Drosophila* was published\(^{16}\). *Drosophila melanogaster* females and *D. simulans* male crosses produce only female hybrids. An incompatibility between an X-linked gene from *D. melanogaster*, Hybrid male rescue (*Hmr*)\(^{22,23}\) and an autosomal gene from *D. simulans*, Lethal hybrid rescue (*Lhr*)\(^{24}\), result in death of F1 hybrid males. However empirical evidence shows that more than just these two genes contribute to hybrid male inviability in *Drosophila*\(^{23}\).

**What types of genes contribute to RI?**

While the first incompatible genes were being isolated, scientists began hypothesizing as to what sort of genes are susceptible to becoming D-M incompatibilities, as the model itself makes no predictions as to what kind of genes contribute to RI. In particular, deleterious epistatic interactions that cause post-zygotic RI phenotypes are specific to the hybrid, and thus are not predictable from genotypes or phenotypes of parental species\(^{13}\). However, with genomic and genetic data, we are able to hypothesize about RI as a quantitative trait. I will discuss those hypotheses relevant to my dissertation, and dedicate a section to regulatory evolution and RI.

An intuitive hypothesis is that incompatible genes are under selection. Genes that are under selection adapt rapidly to their environment, and thus diverge from their homologs more quickly than the
rest of the genome. Most of the known incompatible genes are rapidly evolving or under genetic conflict (see Table 1-1)[5,20,25-29], including Lhr and Hmr[22]. However the list of incompatible genes is not comprehensive enough to fully know how often rapidly evolving genes contribute to RI in comparison to genes evolving under other evolutionary forces.

Recently, it has been hypothesized that rapidly evolving genes that are encoded by the mitochondrial genome (mtDNA) play a disproportionate role in RI[30]. First, mtDNA evolves more quickly than the nuclear genome (gDNA). Second, all proteins that are encoded by mtDNA must interact with the proteins that are encoded by gDNA. Finally, nuclear-encoded proteins are required to start processes of transcription and replication of mtDNA. Thus mtDNA and gDNA must interact, even though they evolve under separate, albeit not necessarily independent, conditions. Consequently an evolutionary environment is formed between mtDNA and gDNA that promotes compensatory mutations, which may be incompatible in a hybrid background. Additionally, all known incompatibilities in yeast involve mtDNA[17,18] and there is evidence that mtDNA contributes to RI in insects[27,31,32].

As with all quantitative traits, we would like to know how many genes and of what effect size contribute to RI. Estimates indicate almost 200 genes contribute to RI between D. melanogaster and D. simulans[33], although there are currently no estimates between other species. However, Orr has hypothesized that once RI has been established between two populations, more and more incompatibilities from either population can evolve independently, as they will not be under selection to maintain gene flow between populations[34]. Under this model, we would expect that RI is comprised of many genetic incompatibilities of small effect, and consistent with this expectation, the incompatible genes that have been identified show small, albeit measurable, effects on RI (e.g. they don’t completely cause RI).

**Regulatory evolution and reproductive isolation**

We would also like to know the contribution of coding and noncoding divergence to RI. Only two noncoding regions have been found to contribute to RI. In Drosophila, Zhr is not a coding region, rather a heterochromic repeat[35,36]. In D. melanogaster male and D. simulans female crosses, the paternal X-linked Zhr satellite block cannot segregate properly in F1 hybrid females causing female inviability.
Additionally, the 5' untranslated region of a mitochondrial gene, \textit{OL1f}, creates an incompatibility in yeast\textsuperscript{17}, which I will describe in detail later in this chapter.

Although only two incompatibilities seem to have noncoding attributes contributing to RI, the relationship between regulatory evolution and RI is still of particular interest. There is an abundance of gene expression divergence between species, and multiple groups have observed extensive hybrid misexpression, which is defined as hybrid expression levels outside the range of either parent’s expression\textsuperscript{26,37-45}. However, the link between gene misexpression and RI is unclear.

Genome expression profiles of interspecific hybrids have revealed suggestive associations between regulatory divergence and post-zygotic isolation. For instance \textit{Drosophila} hybrids disproportionately misregulates genes that are expressed primarily or exclusively in male flies\textsuperscript{46}. Similarly, misexpression of genes on sex chromosomes has been observed in both flies and mice\textsuperscript{43,47-49}. However, misexpression may be a simple consequence of dysgenic phenotypes in the hybrid, such as gonadal atrophy\textsuperscript{37,39}.

Allele-specific analyses show that interactions between divergent \textit{cis}- and/or \textit{trans}- acting factors between the parents can cause misexpression in the hybrid. Mutations that impact expression of its own gene are \textit{cis}-acting factors, while mutations from an outside element that impacts a different gene’s expression are \textit{trans}-acting factors. Both additive and non-additive \textit{cis-/trans}- interactions have been reported to cause novel expression in hybrids\textsuperscript{26,50-52}. For example, parents may have non-additive compensatory mutations in \textit{cis-/trans}- elements that cause gene expression to remain similar between the parents. However, in a hybrid background, these compensatory changes can interact and cause novel expression patterns. Compensatory interactions are of particular interest because they are hypothesized to contribute to RI\textsuperscript{38}.

While there is evidence that dysfunctional regulation in hybrids could influence RI, we do not have a general understanding of gene regulation and RI, let alone a genomic view. It is still unclear if disruptions in hybrid regulation are more commonly due to expression divergence between parental species or due to novel hybrid misregulation in genes for which parental species exhibit no expression divergence. It is also unknown if gene misexpression is a cause or consequence of RI. Despite two
incompatibilities that have been found to have noncoding effects, we still require a genomic view of misexpression and RI.

**Yeast as an evolutionary model**

Significant progress has been made in understanding the contribution of genes to RI in *Drosophila*; however, it has taken 30 years to uncover the seven genes listed in Table 1-1. As such we lack a complete picture of the genetic basis of RI. The budding yeast *Saccharomyces* could serve as a model to more quickly piece together the puzzle of the genetic basis of RI.

In the past decade *Saccharomyces* has emerged as an excellent evolutionary model, since we can manipulate the yeast genome and control its environment arguably more than any other organism. While advances in *Drosophila* genetics propelled the field of speciation genetics forward, simultaneous advances in yeast genetics were made but not applied to speciation genetics. For instance, we can produce yeast plasmid libraries that contain single genes, instead of large genomic regions like P element collections. Likewise, instead of the fly deletion collection that contains deletions of large genomic regions per strain, the yeast deletion collection contains a single gene deletion per strain. The tools for *Drosophila* genetics drove advancements of speciation genetics, even though the tools isolated genomic regions, rather than single genes. As the yeast collections focus on single genes, then using yeast as an evolutionary model could lead to quicker progression in understanding the relationship of single genes to RI.

An abundance of sequence, expression, and phenotypic data already exist for *S. cerevisiae* and its closest relatives. Detailed characterization of yeast expression pathways in various conditions provides us with the means to compare interspecific yeast hybrid and parental gene expression to ultimately deduce which pathways are disrupted in hybrids. Since yeasts are single cell organisms, hybrid misexpression in is not confounded by a variety cell-types or tissues. Furthermore, the progression of gene expression changes that occur over the course of meiosis has been well characterized for *S. cerevisiae*. These genetic and genomic advantages allow us to make interpretations about RI from yeast that might not be possible in other organisms.

**The yeast life cycle**
Saccharomyces can divide in either haploid or diploid form. When in haploid form, a cell is one of two mating types, a or α, which can only fuse with its opposite mating type to form a diploid a/α cell. Diploid cells cannot mate, and under nutrient rich conditions, yeast will preferentially ferment sugar and undergo mitosis. In the presence of nitrogen starvation, a non-fermentable carbon source, and abundant oxygen, diploid cells will sporulate, a specialized version of meiosis. During sporulation yeast form tetrads, which are composed of four haploid spores surrounded by an ascus. I explain details of the pathway below.

As sporulation is analogous to developmental pathways in higher eukaryotes, Saccharomyces has been utilized to test various genome-wide expression profiling techniques, including ribosome profiling. As a result, the regulatory pathway of sporulation is very well understood. Historically, sporulation has been divided into three phases: early, middle and late sporulation. Early sporulation spans pre-meiosis and part of meiosis I, when sister chromosomes are aligned, but have not yet crossed over. Synapsis marks the start of middle sporulation, which spans the rest of meiosis I as well as meiosis II. The actual formation of spore walls around the four nuclei occurs during late sporulation, which begins at the end of meiosis II and spans spore maturation and ascus formation.

Reproductive isolation in yeast

Species of Saccharomyces are both pre- and post-zygotically isolated, although much more is known about post-zygotic isolation. S. cerevisiae and its closest known relative S. paradoxus can fuse to form diploid hybrids in laboratory conditions, but preferentially mate within their own species when given the choice to mate with both species. When S. cerevisiae and S. paradoxus do mate, their interspecific hybrid can undergo mitosis. Under sporulation conditions, the hybrid can form spores, but only 1% of spores are viable, which is a large fitness deficit considering each parental species can have up to 100% spore viability.

Pre-zygotic isolation in yeast

Pre-zygotic isolation exists between S. cerevisiae and S. paradoxus, as both preferentially mate within its own species when given a choice. Mate preference is due to divergence in the timing of germination between S. cerevisiae and S. paradoxus, as experiment synchronization of germination time
leads to random rather than preferential mating. However, the cause of this temporal difference in germination timing is unknown. It is still unclear if pre-zygotic isolation is a form of reinforcement, meaning that pre-zygotic barriers are under selective pressure to build upon the reproductive barriers that are already in place. If true, we would expect sympatric species pairs that are in contact with one another to have more pre-zygotic barriers than allopatric species pairs that are not in contact with one another.

*Post-zygotic isolation in yeast*

Two known mechanisms are known to contribute to post-zygotic isolation between *S. cerevisiae* and *S. paradoxus*, which can form sterile hybrids. First, the anti-recombination pathway causes sterility of F1 hybrids due to abundance of sequence divergence between the homologous chromosomes.

Second, a special Dobzhansky-Muller incompatibility between nuclear and mitochondrial genomes contributes to post-zygotic isolation in F2 hybrids that are derived from rare viable F1 hybrids.

**The anti-recombination pathway and F1 hybrid sterility**

The anti-recombination model proposes that the mismatch repair pathway recognizes a multitude of mismatches between interspecific homologous chromosomes in the hybrid during metaphase of meiosis I and effectively prevents non-homologous chromosomes from crossing over, a phenomenon called anti-recombination. As one crossover per chromosome pair is typically required for proper segregation during meiosis I, this lack of homologous recombination in the hybrid results in improper chromosome segregation. Consistent with this model, the observed lack of homologous recombination in the hybrid results in improper chromosome segregation, and consequently excessive aneuploidy and inviable spores.

The first line of evidence supporting the anti-recombination model came in 1996. Normally, the mismatch repair proteins *PMS1* and *MSH2* recognize mismatches (due to sequence divergence) between hybrid chromosomes and prevent crossover. Deletion of the genes encoding these proteins allows increased recombination between *S. paradoxus* and *S. cerevisiae* chromosomes, seen by the decrease in aneuploidy in viable spores. This results in an increase in hybrid spore viability from less than 1% to 7.2% and 10.2% for deletion of *pms1* or *msh2* respectively.
Further indication that the anti-recombination model contributes to RI comes from the ability for tetraploid hybrids to produce viable spores. Greig *et al.*\textsuperscript{65,66} created *S. cerevisiae* hybrids with five other species in the *sensu stricto* group. When they doubled the hybrid genomes to form tetraploids, spore viability was restored. The authors suggest that this is due to each chromosome in the tetraploid hybrid has a homologous chromosome with which to rearrange, thus allowing proper chromosome disjunction.

**Genetic incompatibilities and F1 hybrid sterility**

Most of the work under the anti-recombination model has claimed that sequence divergence is the sole cause of F1 hybrid sterility. As of yet, no genetic incompatibilities are known to cause hybrid sterility in *Saccharomyces* F1 hybrids. However sequence divergence that causes anti-recombination does not need to be mutually exclusive from causing genetic incompatibilities. Thus there is a possibility that the sequence divergence that causes anti-recombination may also cause genetic incompatibilities, which has yet to be formally tested. Proponents of the anti-recombination model use multiple lines of evidence to argue against the role of genetic incompatibilities. However the arguments are inadequate to disprove that D-M incompatibilities contribute to RI, which I will now discuss.

First, proponents of the anti-recombination model cite that spore viability is a function of sequence divergence both between and within species. The more closely related species or strains are to one another, the more viable spores they are able to produce\textsuperscript{61}. Although this study supports the anti-recombination model, genetic incompatibilities cannot be eliminated from causing RI if they were not directly tested. Sequence divergence underlies the notion of genetic incompatibilities and the study does not differentiate between the two models.

Second, many believe the tetraploid hybrid experiment proves dominant incompatabilities do not exist in yeast\textsuperscript{65,66}. In a simple model of one to one allele dominance and one to one incompatibilities, dominant incompatibilities should be unaffected by ploidy. That is, like their tetraploid progeny that can produce viable spores, F1 hybrids would lack dominant genetic incompatibilities. However the viable spores’ chromosome numbers were not analyzed, and could be more resistant to incompatibilities that are not completely dominant (i.e. a diploid spore can lose a chromosome and still live). Thus a dominant incompatibility that causes either spore death or meiotic chromosome non-disjunction may still exist\textsuperscript{70}. 
Furthermore the experiment does not rule out the possibility that recessive incompatibilities may cause F1 spore inviability or may show affects in F2 diploid hybrids\textsuperscript{17}.

Third, interspecific chromosome replacement assays suggest recessive incompatibilities do not exist in yeast. Grieg\textsuperscript{71} independently replaced nine of the 16 chromosomes of S. \emph{cerevisiae} haploids with their S. \emph{paradoxus} homologues. All strains survived, showing that recessive genetic incompatibilities that cause haploid inviability are not present on these nine chromosomes. Grieg calculated that at most, only five recessive genetic incompatibilities could exist that cause spore inviability. However, Grieg only looked at haploid viability, and the major barrier separating S. \emph{cerevisiae} from S. \emph{paradoxus} seems to be hybrid sterility due to the known lack of recombination. This has yet to remain formally tested\textsuperscript{70}.

Although no incompatibilities have been found to affect F1 hybrids, experimental evolution research has shown that incompatibilities can arise in yeast by selection. If S. \emph{cerevisiae} is grown in either high salt or low glucose conditions for 500 generations, the fitness of their hybrids is decreased in either condition relative to the parental strains. This fitness, measured by growth cycle, can be attributed to specific mutations that consistently arise in the two selection conditions\textsuperscript{72}. Earlier studies by the same group showed that longer selection periods and genetic drift alone can cause complete hybrid sterility due to epistatic interactions, although they did not identify the genetic cause in these studies\textsuperscript{73}.

Additionally a group assessed the genomic composition of 58 rare viable spores from S. \emph{cerevisiae} and S. \emph{paradoxus} hybrids to determine whether D-M incompatibilities exist\textsuperscript{74}. If genetic incompatibilities contribute to the viability of spores after they were produced, there would be an underrepresentation of the incompatibility in the viable F1 spores. The authors find no simple pairs of genetic incompatibilities, but they do identify underrepresented combinations containing more than two loci. Although a lack of aneuploidy would more likely have a stronger effect on the viability of spores, the results imply that more complex genetic interactions could contribute to RI before the production of mature spores.

Genetic incompatibilities and F2 hybrid breakdown

Work surrounding the anti-recombination model supports alternative models to speciation, but do not rule out that genetic incompatibilities could contribute to RI in yeast. Thus Lee \textit{et al.}\textsuperscript{17} directly tested whether recessive incompatibilities could affect hybrid breakdown between S. \emph{cerevisiae} and S. \emph{bayanus}
by isolating 40 viable F2 spores from a homothallic F1 hybrid. The majority of diploid F2 hybrids could not sporulate, indicating recessive incompatibilities contribute to hybrid breakdown in yeast. They individually replaced $S. \text{cerevisiae}$ chromosomes with those from $S. \text{bayanus}$, and mapped the genetic incompatibility to the $S. \text{bayanus} \text{AEP2}$. In $S. \text{cerevisiae}$, AEP2 facilitates translation of $\text{OLI1}$, a subunit of the ATP synthase complex. However the $S. \text{bayanus}$ homolog of AEP2 cannot assist in the translation of the $S. \text{cerevisiae}$ homolog of $\text{OLI1}$. Thus the F2 hybrid ATP synthase complex cannot be built, and the cell cannot respire. Respiration is an essential requirement for sporulation, and thus F2 hybrids cannot sporulate.

The same group expanded their research to include $S. \text{paradoxus}$ in their studies and found incompatibilities causing hybrid breakdown in hybrids deriving from $S. \text{paradoxus}$ and either $S. \text{cerevisiae}$ or $S. \text{bayanus}$. As my thesis work focuses on RI between $S. \text{cerevisiae}$ and $S. \text{paradoxus}$, I will discuss the incompatibility identified between those two species. In short, $\text{MRS1}$ is a splicing protein, which specifically splices introns from $\text{COX1}$ transcripts. This gene encodes subunit I of cytochrome c oxidase, which is required for respiration. The $S. \text{cerevisiae}$ homolog of $\text{MRS1}$ cannot splice the $S. \text{paradoxus}$ homolog of $\text{COX1}$, which prevents respiration, and thus sporulation, in F2 hybrids. These elegant studies have given us clear instances of genetic incompatibilities in $\text{Saccharomyces}$. However, we are still lacking more than these few pet examples of the genes that regulate RI (Table 1-1).

**Focus of dissertation work**

My overarching interest is to understand how genetic divergence between species contributes to RI. The goal of my dissertation is to gain an understanding of the relationship between expression divergence, dysfunctional hybrid expression, and RI in yeast. Specifically, I would like to understand whether hybrid misexpression is a cause or consequence of RI. Additionally I would like to understand whether specific groups of genes are more likely to be misexpressed.

My dissertation is organized as follows. In Chapter 2, I use traditional genetic methods to catalog genetic incompatibilities between $S. \text{cerevisiae}$ and $S. \text{paradoxus}$. The purpose of this chapter is to provide the motivation for my move into next-generation sequencing. I also provide the protocols to verify candidate genes from my RNA-seq study. In Chapter 3, I perform an RNA-Seq study to uncover genes and pathways in which the hybrid regulates genes differently than its parents. In Chapter 4, I interpret my
results in context of what is known about speciation genetics as well as provide future avenues of research in speciation genetics of yeast. My study in hybrid gene regulation not only supports what is known about RI in yeast, but also provides new insights to speciation genetics.
### Table 1. D-M incompatible genes that effect hybrid lethality and sterility between species

<table>
<thead>
<tr>
<th>Loci</th>
<th>Species pair</th>
<th>Phenotype</th>
<th>Interacting loci</th>
<th>Incompatibility mechanism</th>
<th>Putative evolutionary basis</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEP2</td>
<td>S. cerevisiae/S. bayanus</td>
<td>F2 sterility</td>
<td>2; AEP2, OL1</td>
<td>S. bayanus gDNA gene Aep2 cannot aid with translation of S. cerevisiae mDNA gene OL1</td>
<td>Mutation pressure – high mitochondrial rate and/or adaptation to carbon source</td>
<td>17</td>
</tr>
<tr>
<td>OL1</td>
<td>S. cerevisiae/S. bayanus</td>
<td>F2 sterility</td>
<td>2; OL1</td>
<td>S. cerevisiae gDNA gene Prdm9 polymorphisms segregate in both species</td>
<td>Mutation pressure – high mitochondrial rate and/or adaptation to carbon source</td>
<td>18</td>
</tr>
<tr>
<td>MRS1</td>
<td>S. cerevisiae/S. bayanus</td>
<td>F2 lethality</td>
<td>2; MRS1, COX1</td>
<td>S. cerevisiae gDNA gene Mrs1 fails to splice ancestral mDNA gene Cox1</td>
<td>Mutation pressure – high mitochondrial rate and/or adaptation to carbon source</td>
<td>18</td>
</tr>
<tr>
<td>COX1</td>
<td>S. cerevisiae/S. bayanus</td>
<td>F2 lethality</td>
<td>2; OL1</td>
<td></td>
<td>Mutation pressure – high mitochondrial rate and/or adaptation to carbon source</td>
<td>18</td>
</tr>
<tr>
<td>AIM22</td>
<td>S. cerevisiae/S. bayanus</td>
<td>F2 lethality</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Mutation pressure – high mitochondrial rate and/or adaptation to carbon source</td>
<td>18</td>
</tr>
<tr>
<td>Lhr</td>
<td>D. melanogaster/D. simulans</td>
<td>F1 male lethality</td>
<td>3+; Lhr, Hmr, unknown</td>
<td>Undetermined gain of function, both heterochromatin associated proteins</td>
<td>Genetic conflict involving heterochromatin</td>
<td>16,24</td>
</tr>
<tr>
<td>Hmr</td>
<td>D. melanogaster/D. simulans</td>
<td>F1 male lethality; F1 female lethality and sterility</td>
<td>3+; Lhr, Hmr, unknown</td>
<td>Undetermined gain of function, both heterochromatin associated proteins</td>
<td>Genetic conflict involving heterochromatin</td>
<td>22-24</td>
</tr>
<tr>
<td>Zhr</td>
<td>D. simulans/D. melanogaster</td>
<td>F1 female lethality</td>
<td>Unknown</td>
<td>Misregulation of Zhr, satellite DNA unique to D. melanogaster</td>
<td>Genetic conflict involving heterochromatin</td>
<td>35,36</td>
</tr>
<tr>
<td>OdsH</td>
<td>D. mauritiana/D. simulans</td>
<td>F2 backcross male sterility</td>
<td>Unknown</td>
<td>D. mauritiana heterochromatin associate protein OdsH ectopically localizes to D. simulans Y chromosome</td>
<td>Genetic conflict involving heterochromatin</td>
<td>20,21</td>
</tr>
<tr>
<td>Nup96</td>
<td>D. melanogaster/D. simulans</td>
<td>F2 backcross male sterility</td>
<td>2+; Nup96, unknown</td>
<td>Unknown</td>
<td>Host/pathogen genetic conflict</td>
<td>28</td>
</tr>
<tr>
<td>Nup160</td>
<td>D. melanogaster/D. simulans</td>
<td>F2 backcross male sterility and female lethality</td>
<td>2+; Nup160, unknown</td>
<td>Unknown</td>
<td>Host/pathogen genetic conflict</td>
<td>21</td>
</tr>
<tr>
<td>Prdm9**</td>
<td>*Mus musculus musculus/ M. musculus domesticus</td>
<td>F1 male sterility</td>
<td>2+; Prdm9, unknown</td>
<td>Meiotic specific histone 3 lysine 4 trimethyltransferase; number of zinc fingers determines sterility</td>
<td>Genetic conflict involving heterochromatin</td>
<td>19,25</td>
</tr>
</tbody>
</table>

* M. musculus musculus and domesticus may be separate species ** Prdm9 polymorphisms segregate in both species, may not contribute to RI
Chapter 2: Methods to identify and verify incompatible genes
Abstract

Although speciation has been an active area of study since Darwin’s *On the Origin of Species*, we are only just beginning to understand the molecular basis of reproductive isolation (RI). Given that RI is characterized by complex reproductive barriers and the lack of viable or fertile hybrid progeny, identifying the genes that regulate RI and their functions has been difficult. Thus we are lacking a view of the total genetic contribution to RI. In the past decade, *Saccharomyces* has served as an excellent evolutionary model, since abundant sequence, expression, and phenotypic data exist for the model organism *S. cerevisiae* and its closest relatives. Furthermore we can manipulate and control yeast and its environment arguably more than any other model organism. To this end I attempted to develop assays to catalog all genes contributing to both pre- and post-zygotic RI for the budding yeast *Saccharomyces*. In this chapter I discuss assays that have provided insightful knowledge to studying RI in yeast. In Parts 1 and 2, I focus on my pre- and post-zygotic work respectively. In Part 3, I discuss my overall conclusions, and in Part 4, I describe my protocols.

Part 1 – Pre-zygotic isolation in yeast

Introduction

Pre-zygotic barriers exist between *S. cerevisiae* and its closest known relative, *S. paradoxus*. When given a choice to mate within or between species, both *S. cerevisiae* and *S. paradoxus* preferentially mate within its own species. Mate preference is predominantly due to divergence in the timing of germination between *S. cerevisiae* and *S. paradoxus*, and experimental synchronization of germination time leads to a breakdown of pre-zygotic barriers in these species. However, given that these mate choice experiments demand a great deal of time, pre-zygotic RI has only been observed between two allopatric pairs *S. cerevisiae* and *S. paradoxus*. As such we are lacking a complete view of how allopatric and sympatric species pairs discriminate between mating with one another, and furthermore we do not know if strains within a single species distinguish between one another during intraspecific mating. Thus I developed an assay, based on sexual agglutination, which could be utilized to screen through a large number mating pairs to deduce various patterns of differential mating preference between allopatric and sympatric species pairs, as well as between inter- and intra-specific mating pairs.
Sexual agglutination is a mating-specific form of flocculation. In response to pheromones secreted by the opposite mating type, haploid cells produce membrane proteins called agglutinins. Agglutinins are mating-type specific and necessary for mating in conditions that restrict cell-cell contact (e.g. mating in liquid conditions). The two types of agglutinins interact with one another and promote the efficiency of mating by forming conglomerates of cells, which provides cells with the appropriate distance for mating. In culture, agglutinated conglomerates of mating cells sink to the bottom of a culture after a given amount of time while free and non-mating cells remain suspended in media. Therefore I can utilize the percent of agglutinated cells in a culture as a proxy of the mating propensity of two strains. The assay is simple to conduct, and can be used to determine mating propensity for a large number of mating pairs.

**Results and Discussion**

Agglutination assays distinguish ability of strains to mate with one another

Studies in yeast have recently begun to focus on mechanisms of pre-zygotic isolation, and to this end, I developed an assay that measures sexual agglutination (see Materials and Methods) for both intra- and inter-specific mating. For each mating pair, I mixed a MATa and MATα strain in rich liquid media (0 hours) and allowed the samples to incubate for 6 hours, calculating the agglutination indexes at both timepoints (Figure 2-1). I utilized one *S. paradoxus* strain and three *S. cerevisiae* strains (a wine, oak, and laboratory strain). The *S. cerevisiae* strains that I used are allopatric with the *S. paradoxus* strain that I used. As all *S. cerevisiae* strains produce the same results, I will focus on the experiments using the laboratory strain.

Strains that mate with one another have a much higher agglutination index than strains that do not mate (Figure 2-1). At 0 hours, no mating pairs have an agglutination index higher than 0.20. At 6 hours of mating, both inter- and intra-specific mating pairs have an agglutination index of approximately 0.57, and the indexes are significantly different than their 0 hour control (t-test, P < 0.05). By comparison, negative control experiments with species of the same mating type (either both MATa or MATα) have low indexes at both 0 and 6 hours (t-test, P > 0.05). Thus the agglutination assay distinguishes between pairs of strains that do and do not mate.

Intra- and inter-specific mating pairs have comparable agglutination indexes at both 0 and 6 hours (t-test, P > 0.05). Thus agglutinins from the specific strains of *S. paradoxus* and *S. cerevisiae* used
(Table 2-1) interact with one another. As I determined no measurable differences between inter- and intra-specific mating pairs for this particular assay, I chose not focus my thesis work on pre-zygotic isolation between S. paradoxus and S. cerevisiae.

However my agglutination assay has further applications for understanding pre-zygotic isolation in yeast. It could be used to test heterochronic divergence in yeast mating and reinforcement between allopatric and sympatric species. If mating is heterochronic, agglutination indexes may be different between inter- and intra-specific mating pairs, throughout the course of mating. Thus taking more measurements before the six-hour time point may provide details on the speed of mating. Furthermore different species of Saccharomyces can be isolated from the same samples of soil, bark and fruit. Because sympatric species have more opportunities to mate with one another, I would expect sympatric species to have higher levels of pre-zygotic RI, and thus lower levels of agglutination between sympatric species than allopatric species. Similarly sympatric strains of the same species could have higher levels of discrimination than allopatric species, since agglutinins are rapidly evolving and could serve as a genetic barrier during insipient speciation.

Part 2 – Post-zygotic isolation in yeast

Introduction

Along with pre-zygotic isolation, two known post-zygotic mechanisms contribute to RI between S. cerevisiae and S. paradoxus, which can form viable but sterile hybrids. First, the anti-recombination pathway causes F1 hybrids to have 1% spore viability. In the rare case that viable, diploid F2 spores are derived from homothallic F1 hybrids (Chapter 1), a second post-zygotic mechanism enforces RI. Dobzhansky-Müller incompatibility between the nuclear and mitochondrial genomes contributes to hybrid breakdown in F2 hybrids, wherein the F1 generation still sporulates, while the F2 generation cannot. We know individually that anti-recombination contributes to sterility in F1 hybrids, and an incompatibility contributes to hybrid breakdown in F2 hybrids. However we do not know if incompatibilities contribute to sterility in F1 hybrids.

To isolate genes contributing to hybrid sterility in Saccharomyces, I developed two deficiency screens in which I examined the effect of single S. cerevisiae gene deletions in hybrid sterility. The first screen focuses on specifically testing rapidly evolving genes’ effect on RI, as many known genetic
incompatibilities are under selection\textsuperscript{16,19-22,28,35,77}. The second screen uses the original \textit{S. cerevisiae} deletion collection\textsuperscript{55} to unveil novel candidates for recessive incompatibilities.

\textit{Results and Discussion}

Screen to test the effect of rapidly evolving genes on hybrid spore viability

I studied the role of rapidly evolving genes in hybrid sterility, as some known genetic incompatibilities are rapidly evolving. Beginning with a list of rapidly evolving genes\textsuperscript{77}, I chose to focus on transcription factors (TFs), because disruption of a TF’s function could have severe consequences in a hybrid due to pleiotropic effects. I confirmed 14 TFs contain one or more 50 bp regions under higher selective pressure than the rest of the gene based on the ratio of non-synonymous to synonymous mutation rates (data not shown). I focused on five genes for further evaluation: \textit{ZIP2}, \textit{NUP53}, \textit{ACA1}, \textit{RME1}, and \textit{NGG1}.

Although we can hypothesize about what types of genes contribute to RI, we cannot necessarily predict how they affect RI. The simplest two-gene incompatibility model has four options of how genes can contribute to RI: a gene can either be recessive or dominant, and can derive from either parent. I initially tested one of these options, a possible dominant \textit{S. cerevisiae} incompatibility. To this end, I generated hybrids between \textit{S. cerevisiae} and \textit{S. paradoxus} in which \textit{S. cerevisiae} contained a gene deletion for the one of five rapidly evolving TFs identified above. I then determined if hybrids had increased spore viability, indicating the deleted gene acts as a dominant \textit{S. cerevisiae} speciation gene.

I utilized two background strains for each parental species to generate the hybrids (Table 2-1, Materials and Methods). I used a laboratory strain for each species, which I will refer to as either \textit{S. cerevisiae} or \textit{S. paradoxus} laboratory strain. I also used two wild isolates for each species, which I will refer to as either \textit{S. cerevisiae} or \textit{S. paradoxus} wild strain. As a representative sample of the five TFs studied, Table 2-2 depicts the measurements for hybrids derived from the \textit{S. paradoxus} laboratory strain and either \textit{S. cerevisiae} wild or laboratory strains containing a \textit{zip2} deletion (i.e., \textit{zip2\textsubscript{Scer}:KANMX4}).

The first problem I encountered results from the methods utilized in creating the deletion strains, revealed in crosses between the \textit{S. paradoxus} laboratory \textit{MATα} and \textit{zip2::KANMX4} \textit{S. cerevisiae} laboratory \textit{MATα} strains. Biological replicate hybrids containing \textit{zip2\textsubscript{Scer}:KANMX4} have high variability for spore viability (Table 2-2), most likely due to the non-clonal nature of the background mutations in \textit{S.}
cerevisiae deletion collection. Two of the three replicates have very high spore viability, while the third replicate still has low spore viability. To create the deletion collection, one antibiotic resistant and three auxotrophic markers were generated via transformation to ensure proper strain identification. Each of these transformations can create up to three or four background mutations, which could impact hybrid viability multiple ways. The collection is non-clonal in nature; each individual clone within the strain has multiple, variable background mutations, increasing variability in biological replicates (Table 2-2). Furthermore, accumulation of excess mutations could be beneficial for the hybrid. Additionally, QTL mapping for stochastic variation reveals auxotrophic marker ura3 contributes to noisy phenotypes. Any of these mutations could result in a non-specific increase in hybrid spore viability, verified by the experiments detailed below. To avoid an accumulation of mutations from the transformation process and ensure effects are specific to the deleted gene of interest, I regenerated the deletion in an S. cerevisiae wild strain. I backcrossed transformants to their parents, and isolated haploid derivatives from the backcross. The resulting clones produce replicable results as detailed below. However problems still existed due to the laboratory S. paradoxus strain.

The second problem is exemplified by crosses between S. cerevisiae wild and S. paradoxus laboratory strains, in which I see consistency across biological replicates, but a stark contrast between reciprocal crosses (Table 2-2). I only observe an increase in spore viability in crosses with S. paradoxus laboratory strain MATα, not MATa. Furthermore, of the spores that are viable in the high viability crosses, 16.9% are MATa, 55.9% are MATα and 27.2% are diploid. I expect a high number of diploids, as theory suggests viable spores come from autotetraploids (Gennadi Naumov, personal correspondence). However, I would still expect an equal ratio of viable MATa to MATα spores. MAT is located on chromosome III, and sequencing of one hybrid spore revealed a fully diploid spore except for chromosome III, which was S. cerevisiae only (Justin Fay, personal correspondence).

To test the effects of S. paradoxus laboratory strains on spore viability, I utilized haploids without auxotrophic derived from an S. paradoxus wild strain. For the reciprocal crosses between S. cerevisiae and S. paradoxus wild strains, zip2Scer::KANMX4 has no effect on spore viability (data not shown). Additionally reciprocal crosses between the S. cerevisiae laboratory strain and S. paradoxus wild strain, zip2Scer::KANMX4 also has no effect on spore viability (data not shown). The crosses reveal the
importance of utilizing strains with as few auxotrophic markers and point mutations as possible when studying yeast hybrids.

The third problem is the low number of viable spores produced from hybrids. For a significant increase in spore viability due to a dominant incompatible gene (Chi-square, P <0.05), I must dissect 400 tetrads per strain. However a deletion could uncover a recessive incompatible gene, which would decrease spore viability. With the spore viability already being so low, it is inefficient to dissect any more tetrads to identify significant lower spore viability.

The final, and most important, problem to test a specific set of genes is the lack of general knowledge about genetic contribution to RI. First how many genes, if any, contribute to RI and their effect size in yeast is unknown. I suspect that if genes contributed to spore viability in yeast, it would be many genes of small effects based on the hybrids generated from laboratory strains. However spore viability assays are not sensitive enough to measure small effects. Spore viability and gene substitution assays could be useful, if we had actual gene candidates to test, rather than exploring computationally derived sets of genes. For that reason, I developed deficiency screens to identify gene candidates that contribute to RI.

**Deficiency screen to detect S. paradoxus recessive incompatibilities**

To identify novel candidate genes that contribute to RI, I attempted to create multiple genome-wide screens for incompatibilities using the *S. cerevisiae* deletion collection. Although the deletion collection can give variable results, it can still be used to identify candidates. These candidates can then be further analyzed in spore viability assays.

I began with two methods based on standard protocols for random spore analysis. Both methods were highly variable, and therefore not pursued further. Thus, I developed a deficiency screen (Figure 2-2, Materials and Methods), in which I mate the original *S. cerevisiae* deletion strain collection to *S. paradoxus* to uncover recessive *S. paradoxus* gene that cause hybrid sterility and RI.

The *S. cerevisiae* MATα deletion collection was mated to either strains *S. paradoxus* laboratory strain MATα (interspecific) or, as a control, *S. cerevisiae* wild strain MATα (intraspecific, Figure 2-2). This produced F1 inter- and intra-specific hybrids, which were sporulated to produce F2 haploids. F2 haploids were mated to an *S. cerevisiae* MATα tester strain to form F3 diploids. These F3 diploids form
only if three conditions are met: (1) the *S. cerevisiae* gene deletion is complemented by the inter- or intra-specific allele in the F1 hybrid, (2) the F1 hybrid can sporulate, and (3) its F2 progeny can mate back to the *S. cerevisiae* tester strain. The majority of F3 diploids form colonies at the end of the screen. However, candidate genes for recessive *S. paradoxus* incompatibilities may interfere with any of these conditions. This would be evidenced as F3 interspecific diploids that do not grow, though F3 intraspecific diploid controls can grow. In other words, a single copy of the *S. cerevisiae* gene is sufficient for F3 diploids to survive, but a single copy of the *S. paradoxus* gene is not. Based on F3 diploid survival frequencies of the controls (data not shown), I suspected that I would identify 20 – 50 genes with about 2 – 10 false positives.

I identified 114 candidate *S. paradoxus* recessive incompatibilities. However, over half of these candidates are on chromosome III (see below). I identified 47 candidates (Supplemental data file) excluding genes on chromosome III. The genes are involved in chemical sensing (i.e. mating), cell cycle regulation, mitochondrial organization, sporulation, DNA repair, and chromosomal organization. Disruption in any of these processes would disrupt the ability of colonies to form at the end of the experiment. Further, mitochondrial function and DNA repair have previously been linked to RI in yeast\(^{70}\).

Most candidates of this screen are located on chromosome III, the site of the MAT loci. In my experiments analyzing rapidly evolving genes in hybrid spore viability (see above), I determined that the *S. paradoxus* laboratory strain MAT\(\alpha\) likely has a mutation on chromosome III that restores spore viability in the presence of one or more mutations from *S. cerevisiae*. Complementing the above screen with a reciprocal screen starting with *S. cerevisiae* MAT\(\alpha\) deletion collection and *S. paradoxus* MAT\(\alpha\) strains may help further refine *S. paradoxus* recessive incompatibility candidate genes. Further, with the cost of sequencing dramatically decreasing, we could sequence the *S. cerevisiae* and *S. paradoxus* strains that increase spore viability to compare to strains that do not increase spore viability, and deduce mutation(s) that cause RI.

**Part 3 – Conclusions**

My initial thesis work provides valuable insight into speciation genetics in yeast. First, I developed a straightforward agglutination assay that can be utilized to elucidate the heterochronic divergence that isolates yeast species across a large number of yeast strains. Second, my work shows
the importance of having minimal mutations while studying reproductive isolation in yeast. Lastly, some of my screens could be taken forward with the utilization of next-generation sequencing.

I must emphasize the importance of backcrossing transformants to their original parent to minimize the number of background mutations. When yeast geneticists generate mutations, we typically use only biological replicate, which reproduce our phenotype of interest. We discard strains that do not replicate our phenotype, and assume second point mutations are responsible for the lack of phenotype. While this approach has worked for a very long time in yeast genetics, I find that interspecific hybrids are extremely sensitive to second point mutations as seen in the variability of spore viability (Table 2-2). Previous work with pre-zygotic isolation utilizing either wild or genetically manipulated strains has also made the same observation\textsuperscript{62,63,70}. The strains I generated have been backcrossed after every transformation for up to a total of at least two backcrosses, and produce replicable results. That being stated, the spore viability assay may not be sensitive enough to detect an increase in spore viability, and rather a complementation assay, where a species’ gene is replaced with another species’ homolog, may be more efficient at testing a gene’s effect on spore viability.

With the advent of next generation sequencing, I believe we are in a position to more quickly understand the effect of genes on RI. As the cost of sequencing has dramatically dropped, we can compare the genomes of the hybrids that do have high spore viability to those that do not. As these are nearly isogenic strains, there could be up to 30 point mutations (4 point mutations per transformation) that contribute to RI. This approach has been used in yeast experimental evolution studies to deduce mutations that evolved in selective conditions\textsuperscript{72}. We can also use traditional sequencing methods for chromosome III to determine if there is a gross chromosomal abnormality in our original \textit{S. paradoxus} laboratory strain. This would still be of interest, as chromosomal rearrangements are not though to play a role in RI between \textit{S. cerevisiae} and \textit{S. paradoxus}\textsuperscript{61}.

Additionally next generation sequencing allows us to perform comparative genetic studies at a rapid pace, which could isolate pathways or genes that contribute to RI\textsuperscript{80-82}. RNA-Seq makes it possible to measure and to compare gene expression levels across practically any combination of species. As \textit{S. cerevisiae} is one of the most well studied organisms, we have the ability to identify disruption of expression patterns in the hybrid compare to the parents, and trace back the disruption to its origins,
which could be related to RI. Additionally as the genome of *Saccharomyces* is relatively small compared to the number of reads available on a sequencing lane, rather than limiting our observations to one observation, we can observe disruption of hybrid expression across multiple conditions, or more specifically, the different developmental stages of meiosis.

To verify a gene with disrupted expression is an incompatible gene that contributes to RI, we could perform the following analyses. We could perform the gene deletion assays that are described above. Additionally we can generate diploid hybrids containing the gene of interest with two alleles from the same species and compare its spore viability to wild type hybrids that are heterozygous for the gene of interest. If the gene were incompatible, we would expect that a hybrid that is homozygous for the gene would not have high spore viability. Additionally the hybrid containing the homozygous homolog of the gene of interest would increase spore viability. We can also test for the decrease of spore viability in the parental backgrounds, by replacing the gene of interest with its homolog from the other species or using a plasmid containing the candidate gene. If the gene were incompatible, we would expect that in the opposite parental environment, spore viability would decrease. If it were not incompatible, we would not see an affect. These assays must utilize genetic manipulations. Thus they must be generated carefully with backcrosses. These genetic assays could be utilized to verify candidate genes isolated from RNA-Seq. Chapter 3 describes my approach to identify incompatible genes using RNA-Seq.

Part 4 – Materials and Methods

*Strains*

I listed strains that I used in this chapter in Table 2-1. I used lithium acetate transformations for gene manipulations. After each transformation, I backcrossed strains to their parent, and isolate haploid derivatives. For hybrid crosses, I mix haploid cells of opposite mating types on solid rich media plates (YPD –1% yeast extract, 2% peptone, 2% dextrose), and allow them to incubate at 30°C for 24 hours. I isolate for diploids using the appropriate selective media.

For my spore viability assays, I list background strains without zip2Scer::KANMX4 in Table 2-1. All zip2Scer::KANMX4 are listed in Table 2-2. I used strains from the *S. cerevisiae* MATa deletion collection, which was made in the FM391 background (equivalent of BY4791) and I will refer to as *S. cerevisiae* laboratory strain. I also generated deletions of the TF of interest in an oak *S. cerevisiae* strain, YPS163,
which I will refer to as \textit{S. cerevisiae} wild strain. I used \textit{S. paradoxus} laboratory strains YPS16 (MAT\(\alpha\)) and YPS17 (MAT\(a\)), which I will refer to as \textit{S. paradoxus} laboratory strains. I also generated haploid derivatives, with the help of Linda Riles, of \textit{S. paradoxus} strain N17, which I will refer to \textit{S. paradoxus} wild strain. I generated five interspecific crosses: one cross between \textit{S. cerevisiae} laboratory strain MAT\(\alpha\) and \textit{S. paradoxus} laboratory strain MAT\(\alpha\); reciprocal crosses of \textit{S. cerevisiae} oak strain and \textit{S. paradoxus} laboratory strains; and reciprocal crosses of \textit{S. cerevisiae} oak strain and \textit{S. paradoxus} N17. For each cross, I measured spore viability of at least three biological replicates.

\textit{Measurement of Agglutination Index}

For each mating pair, I prepared three 10 ml liquid YPD cultures in test tubes. The first two cultures contained single haploid strains (MAT\(a\) and MAT\(\alpha\)) diluted to \(10^6\) cells/ml, and the last culture contained an equal concentration of the two haploid strains (mixed) for a total concentration of \(2 \times 10^6\) cells/ml. Cultures were incubated at 30\(\degree\)C at 250 rpm. At each time point, I thoroughly mixed 1 ml of culture in a disposable cuvette and measured the OD\(_{600}\). I allowed the cuvette to sit undisturbed for 20 minutes, after which I measured the OD\(_{600}\) again. Flocculated cells sink after 20 minutes, while free cells remain suspended in culture. The expected OD\(_{600}\) (OD\(_E\)) for the mixed culture is the summation OD\(_{600}\) of the haploid cells before 20 minutes of sitting:

\[(1) \text{OD}_E = \text{OD}_{\text{MAT}a} + \text{OD}_{\text{MAT}\alpha}\]

The observed OD\(_{600}\) (OD\(_O\)) of the mixed culture is the different between OD\(_{600}\) before and after 20 minutes of rest:

\[(2) \text{OD}_O = \text{OD}_{t0} + \text{OD}_{t20}\]

I calculated the agglutination index (A.I.) as the difference between OD\(_E\) and OD\(_O\) and normalized the difference OD\(_E\):

\[(3) \text{A.I.} = (\text{OD}_E - \text{OD}_O) / \text{OD}_E\]

I measured four replicates for each mating pair, and used a t-test to test for significant agglutination between time points 0 and 6 hours.

\textit{Recessive screen media and growth}
The following description correlates to Figure 2-2. All replica printing was preformed using Singer RoToR, with 384 well plates and all incubations were for two days at 30°C. To generate F1 interspecific hybrids, I spread a lawn of *S. paradoxus* laboratory or *S. cerevisiae* oak strain MATα onto solid YPD. After the plates dried, I printed the *S. cerevisiae* MATa deletion collection on top of the lawn of MATα cells, and incubated plates to allow the strains to mate. To isolate hybrids, I printed the strains onto solid minimal media (ingredients) + G418 (200ug/ml), and incubated plates. To generate F2 progeny, I replica-printed strains onto solid sporulation media (ingredients) and incubated plates. To generate F3 backcross, I spread a lawn of *S. cerevisiae* tester strain BC187 MATα onto solid YPD. I printed sporulated strains on top of the tester strain lawn, and incubated to allow the strains to mate. Spores are mated only to MATa tester strain, because the majority of viable hybrid spores for this particular interspecific cross are either diploid or MATα. To isolate F3 hybrids, I replica printed the strains onto YPD + G418 (200 ug/ml) + nourseothricin-sulfate (100 ug/ml), and incubated plates.

*Tetrad dissections*

I inoculated liquid YPD with a single hybrid colony, and grew hybrids overnight in a shaking incubator at 30°C. I pipetted 10 µL onto solid sporulation media (ingredients), and incubated plates for two days at 30°C. After which, I scraped a pipette tip of cells into sterilized water. I collected the suspended cells by centrifugation, and wash cells once more with sterilized water. I suspended cells in 100 µL of lyticase (10 mg/ml), and allowed the lyticase to digest the tetrad casings of sporulated cells for 20 minutes at room temperature. I randomly dissected approximately 400 tetrads per sample. I allowed spores from the tetrad dissections to incubate at 30°C for two days. I calculated spore viability as the number of living spores divided by the total number of dissected scores times 100. Fisher’s Exact Test was used to determine the significance of spore viability.
Figures and Tables

Agglutination Indexes of Mating Pairs

![Graph showing agglutination indexes for various mating pairs.]

**Figure 2-1. Pre-zygotic isolation: Agglutination indexes of mating pairs.** Agglutination indexes (y-axis) are shown for various inter- and intraspecific mating pairs at 0 and six hours. Each mating pair is listed as two strains separated by a slash. "C" stands for *S. cerevisiae* and "P" stands for *S. paradoxus*. Each species initial is followed by its mating type "a" or "α."
Figure 2-2. **Post-zygotic isolation: Deficiency screen.** (A) The MATα deletion collection is mated to (B) MATα *S. paradoxus* or *S. paradoxus*, (C) and hybrids are isolated. (D) Strains are then allowed to sporulate. (E) The sporulated strains are mated to MATα BC187 and (F) selected on YPD+G418+NAT. This media selects for spores from hybrids that retained the KAN deletion and have mated to BC187. Strains that do not grow up are candidates as indicated by the red boxes.
**Table 2-1. Spore viability for zip2_{S.cerevisiae::KANMX4} hybrids**

<table>
<thead>
<tr>
<th>S. cerevisiae deletion</th>
<th>S. cerevisiae parent</th>
<th>S. paradoxus parent</th>
<th>Sample size</th>
<th>Viable spores</th>
<th>% spore viability</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>Lab a</td>
<td>Lab α</td>
<td>412</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>wild type</td>
<td>Lab a</td>
<td>Lab α</td>
<td>440</td>
<td>8</td>
<td>1.8</td>
</tr>
<tr>
<td>wild type</td>
<td>Lab α</td>
<td>N17 α</td>
<td>380</td>
<td>3</td>
<td>0.8</td>
</tr>
<tr>
<td>wild type</td>
<td>Lab α</td>
<td>N17 a</td>
<td>380</td>
<td>6</td>
<td>1.6</td>
</tr>
<tr>
<td>zip2 – 1</td>
<td>Lab a</td>
<td>Lab α</td>
<td>400</td>
<td>40</td>
<td>10.0*</td>
</tr>
<tr>
<td>zip2 – 2</td>
<td>Lab a</td>
<td>Lab α</td>
<td>400</td>
<td>137</td>
<td>34.3*</td>
</tr>
<tr>
<td>zip2 – 3</td>
<td>Lab a</td>
<td>Lab α</td>
<td>952</td>
<td>18</td>
<td>1.9</td>
</tr>
<tr>
<td>zip2 – 1</td>
<td>Oak a</td>
<td>Lab α</td>
<td>232</td>
<td>78</td>
<td>33.6*</td>
</tr>
<tr>
<td>zip2 – 2</td>
<td>Oak a</td>
<td>Lab α</td>
<td>400</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>zip2 – 3</td>
<td>Oak a</td>
<td>Lab α</td>
<td>136</td>
<td>49</td>
<td>36.0*</td>
</tr>
<tr>
<td>zip2 – 4</td>
<td>Oak a</td>
<td>Lab α</td>
<td>148</td>
<td>37</td>
<td>25.0*</td>
</tr>
<tr>
<td>zip2 – 1</td>
<td>Oak α</td>
<td>Lab a</td>
<td>200</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>zip2 – 2</td>
<td>Oak α</td>
<td>Lab a</td>
<td>148</td>
<td>5</td>
<td>3.4</td>
</tr>
<tr>
<td>zip2 – 3</td>
<td>Oak α</td>
<td>Lab a</td>
<td>88</td>
<td>1</td>
<td>1.1</td>
</tr>
</tbody>
</table>

* deletion compared to wild type, P < 0.01, Fisher’s Exact Test
### Table 2-2. Strains used in this chapter

<table>
<thead>
<tr>
<th>Strain</th>
<th>Species</th>
<th>Parent(s)</th>
<th>Chapter reference</th>
<th>Genotype</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM391</td>
<td><em>S. cerevisiae</em></td>
<td>FM393</td>
<td>Laboratory</td>
<td>MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</td>
<td>Agglutination</td>
</tr>
<tr>
<td>FM392</td>
<td><em>S. cerevisiae</em></td>
<td>FM393</td>
<td>Laboratory</td>
<td>MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</td>
<td>Agglutination</td>
</tr>
<tr>
<td>YPS17</td>
<td><em>S. paradoxus</em></td>
<td>Laboratory</td>
<td>MATα leu2</td>
<td></td>
<td>Agglutination</td>
</tr>
<tr>
<td>YPS16</td>
<td><em>S. paradoxus</em></td>
<td>Laboratory</td>
<td>MATα lys2</td>
<td></td>
<td>Agglutination, Deficiency Screen</td>
</tr>
<tr>
<td></td>
<td><em>S. cerevisiae</em></td>
<td>YPS163</td>
<td>Wild</td>
<td>MATα trp1::HPMX4</td>
<td>Deficiency Screen</td>
</tr>
<tr>
<td>YJF153</td>
<td><em>S. cerevisiae</em></td>
<td>BC187</td>
<td>Tester</td>
<td>MATα ho::NATMX4</td>
<td>Deficiency Screen</td>
</tr>
<tr>
<td>YJF154</td>
<td><em>S. cerevisiae</em></td>
<td>YPS163</td>
<td>Wild</td>
<td>MATα ho::dsdAMX4</td>
<td>Spore viability</td>
</tr>
<tr>
<td>YJF77</td>
<td><em>S. paradoxus</em></td>
<td>N17</td>
<td>Wild</td>
<td>MATα ho::NATMX4</td>
<td>Spore viability</td>
</tr>
<tr>
<td>YJF78</td>
<td><em>S. paradoxus</em></td>
<td>N17</td>
<td>Wild</td>
<td>MATα ho::NATMX4</td>
<td>Spore viability</td>
</tr>
<tr>
<td>Hybrid</td>
<td></td>
<td>FM391</td>
<td></td>
<td>MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</td>
<td>Spore viability</td>
</tr>
<tr>
<td>Hybrid</td>
<td></td>
<td>YPS166</td>
<td></td>
<td>MATα lys2</td>
<td></td>
</tr>
<tr>
<td>Hybrid</td>
<td></td>
<td>YJF154</td>
<td></td>
<td>MATα ho::dsdAMX4</td>
<td>Spore viability</td>
</tr>
<tr>
<td>YJF843</td>
<td>Hybrid</td>
<td>YJF153</td>
<td></td>
<td>MATα ho::dsdAMX4</td>
<td>Spore viability</td>
</tr>
<tr>
<td>YJF842</td>
<td>Hybrid</td>
<td>YJF842</td>
<td></td>
<td>MATα ho::dsdAMX4</td>
<td>Spore viability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>YJF874</td>
<td></td>
<td>MATα ho::NATMX4</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3: Heterochronic meiotic misexpression in an interspecific yeast hybrid

This chapter was done in collaboration with Linda Riles and Justin C. Fay, and is an adaptation from a manuscript submitted May 2013.
Abstract

Changes in gene regulation rapidly accumulate between species and may contribute to reproductive isolation through misexpression of genes in interspecific hybrids. Hybrid misexpression, defined by expression levels outside the range of both parental species, is thought to be a result of cis- and trans-acting regulatory changes that interact in the hybrid, or arise from changes in the relative abundance of various tissues or cell types due to defects in developmental. Here, we show that misexpressed genes in a sterile interspecific Saccharomyces yeast hybrid result from a heterochronic shift in the timing of the normal meiotic gene expression program. By tracking nuclear divisions, we find that S. cerevisiae initiates meiosis earlier than its closest known relative, S. paradoxus, yet both species complete meiosis at the same time. During meiosis, the hybrid up- and down-regulates genes in a similar manner to both parents. However, the hybrid program occurs earlier than both parents, which results in a heterochronic pattern of misexpression throughout meiosis I and the beginning of meiosis II. Coincident with the timing of misexpression, we find an increase in the relative abundance of opposing cis and trans-acting changes and compensatory changes, as well as a transition from predominantly trans-acting to cis-acting expression divergence over the course of meiosis. However, misexpression does not appear to be a direct consequence of cis- and trans-acting regulatory divergence. Our results demonstrate that hybrid misexpression in yeast results from a heterochronic shift in the meiotic gene expression program.

Introduction

A major goal of evolutionary genetics is to understand the molecular basis of reproductive isolation (RI) between species. The relationship between regulatory evolution and RI is of particular interest, as there is an abundance of gene expression divergence between species. Additionally, extensive misexpression, defined as hybrid expression levels outside the range of either parent’s expression, has been observed in a variety of interspecific hybrids. Direct evidence that RI can arise from changes in gene regulation has been found for some genes. However, the role of most misexpressed genes in RI is not known.

Genome expression profiles of interspecific hybrids have revealed a number of suggestive associations between regulatory divergence and post-zygotic isolation. For instance, disruption of gene expression in both sterile flies and mice has been observed. Similarly, Drosophila hybrids
disproportionately misexpress genes that are mainly or solely expressed in males\(^{39,46}\). In the context of hybrid male sterility, many misexpressed genes are related to spermatogenesis, although misexpression also occurs in other tissues\(^{83}\). However, misexpression can also arise as a consequence of dysgenic phenotypes in the hybrid, such as gonadal atrophy, and so may be a consequence rather than a cause of hybrid inviability or sterility\(^{37,39}\).

Two questions regarding misexpression naturally arise: when does misexpression occur during hybrid development, and what is the cause of misexpression? In *Drosophila*, misexpression was found to be more pronounced in adult compared to larval stages of development\(^{41}\). Similarly, misexpression of a small number of genes in fish was found to increase during development\(^{84}\). However, another *Drosophila* study found more hybrid misexpression in larval and adult stages relative to the pupal stage\(^{85}\), providing evidence against a cascading model of misexpression during development.

While misexpression may be caused by inappropriate hybrid development, studies of allele-specific expression indicate that misexpression is a consequence of divergence in gene regulation. Both additive and non-additive *cis-trans* interactions have been reported to cause novel expression in hybrids\(^{26,38,50,51,86}\). Further, compensatory changes in gene regulation have been associated with hybrid misexpression\(^{38,51}\).

In this study, we examine gene expression during sporulation of the budding yeasts *S. cerevisiae*, *S. paradoxus* and their sterile hybrid. Since yeasts are single cell organisms, hybrid misexpression is not confounded with the relative abundance of cell types or tissues. Furthermore, the progression of gene expression changes that occur over the course of meiosis has been well characterized for *S. cerevisiae*\(^{56,57,59}\). We find that hybrid gene expression, including genes expressed outside the range of both parents, is predominantly related to a heterochronic shift, whereby the meiotic gene expression program proceeds more rapidly than either parent during certain meiotic stages. Coincident with this altered meiotic program, we find that expression divergence between the parental species is initially dominated by *trans*-acting changes and later dominated by *cis*-acting changes, between which we find more *cis-trans* opposing and compensatory changes. We discuss these results in relationship to current models of RI.
Results

Hybrid sporulation is similar to S. paradoxus

To characterize sporulation of S. cerevisiae, S. paradoxus and their hybrid, we monitored the cells’ progression through meiotic stages by DAPI staining of nuclei and fluorescent microscopy (see Methods). The number of nuclei in a cell indicates how many phases of meiosis the cell has completed. Tetra-nucleated cells have completed both meiosis I and II; bi-nucleated cells have only completed meiosis I; and mono-nucleated cells are undifferentiated diploids.

We find that the hybrid progresses through meiosis I and II similarly to S. paradoxus but differently than S. cerevisiae (Figures 3-1 and 3-S1). The parental mono-nucleate curves are significantly different than one another (ANOVA, $P = 0.01$), which indicates that the parents enter meiosis I at different times. However the hybrid mono-nucleate curve is not different than either S. cerevisiae or S. paradoxus (ANOVA, $P = 0.06$ and $P = 0.38$, respectively), which indicates the hybrid initially enters meiosis I intermediate of both parents. Compared to S. cerevisiae, S. paradoxus takes longer to complete both meiosis I and II, as measured by the formation of bi-nucleates and tetra-nucleates, respectively. Bi-nucleate formation is similar between the three strains, with only the hybrid being significantly different from S. cerevisiae (ANOVA, $P = 0.04$). While S. paradoxus and the hybrid generate tetra-nucleates similarly (ANOVA, $P = 0.54$), S. cerevisiae forms tetra-nucleates differently than both the S. paradoxus and the hybrid (ANOVA, $P = 0.003$ and $P = 0.009$, respectively). Thus the hybrid finishes meiosis I and II like S. paradoxus. Based on the observed progression through meiosis, we conclude that the sporulation program of S. paradoxus is largely dominant in the hybrid, with a slight indication of co-dominance at the earliest stages of meiosis (e.g. meiosis I).

The hybrid has a longer transition between meiosis I and II than its parents

To further understand differences between the hybrid and its parents, we measured the area between the bi- and tetra-nucleate curves from the beginning of sporulation to when the two curves intersect. The area between the bi- and tetra-nucleate curves can measure the speed of progression through meiosis, independent of when meiotic divisions begin or end$^{87}$; a larger area signifies a longer delay between meiosis I and II.
We measured the area between the bi- and tetra-nucleate curves for each sample, and find that the hybrid has a larger area between its curves (Figure 3-1). The areas for *S. cerevisiae* and *S. paradoxus* are significantly different (t-test, $P = 0.02$), and the area for the hybrid is significantly different than those of both *S. cerevisiae* and *S. paradoxus* (t-test, $P = 0.002$ and $P = 0.04$, respectively). As the hybrid area is larger than both parents, the hybrid transitions from meiosis I and II over a longer period of time than both parents. To further characterize differences in how the two parents and their hybrid progress through meiosis, we performed an RNA-Seq study of gene expression differences during meiosis.

**RNA-seq profiling during meiosis**

To capture gene expression changes between the hybrid and its parents during meiosis, we defined four developmental stages based on the formation of bi-nucleates and tetra-nucleates (Figure 3-S1). The four stages are defined by: $T_0$ as the time at which we placed cells in sporulation media; $M_1$ as an hour before bi-nucleates appear; $M_1/M_2$ as an hour before tetra-nucleates appear; and $M_2$ as an hour before the tetra-nucleates comprise the majority of the cell population. Thus, the $T_0$, $M_1$, $M_1/M_2$ and $M_2$ stages correspond to: 0, 3, 5 and 8 hours for *S. cerevisiae*; 0, 5, 9 and 11 hours for *S. paradoxus*; and 0, 4, 9 and 12 for their hybrid (Figure 3-1).

We used RNA-Seq to measure gene expression in both parents as well as allele-specific and total gene expression in the hybrid. We obtained a total of $276 \times 10^6$ mapped reads with a median of $4.2 \times 10^6$ per sample (see Methods). We simultaneously mapped hybrid reads to both the *S. cerevisiae* and *S. paradoxus* genome, and found the percentage of mapped reads from the hybrid to be similar to that of both parents. Of the hybrid reads, 49.4% mapped to the *S. cerevisiae* genome and 50.6% mapped to the *S. paradoxus* genome, suggesting minimal read mapping bias in the hybrid.

To characterize our developmental stages, we compared the *S. cerevisiae* expression profile to previously documented changes in gene expression$^{56}$. We found metabolic gene expression is high at $T_0$ and rapidly decreases after $M_1$; early sporulation gene expression begins to increase at $M_1$; middle sporulation gene expression begins to increase at $M_1/M_2$; and late sporulation gene begins to increase at $M_2$ (Figure 3-S2). Thus, our *S. cerevisiae* expression data cover the early and middle stages of meiosis and are consistent with previously reported patterns of gene expression during meiosis.
**Heterochronic changes in the hybrid**

To identify differentially expressed genes across stages and across species, we compared two models (see Methods). Our null model has no explanatory variables and our alternative model has variables for species background (S. cerevisiae, S. paradoxus, and the hybrid), stage (T₀, M1, M1/M2, M2) and an interaction between background and stage. Through this comparison, we found 1,083 out of 3,463 expressed genes are differentially expressed (P < 0.001). To further refine this list to genes that exhibit an interaction between species’ background and developmental stage, we dropped the interaction term and found 352 genes (P < 0.05). The expression of these genes is of particular interest since they change over the course of meiosis in a species- or hybrid-specific pattern and cannot be explained by read mapping bias, which is expected to be constant across developmental stages.

To gain a general view of expression differences, we performed a principle coordinate analysis of the 352 genes differentially expressed across stages and species’ background (Figure 2). The first two coordinates explain 42% and 14% of the variation among samples and separate most of the samples from one another. The first coordinate orders the samples according to developmental stages. At T₀, hybrid expression lies in between those of S. cerevisiae and S. paradoxus. However, at both the M1 and M1/M2 stages, hybrid expression is to the right of both S. cerevisiae and S. paradoxus expression, indicative of a more advanced phase of the meiotic expression program (Figure 2). To a lesser degree, S. paradoxus expression is also more advanced than that of S. cerevisiae, at the M1 and M1/M2 stages. The more advanced phase of S. paradoxus could be explained by our sampling at developmental stages rather than absolute time-points, which were two (M1) and four (M1/M2) hours later compared to S. cerevisiae. However, the more advanced phase of the hybrid cannot be explained by our sampling scheme, since both the M1 and M1/M2 stages were sampled at absolute stages between those of the two parents. At the M2 stage, the expression of the hybrid is similar to both parents and only shows a slight difference due to the second principle coordinate.

For each stage we identified genes whose expression in the hybrid occurs outside the expression range of the parents (Table 3-1, t-test, P < 0.05). Most of the hybrid genes that are expressed outside the range of the parents occur at stages M1 and M1/M2. At M1, the hybrid expresses 21 genes lower than either parent, 11 of which are ribosomal proteins. While the expression of these 21 genes declines during
meiosis for both parents and the hybrid, hybrid expression declines earlier than parental expression (Figure 3-3). At M1/M2, the hybrid expresses 26 genes higher than either parent, 10 of which are involved in middle and late sporulation. While the expression of these 26 genes increases during meiosis for both parents and the hybrid, hybrid expression increases earlier than parental expression (Figure 3-3). Thus, genes that are misexpressed in the hybrid compared to both parents exhibit heterochronic changes, similar to changes seen in the composite analysis of all differentially expressed genes (Figure 2).

Trans-acting factors dominate heterochronic differences between the hybrid and its parents

Gene expression differences can be produced by changes in cis-regulatory sequences, trans-acting factors, or a combination of the two. To determine whether the hybrid’s heterochronic changes are the result of cis- or trans-acting effects, we compared allele-specific expression of the hybrid to both parents’ expression. If gene expression patterns in the hybrid are dominated by trans-acting factors, S. cerevisiae and S. paradoxus alleles should exhibit similar patterns in the hybrid. Using the same set of 352 genes examined above, we find largely overlapping patterns of allele-specific expression in the hybrid (Figure 3-S3). Thus, differences in the timing of gene expression between the hybrid and its parents are dominated by trans-acting factors.

Cis- and trans-acting changes in gene expression

Altered gene regulation in the hybrid may not necessarily be in the form of misexpression; cis- and trans-acting changes can also lead to novel regulatory interactions. To identify genes exhibiting cis- and trans-acting expression differences, we compared allele-specific expression in the hybrid to both parents’ expression. Our null model has no explanatory variables and our alternative model has variables for stage (T₀, M1, M1/M2, M2), species background (S. cerevisiae, S. paradoxus, and the hybrid), parental allele (S. cerevisiae or S. paradoxus) and interactions between at least two terms. Through this comparison, we find 1,102 differentially expressed genes (P < 0.001). Following previous work⁴⁸,⁵¹, we classified each gene’s expression divergence into five categories: cis-only, trans-only, cis + trans, cis*trans and compensatory changes (Figure 3-S4 and Methods). We found 389 genes that were classified into at least one of these categories at one stage (Figure 3-4).
Strikingly, we observe a transition from predominantly trans-acting expression divergence at T₀ to predominantly cis-acting expression divergence at M2 (Figure 3-4, Table 3-S2). At the intermediate stages of M1 and M1/M2, we find similar numbers of cis- and trans-acting expression differences, but also more genes in the other three categories: cis- and trans-effects in the same (cis + trans) or opposite (cis*trans) directions, and compensatory changes, where allele-specific expression differences are present in the hybrid but not between the parents. Of particular interest are genes whose expression is due to opposing cis*trans interactions or compensatory changes, as these genes could be candidates for genetic incompatibilities contributing to sterility in the hybrid. Together, cis*trans and compensatory changes are enriched at the M1 and M1/M2 stages relative to the other categories (Fisher's Exact test, P = 0.003 and P < 0.001, respectively). Table 3-2 lists the genes that show opposing and compensatory cis-trans divergence at each stage. Many of these genes are essential for viability and involved in meiosis, rRNA processing and other translational processes, and mitochondrial functions.

Discussion

Interspecific hybrids often misexpress genes and this misexpression may contribute to reproductive isolation via hybrid sterility or inviability. In this study, we examined expression profiles of S. cerevisiae, S. paradoxus and their sterile hybrid during sporulation. We find that heterochronic changes cause misexpression in the hybrid, which is consistent with the anti-recombination model of yeast speciation. We also find an increased number of cis- and trans-acting changes with compensatory or opposing cis-trans divergence at the same meiotic stages exhibiting a heterochronic shift in hybrid gene expression. Below, we discuss our observations in the context of pathways previously implicated in RI and the role of misexpression in hybrid sterility.

Meiotic differences between the hybrid and its parents

Our microscopy data show that the sterile interspecific yeast hybrid progresses through meiosis differently than both its parents. While the hybrid takes longer to complete meiosis II, in a manner similar to S. paradoxus, it completes meiosis I intermediately of its parents. These meiotic differences result in a longer period between meiosis I and II in the hybrid (Figure 3-1). The cause of the longer meiotic transition in the hybrid may result from divergence between the two parents, whereby the hybrid follows
the early completion of meiosis I in *S. cerevisiae* but the later completion of meiosis II in *S. paradoxus*. However, it is also possible that the longer meiotic transition is a consequence of lack of recombination (see anti-recombination model discussed below).

The shift in how *S. cerevisiae* and *S. paradoxus* progress through meiosis is not the only heterochronic change that has been observed between these two species. *S. paradoxus* both mates and germinates more slowly than *S. cerevisiae*. In addition, mitotic gene expression profiles exhibit heterochronic divergence among *S. cerevisiae* strains and *S. paradoxus*. Given that *S. cerevisiae* is more thermophilic than *S. paradoxus*, at least some of heterochronic changes could be related to temperature preferences.

*Misexpression as a consequence of temporal shifts between the hybrid and its parents*

While misexpression has previously been observed in interspecific hybrids, the causes and consequences of this misexpression are not often known. The number of misexpressed genes we identified is not large due to our conservative analysis of differential expression. However, the overall pattern of differentially expressed genes from our principal coordinate analysis shows a heterochronic shift similar to misexpressed genes. The heterochronic shift does not occur at $T_0$ where we find hybrid expression to lie between the parents’ expression, consistent with previous work. Thus, the simplest explanation for the shift in gene expression is that the normal meiotic expression program is activated earlier in the hybrid than either parent.

Previous studies have found hybrid misexpression changes during development, with adults showing more misexpression than earlier stages of development. An increase in misexpression during development supports a cascading model of misexpression, due to either to evolved differences in gene regulation or changes in development. Regulatory divergence of genes expressed early during development can be propagated to extensive changes in expression later in development. Similarly, early changes in tissue abundance or cell types during development can be propagated to larger differences in adults. The observation that a pupal stage of *Drosophila* hybrids has fewer misexpressed genes than either larval or adult stages supplies evidence against either of the two cascading models and suggests that pupal stages may be more immune to "developmental systems shift", perhaps due to the complexity of gene regulation during metamorphosis. Our results show that there is little to no misexpression outside
the range of the normal meiotic program in an interspecific yeast hybrid. While this shift could be caused by misexpression of even a single master regulator of sporulation, the meiotic program does not appear to be altered other than in its timing.

Surprisingly, we find a near absence of a shift in expression at the final M2 stage, where late sporulation genes involved in meiotic division and spore wall formation are turned on. This observation suggests that the meiotic program in the hybrid is not simply shifted to an earlier time-point, since it closely resembles both parents at the M2 stage. Similarly, even though the completion of meiosis I and the formation of bi-nucleates occurs at different time-points, both parents and their hybrid reach their maximum percentage of tetrads produced at the same time-point, ~15 hours. Thus, both the meiotic divisions and gene expression program complete at the same time.

The anti-recombination model of hybrid sterility

The anti-recombination pathway has been shown to contribute to sterility of yeast hybrids. The model suggests that the mismatch repair pathway recognizes a multitude of mismatches between interspecific homologous chromosomes in the hybrid during metaphase of meiosis I. As a result, the hybrid activates the anti-recombination pathway, which prevents non-homologous chromosomes from crossing over and causes aneuploid spores. Elimination of mismatch repair genes, which start the anti-recombination pathway, significantly increases spore viability in yeast, indicating anti-recombination plays a role in RI. Furthermore spore viability is dramatically improved in hybrid tetraploids, where non-divergent homologous chromosomes can recombine and properly segregate.

Our phenotypic and expression data support the anti-recombination model for yeast hybrid sterility. S. cerevisiae strains that cannot recombine homologous chromosomes progress more quickly through M1 than wild type strains. These mutants produce bi-nucleates quickly, but tetra-nucleates at a wild type pace, which creates a larger area between the bi- and tetra-nucleate curves. Likewise both our phenotypic and expression data is consistent with the hybrid moving through meiosis I but not meiosis II more quickly than both parents. Taking our data in the context of previous yeast speciation studies, we conclude that the hybrid moves more quickly through meiosis I because of a lack of recombination.

Cis-only and trans-only changes between S. cerevisiae and S. paradoxus
We found a transition from trans- to cis-acting expression divergence over the course of meiosis. Previous studies have observed a fairly wide range in the proportion of expression differences attributable to cis-acting elements\(^93\). However, there is a tendency for cis-acting changes to be enriched between species compared to within species\(^86,94\), and the proportion of cis-acting changes in yeast depends on the environment\(^50\). Our observations of a transition from predominantly trans- to cis-acting expression differences between \textit{S. cerevisiae} and \textit{S. paradoxus} add time as another factor contributing to variability.

Previous work has also found associations between genes with both cis- and trans-acting divergence and whether a gene is misexpressed in an interspecific hybrid\(^38,51\). While we find an enrichment of opposing cis-trans and compensatory changes at the M1 and M1/M2 stages, where we also see the most misexpressed genes (Table 3-1), only seven misexpressed genes have opposing cis-trans or compensatory expression divergence. Among these seven, \textit{TEF1} and \textit{EFB1} are both translation elongation factors that help bind aminoacyl-tRNA to ribosomes, and \textit{CDC26} is a subunit of the anaphase promoting complex (APC/C) involved in exit from mitosis. While \textit{CDC26} function in meiosis is not known, the \textit{C. elegans} homologue of \textit{CDC26} is required for the metaphase to anaphase transition through meiosis\(^5\).

\textit{Cis-trans opposing and compensatory changes between the parents and their relationship to reproductive isolation}

During the M1 and M1/M2 stage we observe an increase in the relative abundance of cis*trans and compensatory changes compared to cis-only, trans-only or cis + trans changes. The cis*trans and compensatory changes are of particular interest since they may only be present in the hybrid and could thus contribute to \textit{Ri}\(^38\). A number of genes are noteworthy. Three genes involved in later meiosis (\textit{SPO12}, \textit{HED1}, and \textit{APC11}) have compensatory interactions and \textit{CDC26} shows a cis*trans interaction between the parents. The hybrid expresses both alleles of these genes higher than the same allele in the parents. \textit{APC11} and \textit{CDC26} form the APC/C complex, which regulates the metaphase to anaphase transition of meiosis and the exit from both mitosis and meiosis to G1 phase. During mitosis APC/C regulates \textit{SPO12}, which regulates the exit of mitosis, and \textit{SPO12} may play a similar role during meiosis\(^96\). \textit{HED1} is a suppressor of \textit{RED1}, which is involved in the pachytene checkpoint. Hed1p suppresses Red1p when the
recombination machinery is impaired. We can interpret this result as the hybrid expressing HED1 to prevent and bypass recombination or to turn off early M1 genes.

Previous studies have shown that mitochondrial genes are involved in yeast hybrid sterility. Many of the genes for which we find compensatory changes are involved in mitochondrial maintenance and respiration. Although we did not specifically identify MRS1, which contributes to hybrid sterility between S. cerevisiae and S. paradoxus, we found another mitochondrial RNA splicing gene, MRS3. Overexpression of MRS3 overcomes splicing mutations in S. cerevisiae, and may be involved in nuclear-mitochondrial incompatibilities much like MRS1.

Conclusions

In this study, we show that hybrid misexpression in yeast is a result of a heterochronic shift in the meiotic gene expression program. While the cause of this shift remains unknown, it is consistent with activation of the anti-recombination pathway and bypassing meiotic recombination. No direct relationship is found between genes exhibiting opposing or compensatory cis-trans changes and misexpressed, yet both are enriched at the same stages of development. Also, we cannot exclude the possibility that certain regulatory differences between species contribute to meiotic defects in yeast. The extent to which hybrid misexpression in multicellular organisms is a consequence of changes in development and/or gene regulation remains to be determined.

Methods

Strains

The strains used in this study are listed in Table S1. We derived S. cerevisiae strains from YPS163 and S. paradoxus strains from N17. P. Sneigowski, University of Pennsylvania, provided initial strains. All genetic manipulations were created using a standard lithium acetate transformation and homologous recombination. We replaced the HO locus with dsdAMX4 in S. cerevisiae and NATMX4 in S. paradoxus, and isolated haploid derivatives of the strains. For both parental species, we generated three independent diploid strains containing the double ho deletion. We generated interspecific hybrids by mating haploid strains from the two species and isolating their hybrids that contain the two dominant markers, dsdA and NAT. We generated two independent hybrids for each reciprocal cross.
**Growth and sporulation conditions**

We inoculated 100 ml of YPD (1% Bacto™ yeast extract, 2% Bacto™ peptone and 2% dextrose) in 250 ml Erlenmeyer flasks, and incubated the culture at 30ºC and 340 rpm for 15 hours. To sporulate cells we washed cells with water and resuspended the cells in 250 ml of SPO (1% potassium acetate, 0.1% Bacto™ yeast extract, 0.05% dextrose) for a final concentration of $10^7$ cells/ml. We incubated the cultures in 1 L baffled flasks at 30ºC at 340 rpm for 24 hours. We used distilled water for all media.

**Sampling**

Once we resuspended cells in SPO, we sampled, washed and snap-froze cells at every hour between 0-16 hours, and at 20 and 24 hours. We fixed a subset of sampled cells in formaldehyde and ethanol, and stained fixed cells with DAPI. We counted nuclei using fluorescent microscopy. We defined developmental stages based on nuclei count. $T_0$ is the time at which we placed cells in sporulation media: M1 is defined as an hour before we observe bi-nucleates (when M1 is complete); M1/M2 (the transition between the end of M1 and the beginning of M2) is defined as an hour before tetra-nucleates appear; and the end of M2 is defined as an hour before the tetra-nucleates comprise the majority of the cell population.

**Illumina indexing library, alignments and mapping**

We extracted total RNA from samples using Ambion RiboPure-Yeast Kit. We purified mRNA from total RNA using Ambion Micro PolyPurist Kit, reversed transcribed mRNA into cDNA, and sheered cDNA to 100 bp by sonication using a Bioruptor. We ligated indexed Illumina library adaptors to the sheared cDNA samples, size selected ligated samples (250-350 bp) and then mixed at equal concentrations for a final concentration of 1 nM. The final library was sequenced by GTAC at Washington University using two Illumina HiSeq lanes at a run concentration of 6 pM.

We used bowtie to align each sample’s reads to the *S. cerevisiae* reference genome, S288c, and *S. paradoxus* genome for strain N17, or the two genomes combined for reads from the hybrid. We required unique alignments with up to one mismatch, and set all other options to default. We mapped aligned reads to 6,722 features shared between the annotated S288c and N17 genomes, of which 3,463 had one or more read. The group of 3,463 orthologs is composed of 2,894 ORFs, 189 ARSs, 105 tRNAs,
42 regulatory and chromosome maintenance RNAs (e.g. anti-sense RNA, snRNA), eight telomeres, two variants of 5s ribosomal subunit, and two centromeres. Only ORFs and ARSs have differential expression.

**Differential gene expression measurements**

We used DESeq\textsuperscript{72} to measure differential expression. We estimated dispersions using the pooled-CR method for multivariate designed experiments and a separate model formula for each analysis, described below. As a control we randomized each sample’s label and calculated the number of significant genes across 100 permutations. We calculated an empirical estimate of our false discovery rate for each P-value cutoff used. This empirical FDR was found to be less than 1% for our following analyses.

To identify genes with differential expression between *S. cerevisiae*, *S. paradoxus*, and their hybrid over time, we used the nbinomGLMTest to compare a model without any explanatory variables to one with variables for developmental stage and species’ background: $\text{count} \sim \text{stage} \times \text{species}$, where the asterisk indicates the presence of both additive and interactive terms between the two variables. To identify genes that showed stage-specific differences between species’ background we compared $\text{count} \sim \text{stage} + \text{species}$ to $\text{count} \sim \text{stage} \times \text{species}$. By keeping a non-additive interaction between stage and species, we ensure significant genes’ differential expression is due to biological significance rather than an artifact from any sequencing or read mapping bias between genomes.

To identify allele-specific expression differences in the hybrid and differences in allele expression in the hybrid versus parental background, we expanded our above two variable model to one with variables for developmental stage and species’ background: $\text{count} \sim \text{stage} \times \text{species} \times \text{allele}$. We compared our three-variable model to a null model without any explanatory variables and identified 1,102 genes. To identify genes that showed stage-specific differences between species’ background and allele, we compared $\text{count} \sim \text{stage} + \text{species} + \text{allele}$ to $\text{count} \sim \text{stage} \times \text{species} \times \text{allele}$ and identified 266 genes. Because the set of 266 and 1,102 genes showed very similar patterns of *cis* and *trans* expression categories (defined below and Table 3-S2), we presented the larger set in the results. The differentially expressed genes and P values for these analyses are provided in Supplementary Data.

*Other statistical analyses*
We used variance-stabilized data from the above DESeq analyses for the remainder of our statistical analyses. To compare overall expression differences between the hybrid and parents, we applied principle coordinate analysis (PCA) on the 352 genes whose expression level depended on an interaction between species and stage in our two-variable model. We obtained principle coordinates using the Euclidean distance between each pair of samples and the cmdscale function in the statistical package R. We identified genes expressed by the hybrid outside of the range of the two parents’ expression using a two-tailed t-test with a P value cutoff of 0.05 (Supplementary Data File). For our allele-specific analysis, we applied PCA to the same 352 genes using variance-stabilized data from our three-variable model.

To measure cis- and trans- effects, we followed the analysis from 102 and used the 1,102 genes whose expression level depended on one or more variables in our three-variable model. We tested for significant differences in expression (E) between the parents (P), S. cerevisiae (Sc) and S. paradoxus (Sp) for each gene (i), using a t-test with a P value cutoff of 0.05.

(1) \( E_P = E_{Sc,i} - E_{Sp,i} \)

We tested for significant differences in expression (\( E_H \)) between the S. cerevisiae and S. paradoxus alleles in the hybrid (\( H,Sc \) and \( H,Sp \), respectively), using a t-test with a P value cutoff of 0.05.

(2) \( E_H = E_{H,Sc,i} - E_{H,Sp,i} \)

To test whether the difference in expression between parental genes is of equal size as the allelic differences in the hybrid (\( E_{H,P} \)), we used a t-test with a P value cutoff of 0.05.

(3) \( E_{H,P} = E_H - E_P \)

Using these cutoffs, we define differential gene expression due to cis- and trans- effects in the following paragraph and Figure S4. We define cis-effects as significant expression change between parents (\( E_P \)) and between hybrid alleles (\( E_H \)), but no expression change between parent genes and hybrid alleles (\( E_{H,P} \)).
We define trans-effects as no expression change between parents ($E_P$), but significant expression change between hybrid alleles ($E_H$) and between parent genes and hybrid alleles ($E_{H,P}$). We define additive cis- and trans-effects ($cis + trans$) as significant expression change between parents ($E_P$), hybrid alleles ($E_H$) and between parent genes and hybrid alleles ($E_{H,P}$), wherein the signs of the magnitudes of $E_P$ and $E_H$ are both the same (i.e. both positive or negative values). We define opposing cis- and trans-effects ($cis*trans$) as significant expression change between parents ($E_P$), hybrid alleles ($E_H$) and between parent genes and hybrid alleles ($E_{H,P}$), wherein the signs of the magnitudes of $E_P$ and $E_H$ are the opposite. Compensatory effects are those with no expression change between parents ($E_P$), but significant expression change between hybrid alleles ($E_H$) and between parent genes and hybrid alleles ($E_{H,P}$). Results of these tests are provided in Supplementary Data.

Data access

The sequencing data are available from the NCBI Gene Expression Omnibus under accession number XXXXXXX. Normalized expression levels of differentially expressed genes are also available as part of the supplementary data.

Acknowledgements

We thank Andrew Bergen and Elizabeth Engle for assistance in collecting RNA samples, and Priya Sudarsanam and Bin Wang for advise with RNA-Seq. This work was supported by a National Institutes of Health grant GM080669 to JCF and a training grant to DSL.
Figures and Tables

Figure 3-1. Profiles of meiotic divisions during sporulation. The frequency of bi-nucleates (solid) and tetra-nucleates (dashed) is shown for S. cerevisiae (A), S. paradoxus (B), and their hybrid (C) over a 24 hour time-course. The area between the bi-nucleate and tetra-nucleate curves is shown in gray and numerically labeled. The sampling of T₀, M1, M1/M2 and M2 stages is shown by the arrows above each species’ graph.
Figure 3-2. Principal coordinate analysis of 352 differentially expressed genes. Ovals show the 95% confidence interval of S. cerevisiae (C), S. paradoxus (P), and hybrid (H) principal coordinates at each meiotic stage and are centered on the mean values. Meiotic stages are T₀ (red), M1 (orange), M1/M2 (blue), and M2 (green). The first and second principal coordinate explain 42% and 13% of the variation among samples, respectively.
Figure 3. Temporal changes in misexpressed genes during sporulation. Boxplots of 21 genes expressed lower in the hybrid than either parent at M1 (A), and 26 genes expression higher in hybrid than either parent at M1/M2 (B). Boxes indicate the span of the second and third quartiles and dashed lines indicate an estimate of the 95% confidence of the median. Boxes are shown for *S. cerevisiae* (Scer), *S. paradoxus* (Spar) and the hybrid at each stage (*T*₀, M1, M1/M2 and M2). Expression levels are the normalized log2 number of reads.
Figure 3-4. Genes classified into different categories of cis- or trans-acting expression differences. Each plot shows the difference in the log2 normalized expression level between the parental species, relative to *S. cerevisiae* and *S. paradoxus* alleles in the hybrid (C = *S. cerevisiae* and P = *S. paradoxus*) for the T₀ (A), M1 (B), M1/M2 (C), and M2 (D) stages. Each circle shows one of 389 genes classified into five categories of expression divergence: cis-only (red), trans-only (green), cis+trans (blue), cis*trans* (orange), compensatory (black). Genes without significant expression differences are shown in gray. Inset within each panel shows the number of genes classified into each category.
Table 3-1. Number of hybrid genes significantly different from both parents

<table>
<thead>
<tr>
<th>Hyrid expression relative to parent(^1)</th>
<th>T(_0)</th>
<th>M1</th>
<th>M1/M2</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower</td>
<td>10</td>
<td>21</td>
<td>26</td>
<td>3</td>
</tr>
<tr>
<td>Intermediate</td>
<td>22</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Higher</td>
<td>1</td>
<td>1</td>
<td>26</td>
<td>3</td>
</tr>
</tbody>
</table>

\(^1\)Hybrid expression is significantly different from both parents (t-test P < 0.05).

Table 3-2. Genes with opposing and compensatory *cis-trans* expression divergence.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Meiosis/Cell cycle</th>
<th>Mitochondrial function</th>
<th>rRNA processing</th>
<th>Translation related processes</th>
<th>Essential genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>T(_0)</td>
<td>HST4</td>
<td>IDP1</td>
<td>RPS8B, PWP1</td>
<td></td>
<td>PWP1</td>
</tr>
<tr>
<td>M1</td>
<td>STE5</td>
<td>GCV2, COX26,</td>
<td>RIO2</td>
<td>SUI2, TUF1</td>
<td>SEC14, FAS2, RIO2, SUI2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SHH4, COX6,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GDS1, MRS3,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUF1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1/M2</td>
<td>KAR2, HED1, STE18, SPA2, CDC60, SPO12, APC11, CDC26</td>
<td>MRPL37, ERV1, RSM27, CYT1, ALD6</td>
<td>DBP8, RPS0B, RPS17A, RPL31A</td>
<td>SUI2, TEF1, PLP1</td>
<td>KAR2, ERV1, CDC60, GLN1, DBP8, APC11, RMP1</td>
</tr>
<tr>
<td>M2</td>
<td>--</td>
<td>ECM10</td>
<td>BFR2</td>
<td>--</td>
<td>HIP1, GPI11, BFR2</td>
</tr>
</tbody>
</table>
Figure 3S-1. Sporulation profiles of S. cerevisiae, S. paradoxus, and their hybrid. The frequency of mononucleates (A), binucleates (B), and tetranucleates (C) is shown over a 24 hours period of sporulation for S. cerevisiae (red), S. paradoxus (blue), and their hybrid (green). Bars show the standard deviation from replicates.
Figure 3S-2. *S. cerevisiae* gene expression profiles across four meiotic stages. The average expression level is shown for 12 Metabolic, 15 Early I, 12 Early II, 11 Early-middle, 19 Middle, and 5 Middle-late genes defined by Chu *et al.* (1998) at T0, M1, M1/M2, and M2 stages.
Figure 3S. Principal coordinate analysis of allele-specific expression. Principal coordinates are based on the Euclidean distance of 352 genes. Ovals show the 95% confidence interval of S. cerevisiae (C), S. paradoxus (P) and hybrid coordinates at each meiotic stage (see legend) and are centered on the mean values. Hybrid expression of S. cerevisiae (HC) and S. paradoxus (HP) alleles were treated as separate samples. The first and second principal coordinates explain 42% and 14% of the variation among samples, respectively.
<table>
<thead>
<tr>
<th>Effect</th>
<th>Graph</th>
<th>$E_P^*$</th>
<th>$E_H^*$</th>
<th>$E_{H-P}^*$</th>
<th>Signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>cis</td>
<td><img src="image" alt="Graph" /></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>NA</td>
</tr>
<tr>
<td>trans</td>
<td><img src="image" alt="Graph" /></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>cis + trans</td>
<td><img src="image" alt="Graph" /></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>same</td>
</tr>
<tr>
<td>cis * trans</td>
<td><img src="image" alt="Graph" /></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>opposite</td>
</tr>
<tr>
<td>compensatory</td>
<td><img src="image" alt="Graph" /></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>opposite</td>
</tr>
</tbody>
</table>

**Figure 3S-4. Classifications of cis and trans expression divergence categories.** The definitions of five expression divergence categories along with a graphical representation. $E_P$ is the expression difference between the parents, *S. cerevisiae* and *S. paradoxus*. $E_H$ is the expression difference between the hybrid alleles from *S. cerevisiae* and *S. paradoxus*. $E_{H-P}$ is the expression difference between $E_P$ and $E_H$. Significant and non-significant effects are indicated by ‘+’ and ‘-’.
Table 3S- 1. Strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Species</th>
<th>Parent(s)</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>YJF153</td>
<td>S. cerevisiae</td>
<td>YPS163</td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td>YJF154</td>
<td>S. cerevisiae</td>
<td>YPS163</td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td>YJF850</td>
<td>S. cerevisiae</td>
<td>YPS163</td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td>YJF851</td>
<td>S. cerevisiae</td>
<td>YPS163</td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td>YJF852</td>
<td>S. cerevisiae</td>
<td>YPS163</td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MATα ho::dsdAMX4</td>
</tr>
<tr>
<td>YJF777</td>
<td>S. paradoxus</td>
<td>N17</td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF778</td>
<td>S. paradoxus</td>
<td>N17</td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF853</td>
<td>S. paradoxus</td>
<td>N17</td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF854</td>
<td>S. paradoxus</td>
<td>N17</td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF855</td>
<td>S. paradoxus</td>
<td>N17</td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF843</td>
<td>Hybrid</td>
<td>YJF153 + YJF778</td>
<td>S. cerevisiae MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S. paradoxus MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF876</td>
<td>Hybrid</td>
<td>YJF153 + YJF778</td>
<td>S. cerevisiae MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S. paradoxus MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF842</td>
<td>Hybrid</td>
<td>YJF842 + YJF874</td>
<td>S. cerevisiae MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S. paradoxus MATα ho::NATMX4</td>
</tr>
<tr>
<td>YJF874</td>
<td>Hybrid</td>
<td>YJF842 + YJF874</td>
<td>S. cerevisiae MATα ho::dsdAMX4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S. paradoxus MATα ho::NATMX4</td>
</tr>
</tbody>
</table>

Table 3S- 2. Classification of genes into different categories of expression divergence

<table>
<thead>
<tr>
<th>Geneset*</th>
<th>Stage</th>
<th>cis-only</th>
<th>trans-only</th>
<th>cis+trans</th>
<th>cis*trans</th>
<th>Compensatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>266</td>
<td>T₀</td>
<td>1</td>
<td>51</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>M1</td>
<td>11</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>M1M2</td>
<td>13</td>
<td>12</td>
<td>3</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>14</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1102</td>
<td>T₀</td>
<td>3</td>
<td>172</td>
<td>5</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>M1</td>
<td>23</td>
<td>36</td>
<td>9</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>M1M2</td>
<td>37</td>
<td>38</td>
<td>8</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>40</td>
<td>15</td>
<td>1</td>
<td>0</td>
<td>9</td>
</tr>
</tbody>
</table>

*Geneset refers to 1102 genes found to differ between our null and full model (with terms for species, allele and stage) including all interaction terms, and to 266 genes found to differ between our full model with and without interaction terms
Chapter 4: Conclusion
A major objective of evolutionary genetics is to understand the mechanistic basis of speciation. As such I focused my dissertation on understanding the dysfunctional regulation of genes in a sterile, interspecific yeast hybrid during meiosis, a process that is susceptible to RI. The field of speciation genetics has traditionally used forward genetic approaches to isolate genetic incompatibilities, but due to the lack of sterile of viable F1 hybrid progeny, it has been difficult to gain a general view of how genome-wide divergence between species contributes to RI. I discussed the complications with traditional approaches and yeast speciation in Chapter 2, as well as outlined potential experiments to verify a single gene’s contribution to RI. Fortunately the advent of next-generation sequencing has brought new opportunities to expand our knowledge of the connection between genome-wide divergence and RI. In Chapter 3 I presented my findings from an RNA-Seq assay to specifically understand the relationship between expression divergence and RI between two yeast species, *S. cerevisiae* and *S. paradoxus*, which can form sterile hybrids under laboratory conditions. Here I deepen my discussion about my research in the context of previous literature, consider possibilities of the future direction of yeast speciation genetics and genomics, and conclude with my final thoughts.

**The relationship between dysfunctional hybrid expression and reproductive isolation**

Hybrid misexpression can either be a consequence or a cause of RI, although the connection between regulation and RI is rarely deduced. From previous studies, we know that hybrids have a lack of recombination in hybrids, which in *S. cerevisiae*, leads to a precocious ending of meiosis I. Both my phenotypic and expression data corroborate that the hybrid proceeds through meiosis I more quickly than both of its parents (Chapter 3). However I was not able to identify misexpression of a single master regulator that would lead to either a lack of recombination or an accumulation of misexpression. Thus I could not identify a specific genetic incompatibility that could contribute to RI during meiosis. Despite differences in the timing of each step, the gene expression patterns in the parents and hybrid are largely similar. Hence the simplest explanation is that the hybrid’s lack of recombination results in faster advancement of the hybrid meiotic program. Any further gene misexpression in the yeast hybrid is a consequence of RI and the lack of recombination during occurring in meiosis I.

My RNA-seq study brings us a step closer to understanding the relationship between hybrid gene regulation and RI, but neither answers what causes a lack of recombination in yeast nor distinguishes
whether hybrid sterility is occurring in early or late meiosis I. As stated in the introduction, an incompatibility that causes anti-recombination may still exist. More expression studies and functional assays are needed to determine whether anti-recombination is due to sequence divergence or a genetic incompatibility. Some of the experiments that I propose in this chapter utilize multiple genetic manipulations, and would need to be performed carefully and tested for spore viability and recombination rates beforehand (outlined in Chapter 2) before conclusions could be drawn.

To further study gene expression in meiosis I, I propose two experiments. First, my RNA-Seq experiment can be performed on mismatch repair mutant and tetraploid hybrids as controls (Chapter 1). Both types of hybrids would be predicted to not have a heterochronic shift in their meiotic program in comparison to their parents, as they proceed through recombination normally. Additionally we could completely synchronize cells at G1 to study early meiosis I, and synchronize cells with an inducible promoter for NDT80 to study late meiosis I. NDT80 is expressed after recombination and serves as the master regulator of middle sporulation and the cell’s commitment to meiosis. If RI occurs before the cell’s commitment to meiosis, I would expect to observe an increase in misexpression before NDT80 is induced. On the other hand, if RI occurs after the cell’s commitment to meiosis, I would expect to observe and increase in misexpression after NDT80 is induced. Again the experiment can be performed on mismatch repair mutant and tetraploid hybrids as controls (Chapter 1). Both types of hybrids would be predicted to have no misexpression in their DNA damage and anti-recombination pathways, since they can recombine their chromosomes during meiosis I.

If sequence divergence plays a role in RI, I would specifically expect to see misexpression before NDT80 is induced. I would expect an increase in the hybrid expression of DNA damage and anti-recombination pathways, and a decrease in the hybrid expression of homologous recombination proteins relative to the parents. I was not able to see a change in these pathways in my study detailed in Chapter 3, likely because variations in these pathways are nonexistent, the statistical power was too low, or the time points in my study are too far apart.

The possibility remains that by using completely synchronized hybrids with an inducible version of NDT80, we could identify candidate genes with misexpression before or after the commitment to meiosis. To verify if a candidate gene contributes to RI, functional experiments would still need to be performed.
Assays described in Chapter 2 could reveal that a loss of recombination is caused by either sequence divergence or a genetic incompatibility (Chapter 1). Three specific predictions can be made in either scenario for the effects of manipulating a candidate gene. If sequence divergence activates the anti-recombination pathway, (1) deletions of a candidate gene would not increase hybrid spore viability; (2) generating a hybrid homozygous for either parent’s copy of the gene would not increase hybrid spore viability; and (3) insertion of one parent’s homolog of the candidate would rescue a gene deficiency in the other parental background. On the other hand, if an incompatible genes (rather than sequence divergence) result in the lack of recombination, I would expect the opposite results. Specifically (1) deletions of the candidate gene would increase hybrid spore viability; (2) a hybrid that is homozygous for the candidate gene would not increase hybrid spore viability, while a hybrid that is homozygous for the other parent’s homolog of the candidate gene would increase spore viability; and (3) the candidate gene would not complement its homolog’s deficiency in the other parental background.

**Dysfunctional pathway regulation in hybrids and genetic incompatibilities**

With the massive amount of data from next-generation sequencing experiments, it is difficult to determine candidate genes, which may be involved with RI. Thus as of yet, we have not formed a global view of the relationship between gene regulation and hybrid sterility. Therefore I chose a conservative analysis to narrow down a list of differentially expressed candidate genes that could promote hybrid sterility (Chapter 3). Non-additive cis/trans interactions are of particular interest because they present novel phenotypes only observed in the hybrids, and thus could be genes that contribute to RI. Given my stringency, the number of misexpressed genes and genes with cis/trans interactions is small enough to test for individual genetic contributions to RI. Chapter 2 outlines specific assays to verify a candidate gene, and the previous paragraphs outline how I would expect an incompatible gene to function during those assays.

In Chapter 3 I demonstrated that genes from four major processes’ consistently have non-additive cis/trans and compensatory changes that are uncovered in the hybrid. The four processes are sporulation, mitochondrial function, rRNA processing and translation. Sporulation and rRNA have also been described as rapidly evolving pathways. Additionally sporulation and mitochondrial function are closely involved in meiosis and have been previously associated with RI in multiple species.
Both theoretical and empirical arguments have been made suggesting that incompatibilities are rapidly evolving\textsuperscript{4,12,20,22,25,29,49,104}, and I detect two pathways that are thought to be under positive selection. As such, I detected two pathways that are thought to be under positive selection. In Chapter 3, I discussed genes in the sporulation pathway, such as *HED1* and the APC/C complex as genes of interest. At this point it is not clear whether these genes are misexpressed because of the hybrid’s heterochronic meiotic program, or they are in in incompatibilities. Bullard et al.\textsuperscript{103} detected rRNA processing as a pathway under cis-regulatory selection. I detect the same pathway, and if any of these genes could be verified as incompatibilities, it would lead to further credence of the notion that incompatibilities are rapidly evolving.

As discussed in Chapter 1, mitochondrial genes are predicted to be a class of genes susceptible to become incompatibilities, and interestingly I detect many genes involved in mitochondrial function to be differentially expressed in the hybrid. Mitochondrial genes are associated in RI in plants, copepods, and wasps, and are known to cause hybrid breakdown in *Saccharomyces*. Further genetic analyses of the genes listed in Table 3-2 could provide more insight as to whether mitochondrial-encoded genes contribute to RI more often than nuclear-encoded genes. There is a possibility that we would not observe an affect on hybrid sterility affecting the F1 hybrid generation, but we could observe an affect on hybrid breakdown in the F2 generation. In this case, we would observe dysfunctional gene regulation in the hybrid before a reproductive barrier appeared, whereas in Chapter 3, we found dysfunctional regulation of the whole meiotic program to occur after one barrier, a lack of recombination.

Although I state that it is hard to make predications as to which specific genes could play a role in RI, I cannot miss the opportunity to offer my humble opinion. Given that incompatible genes may prevent recombination in the hybrid, I expect the components of the synaptonemal initiation complex (SIC), which binds homologous chromosomes together for crossover, to be candidates. The proteins Zip1, Zip2, Zip3, Zip4/Spo22, Mer3, Msh4, and Msh5 form the SIC, which is antagonistic of anti-recombination\textsuperscript{105}. Many lines of evidence support my prediction. First Zip2 is a rapidly evolving protein between *S. cerevisiae* and *S. paradoxus*\textsuperscript{77}. Second ZIP1 and ZIP2 have allele specific expression in the hybrid, although they do not show signs of *cis* - or *trans* - interactions (Supplemental data file). Third there is evidence that complex D-M incompatibilities of three or more genes can exist\textsuperscript{74}. As the SIC contains more than three genes, it fits
this model of a complex incompatibility. Lastly silver staining of hybrid chromosomes during meiosis show homologous chromosomes form SIC but also have unpaired axial elements, cohesion containing structures that bind together homologous regions of chromosomes\textsuperscript{67}. Zip1p is integral to the SIC, and \textit{S. cerevisiae} \textit{zip1} mutants also have an increase of unpaired axial elements, much like interspecific yeast hybrids\textsuperscript{106}. Thus the SIC could be dysfunctional in hybrids rather than anti-recombination being active in hybrids. As I mentioned, the same functional genetic studies would need to be performed to verify these genes as incompatibilities. Yet, as I also cautioned in Chapter 2, predictions of incompatible genes is extremely difficult, and more expression data would help clarify whether this is a avenue that is sensible to pursue.

**Future directions for yeast speciation**

\textit{Traditional and modern approaches to speciation genomics studies}

My work in yeast provides a good example of the advantages and disadvantages to both traditional and modern approaches to the field of speciation genomics. While yeast research typically leads the field of genetics and genomics, speciation is one part of the field in which this has not been the case. One potential explanation is that the \textit{Saccharomyces sensu stricto} group is extremely divergent. The two closest relatives \textit{S. cerevisiae} and \textit{S. paradoxus} are 10 – 15% divergent, which is the same amount of divergence between human and mouse. Since such strong RI exists between species of \textit{Saccharomyces}, it is difficult to conduct traditional screens to identify single genes that contribute to hybrid sterility (Chapter 2). Two studies have been able to examine the genetics of hybrid breakdown, although not at a genome-wide level\textsuperscript{17,18}. Thus I do not believe using traditional screening methods to identify genetic incompatibilities between \textit{Saccharomyces} species at a genome-wide level is the best method. However traditional screening methods may prove useful in genome-wide identification of incompatibilities between species’ strains of any one species. A large number of both \textit{S. cerevisiae} and \textit{S. paradoxus} strain have been isolated around the world, and on average, are 1 - 5% divergent from each other. This amount of divergence is similar to \textit{Drosophila}, a model organism that has been successfully utilized to understand the genetic contribution to post-zygotic RI.

Next-generation sequencing gives the opportunity to study RI between more divergent species, and we have the advantage of genomic tools available for yeast to deduce pathways that are disrupted in
the hybrid. I demonstrated in Chapter 3 the utility of RNA-Seq in yeast, as I am able to compile a list of candidate genes that may contribute to RI in *Saccharomyces*. Earlier in this chapter and Chapter 2, I detailed genetic assays, which would be required to verify a single candidate gene’s contribution to RI. Thus the combination of next-generation sequencing and the arsenal of traditional yeast genetic tools provides us with the best opportunity to understand the total contribution of genetic incompatibilities to post-zygotic RI between two species.

**Reassessment of the anti-recombination model**

Although the anti-recombination model of RI is thought to completely isolate *Saccharomyces* species from one another\[^{51,64-66,68}\], scientists continue to search for genetic incompatibilities that contribute to RI. As discussed in the introduction, the numerous caveats of the anti-recombination model allow the possibility that incompatibilities play a role in RI between yeast species. From my research, I have uncovered an additional limitation to the anti-recombination model. I showed that interspecific yeast hybrids are sensitive to background mutations that can artificially increase spore viability. Other authors have also found that genetic manipulations confound results in studies of RI in yeast\[^{62,63,70}\]. Since the original experiments that contributed to the anti-recombination model utilized highly manipulated yeast strains\[^{68}\] and only one study utilized biological replicates\[^{66}\], I would not be surprised if results would differ when strains with at most one mutation are used. Thus the initial experiments that support the anti-recombination model should be revisited.

*Why do null mutants in some laboratory backgrounds increase spore viability?*

As I discussed in Chapter 2, transformations cause background mutations that can artificially rescue hybrid sterility. However with the dropping costs of sequencing, we can compare the genomes of the hybrids that have high spore viability to those that do not. I hypothesize two outcomes, either or both of which could explain the artificial rescue of hybrid sterility. The first is that increased hybrid spore viability occurs via a complex incompatibility between three or more genes. As the hybrids with high spore viability should at most have 40 mismatches than hybrids with low spore viability (Chapter 2), it should be relatively easy to find candidates for a complex incompatibility. The second outcome is that mutations in hybrids with high spore viability have mutations that increase the frequency of
autotetraploidy. Tetraploid hybrids are much more fertile than their diploid counterparts. Flow cytometry experiments could rapidly determine ploidy level in yeast. An increase in tetraploidy in a culture of hybrids with high spore viability would support the anti-recombination model.

What are the evolutionary forces driving dysfunctional regulation in hybrids?

As I have a short list of candidate genes that contribute to RI, I could not quantitatively answer whether certain genes or pathways are more susceptible to dysfunctional regulation. However the data can be analyzed in such a fashion. For my initial analysis in Chapter 3, I filtered genes using a conservative measurement for differential expression, and then measured for misexpression and cis/trans interactions. However the same analyses could be conducted on an unfiltered list of genes to help attain a more general view of hybrid misexpression or cis- and trans-divergence.

To verify a cis- or trans-interaction, we can perform promoter-swapping assays, in which chimeric constructs with one species’ promoter and another species’ genes can be placed in either parental background. If different promoters have different expression levels in the same parental background, this would be a cis-effect. If the promoters have the same expression levels in the same background, this would be a trans-effect. Promoter expression can be compared in the yeast hybrid background and parental backgrounds to confirm more complicated interactions.

As the cells progress through meiosis, I see a shift from trans- to cis-divergence between the parents. With this shift, there is an increase of cis*trans and compensatory interactions in the hybrid. Assuming an analysis on an unfiltered set of genes produces more cis*trans and compensatory interactions, we can determine whether these interactions belong to a rapidly evolving pathway as previously determined. If we are able to increase the list of genes that have cis-only affects, we can test for cis-regulatory evolution of a pathway. A combination of these analyses along with the functional analysis could reveal whether genes that contribute to RI are rapidly evolving and whether it is due to a cis- or trans-only or more complicated interaction.

As stated before, it is difficult to predict what genes or types of divergence contribute to RI; thus interactions besides cis*trans may contribute to RI. For instance I identified nuclear genes, MRS2 and MRS3 that splice mitochondrial genes, and have the same known incompatible gene, MRS1 (Chapter 3). MRS2 has allele specific expression in the hybrid, but does not have a noticeable cis/trans interaction.
Rather cis- only factors appear to effect MRS2 and other mitochondrial genes’ expression (Supplemental Data File). However MRS3 does have cis*trans affects. Assuming both MRS2 and MRS3 have the same results on RI as MRS1, we could not deduce what sorts of factors more often effect regulation and RI. If my analysis were to be performed on an unfiltered set of genes, we could identify more cis- and trans- divergence. We could randomly choose a given number of genes from each interaction (cis-only, trans-only, cis + trans, cis*trans, compensatory), and test their effect on RI in yeast using genetic methods, to ascertain which type of interaction, if any, is more likely to play a role in RI.

Additionally we could determine whether cis- and trans- regulatory changes coevolved for the sporulation pathway. Knowing whether cis- and trans- mutations are coevolving would help understand the evolution forces that drive the regulatory divergence. My allele-specific analysis (Figure 3-4) uncovers a shift from trans- to cis- divergence between parents over the course of meiosis. It has been suggested that the early significant regulatory changes between species arise in trans and have large effects on many genes, and subsequent cis-changes for genes affected by the original trans-effect can compensate for any deleterious effects or further contribute to any fitness advantage\(^\text{109}\). Thus we observe more trans-changes between species, and more cis- changes within species\(^\text{86,94,110}\). It has recently been suggested that cis- and trans- mutations coevolve as compensatory interactions between species to conserve pathways\(^\text{111}\).

My study shows that depending on when in a pathway species’ expression divergence is observed; we can observe either mostly cis- or trans- changes between species. In the context of a simple cascading pathway with a master regulator, such as yeast sporulation, it may not be a surprising result. The two species are differentially regulating their genes during mitosis most likely in accordance to how they have evolved over time, as seen by the abundance of trans-divergence. Once the parents reset their regulatory pathways to proceed through sporulation, they are expressing a similar pathway, which we see by the increase of cis- only divergence over time. Whether these cis- changes evolved after or coevolved with the trans- effects is unknown. If these mutations are polymorphic within each species, we can determine whether these cis- changes evolved after or coevolved with the trans- effects by determining using mutation rates and generation times to estimate the age a mutation evolved\(^\text{112}\). If
trans- mutations occurred first, those mutations will have occurred earlier than cis- mutations. If cis- and trans- mutations coevolved, their mutations on average will be the same age.

Final Thoughts

I believe we can understand the genetic contribute to RI by utilizing Saccharomyces as an evolutionary model. Genome-wide studies are excellent for speciation studies, because it gives us the opportunity to examine reproductive isolation in a powerful way. For the first time, researchers can study genomes and gene expression in organisms whose genomes are not yet sequenced. When studying a model organism like yeast, I have shown through the course of my thesis work that we can relate RI to misexpression in a hybrid and that we can compile a list of genes that could contribute to RI. Unsurprisingly next-generation sequencing has become extremely popular because of the amount of data it produces; however I would like to stress the point that our hypotheses that come from our genome-wide studies must be followed upon with functional assays. Only by examining a multitude of individual examples, will we be able to derive a relationship between genetic incompatibilities and RI. To this end, Saccharomyces serves as an excellent model to study speciation, with outstanding tools to study RI and speciation at both a genome-wide and individual gene level. Through a combination of these studies, we are in a position to more fully comprehend the genetic contribution to RI.
References


27. Shaw, K. L. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species
radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. 


51. McManus, C. J. *et al.* Regulatory divergence in Drosophila revealed by mRNA-seq. *Genome*


