INCOMPATIBILITIES IN MISMATCH REPAIR GENES *MLH1-PMS1* CONTRIBUTE TO A WIDE RANGE OF MUTATION RATES IN HUMAN ISOLATES OF BAKER'S YEAST

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

by

Vandana Raghavan

August 2019

INCOMPATIBILITIES IN MISMATCH REPAIR GENES *MLH1-PMS1* CONTRIBUTE TO A WIDE RANGE OF MUTATION RATES IN HUMAN ISOLATES OF BAKER'S YEAST

Vandana Raghavan, PhD

Cornell University, 2019

The mismatch repair (MMR) pathway maintains genome stability by repairing mutations incorporated in the genome during replication and recombination. While most microorganisms tend to have low mutation rates, a higher mutation rate can provide transient adaptive advantage to stress conditions by promoting adaptive mutations. Variants of the MMR genes, *MLH1* and *PMS1* from different yeast strains can display an incompatibility that results in a high mutation rate. MLH1 and PMS1 function as a heterodimer and the incompatibility is a result of single amino acid polymorphism in each protein. The incompatibility provides an adaptive advantage under stress but does so at the cost of long-term fitness. I identified 18 baker's yeast isolates from 1011 yeast isolates surveyed that contain the incompatible *MLH1-PMS1* genotype in a heterozygous state. I tested the mutation rates of two clinical heterozygous diploid isolates, YJS5885 and YJS5845, and their spore clones. While both of these isolates were non-mutators, their meiotic spore progeny displayed mutation rates that varied over a 340-fold range, and *MLH1-PMS1* incompatibility was the major driver of high mutation rate. The range in mutation rates might be in part because these isolates are heterozygous for several genes that may be modifying the mutation rate. My data are consistent with the variance in mutation rate contributing to adaptation to stress conditions through the acquisition of beneficial mutations, with high mutation rates leading to long-term fitness costs that are buffered by mating, or

eliminated through natural selection. Furthermore, I observed that one of the isolates was aneuploid and generated aneuploid spore clones at a high frequency. Aneuploidy also provides a transient adaptive advantage under stress conditions. Thus, I obtained evidence for mechanisms in clinical yeast isolates that may provide an adaptive advantage in the human body.

BIOGRAPHICAL SKETCH

The author was born in Chennai, India and grew up in New Delhi, India and lived there until she finished her undergraduate degree in Biomedical Sciences at Acharya Narendra Dev College, University of Delhi. The author's mother worked at a bank until retirement, where she was a highly valued employee because of her efficiency and friendliness and warmth towards her co-workers and customers. The author's mother also has a background and a keen interest in botony, which is shared by her daughters. The author's father is an engineer and he worked at Tata Consultancy services before he retired as Vice President, Strategy. Her sister is a microbiologist/immunologist who pursued her PhD in Washington University at St. Louis and currently is a post-doctoral researcher at New York University.

The author's interest in biological research led her to join Tata Institute of Fundamental Research (TIFR), Mumbai, India to pursue her Master's degree in Biological Sciences. At TIFR she was trained as a biochemist in Dr. B.J Rao's laboratory and worked to understand mechanisms of DNA repair in the green alga *Chlamydomonas reinhardtii*. In 2013 the author joined Cornell University as a PhD candidate. There she joined Dr. Eric Alani's laboratory where she continued working on DNA repair and recombination in yeast. The author also further developed her skills as a biochemist and yeast geneticist at the Alani Lab. Throughout her PhD the author took up several opportunities to do science outreach and aims to enter the field of science writing and science communication. The author also enjoys spending time outdoors hiking and gardening.

I would like to dedicate my thesis to my parents Srinivasan Raghavan and Uma Raghavan, my sister Varsha Raghavan and my partner Vishal Chaudhari.

ACKNOWLEDGEMENTS

I want to express my deepest gratitude towards my PhD advisor Dr. Eric Alani. He is a brilliant scientist, an enthusiastic mentor and most important of all, a patient and kind human being. I learn from him every single day. He has encouraged me at every step and has been extremely supportive in my time at Cornell. His leadership style is something to learn from. It amazes me how he is able to identify his student's strengths and weaknesses perhaps even before they do so themselves.

I also want to thank my thesis committee members, Dr. Paula Cohen and Dr. Joe Peters. I am lucky to have mentors like them who are heavily invested in my science as well as in my career. They have been encouraging through the years, and have given me extremely valuable critical suggestions. They were always able to make time to guide me. I really appreciate their support.

I also want to thank my fellow lab members of Alani Lab, both past and present. They have been a fun and smart group to work with. I especially want to thank Carol Manhart who taught me all the biochemistry experiments and patiently helped me in troubleshooting experiments. Also, Najla-Al-Sweel and Duyen Bui with whom I have worked with in different projects, they have been my friends as well as my critics.

I have been lucky to have amazing collaborators from whom I have learnt a lot. I am extremely grateful to Dr. Charles Aquadro and also in awe of his depth of knowledge and clarity of evolutionary biology and population genetics. He patiently explained me fundamental concepts of population genetics and I think I am very fortunate to have interacted with him and learnt from him.

I would also like to thank my friends at Cornell with whom I have shared enjoyable times and I know they will be my friends for life. I also would like to thank our GFAs- Casey, Vic and Ginger for being so friendly, communicative and helpful.

Importantly, I want to thank my family. My partner Vishal, who is also an incredibly smart and talented scientist, for his support, scientific discussions, critical feedback and affection that helped me become a better scientist and person. My sister, Varsha is also an exceptional scientist has been a great source of strength and inspiration and is my advocate for life. Most of all, my parents, who have been with me throughout all my challenges and deeply care for my welfare. They take pride in the smallest of my achievements and remind me to value the good things in life and be cheerful and grateful for what life has offered me.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 4

FOREWORD

The majority of the work described in my thesis is to understand different aspects of mismatch repair of DNA and crossing over during meiotic recombination both using an evolutionary perspective as well as to gain a mechanistic understanding of these basic cellular pathways.

In Chapter 1, I give an overview of adaptation mechanisms in clinical isolates of baker's yeast. This chapter is an introduction to the thesis and is also going to be submitted as a review. I have also included a future directions section based on the introduction and work done in Chapter 2.

In Chapter 2, I present the work on clinical isolates of baker's yeast bearing incompatibilities in the mismatch repair genes *MLH1* and *PMS1*. These isolates were diploids and heterozygous for the *MLH1-PMS1* incompatibility and I found that upon sporulation they generated progeny that had a wide range of mutation rates. *MLH1-PMS1* incompatibility was a major contributor to high mutation rates. Furthermore, I observed high rates of aneuploidies in the isolates and spore clones as well, pointing towards multiple adaptation phenotypes in these isolates. The future directions for Chapter 2 are presented as a part of Chapter 1.

In Chapter 3, I present the work I have done to understand the mechanism of the endonuclease, Mlh1-Mlh3 in yeast and mouse and how it functions in the two different cellular pathways: mismatch repair and meiotic recombination. Biochemical and genetic analyses of separation of function mutants of *mlh3* in yeast indicate that protein-protein interactions are key to directing Mlh1-Mlh3 in mismatch repair and recombination. The chapter includes the work

that I performed in two publications: Al-Sweel, Raghavan *et al.*, PLoS Genetics 2017;13(8); Toledo, Sun, Brieno-Enriquez, Raghavan *et al.,* PLoS Genet. 2019 In Press. It also includes a new protocol for the purification of yeast Mlh1-Mlh3 that I developed to substitute for an older protocol that we have since discontinued.

Chapter 4 includes the future directions for Chapter 3. Here I present my work with another meiotic recombination protein, Exo1 to understand how this protein interacts with Mlh1-Mlh3 and its effect on the endonuclease activity and crossover resolvase function of Mlh1-Mlh3, and I suggest future experiments for this project.

CREDITS

Chapter 2 was published in Genetics: Raghavan V, Bui DT, Al-Sweel N, Friedrich A, Schacherer J, Aquadro CF, Alani E. Incompatibilities in Mismatch Repair Genes *MLH1-PMS1* Contribute to a Wide Range of Mutation Rates in Human Isolates of Baker's Yeast. Genetics. 2018;210(4):1253-66. Epub 2018/10/24. doi: 10.1534/genetics.118.301550. PubMed PMID: 30348651; PubMed Central PMCID: PMC6283166. D.T.B. and E.A. assembled the Figures 2.1 and 2.2, D.T.B. did the experiment for Table 2.4, E.A. and V.R. did the work for Table 2.3, E.A. and NAS did the work for Table 2.1, E.A. did all the work for Figures 2.3A, 2.4 and Tables 2.2, 2.3, 2.9, A.F. did the work for Table 2.6. All other work was performed by V.R..

Chapter 3 was published in two papers: Al-Sweel N, Raghavan V, Dutta A, Ajith VP, Di Vietro L, Khondakar N, Manhart CM, Surtees JA, Nishant KT, Alani E. *mlh3* mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide. PLoS Genet. 2017;13(8):e1006974. doi: 10.1371/journal.pgen.1006974. PubMed PMID: 28827832.

Toledo M, Sun X, Brieno-Enriquez MA, Raghavan V, Gray S, Pea J, Milano CR,Venkatesh A, Patel L, Borst PL, Alani E, Cohen PE. A mutation in the endonuclease domain of mouse MLH3 reveals novel roles for MutLγ during crossover formation in meiotic prophase I. PLoS Genet. 2019 In press

I mainly included the work I performed: the biochemistry of the yeast separation of function mutants of Mlh1-mlh3 and purification of mouse MLH1-MLH3 and MLH1-MLH3-DN complexes.

CHAPTER 1

Clinical isolates of baker's yeast provide a model for how pathogenic yeasts adapt to stress.

This chapter is going to be submitted as a review with co-authors Chip Aquadro and Eric Alani.

Abstract

Recent global outbreaks of drug resistant fungi such as *Candida auris* are thought to be due at least in part to excessive use of antifungal drugs. The baker's yeast *Saccharomyces cerevisiae* has recently gained importance as an emerging opportunistic fungal pathogen that can cause infections in immune-compromised patients. Analyses of over a 1000 *S. cerevisiae* isolates are providing rich resources to better understand how fungi can grow in human environments. A large percentage of clinical *S. cerevisiae* isolates are heterozygous across many nucleotide sites, and a significant proportion are of mixed ancestry and/or are aneuploid or polyploid. Such features potentially facilitate adaptation to new environments. These observations provide strong impetus for expanding genomic and molecular studies on clinical and wild isolates to understand the prevalence of genetic diversity and instability generating mechanisms, how they are selected for, and how they are maintained.

Saccharomyces cerevisiae, a model for studying pathogenic yeast

Saccharomyces cerevisiae, or baker's yeast, is a well-studied single-cell model eukaryote used to understand a wide-variety of cellular pathways. It is found in many natural environments including trees, fruits and soil, is extensively used in industrial fermentation to make bread, beer, and wine (e.g.[1, 2]), and has been found in the respiratory, gastrointestinal, and urinary tracts of healthy individuals ([3-7], reviewed in [8]). It is not thought to adapt quickly to the changing conditions that occur in the human body, and its presence does not normally cause infections because it can be cleared by the immune system and does not cross epithelial barriers [3-7]. However, recent molecular and genomic analyses of *S. cerevisiae* natural **isolates** (see Glossary) have identified loci and mechanisms that promote genetic variability and genomic instability, and consequently aid in adaptation to stressful environments [8-11]. This review provides an overview of genomic processes that can promote adaptation of baker's yeast to changing environments and how such adaptation may lead to virulence in a human host.

S. cerevisiae, an opportunistic pathogen

Over the past 25 years, a significant effort has been made to collect and characterize *S. cerevisiae* isolated from patients (for examples see [11-17]). It has been designated as an emerging opportunistic pathogen that can cause infections in immune compromised patients, and is associated with virulence when the epithelial barrier is breached [5, 18]. The association of baker's yeast with human infection has become better recognized due to improved diagnostic methods. *S. cerevisiae* causes roughly 1-4% of severe fungal infections, with *Candida albicans* (~53-58%) and *Candida glabrata* (~20-23%) causing the majority [19]. There are a few extreme cases where *S. cerevisiae* is the direct cause of mortality, usually by inducing sepsis [19]. Baker's yeast infection is associated with several illnesses including pneumonia, peritonitis, esophagitis, and liver abscesses (reviewed in [20]). Furthermore, it can cause infections ranging from vaginitis in healthy patients, cutaneous infections, systemic bloodstream infections, and

infections of essential organs in immunocompromised and critically ill patients [5, 20-22]. Based on these reports, *S. cerevisiae* is considered a low-virulence human pathogen [8, 22].

Challenges for baker's yeast in the human host and clinical environment

The human body is a stressful growth environment for baker's yeast due to the presence of high temperature, antifungal agents, antimicrobials, and competing commensals and/or infectious microbes. These stresses are analogous to what yeast might experience in a wine fermentation environment, which is constantly fluctuating [23]. The patient in a clinical setting may thus pose a particularly challenging environment compared to yeast growing in a typical environment such as the bark of a tree or in a vineyard. Major factors that allow for opportunistic infection by *S. cerevisiae* are impaired host immune response, use of invasive and infected catheters in hospitals, antibiotic therapies to treat some infections, and probiotics that *S. cerevisiae* [20, 21]. Thus, it is not a surprise that the majority of infections caused by *S. cerevisiae* are hospital acquired. Apart from being a commensal in patients and health-care workers, it may be acquired from other infected patients and may appear in the hospital environment through probiotics and other food sources [5, 20-22].

Can human behavior impact adaptation mechanisms?

Indiscriminate use of antimicrobials is thought to be a leading cause of the generation of multi-drug resistant bacterial pathogens [24]. Recent work has suggested a similar situation in fungi [25, 26]. Excessive use of antifungals in farm animals and crops is thought to be a major source of drug-resistant fungi that cause infections. This is exemplified by global outbreaks of *Candida auris* within the last five years [25]. This pathogen is resistant to major antifungal medications and causes infections in patients with weak immunity such as critically ill patients, infants and patients on immune-suppressive medications. The commensal fungal population is important to maintain gut health and promote anti-fungal immunity [26]. However, the use of probiotics, especially in immune-compromised patients, is also thought to provide an advantage for specific

fungi to grow in a clinical environment. For example, *Saccharomyces boulardii*, a subtype of *S. cerevisiae* is used in probiotics for the treatment of diarrheal disease such as those caused by *Clostridium difficile* [5, 21]. These live preparations are often consumed at high doses for long periods of time and have been associated with invasive *S. cerevisiae* infections in immunosuppressed individuals [5, 21].

Phenotypes associated with virulence

Some of the earliest studies analyzing baker's yeast infections in mammals were performed by Clemons *et al*. [12], who compared the virulence of clinical and non-clinical human isolates by injecting them into immune-compromised mice. They found virulence was likely to be a dominant trait involving multiple loci [12]. Outlined below and in Table 1.1 are summaries of phenotypes that are important for growth in clinical conditions.

Antifungal resistance. The most common class of antifungals, azoles, targets enzymes in the ergosterol biosynthesis pathway. Ergosterol is the major sterol in the plasma and mitochondrial membranes of fungi. It is not present in mammals and is an established target of antifungals [27]. Resistance to antifungals can occur through the deletion or overexpression of genes that regulate membrane transporters that cause efflux of the drug, or modulate the expression of targets in the ergosterol pathway (Table 1.1; [28-32]). For example, in experimentally evolved populations (defined as cultures grown in a laboratory condition for a large number of generations) of *S. cerevisiae*, resistance to the antifungal fluconazole resulted from mutations in several targets [28]. At low drug concentrations, diploids acquired resistance more rapidly than haploids. This advantage was thought to be due to resistance resulting from dominant mutations and diploids having twice the number of mutational targets. At high drug concentrations, resistance occurred more quickly in haploids due to the acquisition of recessive mutations. In experimentally evolved populations of *C. albicans*, resistance to fluconazole resulted from an additional copy of isochromosome 5L, which became fixed in multiple independent populations [31].

High temperature growth. The ability to grow at high temperature aids in virulence of *S. cerevisiae*. Four genes have been linked to high temperature growth in *S. cerevisiae: NCS2, MKT1, END3* and *RHO2* [33, 34]. These, along with other unmapped loci, were identified through a targeted backcross mapping strategy and reciprocal hemizygosity analysis [33, 34].

C*olony phenotype switching.* Many clinical isolates display several colony phenotypes; these can include differences in smoothness, color, size and colony border color [13, 35]. Clemons *et al*. [35] observed that clinical isolates showed different colony phenotypes based on the media grown in, and such phenotypes were reversible. Colony phenotype switching was observed more often in clinical compared to non-clinical isolates [35]. Such a property likely facilitates colonizing different parts of the human body that provide different stress conditions and nutrient availability. It is also important to note that meiotic progeny of clinical isolates can display a wide variety of colony phenotypes in a single growth condition (for example, see [10]).

Pseudohyphal growth. In this condition, cells bud and become elongated but the buds do not separate, creating chains of cells that can invade the growth substrate [36]. This unipolar growth may be important under nutrient deprived conditions to identify food sources. Virulent isolates show significantly higher pseudohyphal growth compared to avirulent isolates [37].

Table 1.1: Phenotypes observed in fungi grown in clinical environments or conditions.

Phenotype	Description
Antifungal resistance	**1. Genes that act in drug efflux pathways.** A. Overexpression of genes in clinical isolates of *C. albicans* [32]. i. *CDR1* and *CDR2*: ABC (ATP-binding cassette) transporters ii. *MDR1*: Regulates intracellular protein transport B. Experimentally evolved *S. cerevisiae* [28]. i. *PDR1*: Mutations in this transcription factor cause overexpression of *PDR5* and *SNQ2* ii. *PDR3*: Mutations in this transcriptional activator alter expression of ABC transporters. **2. Overexpression of genes observed in experimentally evolved *S. cerevisiae* treated with fluconazole or amphotericin B.** Overexpression of the following genes was reported to cause increased drug efflux [29]: *ICT1* (phosphatidic acid biosynthesis), *YOR1* (ABC transporter), *GRE2* (catabolism of some sugars), *PDR16* (lipid synthesis), *YGR035C and YPL088W* (unknown).
	3. Modulation of expression of ergosterol biosynthesis pathway genes in experimentally evolved *S. cerevisiae* [28, 30]. A. Loss of function mutation in *ERG3* causes overexpression of *ERG11* which causes resistance to fluconazole B. Loss of function of *ERG6* causes resistance to fluconazole
	4. Aneuploidy of isochromosome 5L and trisomy of chromosome 7 in experimental evolution of *C. albicans* [31].
High temperature growth	Mutations in *NCS2, MKT1, END3* and *RHO2* in *S. cerevisiae* [33, 34].
Colony phenotype switching	Linked to clinical isolates of *S. cerevisiae* [13, 35]**.**
Pseudohyphal growth	Linked to clinical isolates of *S. cerevisiae* [37]**.**

Genetic mechanisms that contribute to baker's yeast virulence

Studies on natural and clinical isolates have identified different factors that aid in the adaptation of baker's yeast to stress environments (Figure 1.1). Many of these factors have been associated with exposure to human associated environments. These include: 1. **Multi-site heterozygosity,** which offers a wider exploration of phenotypes in the progeny of an organism. 2. **Mosaicism**, created when multiple isolates in a single environment outcross to create a novel isolate. 3. High mutation rates, which provide a source of mutations that can accelerate adaptation. 4. **Aneuploidy,** or change in the number of one or more chromosomes, provides a transient adaptive mechanism. 5. **Polyploidy**, which is also observed in natural isolates, and can generate genome instability and variability associated with rapid adaptation.

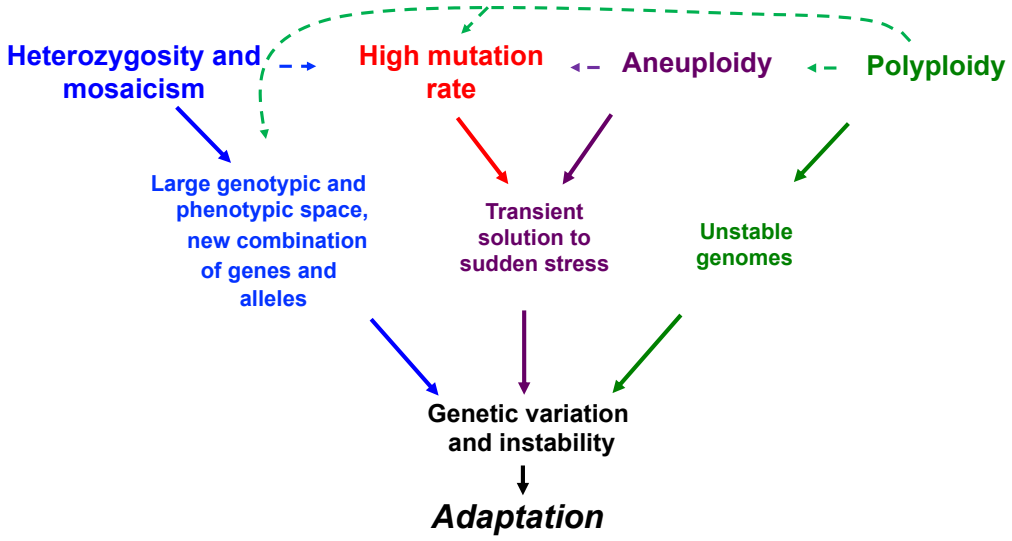


Figure 1.1. Factors aiding in adaptation of *S. cerevisiae* clinical isolates. Several interrelated factors that lead to genetic variation and genomic instability help in yeast adaptation to clinical environments. Heterozygosity and mosaicism of isolates results in spore clones accessing larger genotypic and phenotypic space and also give rise to new combinations of genes and alleles, which aid in adaptation to new environments. The allelic variation and hybrid incompatibility resulting from heterozygosity and mosaicism can also cause variation in mutation rates. Aneuploidy and high mutation rates provide a transient advantage in sudden stress conditions and aneuploidy generates huge phenotypic variations which may lead to high mutation rates. Polyploids, show genome instability phenotypes that include an increase in the frequencies of chromosome mis-segregation causing aneuploidies, a higher generation of mutations, and also provides a larger genotypic and phenotypic space, consequently displaying rapid adaptation [38].

1. Multi-site heterozygosity

Multi-site heterozygosity is defined as the presence of allelic variation at multiple loci in the same yeast strain. In the 1011 *S. cerevisiae* genome project [9], heterozygous isolates contained 2,000-78,000 single nucleotide sites that were heterozygous. More specifically, 63% of 794 diploid natural isolates and 49 of 107 clinical isolates were heterozygous at multiple nucleotide positions (Table 1.2). In general, higher levels of heterozygosity were associated with isolates recovered from human-associated environments [9, 11, 15, 39]. A simple explanation for the above findings is that actions by humans, such as consuming yeast and culturing them for beer and wine, increases the possibility of outcrossing by increasing stress conditions and providing proximity to other isolates [15, 39]. Thus, by inference, outcrossing between different isolates would lead to multi-site heterozygosity. With respect to clinical isolates, Muller and McCusker [39] showed that levels of heterozygosity at 12 microsatellite markers were higher in clinical compared to non-clinical isolates, and Magwene *et al*. [15] showed in a small sample that clinical isolates have a larger proportion of heterozygosity in their genomes compared to non-clinical isolates. However, it is important to note that Peter *et al.* [9] found a wide variation of heterozygosity in the 49 clinical isolates.

How is heterozygosity generated? Baker's yeast can reproduce asexually as well as enter meiosis infrequently for sexual reproduction to generate haploid spores. Once spores are formed, *S. cerevisiae* haploid cells have the potential to switch mating type (if **homothallic**) and **autodiploidize** to create a homozygous diploid or mate with other haploid progeny in the vicinity, termed as outcrossing. Heterozygosity at multiple loci can result by recent outcrossing events or from the accumulation of mutations during vegetative growth. Outcrossing in baker's yeast, which depends on environment and proximity to other isolates, has been estimated to occur between 1 in 100, to 1 in 50,000 vegetative divisions [23, 40], with high rates observed in the gut of social wasps [41]. A closely related yeast, *Saccharomyces paradoxus,* is unable to survive in the gut of social wasps, but conditions in the gut favor their sporulation and germination and most importantly, mating with *S. cerevisiae* to form hybrids that can survive

[41]. Mixed ancestry from two or more populations due to outcrossing would result in a larger proportion of heterozygosity in the genome and is termed mosaicism. While it is difficult to distinguish heterozygosity created by outcrossing from that created by mutation accumulation, the finding that 63% of wild isolates are heterozygotes despite a prevalence of asexual reproduction suggests that the heterozygous state is advantageous [9].

Advantages of heterozygosity. High levels of heterozygosity, especially in clinical isolates, points toward a **heterosis**-like advantage that allows for adaptation and survival in particular environments. Heterozygous genotypes allow meiotic progeny to explore larger genotypic and consequently phenotypic spaces, which can lead to survival under a variety of stress conditions [39, 42]. This advantage is demonstrated by the variation seen in the phenotypes of meiotic progeny of the clinical isolate YJM311 [37, 42, 43], and in the variation in mutation rate of meiotic progeny of the clinical isolates YJS5845 and YJS5885 ([10]; Table 1.3).

How is heterozygosity maintained? Baker's yeast is primarily homothallic and capable of self-mating, and so it seems surprising that high levels of heterozygosity are maintained in natural isolates. The frequency of **heterothallism** (inability to switch mating-type to form diploids) was reported to be higher in clinical isolates than non-clinical isolates [39]. For example, based on genotyping of the *HO* locus, Muller and McCusker [39] observed that four of eight clinical heterozygous isolates were heterothallic, indicating they would thus support maintenance of heterozygosity. The heterozygote advantage may select for the heterothallic phenotype in these isolates. Raghavan *et al.* [10] showed that one of two clinical isolates studied was functionally heterothallic, with its meiotic spore progeny remaining haploid. This isolate did not have any defects in the open reading frame of the *HO* gene, so there are likely to be mutations in other loci in this isolate that confer a heterothallic phenotype. Interestingly, Magwene *et al*. [15] reported in a set of 28 natural isolates that there was a seven-fold higher sporulation efficiency in homozygous compared to heterozygous isolates. Thus, lower sporulation efficiency may alsoplay a role in maintaining heterozygosity. It would be interesting to see if this correlation holds for the 1011 isolates studied in Peter *et. al.* [9].

Table 1.2: Adaptive mechanisms observed in *S. cerevisiae* clinical isolate studies.

Adaptive mechanism	Description
Multi-site heterozygosity	794 natural isolates [9] • 46% of 107 clinical origin isolates were heterozygous • 66% of 687 non-clinical origin isolates were heterozygous
	Higher levels of heterozygosity in clinical isolates [15] • 37%-59% of total SNPs were heterozygous in clinical isolates • 0.9-15% of total SNPs were heterozygous in other isolates
Mosaicism	1011 natural isolates [9] • 47% clinical isolates belong to mosaic clades • 11% non-clinical isolates belong to mosaic clades
	144 natural isolates [8] • Approximately 19% of the 132 clinical isolates belong to mosaic clades • None of the 12 non-clinical isolates are mosaics
	100 natural isolate segregants [11] • 63% of 43 clinical strains belong to mosaic clades • 33% of 57 non-clinical strains belong to mosaic clades
High mutation rate	Clinical isolates YJS5885 and YJS5845 with a genetic incompatibility in mismatch repair genes *MLH1* and *PMS1* generate spore clones with a range of mutation rates [10].
Aneuploidy	1011 natural isolates [9] • 19% of 904 non-clinical isolates are aneuploid, 7.2% of which have aneuploidies in more than one chromosome (two to nine). • 17% of 107 clinical isolates are aneuploid, 50% of which have aneuploidies in more than one chromosome (two to seven).
	144 natural isolates [8]. 36% contained aneuploidies, 132 of which are clinical isolates. Eight clinical isolates contained multiple aneuploidies.
	93 natural isolate segregants [11] • 5% of 47 clinical strains were aneuploid. • 10% of 50 non-clinical strains were aneuploid. Two contained multiple aneuploidies (one with two, the other with three).
	47 natural isolates [44] • 23.6% of 37 non-clinical natural isolates are aneuploid. Two of them had aneuploidies in multiple chromosomes. One had aneuploidy in two chromosomes and the other in four chromosomes. • 30% of 10 clinical isolates are aneuploid. One had aneuploidy in two chromosomes.
Polyploidy	794 natural isolates [9] 16% of 505 non-clinical isolates are polyploid 8% of 107 clinical isolates are polyploid

	32% of 144 natural isolates are polyploid, 132 of which are clinical isolates [8].
	30% of both clinical and non-clinical isolates are polyploid in a total of 137 isolates [39].

Table 1.3: Heterozygosity in an isolate results in phenotypic range in its progeny: a few examples.

Isolate	Phenotype of spore clones
YJM311	1. Varied resistance to the antifungal drug fluconazole [42]. Determined by measuring MIC_{50} (minimum concentration of fluconazole that inhibits 50% growth) for 288 homozygous diploid spore clones. 2. Colony biofilm complexity [42, 43]. Colony morphologies of 288 spore clones were categorized visually ranging from simple, non-biofilm phenotype to highly complex biofilm colony morphology. 3. Varied invasive growth on agar [42]. Assessed by growing yeast on agar containing low ammonium, washing the surface growing cells and quantifying the levels of invasive growth. 4. Varied growth at 42ºC [37, 42]. Determined by measuring growth on agar plates at 42ºC for the 288 spores.
YJS5885	1. 120-fold variation in mutation rate in 12 spore clones [10]. Determined by a frameshift reversion assay involving a plasmid containing a frameshift mutation in the gene encoding resistance for Geneticin. 2. Variation in colony sizes in 12 spore clones as determined by growing on agar plates at 30ºC for 2 days [10].
YJS5845	1. 340-fold variation in mutation rate in 11 spore clones [10]. 2. Variation in colony sizes in 11 spore clones [10].

2. Mosaicism

Different isolates can often colonize and cause infection in individual immune-deficient patients. Such multiple isolate infections can provide an opportunity for isolates to outcross and generate mosaics [8, 11, 16, 45]. Isolates are classified as mosaics if their genomes contain sequences derived from more than one genetically diverse ancestor. Importantly, they may also contain high levels of genomic heterozygosity if the outcrossing events occurred relatively recently. Human influence is likely to play a major role in bringing different isolates together and generating stress conditions that promote outcrossing between such isolates. Thus, it is not a surprise that a majority of the isolates that are mosaics are isolated from human-related (e.g., clinical and wine isolates) rather than natural environments ([8, 9, 11], Table 1.2). Additionally, a high proportion of clinical isolates belong to mosaic clades; 47% of clinical isolates compared to 11% of non-clinical isolates, belong to mosaic clades, though it is important to note that a significant number of clinical isolates are part of the wine clade ([8, 9, 11], Table 1.2).

Outcrossing between different isolates with varying degrees of adaptive potential will likely lead to the generation of mosaic isolates with higher adaptive potentials because they would contain new and unexplored combinations of genes and alleles that facilitate adaptation to new environments. Mosaicism may also aid in adaptation by providing a phenotypic range in the progeny as mentioned above for heterozygosity.

3. High mutation rates

High mutation rates can accelerate adaptation to stress conditions because they provide an elevated mutation supply that can more rapidly yield beneficial mutations. Bacteria that display high mutation rates are frequently found in nature [46-53]; however, modeling analyses and molecular studies indicate that bacteria prevent the long-term fitness cost of accumulating deleterious mutations through the horizontal transfer of genes that restore a low mutation rate [48, 52, 54]. Horizontal gene transfer events are rare in fungi [55, 56], and baker's yeast active mutators have yet to be isolated in natural environments. Recently, active mutators were

identified in clinical isolates of the human fungal pathogen *Cryptococcus* that contain mutations in the mismatch repair gene *MSH2* [57, 58], demonstrating that a high mutation rate can provide beneficial mutations for adaptation in fungi despite the associated fitness costs.

Outcrossing between isolates with high sequence divergence can result in the creation of novel mutator combinations of non-mutator variants in different genes. For example, mating between two laboratory baker's yeast strains can yield progeny that display mutator phenotypes due to the presence of an incompatible combination of the *MLH1* and *PMS1* mismatch repair genes which act in a highly conserved pathway to remove DNA replication errors [10, 59]. The proteins Mlh1 and Pms1 function as a heterodimer, and a specific incompatible combination of single amino-acid polymorphisms in Mlh1 and Pms1 results in elevated mutation rates [60], which can provide an adaptive advantage to stress [61]. In the heterozygous clinical isolates YJS5885 and YJS5845, which are not mutators, *MLH1-PMS1* incompatibility acts as a major contributor to high mutation rates seen in spore clones derived from either isolate [10]. Interestingly, these spore clones displayed a wide range of mutation rates, indicating the presence of extragenic suppressors and enhancers of mutation rate ([10], Figure 1.2A).

MLH1-PMS1 incompatibility allele combinations are rare in nature, most likely due to the detrimental effects of defects in mismatch repair on fitness [59]. Analysis of the patterns of sequence polymorphisms in DNA encompassing the *PMS1* locus provided evidence that recombination to generate incompatible genotypes had occurred in the past, suggesting that natural isolates of baker's yeast do mate and occasionally produce the incompatible genotype [60]. Only one isolate among 1011 natural isolates was found to be homozygous for the incompatibility genotype, but a spore clone of this isolate had acquired suppressor mutations and was not a mutator [9, 45, 59]. However, a diploid strain containing an *MLH1-PMS1* incompatibility in the heterozygous state may have an advantage because the incompatibility is recessive and the effect of incompatibility is only observed in spore progeny [10]. The presence of the *MLH1-PMS1* incompatibility could thus provide a transient advantage for adaptation, with

mutator spore clones adapting to a stress condition and escaping fitness costs by acquiring suppressors, mating to nearby spore clones, or outcrossing to become non-mutators [10].

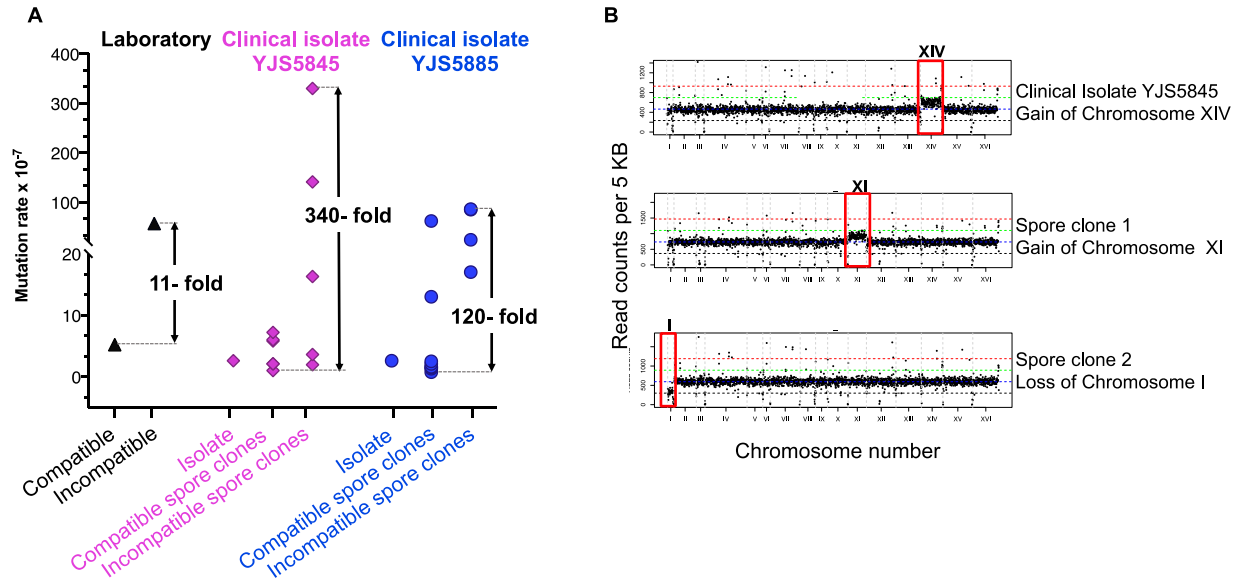


Figure 1.2. Mutation rate variation and aneuploidy in a clinical isolate. A. Mutation rates were determined using a plasmid-based frameshift reversion assay of an *MLH1-PMS1* compatible laboratory strain (S288c) and an isogenic derivative containing an incompatible *MLH1-PMS1* combination. The mutation rates of the heterozygous diploid clinical isolates YJS5845 and YJS5885 were also determined, as well as the rates for six compatible and five incompatible spore clones of YJS5845 and eight compatible and four incompatible spore clones of YJS5885. Larger variations in mutation rate were observed between incompatible and compatible spore clones of YJS5845 (340-fold) and YJS5885 (120-fold) compared to the S288c laboratory strain (11-fold). B. Whole genome sequencing of clinical isolate YJS5845 and 16 spore clones identified aneuploidy in the isolate and two of the spore clones. This isolate had a trisomy of chromosome XIV, one spore clone had a trisomy of chromosome XI, and the other spore clone had a monosomy of chromosome I. Figure adapted from Raghavan *et al*. [10].

4. Aneuploidy

A change in the chromosome number from the euploid set is referred to as aneuploidy; in most cases such events are deleterious to the cell. Effects of aneuploidy could be mediated by a direct change in the expression of genes on the aneuploid chromosome due to copy number variation, or be an indirect effect due to a change in expression of a gene that regulates targets located throughout the genome [62]. Additionally, there could be a general effect of aneuploidy that is not specific to a particular chromosome. As shown by Torres *et al.* [63], an extra copy of almost any yeast chromosome caused a reduction in cellular proliferation, which was attributed to altered levels of gene products encoded by genes that reside on the extra chromosome. In organisms such as *Drosophila*, *C. elegans*, mice, plants and humans, most aneuploidies are lethal [64]. Aneuploidies are observed in almost all cancerous cells, and it is debated whether they are a consequence of chromosome instability and segregation defects, or they are direct drivers of cellular transformation [64].

Despite conferring negative fitness effects, aneuploidy has been commonly observed in natural baker's yeast and may provide an important route to natural genetic variation. Aneuploidy has been suggested to help in adaptation in environments with human association such as in brewing, baking and wine strains of yeast [65]. In a study of 1011 natural isolates, Peter *et al.* [9] observed aneuploidy in 19% of isolates, and Strope *et al*. [11] reported 7.5% aneuploidy in 100 *S. cerevisiae* isolates obtained from various different geographical sources (Table 1.2, with other examples provided). Furthermore, aneuploid isolates may be more likely to generate aneuploid progeny, as seen in the clinical isolate, YJS5845 ([10], Figure 1.2B). This isolate was a mix of aneuploid and euploid cells, probably generated due to mitotic chromosome segregation defects. When sporulated in the laboratory, two of sixteen spore clones displayed aneuploidies in different chromosomes [10].

In certain circumstances, the beneficial effects of aneuploidy could offset negative effects on cell survival [66, 67]. Aneuploidy appears to be one of the first lines of defense to stress that increases chances for survival under strong and abrupt selective pressures [67]. In baker's yeast,

sudden stress conditions induced in the laboratory selected for aneuploidy, with loss of aneuploidy occurring when the stressor was removed [67]. When the stress condition was maintained, aneuploidy was lost in about 2,000 generations, with a more stable solution obtained through mutations that modified the expression of stress resistance genes, as seen by Yona *et al*. for heat stress [44, 67]. Thus, aneuploidy may be a valuable mechanism for adaptation in clinical isolates, which live in an environment with variable and sudden stressors.

How is aneuploidy tolerated? In many situations, an extra chromosomal copy provides tolerance to a particular stress condition [66]. For example, extra chromosomal copies confer resistance to heat (Chr. III), high pH (Chr. V), and the ultra violet light mimetic mutagen 4-Nitroquinoline 1-oxide (Chr. XIII) [66, 67]. Aneuploid isolates appear to better tolerate chromosome gains or losses due to dosage compensation mechanisms [44, 68]. In support of this, Hose *et al.* [44] analyzed twelve isolates containing an extra chromosome. They found that for 40% of genes located on such a chromosome, gene expression levels were lower than expected based on their dosage [44]. Furthermore, in contrast to artificially created aneuploids in the laboratory, natural aneuploids had growth rates similar to closely related euploids [44].

5. Polyploidy

Baker's yeast is most commonly diploid in nature, but polyploidy has been associated with human interference in natural isolates of yeast [8, 9]. Peter *et al*. [9] found that 87% of 794 natural isolates were diploid and 11.4% were polyploid. These polyploid isolates were enriched in human associated environments including beer, mixed-origin and African palm wine clades, although there was no significant enrichment in clinical isolates ([9], Table 1.2). Muller and McCusker [39] analyzed a diverse natural population and estimated that 70–80% of their isolates were diploid, with the remaining 20–30% of isolates being triploid or tetraploid, and Cubillos *et al*. [69] found that 95% of a group of more than 200 wine strains were diploid. Polyploidy has also been observed in clinical isolates [8, 39]. Zhu *et al*. [8] showed in a study of 144 isolates, the majority of which were clinical (132), that 34% were polyploid.

In baker's yeast, polyploid genomes are less stable than haploid and diploid genomes, and have been hypothesized to act as drivers of adaptation [39]. Polyploid baker's yeast display genetic instability phenotypes that are thought to arise as a result of their tendency to mis-segregate chromosomes [38, 70-72]. Consistent with this idea, polyploid baker's yeast genomes are associated with a more than a two-fold increase in aneuploidy [8] and also display higher mutation rates [38, 73]. Such ploidy-specific effects are thought to facilitate adaptation by creating larger genotypic and phenotypic spaces, and the higher numbers of chromosomes in polyploids are thought to buffer the effects of deleterious mutations [38, 73]. Laboratory evolution experiments support these ideas, but such support depends on the stress condition used [38, 74]. For example, Selmecki *et al.* [38] compared adaptation of tetraploid, diploid, and haploid *S. cerevisiae* to a low carbon environment and found that tetraploids adapted significantly faster. This more rapid adaptation was due to a higher rate of beneficial mutations as well as a higher fitness conferred by the mutations [38]. In contrast, Gerstein *et al.* showed in *C. albicans* that both higher and lower ploidy states can be advantageous under different stress conditions [74]. Since polyploidy can be beneficial in adaptation by multiple mechanisms, it may be used by clinical isolates to adapt to the unpredictable environment of the human host.

Conclusions and Future Directions

Adaptation to novel stress environments often requires organisms to incur genetic changes. Such changes can be accelerated in organisms by high mutation rates, acquiring specific stress resistance genes through horizontal gene transfer, or genomic instabilities that lead to chromosome gains and losses. Clinical yeast isolates have been subjected to novel and repeated stresses aimed at suppressing their growth and survival. Thus, they must be able to rapidly and repeatedly adapt to stress conditions to survive such conditions. As summarized in this review, multi-site heterozygosity, mosaicism, transient high mutation rates, aneuploidy, and polyploidy appear to be major sources of genetic raw material for adaptation in baker's yeast.

Analyzing more extensive collections of *S. cerevisiae* isolates from hospitals will be critical to determine more precisely the prevalence of phenotypes and genetic mechanisms that lead to these phenotypes. A plasmid-based mutator assay based on reversion to antibiotic resistance can allow for a more rapid analysis of mutator phenotypes in natural/clinical isolates and their spore clones [10, 59]. Such an approach may also be used to identify loci in natural isolates that impact mutation rate [75-77].

Molecular evolution studies performed with baker's yeast grown under different stress conditions have identified beneficial mutations that drive adaptation and propagate the population [75, 76]. These studies were performed by barcoding individual cells and tracking their lineages through whole genome sequencing to identify adaptive mutations. Such methods will be helpful to identify adaptive mutations in clinical yeast isolates grown in a mouse model.

It is clear that excessive use of antifungals plays an important role in the generation of fungal infections. Their use can result in the development of multi-drug resistant fungi [25, 26] and also affect the protective commensal fungal population that aids in immunity [26]. The use of live cultures of *S. cerevisiae* and *S. boulardii* in probiotics has also resulted in infections in immunocompromised patients [5, 21, 77]. The indiscriminate use of these antifungals should be restricted; we need to re-think the use of *S. cerevisiae* in probiotic preparations. Additionally,

learning more about adaptation mechanisms and genetic signatures of clinical isolates would help in generation of anti-fungal drugs to fight these infections.

Acknowledgements

We are grateful to Joseph Schacherer for discussions and comments on the manuscript and Michael McGurk for discussions. V.R. and E.A. were supported by National Institutes of Health (NIH) grant GM-053085. C.F.A. was supported by NIH grant GM-095793.

Outstanding questions

How prevalent are phenotypes such as high temperature growth, resistance to antifungal drugs, colony phenotype switching, and pseudohyphal growth in clinical yeast isolates? Have all loci been identified that are responsible for these phenotypes?

How frequently are heterozygosity, mosaicism, high mutation rates, aneuploidy and polyploidy seen in clinical isolates that facilitate adaptation?

Can loci be mapped that enhance or suppress mutation rates in natural isolates and their spore clones?

Adaptive evolution experiments have yet to be performed for clinical yeast isolates grown in an immune-compromised mouse model. What are the evolutionary landscapes and mutational signatures that accompany adaptation of clinical yeast isolates in a mammalian host environment?

Glossary

Isolate	An isolate is found in nature, whereas a strain has been manipulated in the laboratory.
Outcrossing	Mating between genetically different strains/isolates.
Mosaic	Mixed genotype derived from two or more genetically different populations. An isolate is classified as mosaic when it has multiple sources of ancestry and when it has less than 60% ancestry from any one population [11]. Mosaic isolates are polymorphic for a majority of segregating sites and are derived from outbreeding [9, 14]. They frequently manifest as isolated branches on a phylogenetic tree [9].
Homothallic	Baker's yeast cells that have the ability to undergo mating type switching. This allows mating (auto-diploidization) between mother and daughter cells.
Heterothallic	Baker's yeast cells that cannot undergo mating type switching and can be maintained as stable haploids.
Autodiploidization	Mating between mother and daughter cells after mating type switching.
Heterosis	Higher fitness of the hybrid resulting from mating between different strains/isolates. Also referred to as hybrid vigor.
Multi-site heterozygosity	Allelic variation seen in the same individual at multiple genetic loci. Peter *et al*. [9] classified heterozygous isolates having more than 5% heterozygous sites out of a total number of SNPs (single nucleotide polymorphisms) compared to the reference genome S288c. In 1011 isolates, heterozygosity ranged from 0.63-6.56 SNP sites per kilobase in heterozygous isolates [9]. There were 2,000-78,000 total heterozygous SNPs in the 1011 isolates [9].
Aneuploidy	Loss or gain of one or few chromosomes that causes a change in the chromosome number from a multiple of the haploid set.

Polyploidy	Having more than two complete sets of chromosomes, e.g.: triploid has three copies of each chromosome.

References

1. Cavalieri, D., McGovern, P.E., Hartl, D.L., Mortimer, R. and Polsinelli, M. (2003) Evidence for *S. cerevisiae* fermentation in ancient wine. J Mol Evol 57, S226-32 DOI: 10.1007/s00239-003-0031-2.

2. McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R. *et al.* (2004) Fermented beverages of pre- and proto-historic China. Proc Natl Acad Sci U S A 101, 17593-8 DOI: 10.1073/pnas.0407921102.

3. Ackerman, A.L. and Underhill, D.M. (2017) The mycobiome of the human urinary tract: potential roles for fungi in urology. Ann Transl Med 5, 31 DOI: 10.21037/atm.2016.12.69.

4. Cui, L., Lucht, L., Tipton, L., Rogers, M.B., Fitch, A. *et al.* (2015) Topographic diversity of the respiratory tract mycobiome and alteration in HIV and lung disease. Am J Respir Crit Care Med 191, 932-42 DOI: 10.1164/rccm.201409-1583OC.

5. de Llanos, R., Llopis, S., Molero, G., Querol, A., Gil, C. *et al.* (2011) *In vivo* virulence of commercial *Saccharomyces cerevisiae* strains with pathogenicity-associated phenotypical traits. Int J Food Microbiol 144, 393-9 DOI: 10.1016/j.ijfoodmicro.2010.10.025.

6. Nash, A.K., Auchtung, T.A., Wong, M.C., Smith, D.P., Gesell, J.R. *et al.* (2017) The gut mycobiome of the Human Microbiome Project healthy cohort. Microbiome 5, 153 DOI: 10.1186/s40168-017-0373-4.

7. Pillai, U., Devasahayam, J., Kurup, A.N. and Lacasse, A. (2014) Invasive *Saccharomyces cerevisiae* infection: a friend turning foe? Saudi J Kidney Dis Transpl 25, 1266-9.

8. Zhu, Y.O., Sherlock, G. and Petrov, D.A. (2016) Whole Genome Analysis of 132 Clinical *Saccharomyces cerevisiae* Strains Reveals Extensive Ploidy Variation. G3 6, 2421-34 DOI: 10.1534/g3.116.029397.

9. Peter, J., De Chiara, M., Friedrich, A., Yue, J.X., Pflieger, D. *et al.* (2018) Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature 556, 339-344 DOI: 10.1038/s41586-018-0030-5.

10. Raghavan, V., Bui, D.T., Al-Sweel, N., Friedrich, A., Schacherer, J. *et al.* (2018) Incompatibilities in Mismatch Repair Genes MLH1-PMS1 Contribute to a Wide Range of

Mutation Rates in Human Isolates of Baker's Yeast. Genetics 210, 1253-1266 DOI: 10.1534/genetics.118.301550.

11. Strope, P.K., Skelly, D.A., Kozmin, S.G., Mahadevan, G., Stone, E.A. *et al.* (2015) The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. Genome Res 25, 762-74 DOI: 10.1101/gr.185538.114.

12. Clemons, K.V., McCusker, J.H., Davis, R.W. and Stevens, D.A. (1994) Comparative pathogenesis of clinical and nonclinical isolates of *Saccharomyces cerevisiae*. J Infect Dis 169, 859-67.

13. Diezmann, S. and Dietrich, F.S. (2009) *Saccharomyces cerevisiae*: population divergence and resistance to oxidative stress in clinical, domesticated and wild isolates. PLoS One 4, e5317 DOI: 10.1371/journal.pone.0005317.

14. Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L. *et al.* (2009) Population genomics of domestic and wild yeasts. Nature 458, 337-41 DOI: 10.1038/nature07743.

15. Magwene, P.M., Kayikci, O., Granek, J.A., Reininga, J.M., Scholl, Z. *et al.* (2011) Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 108, 1987-92 DOI: 10.1073/pnas.1012544108.

16. Schacherer, J., Shapiro, J.A., Ruderfer, D.M. and Kruglyak, L. (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. Nature 458, 342-5 DOI: 10.1038/nature07670.

17. Muller, L.A., Lucas, J.E., Georgianna, D.R. and McCusker, J.H. (2011) Genome-wide association analysis of clinical vs. nonclinical origin provides insights into *Saccharomyces cerevisiae* pathogenesis. Mol Ecol 20, 4085-97 DOI: 10.1111/j.1365-294X.2011.05225.x.

18. Perez-Torrado, R., Llopis, S., Jespersen, L., Fernandez-Espinar, T. and Querol, A. (2012) Clinical *Saccharomyces cerevisiae* isolates cannot cross the epithelial barrier in vitro. Int J Food Microbiol 157, 59-64 DOI: 10.1016/j.ijfoodmicro.2012.04.012.

19. Piarroux, R., Millon, L., Bardonnet, K., Vagner, O. and Koenig, H. (1999) Are live *Saccharomyces* yeasts harmful to patients? Lancet 353, 1851-2 DOI: 10.1016/S0140-6736(99)02001-2.

20. Munoz, P., Bouza, E., Cuenca-Estrella, M., Eiros, J.M., Perez, M.J. *et al.* (2005) *Saccharomyces cerevisiae* fungemia: an emerging infectious disease. Clin Infect Dis 40, 1625-34 DOI: 10.1086/429916.

21. Enache-Angoulvant, A. and Hennequin, C. (2005) Invasive *Saccharomyces* infection: a comprehensive review. Clin Infect Dis 41, 1559-68 DOI: 10.1086/497832.

22. Perez-Torrado, R. and Querol, A. (2015) Opportunistic Strains of *Saccharomyces cerevisiae*: A Potential Risk Sold in Food Products. Front Microbiol 6, 1522 DOI: 10.3389/fmicb.2015.01522.

23. Marsit, S. and Dequin, S. (2015) Diversity and adaptive evolution of *Saccharomyces* wine yeast: a review. FEMS Yeast Res 15, pii: fov067 DOI: 10.1093/femsyr/fov067.

24. Davies, J. and Davies, D. (2010) Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev 74, 417-33 DOI: 10.1128/MMBR.00016-10.

25. Chowdhary, A., Sharma, C. and Meis, J.F. (2017) Candida auris: A rapidly emerging cause of hospital-acquired multidrug-resistant fungal infections globally. PLoS Pathog 13, e1006290 DOI: 10.1371/journal.ppat.1006290.

26. Leonardi, I., Li, X., Semon, A., Li, D., Doron, I. *et al.* (2018) CX3CR1(+) mononuclear phagocytes control immunity to intestinal fungi. Science 359, 232-236 DOI: 10.1126/science.aao1503.

27. Kodedova, M. and Sychrova, H. (2015) Changes in the Sterol Composition of the Plasma Membrane Affect Membrane Potential, Salt Tolerance and the Activity of Multidrug Resistance Pumps in *Saccharomyces cerevisiae*. PLoS One 10, e0139306 DOI: 10.1371/journal.pone.0139306.

28. Anderson, J.B., Sirjusingh, C. and Ricker, N. (2004) Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. Genetics 168, 1915-23 DOI: 10.1534/genetics.104.033266.

29. Anderson, J.B., Sirjusingh, C., Syed, N. and Lafayette, S. (2009) Gene expression and evolution of antifungal drug resistance. Antimicrob Agents Chemother 53, 1931-6 DOI: 10.1128/AAC.01315-08.

30. Bhattacharya, S., Esquivel, B.D. and White, T.C. (2018) Overexpression or Deletion of Ergosterol Biosynthesis Genes Alters Doubling Time, Response to Stress Agents, and Drug Susceptibility in *Saccharomyces cerevisiae*. MBio 9, DOI: 10.1128/mBio.01291-18.

31. Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B. and Berman, J. (2009) Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. PLoS Genet 5, e1000705 DOI: 10.1371/journal.pgen.1000705.

32. White, T.C., Holleman, S., Dy, F., Mirels, L.F. and Stevens, D.A. (2002) Resistance mechanisms in clinical isolates of Candida albicans. Antimicrobial Agents and Chemotherapy 46, 1704-1713 DOI: 10.1128/Aac.46.6.1704-1713.2002.

33. Sinha, H., David, L., Pascon, R.C., Clauder-Munster, S., Krishnakumar, S. *et al.* (2008) Sequential elimination of major-effect contributors identifies additional quantitative trait loci conditioning high-temperature growth in yeast. Genetics 180, 1661-70 DOI: 10.1534/genetics.108.092932.

34. Steinmetz, L.M., Sinha, H., Richards, D.R., Spiegelman, J.I., Oefner, P.J. *et al.* (2002) Dissecting the architecture of a quantitative trait locus in yeast. Nature 416, 326-30 DOI: 10.1038/416326a.

35. Clemons, K.V., Hanson, L.C. and Stevens, D.A. (1996) Colony phenotype switching in clinical and non-clinical isolates of *Saccharomyces cerevisiae*. J Med Vet Mycol 34, 259-64.

36. Halme, A., Bumgarner, S., Styles, C. and Fink, G.R. (2004) Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. Cell 116, 405-15.

37. McCusker, J.H., Clemons, K.V., Stevens, D.A. and Davis, R.W. (1994) *Saccharomyces cerevisiae* virulence phenotype as determined with CD-1 mice is associated with the ability to grow at 42 degrees C and form pseudohyphae. Infect Immun 62, 5447-55.

38. Selmecki, A.M., Maruvka, Y.E., Richmond, P.A., Guillet, M., Shoresh, N. *et al.* (2015) Polyploidy can drive rapid adaptation in yeast. Nature 519, 349-52 DOI: 10.1038/nature14187.

39. Muller, L.A. and McCusker, J.H. (2009) Microsatellite analysis of genetic diversity among clinical and nonclinical *Saccharomyces cerevisiae* isolates suggests heterozygote advantage in clinical environments. Mol Ecol 18, 2779-86 DOI: 10.1111/j.1365-294X.2009.04234.x.

40. Ruderfer, D.M., Pratt, S.C., Seidel, H.S. and Kruglyak, L. (2006) Population genomic analysis of outcrossing and recombination in yeast. Nat Genet 38, 1077-81 DOI: 10.1038/ng1859.

41. Stefanini, I., Dapporto, L., Berna, L., Polsinelli, M., Turillazzi, S. *et al.* (2016) Social wasps are a *Saccharomyces* mating nest. Proc Natl Acad Sci U S A 113, 2247-51 DOI: 10.1073/pnas.1516453113.

42. Magwene, P.M. (2014) Revisiting Mortimer's Genome Renewal Hypothesis: heterozygosity, homothallism, and the potential for adaptation in yeast. Adv Exp Med Biol 781, 37-48 DOI: 10.1007/978-94-007-7347-9_3.

43. Granek, J.A., Murray, D., Kayrkci, O. and Magwene, P.M. (2013) The genetic architecture of biofilm formation in a clinical isolate of *Saccharomyces cerevisiae*. Genetics 193, 587-600 DOI: 10.1534/genetics.112.142067.

44. Hose, J., Yong, C.M., Sardi, M., Wang, Z., Newton, M.A. *et al.* (2015) Dosage compensation can buffer copy-number variation in wild yeast. Elife 4, e05462 DOI: 10.7554/eLife.05462.

45. Skelly, D.A., Magwene, P.M., Meeks, B. and Murphy, H.A. (2017) Known mutator alleles do not markedly increase mutation rate in clinical *Saccharomyces cerevisiae* strains. Proc Biol Sci 284, pii: 20162672 DOI: 10.1098/rspb.2016.2672.

46. Boe, L., Danielsen, M., Knudsen, S., Petersen, J.B., Maymann, J. *et al.* (2000) The frequency of mutators in populations of *Escherichia coli*. Mutat Res 448, 47-55.

47. Chao, L. and Cox, E.C. (1983) Competition between High and Low Mutating Strains of *Escherichia coli*. Evolution 37, 125-134 DOI: 10.1111/j.1558-5646.1983.tb05521.x.

48. Denamur, E., Lecointre, G., Darlu, P., Tenaillon, O., Acquaviva, C. *et al.* (2000) Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. Cell 103, 711-21.

49. LeClerc, J.E., Li, B., Payne, W.L. and Cebula, T.A. (1996) High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science 274, 1208-11.

50. Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H. *et al.* (1997) Role of mutator alleles in adaptive evolution. Nature 387, 700-2 DOI: 10.1038/42696.

51. Tanaka, M.M., Bergstrom, C.T. and Levin, B.R. (2003) The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. Genetics 164, 843-54.

52. Townsend, J.P., Nielsen, K.M., Fisher, D.S. and Hartl, D.L. (2003) Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. Genetics 164, 13-21.

53. Giraud, A., Radman, M., Matic, I. and Taddei, F. (2001) The rise and fall of mutator bacteria. Curr Opin Microbiol 4, 582-5.

54. Giraud, A., Matic, I., Tenaillon, O., Clara, A., Radman, M. *et al.* (2001) Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. Science 291, 2606-8 DOI: 10.1126/science.1056421.

55. Fitzpatrick, D.A. (2012) Horizontal gene transfer in fungi. FEMS Microbiol Lett 329, 1-8 DOI: 10.1111/j.1574-6968.2011.02465.x.

56. Hall, C., Brachat, S. and Dietrich, F.S. (2005) Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. Eukaryot Cell 4, 1102-15 DOI: 10.1128/EC.4.6.1102-1115.2005.

57. Billmyre, R.B., Clancey, S.A. and Heitman, J. (2017) Natural mismatch repair mutations mediate phenotypic diversity and drug resistance in Cryptococcus deuterogattii. Elife 6, e28802 DOI: 10.7554/eLife.28802.

58. Boyce, K.J., Wang, Y., Verma, S., Shakya, V.P.S., Xue, C. *et al.* (2017) Mismatch Repair of DNA Replication Errors Contributes to Microevolution in the Pathogenic Fungus Cryptococcus neoformans. MBio 8, e00595-17 DOI: 10.1128/mBio.00595-17.

59. Bui, D.T., Friedrich, A., Al-Sweel, N., Liti, G., Schacherer, J. *et al.* (2017) Mismatch Repair Incompatibilities in Diverse Yeast Populations. Genetics 205, 1459-1471 DOI: 10.1534/genetics.116.199513.

60. Heck, J.A., Argueso, J.L., Gemici, Z., Reeves, R.G., Bernard, A. *et al.* (2006) Negative epistasis between natural variants of the *Saccharomyces cerevisiae* MLH1 and PMS1 genes results in a defect in mismatch repair. Proc Natl Acad Sci U S A 103, 3256-61 DOI: 10.1073/pnas.0510998103.

61. Bui, D.T., Dine, E., Anderson, J.B., Aquadro, C.F. and Alani, E.E. (2015) A Genetic Incompatibility Accelerates Adaptation in Yeast. PLoS Genet 11, e1005407 DOI: 10.1371/journal.pgen.1005407.

62. Cromie, G.A., Tan, Z., Hays, M., Jeffery, E.W. and Dudley, A.M. (2017) Dissecting Gene Expression Changes Accompanying a Ploidy-Based Phenotypic Switch. G3 7, 233-246 DOI: 10.1534/g3.116.036160.

63. Torres, E.M., Sokolsky, T., Tucker, C.M., Chan, L.Y., Boselli, M. *et al.* (2007) Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science 317, 916-24 DOI: 10.1126/science.1142210.

64. Gordon, D.J., Resio, B. and Pellman, D. (2012) Causes and consequences of aneuploidy in cancer. Nat Rev Genet 13, 189-203 DOI: 10.1038/nrg3123.

65. Rancati, G. and Pavelka, N. (2013) Karyotypic changes as drivers and catalyzers of cellular evolvability: a perspective from non-pathogenic yeasts. Semin Cell Dev Biol 24, 332-8 DOI: 10.1016/j.semcdb.2013.01.009.

66. Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A. *et al.* (2010) Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature 468, 321-5 DOI: 10.1038/nature09529.

67. Yona, A.H., Manor, Y.S., Herbst, R.H., Romano, G.H., Mitchell, A. *et al.* (2012) Chromosomal duplication is a transient evolutionary solution to stress. Proc Natl Acad Sci U S A 109, 21010-5 DOI: 10.1073/pnas.1211150109.

68. Cromie, G.A. and Dudley, A.M. (2015) Aneuploidy: Tolerating Tolerance. Curr Biol 25, R771-3 DOI: 10.1016/j.cub.2015.06.056.

69. Cubillos, F.A., Vasquez, C., Faugeron, S., Ganga, A. and Martinez, C. (2009) Self-fertilization is the main sexual reproduction mechanism in native wine yeast populations. FEMS Microbiol Ecol 67, 162-70 DOI: 10.1111/j.1574-6941.2008.00600.x.

70. Mayer, V.W. and Aguilera, A. (1990) High levels of chromosome instability in polyploids of *Saccharomyces cerevisiae*. Mutat Res 231, 177-86.

71. Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R. *et al.* (2008) Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. Cell 135, 879-93 DOI: 10.1016/j.cell.2008.09.039.

72. Storchova, Z., Breneman, A., Cande, J., Dunn, J., Burbank, K. *et al.* (2006) Genome-wide genetic analysis of polyploidy in yeast. Nature 443, 541-7 DOI: 10.1038/nature05178.

73. Scott, A.L., Richmond, P.A., Dowell, R.D. and Selmecki, A.M. (2017) The Influence of Polyploidy on the Evolution of Yeast Grown in a Sub-Optimal Carbon Source. Mol Biol Evol 34, 2690-2703 DOI: 10.1093/molbev/msx205.

74. Gerstein, A.C., Lim, H., Berman, J. and Hickman, M.A. (2017) Ploidy tug-of-war: Evolutionary and genetic environments influence the rate of ploidy drive in a human fungal pathogen. Evolution 71, 1025-1038 DOI: 10.1111/evo.13205.

75. Blundell, J.R., Schwartz, K., Francois, D., Fisher, D.S., Sherlock, G. *et al.* (2019) The dynamics of adaptive genetic diversity during the early stages of clonal evolution. Nat Ecol Evol 3, 293-301 DOI: 10.1038/s41559-018-0758-1.

76. Levy, S.F., Blundell, J.R., Venkataram, S., Petrov, D.A., Fisher, D.S. *et al.* (2015) Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature 519, 181-6 DOI: 10.1038/nature14279.

77. Herbrecht, R. and Nivoix, Y. (2005) *Saccharomyces cerevisiae* fungemia: an adverse effect of *Saccharomyces boulardii* probiotic administration. Clin Infect Dis 40, 1635-7 DOI: 10.1086/429926.

CHAPTER 2

Incompatibilities in mismatch repair genes *MLH1-PMS1* contribute to a wide range of mutation rates in human isolates of baker's yeast

This chapter was published in Genetics on 10/24/2018: Raghavan V, Bui DT, Al-Sweel N, Friedrich A, Schacherer J, Aquadro CF and Alani, E. Incompatibilities in Mismatch Repair Genes *MLH1-PMS1* Contribute to a Wide Range of Mutation Rates in Human Isolates of Baker's Yeast. Genetics. 2018;210(4):1253-66. Epub 2018/10/24. doi: 10.1534/genetics.118.301550. PubMed PMID: 30348651; PubMed Central PMCID: PMC6283166.

Abstract

Laboratory baker's yeast strains bearing an incompatible combination of *MLH1* and *PMS1* mismatch repair alleles are mutators that can adapt more rapidly to stress but do so at the cost of long-term fitness. We identified 18 baker's yeast isolates from 1011 surveyed that contain the incompatible *MLH1-PMS1* genotype in a heterozygous state. Surprisingly, the incompatible combination from two human clinical heterozygous diploid isolates, YJS5845 and YJS5885, contain the exact *MLH1* (S288c derived) and *PMS1* (SK1 derived) open reading frames originally shown to confer incompatibility. While these isolates were non-mutators, their meiotic spore clone progeny displayed mutation rates in a DNA slippage assay that varied over a 340-fold range. This range was 30-fold higher than observed between compatible and incompatible combinations of laboratory strains. Genotyping analysis indicated that *MLH1-PMS1* incompatibility was the major driver of mutation rate in the isolates. The variation in the mutation rate of incompatible spore clones could be due to background suppressors and enhancers as well as aneuploidy seen in the spore clones. Our data are consistent with the observed variance in mutation rate contributing to adaptation to stress conditions (e.g. in a human host) through the acquisition of beneficial mutations, with high mutation rates leading to long-term fitness costs that are buffered by mating, or eliminated through natural selection.

Introduction

Loss of DNA mismatch repair (MMR) functions is often seen in bacteria grown in stressful environments. MMR defective bacteria display an increased mutation supply and adapt to stress by acquiring beneficial mutations; however, their high mutation rate ultimately results in the accumulation of deleterious mutations that reduce fitness [1-8]. Bacteria defective in MMR can overcome fitness costs associated with high mutation rate by regaining MMR functions through horizontal gene transfer [5]. In baker's yeast loss of MMR functions in the laboratory provides an adaptive advantage to stress [9-11]. However, there is no evidence that baker's yeast defective in MMR can undergo horizontal transfer in the wild [12-14]. We and others hypothesized that MMR defective baker's yeast isolates in the wild could potentially mate to MMR proficient isolates to become non-mutators and thus eliminate long term fitness costs (reviewed in Bui *et al*., [15]).

S. cerevisiae are usually non-pathogenic and are infectious only in immunocompromised individuals, causing opportunistic infections [16, 17]. Phylogenetic analyses of sixteen human clinical isolates indicated that these isolates were not derived from a common ancestor or single strain but may be a consequence of multiple independent origins arising from chance events involving opportunistic colonization of human tissues by different strains [18-21]. Yeast strains that infect humans must survive unfavorable growth conditions such as elevated temperature and exposure to antifungal agents. Interestingly, recent studies have suggested that ploidy changes and chromosome aneuploidy can promote phenotypic diversity; in some cases, such events are thought to increase the rate of acquiring adaptive mutations in the face of stress and have been seen at higher frequencies in clinical isolates [20, 22-31]. At present it is unclear if baker's yeast

strains associated with humans can be MMR defective; however, two recent reports identified *Cryptococcus* fungal human pathogens that are defective in MMR and are mutators [32, 33].

We showed previously that the *MLH1* and *PMS1* mismatch repair genes from S288c and SK1 yeast strains can display negative epistasis; strains bearing the S288c *MLH1*-SK1 *PMS1* genotype are mutators [34]. Single amino acid changes in each protein, MLH1 D761 from S288c, PMS1 K818/822 from SK1, were sufficient to cause this incompatibility (Figure 2.1; [34]). The *MLH1-PMS1* combinations are compatible in the S288c and SK1 group strains as well as in the inferred ancestral strain from which they diverged (Figure 2.1, [11, 34]). The incompatible combination of S288c *MLH1*-SK1 *PMS1* in laboratory strains (abbreviated as *cMLH1-kPMS1*) confers an ~100-fold increase in the mutation rate in an assay in which null mutants display a 10,000-fold higher rate [11, 34]. Incompatible strains have an adaptive advantage in high salt stress [11]. However, they also display a long-term fitness cost due to the accumulation of deleterious mutations. This was demonstrated by a fitness decline of incompatible cells in competition experiments between compatible and incompatible cells in rich media [15]. To our knowledge the *cMLH1-kPMS1* genotype is the only incompatibility involving mismatch repair genes that has been characterized. However, recessive mutations in mismatch repair genes *MSH2* and *MSH6* found in colon cancer patients were postulated to be due to a compound effect of mutations in the two genes, which could reflect negative epistasis [35].

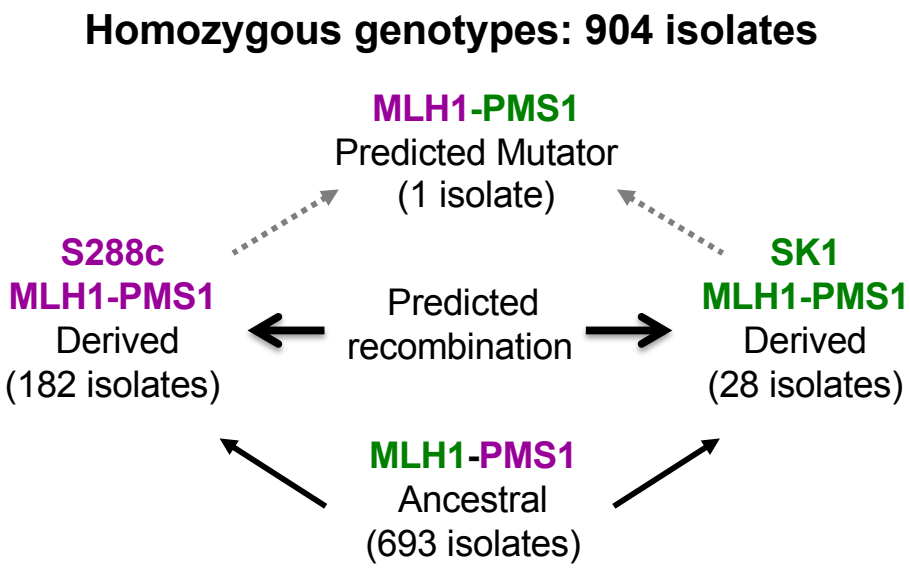


Figure 2.1: Incompatibility involving the *MLH1* and *PMS1* MMR genes. In the model presented, ancestral isolates bearing MLH1 Gly 761 and PMS1 Arg 818/822 alleles acquire neutral or beneficial mutations that lead to the derived S288c (purple, Asp 761, Arg 818/822) and SK1 (green, Gly 761, Lys 818/822) group isolates. Matings between derived isolates yield an allele combination (MLH1 Asp 761, PMS1 Lys 818/822) that displays negative epistasis, revealed as a mutator phenotype. Previous work suggested that recombination occurred between the two derived classes, leading to a hybrid mutator genotype that can also remate to reconstruct derived or ancestral genotypes [34]. In this figure, sequences of *MLH1* and *PMS1* genes from 1,011 *S. cerevisiae* from worldwide genome collections are grouped according to their amino acid residues 761 (G or D) in MLH1 and 818 (R or K) in PMS1 [15]. One isolate maps to the predicted hybrid mutator category. Adapted from Bui *et al.* [15].

In a search for yeast bearing the incompatible c*MLH1-kPMS1* combination, we screened a collection of 1011 naturally occurring worldwide isolates of *Saccharomyces cerevisiae* [15, 36]. Only YJM523, a human clinical isolate of *S. cerevisiae,* is homozygous for *cMLH1-kPMS1* incompatibility [15, 19, 21]. A spore clone of YJM523, YJM555, was a non-mutator that had accumulated multiple suppressors of the incompatibility phenotype [15, 21]. To our knowledge baker's yeast isolates that are mutators have not been identified in the wild [15, 21]. However, yeast heterozygous for the c*MLH1-kPMS1* incompatibility, a recessive trait [34], might have an advantage because they can be non-mutators in a heterozygous diploid background. Under stress conditions the diploids can sporulate and give rise to mutator spore clones. These mutators would be able to provide a transient adaptive advantage to stress and can potentially mate back to compatible strains and become non-mutators to prevent long-term fitness costs.

In eukaryotes, a complete loss of MMR functions would likely have dramatic effects on organism fitness. In changing environments, as would likely be the case for yeast growing in a human host, adaptation to stress would involve a compromise between efficient growth and long-term survival. One way to deal with such variable stress is for progeny of an organism to display variations in mutation supply that can yield adaptation phenotypes and thus prevent the population from undergoing extinction under stress (see [37]). We examined two human clinical diploid isolates, YJS5845 and YJS5885, that are each heterozygous for c*MLH1-kPMS1* incompatibility polymorphisms. In both isolates, one of the two copies of MLH1 is identical in amino acid sequence to S288c MLH1, and one of two copies of PMS1 is identical to SK1 PMS1. The isolates were non-mutators but derived spore clones displayed a 340-fold range of mutation rates in a DNA slippage assay, with *cMLH1-kPMS1* incompatibility being the major contributor to increased mutation. In contrast, the difference in mutation rates between laboratory-

compatible and incompatible strains was only eleven-fold, indicating that the human isolate backgrounds contained suppressors and enhancers of mutation rate. Curiously, high levels of aneuploidy were observed in the spore clones from YJS5845, suggesting additional mechanisms that may contribute to modifying mutation rate and/or adaptation. These findings suggest that there is a variation in mutation supply that is balanced with associated fitness costs. Such variation could provide a greater chance for yeast to survive stress conditions present in a human host.

Materials and Methods

S. cerevisiae strains and media

The natural isolates YJS5845, YJS5885, YJS4806, YJS4810, YJS5882, YJS5678, YJS5512, and YJS4970 (Table 2.3) were obtained from the 1011 Yeast Genomes Project [36]. YJM521, YJM523 and YJM555 were obtained from the FGSC collection ([19]; http://www.fgsc.net/). Isolates were sporulated on plates by streaking cells from frozen stocks onto YPD (2% peptone, 1% yeast extract, 2% dextrose) media. Cells were grown for one day at 30°C, after which they were patched onto sporulation medium (1% (g/L) potassium acetate, 2% (g/L) agar) and incubated at 30°C for 3 days. Tetrads were dissected and germinated on minimal complete media [38]. Isolates were sporulated in liquid by growing overnights in YPD. 50 µl of each overnight was then inoculated into 5 ml YPA (2% peptone, 1% yeast extract, 2% potassium acetate) media and grown with shaking at 30°C for 16 hours, after which each culture was washed with 3 ml of sterile water and then transferred to 1% potassium acetate and incubated at 30°C with shaking for 3 days. In order to test for growth of spore clones in lactate media, spore clones were patched on YPL plates (2% peptone, 1% yeast extract, 2% lactate, 2% agar) and grown for 4 days at 30°C (Table 2.8).

The S288c derived strains EAY1365 (*MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2::insE-A14, mlh1Δ::KanMX4, pms1Δ::KanMX4)***,** EAY1369 (*MATalpha*, *ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2::insE-A14, cPMS1::HIS3, cMLH1)*, EAY1370 (*MATalpha*, *ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2::insE-A14, kPMS1::HIS3, cMLH1),* EAY1372 *(MATa*, *ura3-52, leu2Δ1, trp1Δ63, lys2::insE-A14, msh2*D*::hisG),* FY90 (*MATa, ade8;* Winston *et al*. 1995) and EAY4087 (*MATa, ade8, mlh1*D*::KanMX)*, and the natural isolates YJS5845 and YJS5885 and their spore clone derivatives were analyzed for mutator phenotypes (see below for details). Transformation of plasmids into strains, isolates and spore clones was performed as described [39]. In this manuscript, genes derived from the S288c background are designated with a "c" (e.g. *cMLH1*) and those derived from SK1 with a "k" (e.g. *kMLH1*).

Plasmids

pEAA213 (c*MLH1, ARSH4 CEN6, LEU2*) and pEAA214 (k*MLH1, ARSH4 CEN6, LEU2*) were described previously [34, 40]. The *cMLH1* gene from YJM521 was cloned into pEAA213 as described previously [15]. All of the constructs expressed *MLH1* via the S288c *MLH1* promoter. pEAA238 (c*PMS1, ARSH4, CEN6, HIS3*) and pEAA239 (k*PMS1*, *ARSH4, CEN6, HIS3*) were described previously [34, 40]. The *kPMS1* gene from YJM521 was cloned into pEAA238 as described previously [15]. All of the constructs expressed *PMS1* via the S288c *PMS1* promoter.

pEAA613 (*ARS-CEN, NATMX*) contains the *URA3 promoter-KanMX::insE-A14* reversion reporter [15]. This reporter is expressed via the *URA3* promoter (-402 to the ATG start site). A 55 bp sequence containing a +1 frameshift in the 14 bp homopolymeric A run (*insE-A14;* [41]) was inserted immediately after the *URA3* ATG, followed by codons 18 to 269 of the *KANMX*

open reading frame derived from pFA6-KANMX. pEAA611 (*ARS-CEN, NATMX)* contains the *URA3 promoter-KanMX::insE-A10* in frame reporter [15].

Genotyping analysis

The 2.3 kb *MLH1* and 2.6 kb *PMS1* open reading frames from YJM521, YJS5845, YJS5885, YJS4806, YJS4810, YJS5882, YJS5678, YJS5512, YJS4970 and derived spore clones were determined by sequencing PCR amplified DNA from chromosomal DNA (Table 2.1; Table 2.2; [42]) using Expand High Fidelity Polymerase (Roche Life Sciences). Primers AO324 (5'ATAGTGTAGGAGGCGCTG) and AO821 (5'AACTTTGCGGCCGCGGATCCAGCCAAAACGTTTTAAAGTTA) were used to amplify the *MLH1* open reading frame and primers AO481 (5'CCACGTTCATATTCTTAATGGCTAAGC) and AO548 (5'CGATTCTAATACAGATTTTAATGACC) were used to amplify the *PMS1* open reading frame. PCR products were sequenced by the Sanger method in the Cornell Biotechnology Resource Center.

The diploid isolates YJM521, YJS5845, and YJS5885 were shown previously to be heterozygous for genetic information at both the *MLH1* and *PMS1* incompatibility sites (Gly/Asp at amino acid 761 in MLH1, Arg/Lys at amino acid 818/822 in PMS1; [15, 21]). For each diploid isolate, the sequences of the two *MLH1* and *PMS1* alleles were determined by sequencing the *MLH1* and *PMS1* genes from two ancestral (MLH1 Gly 761, PMS1 Arg 818/822) and two incompatible (MLH1 Asp 761, PMS1 Lys 818/822) haploid spore clones. In all cases the DNA sequences of the two different spore clone isolates of the same *MLH1-PMS1* genotype were

identical, thus allowing us to assign the *MLH1* and *PMS1* sequences present in each parental chromosome (Table 2.2).

To determine the compatibility genotype (Gly/Asp at amino acid 761 in MLH1, Arg/Lys at amino acid 818/822 in PMS1; Figure 2.1; Figure 2.2) of YJM521, YJS5845, and YJS5885 spore clones, *MLH1* and *PMS1* open reading frame PCR products were sequenced with primers AO328 (5' GACGAGTTAAATGACGATGCTTCC) and AO485 (5' AAAGTATCTGACGTTAACAGTTTC), respectively. To test for the presence of the proline 271 suppressor polymorphism in MLH1 incompatible spore clones [43] in YJS5845 and YJS5885, *MLH1* PCR products were also sequenced using primer AO325 (5' CATGTGGCAACAGTCACAGTAACG). None of the spore clones (10 from YJS5845, 11 from YJS5885) contained this polymorphism, displaying instead the leucine residue.

Table 2.1: Genotyping of spore clones obtained by dissection of isolate tetrads

			Number of spore clones genotyped for *MLH1-PMS1* as:			
Lab name	***MLH1* genotype**	***PMS1* genotype**	**Ancestral**	**SK1**	**S288c**	**Incompatible**
YJS5845	SK1/S288c	SK1/S288c	5	17	1	10
YJS5885	SK1/S288c	SK1/S288c	5	6	11	11
YJM521	SK1/S288c	SK1/S288c	5	7	7	5
YJS4806	SK1/S288c (3:1)	SK1	not relevant			
YJS4810	SK1/S288c (3:1)	SK1	not relevant			
YJS5882	SK1/S288c (3:1)	SK1-S288c (2:2)	not relevant			
YJS5678	SK1/S288c (3:1)	SK1/S288c (2:2)	not relevant			
YJS5512	SK1/S288c (3:1)	SK1/S288c (2:2)	not relevant			
YJS4970	SK1/S288c (2:2)	SK1/S288c (2:2)	not relevant			

MLH1 and *PMS1* genes were PCR amplified from isolates and derived spore clones and sequenced as described in the Materials and Methods. For YJS5845, three spore clones were genotyped from random spores and 30 were genotyped from spores isolated after tetrad dissection. For YJS5885 and YJM523, all spore clones were genotyped from tetrad dissection. None of the incompatible YJS5845 and YJS5885 spore clones contained the Pro 271 suppressor polymorphism in MLH1[43]. For YJM521, 24 spore clones were genotyped from six four-spore viable tetrads. For YJS4806, 24 spore clones were genotyped from random spores with 22 showing the parental genotype for *MLH1*, and two showing a different segregation pattern (2G:2A, 4G:0A). For YJS4810, 24 spore clones were genotyped from random spores, with all showing the parental genotype for *MLH1*.

Table 2.2: Genotyping of *MLH1* and *PMS1* loci in YJM and YJS isolates and derived spore clones

base pair	aa	S288c	SK1	YJS5845c	YJS5845k	YJS5885c	YJS5885k	YJM523	YJM521c	YJM521k
MLH1										
486	162	C, ALA	T, ALA	C	C	C	C	C	C	C
552	184	C, SER	C, SER	C	C	C	C	C	T, SER	C
720	240	C, SER	A, ARG	C	C	C	C	C	C	C
812	271	T, LEU	**C, PRO**	T	C	T	C	T	C	C
834	278	C, SER	T, SER	C	T	C	T	C	C	T
997	333	G, GLU	A, LYS	G	G	G	G	G	G	G
1044	348	T, ILE	C, ILE	T	C	T	C	T	T	C
1237	413	T, LEU	C, LEU	T	C	T	C	T	T	C
1393	465	G, ASP	G, ASP	G	G	G	G	G	A, ASN	G
1875	625	T, SER	C, SER	T	C	T	C	T	T	C
2032	678	G, ASP	A, ASN	G	A	G	A	G	G	A
2108	703	C, PRO	T, LEU	C	T	C	T	C	C	T
2282	761	A, ASP	G, GLY	A	G	A	G	A	A	G
PMS1										
122	41	A, ASN	G, SER	G	G	G	G	A	A	A
162	54	T, SER	T, SER	T	T	T	C	T	T	T
177	59	T, ASP	T, ASP	T	T	T	T	C, ASP	T	C, ASP
210	70	G, GLU	G, GLU	G	G	G	G	A, GLU	G	A, GLU
213	71	C, PHE	T, PHE	T	T	T	T	T	C	T
258	86	T, ASP	C, ASP	C	C	C	C	C	C	C
333	111	G, VAL	C, VAL	C	C	C	C	C	C	C
335	112	T, ILE	C, THR	C	C	C	C	T	T	T
465	155	C, PRO	T, PRO	T	T	T	T	T	C	T
552	184	T, ALA	A, ALA	T	A	T	A	T	T	T
558	186	T, ILE	C, ILE	C	C	C	C	C	T	C
708	236	A, LEU	G, LEU	G	G	G	G	G	A	G
711	237	T, ASN	C, ASN	C	C	C	C	C	T	C
810	270	G, SER	C, SER	G	C	G	C	G	G	G
855	285	G, VAL	A, VAL	G	A	G	A	G	G	G
858	286	T, ASN	T, ASN	T	T	T	T	C, ASN	T	C, ASN
918	306	C, PHE	C, PHE	C	C	C	C	C	T, PHE	C
925	309	G, VAL	G, VAL	G	G	G	G	G	T, PHE	G

939	313	T, ALA	A, ALA	A	A	A	A	A	T	A
base pair	**aa**	**S288c**	**SK1**	**YJS5845c**	**YJS5845k**	**YJS5885c**	**YJS5885k**	**YJM523**	**YJM521c**	**YJM521k**
1150	384	T, PHE	G, VAL	G	G	G	G	G	T	G
1175	392	A, GLU	A, GLU	A	A	A	A	T, ASP	A	T, ASP
1191	397	C, ASN	C, ASN	T, ILE	C	T, ILE	C	T, ILE	C	T, ILE
1199	400	C, THR	G, SER	G	G	G	G	G	C	G
1201	401	G, ALA	G, ALA	T, SER	G	T, SER	G	T, SER	G	T, SER
1249	416	NO INS	INS	INS	INS	INS	INS	INS	NO INS	INS
1329	443	C, ILE	C, ILE	T, ILE	C	C	C	T, ILE	C	T, ILE
1538	513	A, TYR	T, PHE	A	T	A	T	A	A	A
1575	525	G, ALA	C, ALA	G	C	G	C	G	G	G
1691	564	C, ALA	C, ALA	C	C	C	C	T, VAL	C	T, VAL
1782	594	T, TYR	C, TYR	C	C	C	C	C	T	C
1821	607	A, GLU	G, GLU	G	G	G	G	G	A	G
2076	692	T, ASP	T, ASP	C, ASP	T	T	T	T	T	T
2303	768	A, LYS	A, LYS	A	A	A	A	G, ARG	A	G, ARG
2322	774	T, THR	G, THR	G	G	G	G	G	T	G
2364	788	G, LEU	A, LEU	A	A	A	A	A	G	A
2453	818	G, ARG	A, LYS	G	A	G	A	A	G	A

The sequences for the *MLH1* and *PMS1* open reading frames for each of the two parental chromosomes are shown relative to the S288c and SK1 sequences for YJS5845, YJS5885, and YJM521. The parental chromosomes were genotyped as "c" (S288c) or "k" (SK1) based on the amino acid polymorphisms seen at the incompatibility loci (bp 2282 in *MLH1*, bp 2453 in *PMS1*) in the S288c and SK1 sequences (Materials and Methods). The MLH1-271P suppressor allele is highlighted at bp 812 in *MLH1*. INS = 12 bp insertion in *PMS1*; NO INS = lacking the insertion.

Efficiency of plating

EAY1369 (*cMLH1-cPMS1*, compatible) and EAY1370 (*cMLH1-kPMS1*, incompatible) strains, and YJM521, YJM555, YJS5845, and YJS5885 isolates/spore clone derivatives were transformed with pEAA611 (*URA3 promoter-KanMX::insE-A10* in frame reporter; [15]), and grown on YPD media containing clonNAT (100 μg/ml). Independent transformants were grown overnight on YPD-clonNAT (50 μg/ml). 10 μl of 10^{0}, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-3} and 10^{-4} dilutions were plated on YPD-clonNAT (50 μg/ml) and YPD-clonNAT (50 μg/ml) geneticin (200 μg/ml) plates. Growth on YPD-clonNAT and YPD-clonNAT, geneticin plates was compared and strains that grew to the same extent on both plates (within a few fold variation) were considered to have an efficiency of plating of ~1. Reversion assays were performed only on spore clones showing an efficiency of plating of ~1 (Figure 2.5).

kanMX::insE-A14 reversion assay

The S288c strain EAY1369 (*cMLH1-cPMS1*, compatible) and EAY1370 (*cMLH1-kPMS1*, incompatible), and YJS5845 and YJS5885 isolates and spore clone derivatives were transformed with pEAA613 and grown on YPD media containing clonNAT (100 μg/ml). Independent transformants were subsequently grown overnight in YPD + clonNAT and then plated on to YPD + clonNAT (50 μg/ml) and YPD + clonNAT (50 μg/ml), geneticin (200 μg/ml). These strains were analyzed for reversion to resistance to geneticin using methods described previously [11, 15, 41, 44]. Previously we measured resistance to geneticin (G418) for EAY1369 and YJM555 lacking the *kanMX::insE-A14* reporter plasmid [15]. We estimated the mutation rates to be $< 2 \times 10^{-10}$, indicating that spontaneous reversion to $G418^{r}$ would not interfere with the detection of $G418^{r}$ using the pEAA613 *kanMX::insE-A14* reporter plasmid. The 95% confidence intervals were determined as described by Dixon and Massey [45]. Mann-

Whitney U tests were performed using Prism (GraphPad Prism 7.00 for Mac OS X, GraphPad Software, La Jolla California USA, www.graphpad.com) to determine the significance of the differences in median reversion rates [46, 47].

Sequencing of homopolymeric A repeats in kanMX::insE-A14 G418^r clones

A 2.2 KB region of the plasmid pEAA613 containing *kanMX::insE-A14* was PCR amplified from total DNA [42] isolated from revertant colonies (NAT^r, G418^r) or non- revertant controls (NAT^r, G418^S) using Phusion Hot Start II DNA polymerase (Thermo Fischer Scientific) and primers AO3879 (5'CTCGTTTTCGACACTGGATGGC) and AO3880 (5'GCGTGAGCTATGAGAAAGCGC). Primer AO3838 (5'TGGTCGGAAGAGGCATAAATTC) was used to sequence the PCR product in the region surrounding the homopolymeric A sequence.

lys2-A14 reversion assay

Independent colonies of EAY1365 (relevant genotype *lys2-A14*) containing the *ARS-CEN, MLH1* and *ARS-CEN, PMS1* plasmids presented in Table 2.4 were inoculated into YPD liquid media, grown overnight at 30°C, and then plated onto LYS, HIS, LEU dropout and HIS, LEU dropout synthetic plates. These strains were analyzed for reversion to Lys$^+$ as described previously [11, 41]. The 95% confidence intervals were determined as described by Dixon and Massey [45]. The Mann-Whitney U test [47] was performed to determine the significance of the differences in median reversion rates.

5-fluoroorotic acid resistance assay

Resistance to 5-fluoroorotic acid (5-FOA) was measured in FY90, EAY4087 (*mlh1*Δ derivative of FY90), and haploid spore clones of YJS5885 using a protocol similar to Lang and Murray [48]. Single colonies from a synthetic complete media were grown overnight in 2 ml liquid synthetic complete media (0.7% yeast nitrogen base, 0.087% complete amino acid mix, 2% dextrose); diluted 1:500 into synthetic complete media and grown in 5 ml for 2 days at 30°C with shaking. Appropriate dilutions were then plated onto 5-FOA (0.1% 5-FOA, 0.7% yeast nitrogen base, 0.087% amino acid mix without uracil, 0.005% uracil, 2% dextrose, 2% agar) and synthetic complete plates, and mutation rates were determined using previously published methods [11, 41, 44]. 95% confidence intervals were determined as described by Dixon and Massey [45]. Mann-Whitney U tests were performed using Prism (GraphPad Prism 7.00 for Mac OS X, GraphPad Software, La Jolla, California USA; www.graphpad.com) to determine the significance of the differences in median reversion rates [46, 47]. The *URA3* open reading frame was PCR amplified from genomic DNA of independent 5- FOA resistant spore clones using Phusion Hot Start II DNA polymerase (Thermo Fischer Scientific) and primers AO1115 (AGAAGAGTATTGAGAAGGGCAA) and AO3784 (TTAGTTTTGCTGGCCGCATC). Primers AO3156 (GGTGAAGGATAAGTTTTGACCATCAAAGAA) and AO3788 (CTGGAGTTAGTTGAAGCATTAGGTC) were used to sequence the PCR product.

Whole genome sequencing

Single colonies of isolates, spore clones and transformants were grown overnight in 2 ml YPD media and genomic DNA was isolated using the YeaStar Genomic DNA kit (Zymo Research). DNA was barcoded using Illumina Nextera XT and high throughput DNA sequencing was performed on an Illumina NextSeq500 at the Cornell Biotechnology Resource Center,

achieving a 50-fold mean coverage. Sequences were aligned to the reference S288c genome sequence (SGD: https://www.yeastgenome.org/) using HISAT2 to create SAM (Sequence alignment map) files. SAM files were converted to binary version BAM files. BAM files were sorted and indexed, and duplicates were removed using SAMTools-1.7 (http://samtools.sourceforge.net/). Aneuploidy plots (Figure 2.8 and 2.10) were constructed using a custom made script (provided by V.P. Ajith, IISER Trivandrum): read counts were tabulated for 5000 bp windows using the GenomicRanges and GenomicAlignments Bioconductor packages in a custom R-script. SnpEff was used to annotate variants using sequence information for YJS5845 and YJS5885 described in Peter *et al.* [36].

Flow cytometry

Cells were prepared using a protocol modified from Rosebrock [49]. Cells were grown overnight in rich medium, washed and fixed in 70% ethanol in -20°C for 24 hours. Fixed cells were washed, resuspended in 50 mM sodium citrate, pH 7.4 and allowed to rehydrate. RNase A (Thermo Scientific) was added to a final concentration of 0.2 mg/ml and incubated at 37°C overnight. Cells were then treated with Proteinase K (NEB) to a final concentration of 16 units/ml at 37°C for 1 hour, sonicated for 10 sec at 30% power. Cells were then pelleted, supernatant was removed, and fresh 50 mM sodium citrate pH 7.4 was added, followed by staining of cells using Sytox Green at a final concentration of 1 μM. Flow cytometry was performed on BD FACS Aria at Cornell University flow cytometry core laboratory. YJS5885 spore clones showed multiple peaks even after gating out cell clumps, therefore the peaks were sorted and visualized under the microscope to determine the composition of cells under each peak. Peak one had single cells and peak two had a combination of single cells and small budded

cells, peaks three and four had large budded cells or multiple budded cells. The population from last two peaks were removed during analysis, because they were present as cell clumps and were thus not a correct representation of the DNA content in each cell, using FlowJo 10.4.2 software. The unsorted cells were counted under the microscope to determine the fraction of single cells, small-budded cells and large budded cells. The buds with diameter less than half of mother cell were classified as small-budded cells [50].

Data Availability

Strains and plasmids are available upon request, and the entire DNA sequences of the *MLH1* and *PMS1* genes from 1011 isolates are presented in Bui *et al.* [15]. Supporting information contains all detailed descriptions of supplemental files. All wet lab experiments presented (*lys2- A14* reversion, *kanMX::insE-A14* reversion, 5-FOA resistance) were repeated on at least two separate days.

Results

Genotyping of yeast isolates containing incompatible cMLH1-kPMS1 combinations

We identified among 1011 yeast isolates one clinical isolate (YJM523) that was homozygous and eighteen that were heterozygous for an S288c *MLH1*-SK1 *PMS1* incompatibility that confers a mutator phenotype in laboratory strains (Materials and Methods; Figure 2.1, 2.2; Table 2.1; [15]). This phenotype is abbreviated as *cMLH1-kPMS1* and was assigned based on the amino acid at position 761 in MLH1 and 818/822 in PMS1. For YJM523, analysis of a derived spore clone (YJM555), and DNA sequencing and phylogeny analysis suggested that it is homothallic and homozygous for *MLH1* and *PMS1* sequence information (Table 2.2; [19]). Genetic analysis indicated that YJM555 is not a mutator but contains multiple

mutations that suppress and enhance the c*MLH1*-k*PMS1* incompatibility [15, 21]. Of the 18 isolates heterozygous for *cMLH1-kPMS1* incompatibility, five are diploid, three of which can sporulate (Figure 2.2; Table 2.3). The remainder are triploid or tetraploid. Several tetraploids were genotyped (Table 2.1; Figure 2.4), showing a variety of c*MLH1*:*kMLH1* genotypes (4:0, 3:1, 2:2).

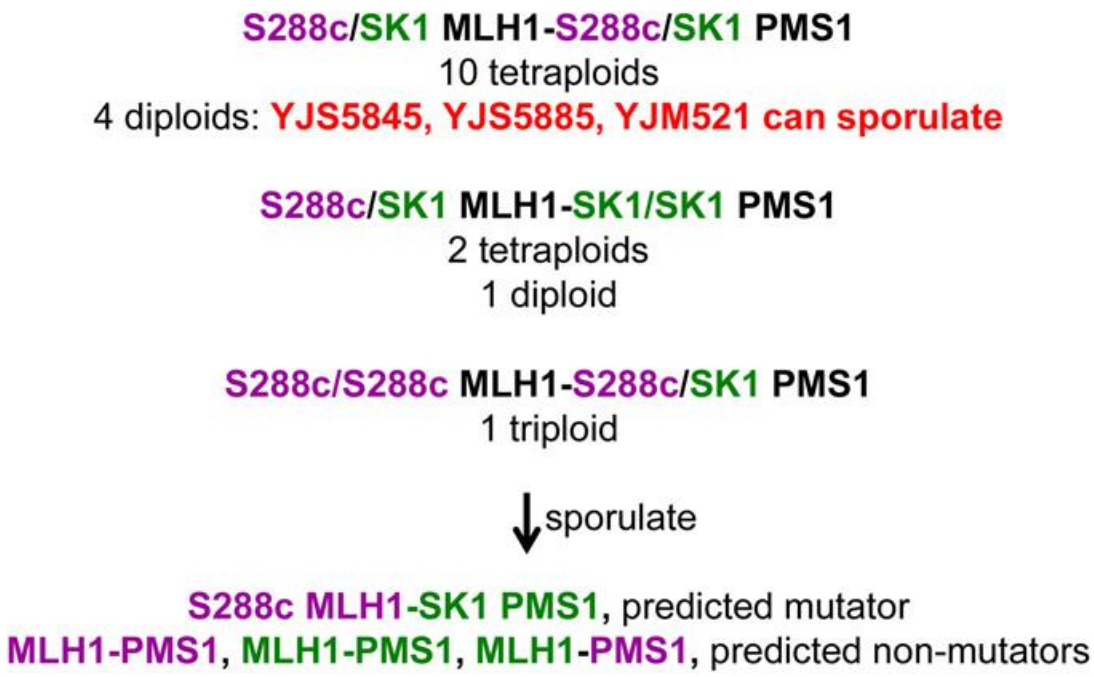


Figure 2.2: Isolates containing heterozygous *MLH1-PMS1* genotypes predicted to form mutator spore progeny. Eighteen isolates were identified from the 1011 yeast genome project that are heterozygous for the MLH1 761 and/or PMS1 818/822 genotypes, and are predicted to yield incompatible spore progeny [15].

Table 2.3: Yeast isolates analyzed in this study

Lab, standard, isolate ID	Origin	ploidy	SNPs, (singletons)	spore form. (%) (n)	% 3,4 spore	% 2 spore	% viable spores (tetrads)
YJS5845, CIC, Ponton 11	Human, mouth, Spain	2	68,564 (825)	93 (147)	53	39	20 (40)
YJS5845 + CHR XIV				96 (174)	60	35	19 (39)
YJS5885, CKN, CLI_19	Human, feces, France	2	76,104 (955)	57 (139)	31	24	30 (28)
YJM521	Human clinical, CA, USA	2	53,947 (22)	78 (221)	69	5.9	94 (10)
YJS4806, CFI, WLP013	Beer, UK	4*		nt			15 (10)
YJS4810, CFN, WLP006	Beer, unknown	4*		nt			5.0 (10)
YJS5882, CKK, CLI_16	Human clinical, France	4		nt			72 (8)
YJS5678, CBF, SD-15	Bakery, Italy	4*		nt			70 (10)
YJS5512, BML, NCYC_2780	Human clinical, Belgium	4*		nt			nt
YJS4970, CGC, UCD_06-645	Fruit, Davis, CA, USA	4		no spores			

19 of 1011 baker's yeast isolates contain the incompatible *MLH1-PMS1* genotype in heterozygous (18 isolates) or homozygous (YJM523) combinations (Bui *et al*. 2017). One isolate is triploid and 12 are tetraploid (six are shown here), two of which cannot sporulate. The remaining six are diploid, and the four diploid isolates that sporulate, YJS5845, YJS5885, YJM521, YJM523, all belong to the MR3 mosaic clade (113 members) that are admixed with ancestry from two or more populations. In the MR3 clade, the mean number of singletons was 721 (+/-1150 SD, +/-108 SE, median = 163). The diploid isolates YJM521 [51] and YJS5845 are homothallic and YJS5885 appears to be functionally heterothallic (see text). The % of cells forming at least one spore (n is the cells examined) is shown under spore formation, and the distribution of these cells into 3,4 or 2 spore asci is shown, along with the % spore viability seen in dissected tetrads (n= number dissected). Our original stock of YJS5845 contained euploid and aneuploid (additional copy of Chr XIV) cells. The euploid and aneuploid isolates were analyzed for spore formation and viability separately. * Contains chromosomal aneuploidy [36]. nt, not tested.

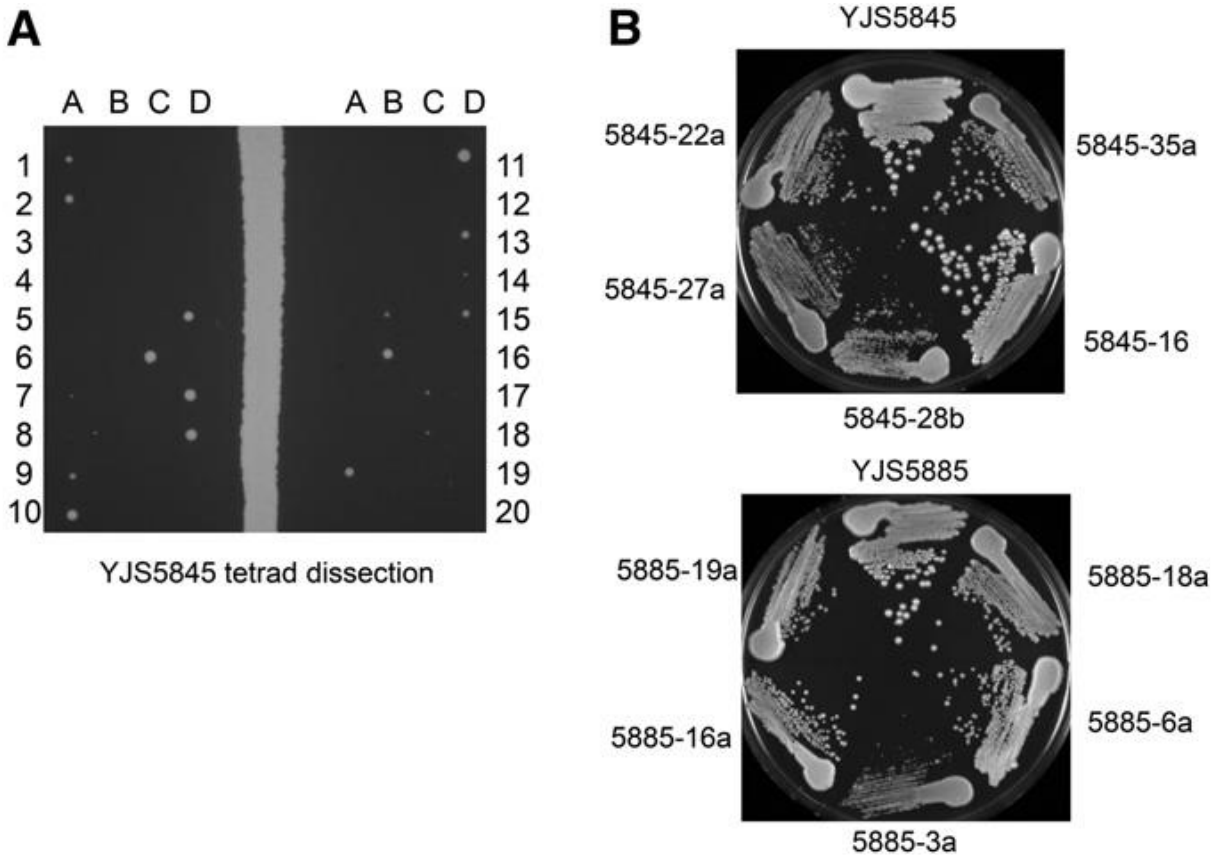
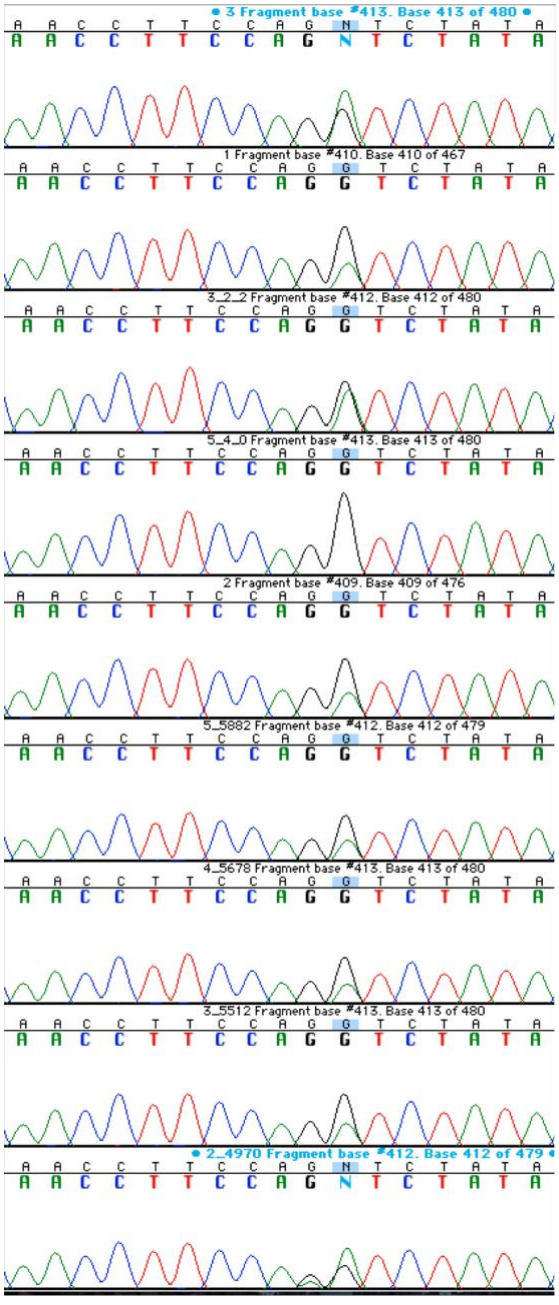


Figure 2.3: YJS5845, YJS5885, and spore clones have different colony sizes/growth properties. (A) YJS5845 was sporulated and then tetrad- dissected on minimal complete plates. Germinated spore clones were photographed after a 48-hr incubation at 30ºC . (B) YJS5845 and YJS5885 isolates, and representative spore clones, were struck to single colonies on YPD media and photographed after a 48-hr incubation at 30ºC.

Characterizing the yeast isolates that contain incompatible cMLH1-kPMS1 combinations

Three heterozygous diploid clinical isolates (YJM521, YJS5845, YJS5885) displayed rapid vegetative growth and formed colonies from single cells on YPD plates in 24 hrs. They also showed a high efficiency of sporulation, with variable spore viability (20-94%) and spore clone growth (Figure 2.3; Table 2.3). At least one of the YJS5845 spore clones genotyped displayed what appeared to be invasive growth properties on YPD. We also saw some deviation from 1:1:1:1 for *MLH1* and *PMS1* genotypes in YJS5845 and YJS5885, which we had not seen in the high spore viability isolate YJM521 (Table 2.1). We sporulated some of the tetraploids with the goal of identifying homozygous incompatible spore clones. However, as shown in Table 2.1 and Figure 2.4, such clones were not identified.

Genotyping at bp 2282 in *MLH1* (G/D at amino acid 761)



YJS5845	1:1 G:A
YJS4806	3:1 G:A
YJS4806 spore clone	2:2 G:A
YJS4806 spore clone	4:0 G:A
YJS4810	3:1 G:A
YJS5882	3:1 G:A
YJS5678	3:1 G:A
YJS5512	3:1 G:A
YJS4970	2:2 G:A

Figure 2.4: DNA sequence, as shown by chromatogram traces, of the *MLH1* incompatibility site (bp 2282, Gly or Asp at amino acid 761) in the indicated isolates and spore clones.

The YJS5845, YJS5885 and YJM521 are all human clinical isolates, belong to an admixture clade, and have different geographic locations (Table 2.3). For *MLH1*, YJS5845 and YJS5885 contain the exact amino acid sequences for the S288c *MLH1* allele on one chromosome and the exact amino acid sequences for the YJM521 *MLH1 k*-allele on the other chromosome (Table 2.2). For *PMS1*, YJS5845 and YJS5885 contain the exact amino acid sequences for the SK1 *PMS1* allele on one chromosome. The other *PMS1* bearing chromosome in both isolates contains the S288c *PMS1 K818* allele; however, this chromosome contains unique variants for the two isolates (Table 2.2). YJM521 is homozygous for the *MLH1- P271* suppressor allele (Table 2.2 and 2.4).

Mutator phenotypes exhibited by incompatible S288c MLH1-SK1 PMS1 combinations of the YJM521, YJS5845 and YJS5885 isolates in the S288c strain background

The incompatible combinations of *MLH1* and *PMS1* present in YJS5845, YJS5885, and YJM521 were tested for their ability to confer a mutator phenotype in the S288c background (Table 2.4). The incompatible combination from YJM521 was cloned; the incompatible combinations in YJS5845 and YJS5885 are represented by *cMLH1-kPMS1* because the amino acid sequences are identical to the cMLH1-kPMS1 combination. The *cMLH1-kPMS1* combination representing YJS5845 and YJS5885 conferred an incompatible mutator phenotype (75-fold higher than compatible), while the *cMLH1-kPMS1* combination from YJM521 displayed a suppressed incompatible mutator phenotype (19-fold higher than compatible) that was expected because YJM521 is homozygous for the *MLH1-P271* suppressor allele [43].

Table 2.4: Mutation rates in an S288c strain containing *MLH1* and *PMS1* genes combinations identical in amino acid sequence to those present in S288c, SK1, YJS5845, YJS5885, and YJM521.

***MLH1-PMS1* genotype**	**Lys^+ reversion rate (10^{-7}), (95% CI)**	**Relative rate**	**(n)**
S288c-S288c, compatible	4.1 (1.7-13.8)	1	13
S288c-SK1, incompatible	311 (111-919)**	75	16
YJM521c-YJM521k	92 (78.4-690)*	22	29
YJM521c-SK1	80 (47-182)*	19	15
mlh1Δ, pms1Δ	45,300 (13,170-126,800)**	10,970	10

EAY1365 (relevant genotype *mlh1Δ::KanMX4, pms1Δ::KanMX4)* was transformed with *ARS-CEN* plasmids containing the *MLH1* and *PMS1* genes obtained from the indicated strains and isolates. Independent cultures (n) were examined for reversion to Lys^+. Median mutation rates are presented with 95% confidence intervals, and relative mutation rates compared to the wild type strain are shown. Data for S288c-S288c compatible, S288c-SK1 incompatible, and *mlh1Δ, pms1Δ* were reported previously [15]. Note that *MLH1* open reading frames in YJS5845c and YJS5885c are identical to S288c *MLH1* and the *PMS1* open reading frames in YJS5845k and YJS5885k are identical to SK1 *PMS1* (Table S2).

*Significantly different from S288c-S288c, compatible ($p < 0.01$, Mann Whitney test)

**Significantly different from S288c-S288c ($p < 0.001$, Mann Whitney test)

Spore clones from YJS5845 and YJS5885 display a wide range of mutator phenotypes, with MLH1-PMS1 genotype being the major contributor to mutation rate

We examined the mutation rate in isolates and spore clone derivatives that had an approximately 100% efficiency of growth in YPD media containing geneticin when transformed with the in-frame plasmid pEAA611 (Figure 2.5). This was done to remove spore clones that were highly sensitive to aminoglycoside antibiotics (often observed in yeast isolates; [52, 53]). It is important to note that as a result of this analysis we assayed mutator phenotypes in roughly half of the spore clones obtained; the other half displayed poor efficiencies of plating on YPD media containing geneticin (Figure 2.5; Materials and Methods; no linkage was seen between efficiency of plating and *MLH1, PMS1* compatible or incompatible genotypes). Using SNPeff (Materials and Methods), we confirmed the sequences of previously identified genes involved in antibiotic sensitivity, *STP22* and *PHO80*, in YJS5845 and YJS5885 but did not find any disruptions or predicted deleterious alleles (Table 2.6; Ernst and Chan 1985; Wickert *et al.* 1998).

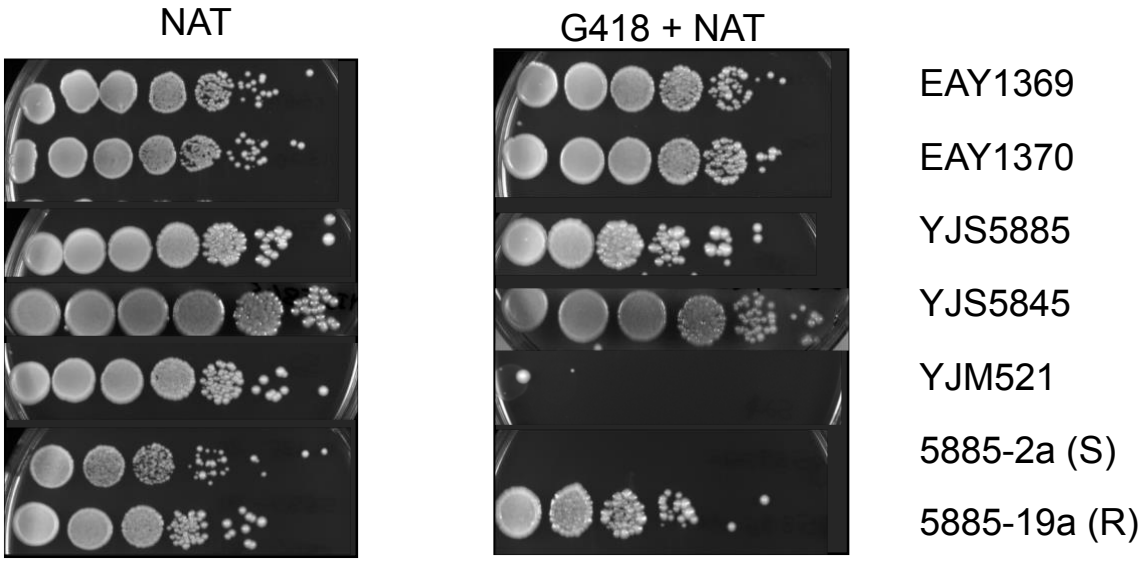


Figure 2.5: Efficiency of plating of strains transformed with pEAA611, comparing growth on clon-NAT and clon-NAT + G418 plates. Representative images of EAY1369, EAY1370, YJM521, YJS5845 and YJS5885 isolates and YJS5885 spore clones are shown.

Table 2.5: Reversion assay using the *URA3 promoter-KanMX::insE-A$_{14}$* plasmid.

Strain or Isolate	Genotype *MLH1-PMS1*	Incompatible/ Compatible	rate* G418^r (10^{-7}), (95% CI), n	rate** G418^r (10^{-7}), (95% CI), n	relative rate
EAY1369	c-c	C	5.2 (3.3-7.9), 15		1
EAY1370	c-k	I	57[a] (38-89), 19		11
EAY1372	*msh2*Δ	not applicable	9540[a, c] (6640-24800), 10		1840
YJS5845	c/k, c/k	Parental (C/I)	2.6[a, c] (2.07-2.8), 16		0.50
5845-7a	k-c	C	0.99[b, d] (0.56-1.1), 10		0.19
5845-16	k-k	C	2.1[a, c] (1.6-2.7), 10		0.40
5845-27a	k-k	C	2.1[a, c] (1.2-3.7), 10		0.41
5845-35a	k-k	C	5.8[c] (3.1-14), 15		1.1
5845-19a	k-k	C	6.0[c] (3.8-7.4), 15		1.2
5845-22a	k-k	C	7.2[c] (5.2-9.4), 15		1.3
5845-21a	c-k	I	2.7[b, c] (0.53-10), 16	0.38[b, c] (0.32-0.94), 36	0.51
5845-20a	c-k	I	3.6[c] (0.9-8.4), 11		0.69
5845-30a	c-k	I	16[a, c] (7.6-26), 16		3.2
5845-19	c-k	I	141[a, c] (86-320), 16		27
5845-18a	c-k	I	330[a, c] (253-1300), 10		64
YJS5885	c/k, c/k	Parental (C/I)	2.6[b, c] (1.4-6.5), 15		0.51
5885-1a	k-c	C	0.74[a, c] (0.54-1.2), 15		0.14
5885-14a	c-c	C	1.1[a, c] (0.44-3.4), 10		0.21
5885-20b	k-c	C	1.5[a, c] (1-1.6), 13		0.29
5885-5b	k-k	C	1.6[a, c] (0.53-2.2), 15		0.31
5885-10a	k-k	C	2.1[a, c] (1.2-3.5), 25	1.6[a, c] (0.65-2.6), 34	0.41
5885-19b	k-c	C	2.5[a, c] (1.4-3.2), 25	1.4[a, c] (0.75-2.6), 34	0.48
5885-15b	c-c	C	13[a, c] (8.6-18), 25	8.6[b, c] (5.8-13), 36	2.5
5885-6a	c-c	C	62[a] (46-97), 10		12
5885-9a	c-k	I	17[a, c] (11-23), 15		3.3
5885-4b	c-k	I	24[a, c] (16-29), 15		4.6
5885-16a	c-k	I	85[a] (37-222), 10		16
5885-19a	c-k	I	86[a] (20-130), 10		17

Table 2.5 (continued)

The isolates YJS5845 and YJS5885 and derived spore clones were transformed with pEAA613 (*ARS–CEN URA3* promoter-*kanMX::insE-A14*). 4 to 6 independent cultures of each transformant were examined for reversion to geneticin resistance as described in the Materials and Methods. Median mutation rates are presented with 95% confidence intervals, and relative mutation rates compared to EAY1369 (S288c compatible) are shown. n= number of independent repetitions. The genotype of the isolates and spore clones is presented with respect to the S288c (c) or SK1 (k) amino acid position in amino acid 761 in Mlh1 and 818/822 in Pms1 (Figure 2.2). Heterozygous genotypes are indicated by the "/". In this nomenclature c-c =S288c genotype, k-k = SK1, k-c = ancestral, and c-k = incompatible. **For 5845-21a, 5845-41a, 5885-10a, 5885-15b and 5885-19b spore clones there were 4 (of 8 tested), 4 (of 8 tested), 2 (of 7 tested), 2 (of 7 tested), 2 (of 6 tested) transformants, respectively, that gave low reversion rates (see Materials and Methods for details). The rate of reversion to $G418^r$ in this column include the data from the low reversion transformants; however, the column marked with * does not include these transformants, nor does the relative rate column. [a] Significantly different from EAY1369 ($p < 0.001$, Mann–Whitney test); [b] Significantly different from EAY1369 ($p < 0.01$, Mann–Whitney test); [c] Significantly different from EAY1370 ($p < 0.001$, Mann–Whitney test); [d] Significantly different from EAY1370 ($p < 0.01$, Mann–Whitney test). YJS5845 compatible and incompatible spore clones ($p < 0.001$) and YJS5885 compatible and incompatible spore clones ($p < 0.001$) are significantly different from each other (Mann-Whitney test).

Table 2.6: Analysis of *HO*, *PHO80* and *STP22* genes in YJS5845 and YJS5885 for variants using SnpEff.

Gene Name	Mutation	Type of mutation	Amino acid change	Predicted effect of missense mutation by SnpEff
YJS5845				
HO gene in Chromosome IV	1756T>C	missense	Cys586Arg	Deleterious
	1740C>T	synonymous	Asn580Asn	
	1722T>C	synonymous	His574His	
	1718C>T	missense	Pro573Leu	Tolerated
	1710C>T	synonymous	Val570Val	
	1635C>T	synonymous	Gly545Gly	
	1214C>T	missense	Ser405Leu	Tolerated
	1059T>C	synonymous	Val353Val	
	1026C>A	synonymous	Gly342Gly	
	789A>G	synonymous	Leu263Leu	
	667A>G	missense	Ser223Gly	Tolerated
	565G>A	missense	Ala189Thr	Tolerated
	369G>A	synonymous	Arg123Arg	
PHO80 in Chromosome XV	21A>C	missense	Glu7Asp	Tolerated
	111G>A	synonymous	Val37Val	
STP22 in Chromosome III	1000C>A	missense	Gln334Lys	Tolerated
	546A>G	synonymous	Pro182Pro	
	528T>C	synonymous	Asn176Asn	
	525G>A	synonymous	Gln175Gln	
	492C>G	synonymous	Pro164Pro	
	36G>A	synonymous	Ala12Ala	
YJS5885				
HO gene Chromosome IV	1740C>T	synonymous	Asn580Asn	
	1710C>T	synonymous	Val570Val	
	1635C>T	synonymous	Gly545Gly	
	1424T>A	missense	Leu475His	Tolerated
	1214C>T	missense	Ser405Leu	Tolerated
	667A>G	missense	Ser223Gly	Tolerated
	369G>A	synonymous	Arg123Arg	
PHO80 gene in Chromosome XV	21A>C	missense	Glu7Asp	Tolerated
	111G>A	synonymous	Val37Val	
	266C>T	missense	Ser89Phe	Tolerated
	375A>G	synonymous	Thr125Thr	
	739C>T	missense	Pro247Ser	Tolerated

Table 2.6 (continued)

Gene Name	Mutation	Type of mutation	Amino acid change	Predicted effect of missense mutation by SnpEff
YJS5885				
STP22 gene in Chromosome III	1000C>A	missense	Gln334Lys	Tolerated
	528T>C	synonymous	Asn176Asn	
	525G>A	synonymous	Gln175Gln	
	492C>G	synonymous	Pro164Pro	
	123T>A	missense	Asn41Lys	Tolerated
	78C>A	missense	Asn26Lys	Tolerated
	36G>A	synonymous	Ala12Ala	

To measure mutation rate, isolates and spore clones were transformed with pEAA613, a mutation rate reporter vector that measures reversion of an A14 sequence inserted into the *KANMX* gene (Figure 2.7A). We confirmed that in these isolates and spore clones resistance to G418 resulted in frameshift mutations in the A14 sequence that restored the *KANMX* reading frame (Figure 2.6). Mutation rates in EAY1369 (compatible) and EAY1370 (incompatible) controls were highly reproducible when measured with independent transformations and repetitions on different days. YJS5885 and YJS5845 were non-mutators but yielded spore clones with a range of mutation rates in a DNA slippage assay that varied over ~340-fold (Table 2.5; Figure 2.7). Mutation rates of incompatible spore clones were significantly different from compatible spore clones in both YJS5845 and YJS5885 as determined by Mann- Whitney U test ($p<0.001$, Figure 2.7B; Table 2.5). This indicates that the *cMLH1-kPMS1* genotype is the major contributor to high mutation rates, though there were examples where spore clones compatible

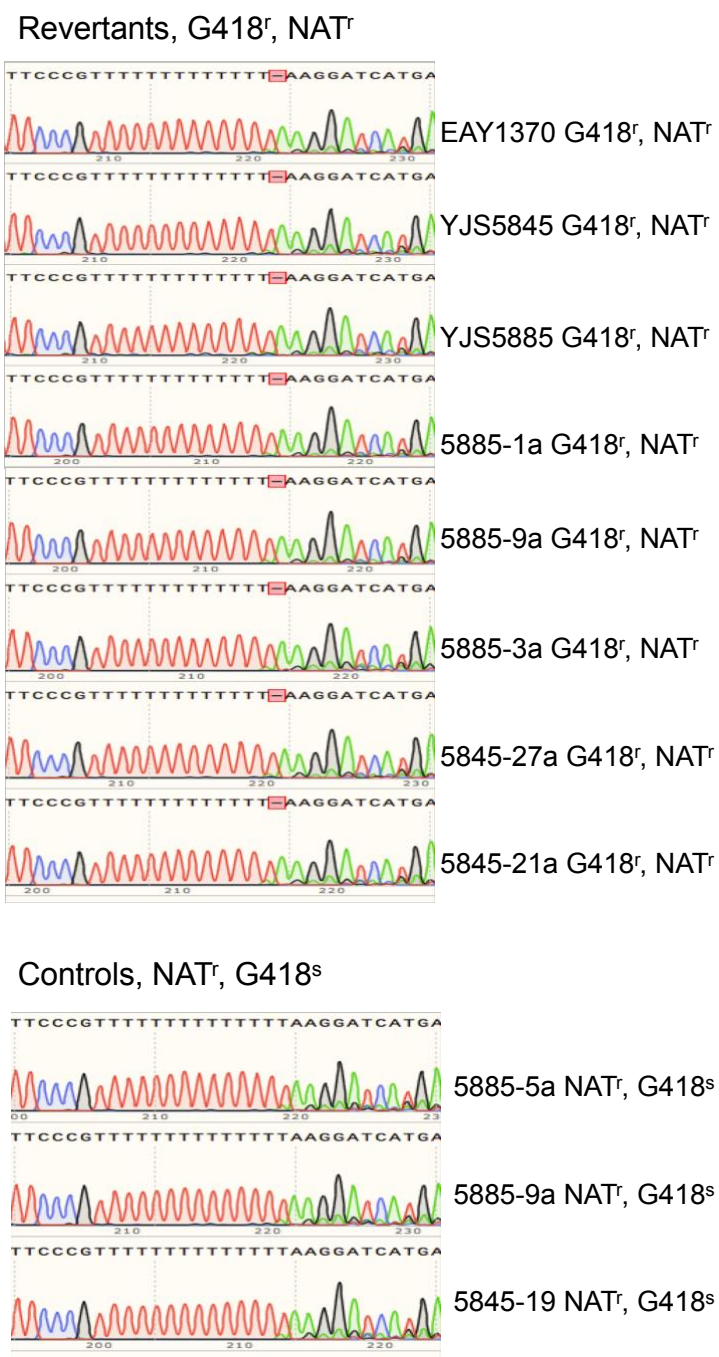


Figure 2.6: Sequencing analysis of G418 resistant revertants and sensitive control colonies. The homopolymeric A-runs in isolates and spore clones transformed with pEAA613 were sequenced (Materials and Method). The sequencing data from G418 resistant (top) and sensitive (bottom) colonies are presented. Only G418 resistant colonies displayed A_{14} to A_{13} frameshift mutations.

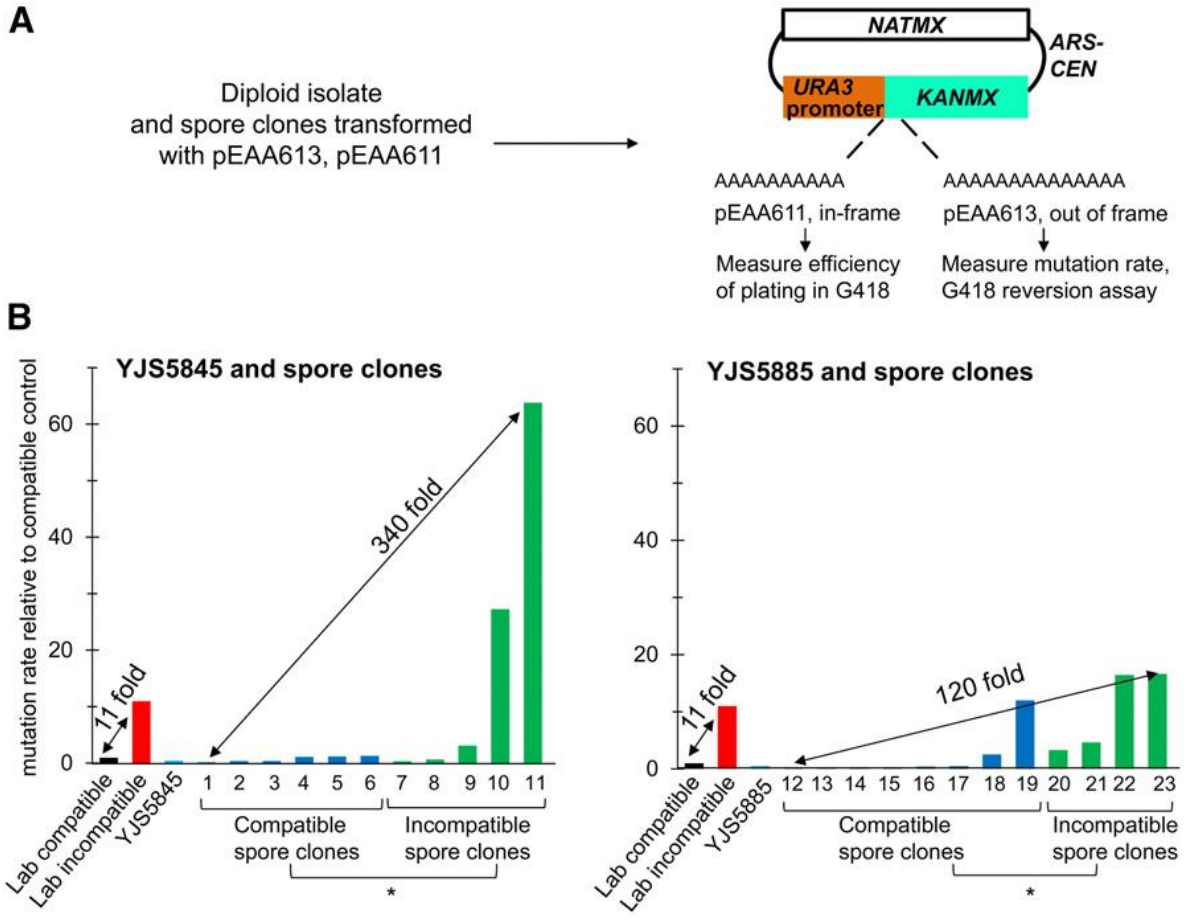


Figure 2.7: Mutation rate assay. (A) Isolates and spore clones were transformed with pEAA613, an ARS-CEN, URA3 promoter-KanMX::insE-A14 plasmid used to measure mutation rates in natural yeast isolates that contains a NatMX selectable marker and a frameshift reporter in which the insE-A14 sequence from Tran et al. (1997) was inserted immediately after methionine 17 in the KanMX open reading frame. pEAA611 is in-frame for KanMX whereas pEAA613 contains a +1-frameshift mutation that disrupts KanMX function. Frameshift mutation events (e.g., a 21 deletion in the homopolymeric A run) are detected on YPD plates containing clonNAT and geneticin (G418). Spore clones were first screened for efficiency of plating in G418 by transforming with pEAA611 that contains KanMX::insE-A10. (B) Mutation rates of YJS5845, YJS5885, and their spore clones in a G418 reversion assay relative to the compatible S288c-derived strain EAY1369. The parental isolates YJS5845 and YJS5885 are in light blue, compatible spore clones are in blue, incompatible spore clones are in green, the EAY1369 compatible laboratory strain is in black, and the EAY1370 incompatible laboratory strain is in red. Compatible and incompatible spore clones from YJS5845 and YJS5885 were significantly different from each other, as determined by Mann–Whitney U-test (* P , 0.001). The spore clones are named as follows: 1, 5845-7a; 2, 5845-16; 3, 5845-27a; 4, 5845-35a; 5, 5845-19a; 6, 5845-22a; 7, 5845-21a; 8, 5845-20a; 9, 5845-30a; 10, 5845-19; 11, 5845-18a; 12, 5885-1a; 13,

Figure 2.7 (continued)

5885-14a; 14, 5885-20b; 15, 5885-5b; 16, 5885-10a; 17, 5885-19b; 18, 5885-15b; 19, 5885-6a; 20, 5885-9a; 21, 5885-4b; 22, 5885-16a; and 23, 5885-19a

for *MLH1-PMS1* showed mutation rates similar to that observed in incompatible strains, and vice-versa (for example, see analyses of spore clones 5885-6a and 5845-21a in Figure 2.7 and Table 2.5). Several compatible and incompatible spore clones had significantly different mutation rates compared to EAY1369 (laboratory compatible) and EAY1370 (laboratory incompatible). As shown in Figure 2.7 and Table 2.5, up to a seven-fold lower mutation rate was seen between compatible spore clones and strain EAY1369 and up to a six-fold higher mutation rate was seen between incompatible spore clones and strain EAY1370. These data indicate that modifiers are present in the isolates that impact mutation rate (Figure 2.7B; Table 2.5). YJM521 displayed poor efficiency of plating with pEAA611 (a plasmid containing an in- frame homopolymeric A_{10} run in the *KANMX* gene; Figure 2.5). We also found that the previously characterized isolate YJM555 [15], a spore clone of YJM523, also displayed inefficient plating with pEAA611. Thus, it was not possible to accurately measure mutation rates in these isolates or their spore clones using pEAA613 (Table 2.5). However, YJM521 is unlikely to be a strong mutator because it is homozygous for the *MLH1-P271* suppressor allele (Table 2.4), and YJM555 was shown not to be a mutator using a 5-FOA based reversion assay [21], and had acquired suppressor mutations [15].

We also measured the rate of resistance to 5-fluoroorotic acid (FOA^r) in haploid YJS5885 spore clones (see FACS analysis below). This resistance results primarily from recessive base substitution mutations in the *URA3* gene that confer loss of function [48]. Zeyl and DeVisser [54] observed a 152-fold difference in the rate of 5-FOA^r between wild-type and *msh2* haploid strains and Thompson *et al.* [9] reported a 10-fold difference between wild type and mutator strains. We observed a 20-fold difference between the haploid wild type in the S288c background (7.9×10^{-8}) and an *mlh1*Δ derivative (1.6×10^{-6}; Table 2.7). Interestingly, we

observed a thirteen-fold range in mutation rate in a set of seven YJS5885 spore clones (4.8×10^{-8} to 6.4×10^{-7}; Table 2.7). We sequenced the *URA3* open reading frame from thirteen 5-FOA resistant colonies (Table 2.7). Ten independent *URA3* mutations were detected, three missense, five nonsense, and two single-base deletions. There were no mutations identified in three mutants. Similar percentages of 5-FOA^r colonies containing mutations in the *URA3* gene were seen in our study (77%) and an earlier one (87%; [48]). The rank order of mutation rates for the YJS5885 spore clones in the 5- FOA^r and *kanMX::insE-A14* reversion assays did not precisely correlate (Table 2.5; Table 2.7). This result is not surprising because the two assays measure different mutation spectra, with the *kanMX::insE-A14* reversion assay detecting DNA slippage events in homopolymeric repeats, and the 5-FOA^r assay detecting primarily base substitutions [48].

Table 2.7: Analysis of resistance to 5-FOA in YJS5885 spore clones.

Strain or spore	Incompatible/ Compatible	Rate 5-FOA^r (10^{-7}), (95% C.I.), n	Relative rate
FY90	C	0.79 (0.26-2.6), 22	1
EAY4087 (*mlh1Δ)*	Not applicable	16[a] (9.5-18), 20	20
5885-1a	C	1.9[b] (0.72-5.2), 15	2.4
5885-14a	C	0.98[b] (0.53-2.1), 15	1.2
5885-15b	C	0.48[b] (0.37-3.2), 15	0.61
5885-6a	C	0.68[b] (0.32-1.5), 15	0.86
5885-9a	I	6.0[a, b] (4.7-9.7), 15	7.6
5885-16a	I	1.5[b] (0.92-1.9), 15	1.9
5885-19a	I	6.4[a, b] (2.9-7.9), 15	8.1

The rate of resistance to 5-FOA, presented with 95% confidence intervals (95% C.I.), was determined for n independent cultures of FY90 and the indicated spore clones of YJS5885 as described in the Materials and Methods. The *URA3* open reading frame (ATG =+1) was sequenced from 7, 2, 1, 1, 1 and 1 independent 5-FOA^r colonies from 5885-9a, 5885-15b, 5885-1a, 5885-14a, 5885-16a, and 5885-19a, respectively. [a] Significantly different from FY90 ($p<0.001$, Mann-Whitney test); [b] Significantly different from EAY4087 ($p<0.001$, Mann-Whitney test). YJS5885 compatible and incompatible spore clones are significantly different from each other ($p<0.001$, Mann-Whitney test).

10 of the 13 spore clones contained single mutations in *URA3*, with the following distribution:
5885-9a: Two missense (bp287,A>T; bp542, G>A), One nonsense (bp577, G>T), Two single nucleotide deletions (bp178, A deleted; bp629, G deleted), no changes in ORF for two 5-FOA^r mutants.
5885-15b: One missense (bp205, T>C), one nonsense (bp345 G>A).
5885-1a: One nonsense (bp223 A>T).
5885-14a: One nonsense (bp593 T>A).
5885-16a: No changes in ORF for one 5-FOA^r mutant.
5885-19a: One nonsense (bp310 C>T).

YJS5845 and YJS5885 spore clones are diploid and haploid, respectively

Wild yeasts are primarily homothallic, indicating that daughter cells can switch mating type and mate with mother cells [55, 56]. YJM521 and YJM523 are homothallic [51] and we hypothesized that YJS5845 is also homothallic because it yielded spore clones (12 of 15) that can sporulate (Table 2.8). It appears that YJS5885 is functionally heterothallic because most of its spore clones (24 of 26) were unable to sporulate. We tested the ploidy of YJS5845 and YJS5885 spore clones for which we measured mutation rates by flow cytometry and showed that YJS5845 spore clones were diploid while those of YJS5885 were haploid (Figure 2.8A; Figure 2.9). There are no disruptions in the open reading frame of the *HO* gene, which codes for the endonuclease involved in mating type switching in both YJS5845 and YJS5885 ([57]; Table 2.6). However, there is a predicted deleterious variant in the *HO* gene of YJS5845 but not in YJS5885 (Table 2.6). To determine if defects in sporulation correlated to growth deficiency or defects in mitochondria (normal mitochondrial function is necessary for meiosis; [58], spore clones.

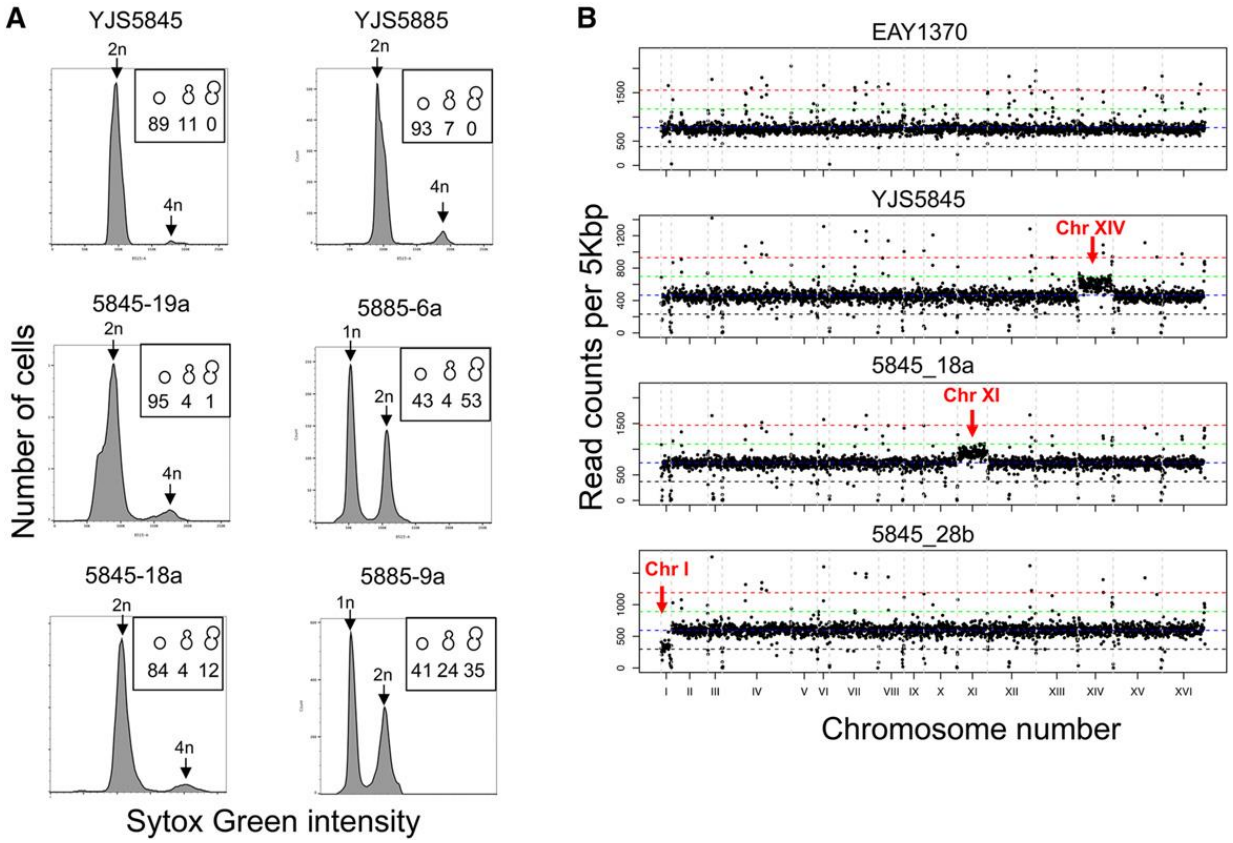


Figure 2.8: Ploidy of isolates and spore clones. (A) Representative flow cytometry plots of YJS5845 [59], YJS5885 (right), and derived spore clones. All spore clones of YJS5845 and YJS5885 tested were diploid and haploid, respectively. The black arrows show the position of 1n, 2n, and 4n DNA content. Inset shows percentage of single cells, small-budded cells, and large-budded cells assessed by light microscopy. (B) Whole-genome sequencing was performed for YJS5845 and spore clones (Materials and Methods). YJS5845, and spore clones 5845-18a and 5845-28b, displayed aneuploidy for chromosomes (Chr) XIV, XI, and I, respectively.

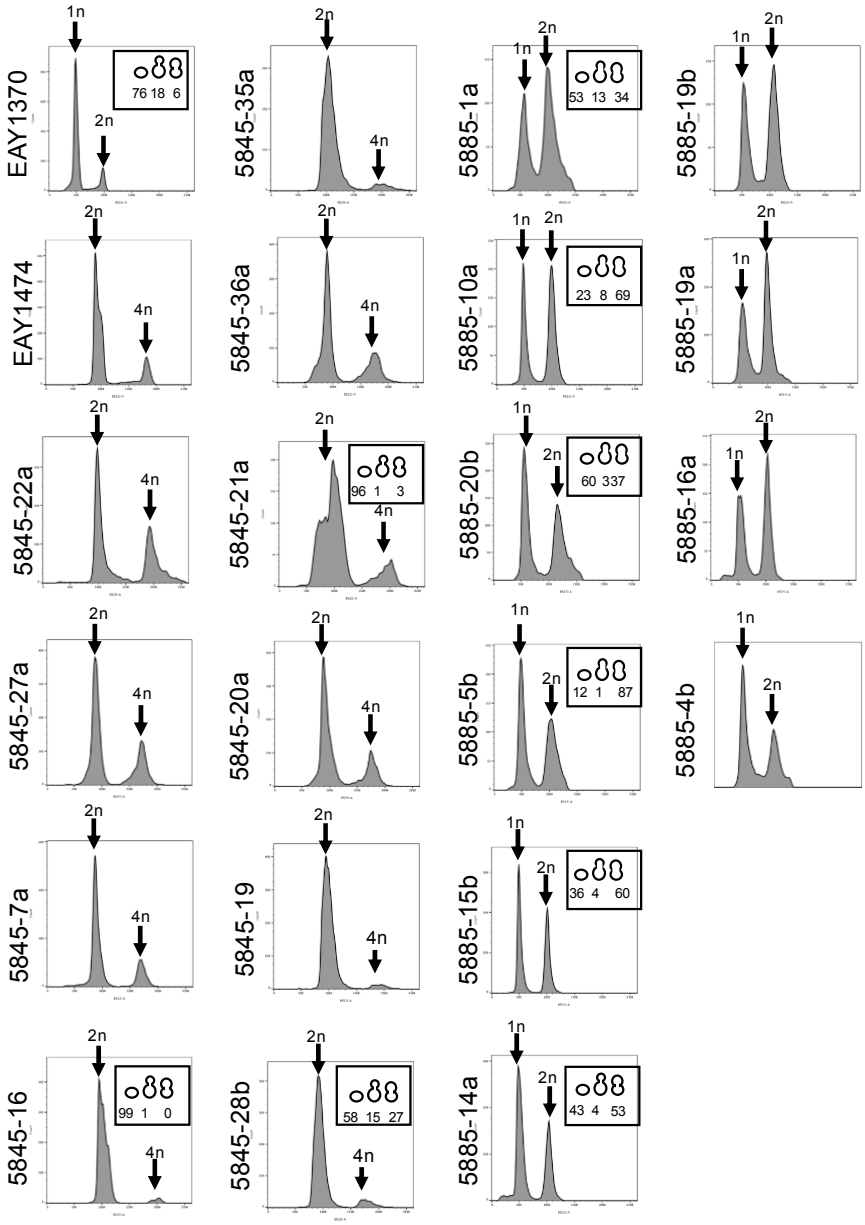


Figure 2.9: Flow cytometry of spore clones. Spore clones of YJS5845 and YJS5885 were prepared for flow cytometry as described in the Materials and Methods. YJS5845 derived spore clones were diploid and YJS5885 derived spore clones were haploid. The black arrows show the position of 1n, 2n, and 4n DNA content. Inset shows the percentage of single cells, small budded and large budded cells in the indicated samples.

Table 2.8: Sporulation and lactate growth phenotype

Spore clone	Sporulation	Lactate+ or -		Spore clone	Sporulation	Lactate + or -
5885-1a	-	+		5845-19a	-	-
5885-6a	-	+		5845-22a	+	+
5885-10a	-	-		5845-27a	+	+
5885-20b	-	+		5845-7a	+	+
5885-5b	-	+		5845-28b	+	+
5885-15b	-	-		5845-29a	+	+
5885-14a	-	+		5845-41a	-	-
5885-19b	-	-		5845-16	+	+
5885-6b	-	+		5845-35a	poor growth- few dyads	+
5885-5a	-	+		5845-18a	+	+
5885-11a	-	+		5845-21a	-	+
5885-12a	-	-		5845-20a	+	+
5885-18a	-	-		5845-30a	+	-
5885-9a	-	-		5845-19	+	+
5885-4b	-	+		5845-36b	+	NT
5885-19a	-	-				
5885-16a	-	-				
5885-3a	-	+				
5885-12b	-	-				
5885-1b	-	NT				
5885-2a	-	NT				
5885-4a	-	NT				
5885-7a	+	NT				
5885-8a	-	NT				
5885-13a	-	NT				
5885-17a	+	NT				

Spores were patched on sporulation media and incubated at 30°C for 6 days after which they were examined for evidence of sporulation by light microscopy. Any samples with dyads, triads and tetrads were marked as being able to sporulate (+). Spores were also patched on YP-lactate media and scored as able to grow or not (Lactate^{+} or $^{-}$) after 4 days in 30°C. NT: not tested

from YJS5845 and YJS5885 were analyzed for their ability to grow on media containing lactate as a carbon source. Three of fourteen YJS5845 spore clones and nine of nineteen YJS5885 spore clones failed to grow with lactate as a carbon source (Table 2.8). This explains why two of the YJS5845 spore clones were unable to sporulate, but we believe that YJS5885 is functionally heterothallic because its spore clones did not diploidize (Figure 2.8A and 2.9).

Spore clones from saturated cultures were prepared for flow cytometry and subsequently examined by light microscopy to determine the populations of single, small budded and large budded cells in each sample. We observed that YJS5885 spore clones had a higher proportion of large budded cells (34-87% large budded cells as compared to <1% large budded cells in the isolate YJS5885; Figure 2.8A and 2.9) suggesting a possible activation of a G2/M checkpoint or a cell division defect. Some cells of spore clones were much larger than the cells from the original isolate (5885-9a, 20b, 5b), which might be due to higher DNA content in the spore clones or the possible defects outlined above. Spore clones of YJS5845 were primarily single unbudded cells and appeared similar to the parental isolate.

Spore clones of YJS5845 display chromosome gain and loss

Whole genome sequence analyses of YJS5845 and YJS5885 and their spore clones indicated that most of the spore clones are euploid (Figure 2.8B and 2.10). Interestingly, YJS5845 and two of 16 spore clones analyzed, 5845-18a and 5845-28b, displayed aneuploidy, with gains of a chromosome in YJS5845 (XIV) and 5845-18a [60], and a loss of a chromosome in 5845-28b (I). The YJS5845 stock that we obtained is a mixture of euploid and aneuploid cells (trisomy in chromosome XIV that contains *PMS1*) as determined by sequencing the *PMS1* gene from several single colonies struck out from our original stock. Based on the heterozygosity in

the *PMS1* sequence, we determined that of the 14 single YJS5845 colonies sequenced for *PMS1*, six had three copies (indicating aneuploidy) and eight had two copies of *PMS1* (indicating euploidy). However, the sporulation efficiency and viability of euploid and aneuploid YJS5845 were very similar (Table 2.3). The majority of the spore clones (except for 5845-7a, 16, 19) were isolated from the euploid YJS5845. YJS5885 is euploid and all spore clones derived from YJS5885 were also euploid (Figure 2.10).

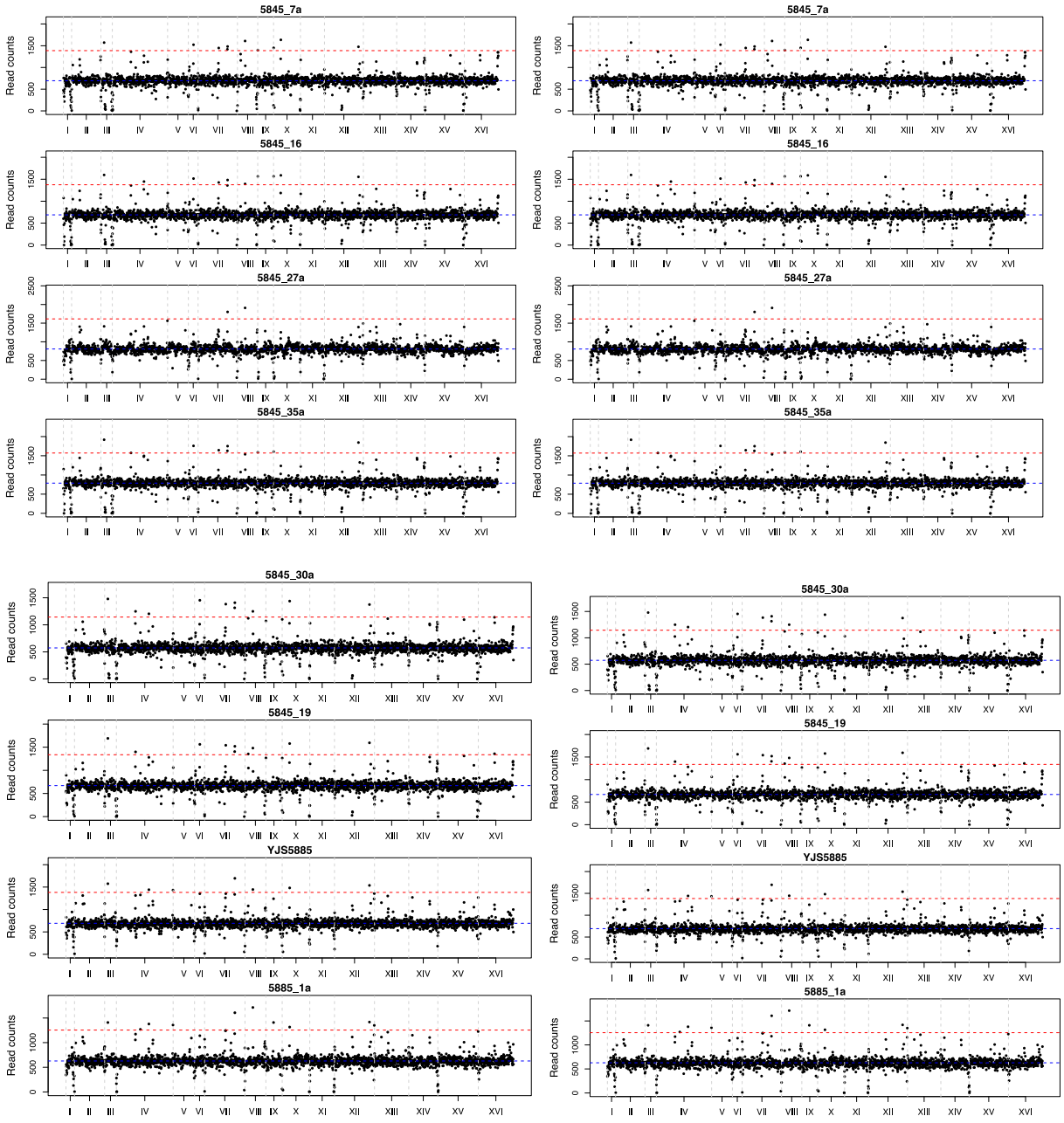


Figure 2.10: Ploidy of YJS5845 and YJS5885 isolates and their spore clones. Whole genome sequencing is presented for YJS5845, YJS5885 and derived spore clones (Materials and Methods). The entire set can be found in Figure 2.8 and Figure 2.10.

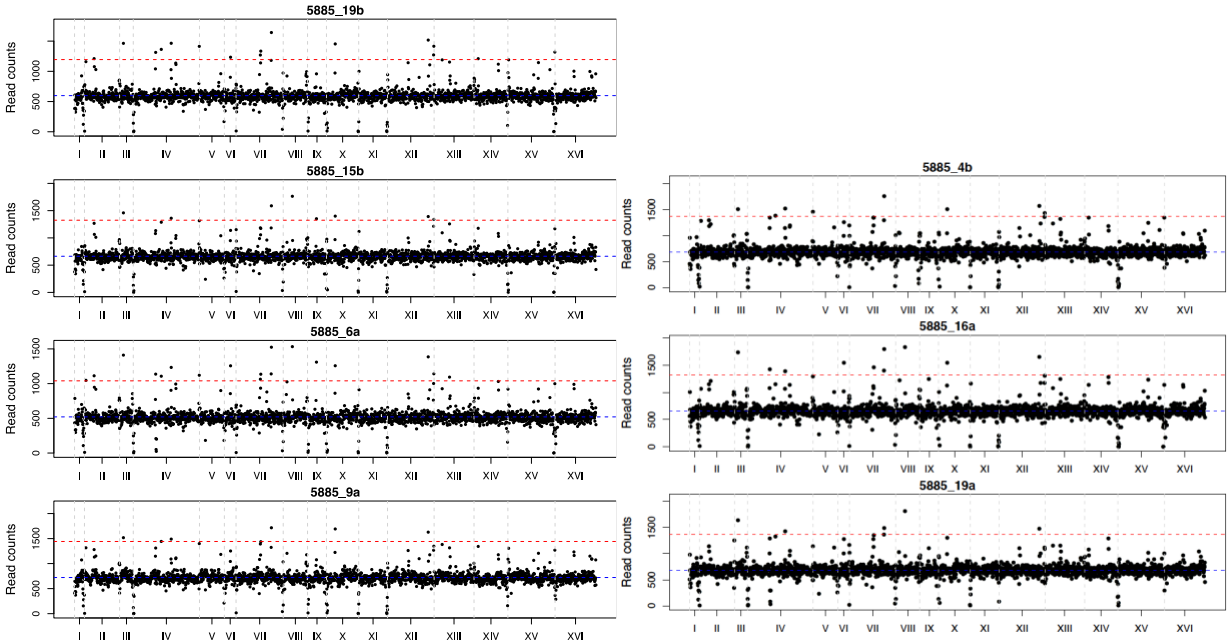


Figure 2.10 (continued)

Discussion

Spore clones of human clinical isolates yield a wide range of mutator phenotypes

Mutator phenotypes can be challenging to analyze in natural isolates because they lack genetic markers, display differences in growth, show colony variation, and can have differential resistance to antibiotics. We overcame these hurdles to measure mutation rates in YJS5845 and YJS5885 isolates and their spore clone derivatives using a recently developed frameshift mutation reporter. We found that while the isolates themselves were non-mutators, they sporulated at high efficiency and their spore clones displayed a wide range of mutation rates, with a 340-fold difference between the lowest and highest mutator using a DNA slippage assay. *cMLH1-kPMS1* incompatibility was the major contributor of the high mutation rate (Figure 2.7; Table 2.5). This study provides a practical and efficient strategy to characterize the phenotypic properties of human clinical yeast isolates. Other strategies such as bottleneck analysis of large numbers of independent isolates are also possible but are time consuming, and would require a significant effort to analyze a large amount of whole genome sequencing data.

Opposing forces of genetic drift and natural selection ensure that most individuals in a population evolve to a basal non-zero mutation rate (reviewed in [61]). For the natural isolates analyzed in this study, we hypothesize that stress can induce sporulation, giving rise to individuals that display high deviations from the basal mutation rate. Under unpredictable and changing stress conditions in the human host, it is likely to be valuable to produce spores with a broad range of mutation rates that will ensure population survival, but such a range might not be optimal for individual cells. Spore clones with a high mutation rate might rapidly gain beneficial mutations but will also acquire deleterious mutations. Thus, having a range of mutation rates gives the population an opportunity to survive changing stress conditions. To prevent long-term

fitness costs these strains can either acquire suppressor mutations to become non-mutators or mate back with the compatible strains or outcross to become diploid non-mutators. The compatible mutants surviving stress would also have a higher mutation rate, but it is reasonable to assume that diploids resulting from mating of compatible and incompatible spore clones would have complementary recessive mutations and thus a non- mutator phenotype (as seen for the low spore viability in YJS5845 and YJS5885).

Outcrossing with different isolates is thought to occur at a low frequency in the wild, once in every 50,000 generations [13, 62]. Stressful environments likely influence this rate, with levels estimated to be as high as one in every hundred to one in every two generations [63]. While such estimates are valuable, we hypothesize that the YJS5845 and YJS5885 isolates sporulated in stress conditions could yield mutator clones that are capable of mating with siblings, and thus not require outcrossing, to become compatible.

Modifiers of incompatibility genotypes in clinical isolates

The mutation rates of compatible and incompatible spore clones of YJS5845 and YJS5885 vary significantly from their corresponding laboratory strains (Figure 2.7; Table 2.5). This finding indicates the presence of enhancers and suppressors of mutation rate in the YJS5885 and YJS5845 isolate backgrounds. Previously we identified both intragenic (MLH1-P271; [43]) and extragenic modifiers [15] of the *cMLH1-kPMS1* incompatibility. In the case of the homozygous incompatible spore clone, YJM555, we observed what appeared to be multiple intergenic modifiers of the *cMLH1-kPMS1* incompatibility [15, 21].

Is there evidence for SNPs in *MLH1* conferring a mutator phenotype or acting as intragenic modifiers of the *cMLH1-kPMS1* incompatibility? 107 of the 1011 yeast isolates analyzed by Bui

et al. [15] display heterozygosities in the *MLH1* and *PMS1* genes. This information, and the fact that the yeast *MLH1* gene has been extensively mutagenized through alanine-scanning, random mutagenesis, and site-specific mutagenesis (both in known domains and in homology modeling to HNPCC alleles; see Table 2.9) encouraged us to determine if any MMR alleles exist in heterozygous isolates that could confer a deleterious phenotype. As shown in Tables 2.9, we mapped heterozygosities present in 107 (including a newly identified one) isolates onto the MLH1 structure-function map and found that some cluster to regions in MLH1 predicted to affect function. These observations provide support for the idea that mutators in yeast grown in stressful conditions could be obtained through incompatibilities as well as the presence of recessive deleterious alleles.

Table 2.9: Assigning MLH1 polymorphisms found in heterozygous genotypes onto the MLH1 structure-function map.

***mlh1* allele (Reference)**	**Amino acid position: isolate(s) with heterozygous genotype**
N-terminal/ATP binding	
I22T [64]	22-ILE/LEU: APL, CFI, CFN
Linker [40]	
R390A,K391A	391-LYS/ASN: BHL, BMQ
K393A,R394A [40]	393-LYS/GLU: CPN, CPR
R401A,D403A	402-ILE/LEU: ASN, BGB, BGI, BGM, BGS
C-terminal interaction [40]	
E603A,D605A,E606A	605-ASP/ASN: BHL, BMQ
K648A,K650A	650-LYS/THR: ADL, AKT, BFE, BFG, BFM, BML, BMM, BTD, BTE, CKK, CKS, CLA
E680A,D681A,E682A	681-ASP/ASN: CPN, CPR

A structure function map for MLH1 was created from an analysis of *MLH1* alanine scan and site-specific mutations, and *mlh1* alleles generated based on homology to HNPCC mutations [64-72]. Alleles that conferred a mutator phenotype in a variety of reporter assays are shown. In MLH1, amino acids 1-335 is referred to as that N-terminal/ATP binding domain, 335-509 as the linker domain, and 510-769 as the C-terminal interaction domain [73].

Yeast growing in a human environment are likely to encounter multiple stresses

Human clinical isolates of the fungal pathogen *Cryptococcus* have been identified that display a mutator phenotype due to mutations in the *MSH2* mismatch repair gene [32, 33]. The authors of these studies suggest that pathogens undergo a significant change in environment when entering a human host and that challenges specific to the human host, such as survival in the presence of antifungals, provide an adaptive advantage for *Cryptococcus* isolates that are mutators [32, 33]. In a human environment *S. cerevisiae* is likely to deal with challenges that are analogous to those faced by *Cryptococcus,* including growth at high temperature and exposure to fungicides. In fact, several causative alleles were identified in clinical and laboratory yeast strains that provide growth advantages at high temperature [74, 75]. As summarized in the Introduction, many human clinical isolates of *S. cerevisiae* are mosaics (including YJS5845 and YJS5885) that contain a mixture of alleles from different subpopulations [16, 19, 21, 76]. Our study of YJS5845 and YJS5885 highlights how isolates can adapt to stressful human environments. The incompatible alleles do not confer a mutator phenotype in these heterozygotes but may provide an advantage as they appear to be poised to adapt to stress through the variable mutation rates in the progeny (Figure 2.7B, Table 2.5). This variation in mutation rates is due to the presence of modifiers in the background of these strains as a consequence of being heterozygotes. Furthermore, one of the isolates and its spore clones show evidence of losing and gaining chromosomes at high frequency (Figure 2.8B), which could also be highly beneficial for rapid adaptation in the clinical setting by affecting mutation rates. Variance in mutation rate, at least for baker's yeast, is likely to provide only a transient advantage due to fitness costs, and changes in environment would need to be accompanied by the acquisition of suppressor

mutations, mating with nearby spore clones, return to euploidy, or by outcrossing to become non-mutators.

Ploidy, mutation rate and adaptation

Chromosome aneuploidy and increases in ploidy have been shown to provide faster routes to adaptation by increasing mutation rate and increasing the likelihood of gaining a beneficial mutation [22-26, 77]. For example, compared to haploid and diploid baker's yeast, tetraploid yeasts display more rapid adaptation, suggesting that increased ploidy impacts the rate of adaptation by providing a broad spectrum of adaptive mutations [29, 78]. Consistent with this observation, diploid mutators display a growth advantage over diploid and haploid non-mutators and haploid mutators in several stress conditions [9]. In our study we identified both haploid (YJS5885 derived) and diploid (YJS5845 derived) spore clone progeny, suggesting that ploidy may also play a role in regulating mutation rates of spore clones.

Zhu *et al*. [20, 26] analyzed 132 clinical isolates of *S. cerevisiae* by whole genome sequencing and found that roughly one third had higher levels of ploidy (3n and 4n genome copy number), one quarter had partial chromosome copy-number variations and one third were aneuploid. We observed that our YJS5845 stock consisted of a mix of euploid and aneuploid cells. Furthermore, euploid YJS5845 yielded aneuploid spore clones, either through meiotic or mitotic chromosome segregation defects (Figure 2.8B). Interestingly, YJS5845 belongs to an admixture clade; thus, it is reasonable to assume that the chromosome segregation defects seen in this isolate are due to incompatibilities in processes involving a large number of components. While we do not have direct evidence that the aneuploidy seen in YJS5845 impacts mutation rate, work from the studies referenced above suggest that it is likely to play a significant role.

Acknowledgments

We thank members of Alani Lab for their helpful comments and advice, V.P. Ajith and K.T. Nishant at IISER Trivandrum for sharing the aneuploidy script and for their help in analyzing whole genome sequencing data, Christopher Donahue at Flow Cytometry Core Laboratory for his help in running the flow cytometry and sorting experiments, and Peter Schweitzer and Jeff Glaubitz at the BRC sequencing facility for performing whole genome sequencing. J.S. is a member of the Institut Universitaire de France. V.R., D.T.B., N.A.-S., and E.A. were supported by National Institutes of Health (NIH) grant GM53085. C.F.A. was supported by NIH grant GM095793. The 1011 Yeast Genomes Project (J.S. and A.F.) was funded by France Génomique (ANR-10-INBS-09-08). D.T.B. was a fellow of the Vietnam Education Foundation. N.A.-S. was supported by a scholarship from the Saudi Arabian Cultural Mission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

References

1. Chao, L. and Cox, E.C. (1983) Competition between High and Low Mutating Strains of *Escherichia coli*. Evolution 37, 125-134 DOI: 10.1111/j.1558-5646.1983.tb05521.x.

2. LeClerc, J.E., Li, B., Payne, W.L. and Cebula, T.A. (1996) High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science 274, 1208-11.

3. Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H. *et al.* (1997) Role of mutator alleles in adaptive evolution. Nature 387, 700-2 DOI: 10.1038/42696.

4. Boe, L., Danielsen, M., Knudsen, S., Petersen, J.B., Maymann, J. *et al.* (2000) The frequency of mutators in populations of *Escherichia coli*. Mutat Res 448, 47-55.

5. Denamur, E., Lecointre, G., Darlu, P., Tenaillon, O., Acquaviva, C. *et al.* (2000) Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. Cell 103, 711-21.

6. Giraud, A., Matic, I., Tenaillon, O., Clara, A., Radman, M. *et al.* (2001) Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. Science 291, 2606-8 DOI: 10.1126/science.1056421.

7. Tanaka, M.M., Bergstrom, C.T. and Levin, B.R. (2003) The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. Genetics 164, 843-54.

8. Townsend, J.P., Nielsen, K.M., Fisher, D.S. and Hartl, D.L. (2003) Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. Genetics 164, 13-21.

9. Thompson, D.A., Desai, M.M. and Murray, A.W. (2006) Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. Curr Biol 16, 1581-90 DOI: 10.1016/j.cub.2006.06.070.

10. Raynes, Y., Gazzara, M.R. and Sniegowski, P.D. (2011) Mutator dynamics in sexual and asexual experimental populations of yeast. BMC Evol Biol 11, 158 DOI: 10.1186/1471-2148-11-158.

11. Bui, D.T., Dine, E., Anderson, J.B., Aquadro, C.F. and Alani, E.E. (2015) A Genetic Incompatibility Accelerates Adaptation in Yeast. PLoS Genet 11, e1005407 DOI: 10.1371/journal.pgen.1005407.

12. Liti, G. and Louis, E.J. (2005) Yeast evolution and comparative genomics. Annu Rev Microbiol 59, 135-53 DOI: 10.1146/annurev.micro.59.030804.121400.

13. Ruderfer, D.M., Pratt, S.C., Seidel, H.S. and Kruglyak, L. (2006) Population genomic analysis of outcrossing and recombination in yeast. Nat Genet 38, 1077-81 DOI: 10.1038/ng1859.

14. Nishant, K.T., Chen, C., Shinohara, M., Shinohara, A. and Alani, E. (2010) Genetic analysis of baker's yeast Msh4-Msh5 reveals a threshold crossover level for meiotic viability. PLoS Genet 6, DOI: 10.1371/journal.pgen.1001083.

15. Bui, D.T., Friedrich, A., Al-Sweel, N., Liti, G., Schacherer, J. *et al.* (2017) Mismatch Repair Incompatibilities in Diverse Yeast Populations. Genetics 205, 1459-1471 DOI: 10.1534/genetics.116.199513.

16. Perez-Torrado, R. and Querol, A. (2015) Opportunistic Strains of *Saccharomyces cerevisiae*: A Potential Risk Sold in Food Products. Front Microbiol 6, 1522 DOI: 10.3389/fmicb.2015.01522.

17. Hall, R.A. and Noverr, M.C. (2017) Fungal interactions with the human host: exploring the spectrum of symbiosis. Curr Opin Microbiol 40, 58-64 DOI: 10.1016/j.mib.2017.10.020.

18. Schacherer, J., Shapiro, J.A., Ruderfer, D.M. and Kruglyak, L. (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. Nature 458, 342-5 DOI: 10.1038/nature07670.

19. Strope, P.K., Skelly, D.A., Kozmin, S.G., Mahadevan, G., Stone, E.A. *et al.* (2015) The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. Genome Res 25, 762-74 DOI: 10.1101/gr.185538.114.

20. Zhu, Y.O., Sherlock, G. and Petrov, D.A. (2016) Whole Genome Analysis of 132 Clinical *Saccharomyces cerevisiae* Strains Reveals Extensive Ploidy Variation. G3 6, 2421-34 DOI: 10.1534/g3.116.029397.

21. Skelly, D.A., Magwene, P.M., Meeks, B. and Murphy, H.A. (2017) Known mutator alleles do not markedly increase mutation rate in clinical *Saccharomyces cerevisiae* strains. Proc Biol Sci 284, pii: 20162672 DOI: 10.1098/rspb.2016.2672.

22. Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B. and Berman, J. (2009) Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. PLoS Genet 5, e1000705 DOI: 10.1371/journal.pgen.1000705.

23. Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A. *et al.* (2010) Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature 468, 321-5 DOI: 10.1038/nature09529.

24. Rancati, G. and Pavelka, N. (2013) Karyotypic changes as drivers and catalyzers of cellular evolvability: a perspective from non-pathogenic yeasts. Semin Cell Dev Biol 24, 332-8 DOI: 10.1016/j.semcdb.2013.01.009.

25. Zorgo, E., Chwialkowska, K., Gjuvsland, A.B., Garre, E., Sunnerhagen, P. *et al.* (2013) Ancient evolutionary trade-offs between yeast ploidy states. PLoS Genet 9, e1003388 DOI: 10.1371/journal.pgen.1003388.

26. Zhu, Y.O., Siegal, M.L., Hall, D.W. and Petrov, D.A. (2014) Precise estimates of mutation rate and spectrum in yeast. Proc Natl Acad Sci U S A 111, E2310-8 DOI: 10.1073/pnas.1323011111.

27. Cromie, G.A. and Dudley, A.M. (2015) Aneuploidy: Tolerating Tolerance. Curr Biol 25, R771-3 DOI: 10.1016/j.cub.2015.06.056.

28. Hose, J., Yong, C.M., Sardi, M., Wang, Z., Newton, M.A. *et al.* (2015) Dosage compensation can buffer copy-number variation in wild yeast. Elife 4, e05462 DOI: 10.7554/eLife.05462.

29. Selmecki, A.M., Maruvka, Y.E., Richmond, P.A., Guillet, M., Shoresh, N. *et al.* (2015) Polyploidy can drive rapid adaptation in yeast. Nature 519, 349-52 DOI: 10.1038/nature14187.

30. Sirr, A., Cromie, G.A., Jeffery, E.W., Gilbert, T.L., Ludlow, C.L. *et al.* (2015) Allelic variation, aneuploidy, and nongenetic mechanisms suppress a monogenic trait in yeast. Genetics 199, 247-62 DOI: 10.1534/genetics.114.170563.

31. Sunshine, A.B., Payen, C., Ong, G.T., Liachko, I., Tan, K.M. *et al.* (2015) The fitness consequences of aneuploidy are driven by condition-dependent gene effects. PLoS Biol 13, e1002155 DOI: 10.1371/journal.pbio.1002155.

32. Billmyre, R.B., Clancey, S.A. and Heitman, J. (2017) Natural mismatch repair mutations mediate phenotypic diversity and drug resistance in Cryptococcus deuterogattii. Elife 6, e28802 DOI: 10.7554/eLife.28802.

33. Boyce, K.J., Wang, Y., Verma, S., Shakya, V.P.S., Xue, C. *et al.* (2017) Mismatch Repair of DNA Replication Errors Contributes to Microevolution in the Pathogenic Fungus Cryptococcus neoformans. MBio 8, e00595-17 DOI: 10.1128/mBio.00595-17.

34. Heck, J.A., Argueso, J.L., Gemici, Z., Reeves, R.G., Bernard, A. *et al.* (2006) Negative epistasis between natural variants of the *Saccharomyces cerevisiae* MLH1 and PMS1 genes results in a defect in mismatch repair. Proc Natl Acad Sci U S A 103, 3256-61 DOI: 10.1073/pnas.0510998103.

35. Kariola, R., Otway, R., Lonnqvist, K.E., Raevaara, T.E., Macrae, F. *et al.* (2003) Two mismatch repair gene mutations found in a colon cancer patient--which one is pathogenic? Hum Genet 112, 105-9 DOI: 10.1007/s00439-002-0866-4.

36. Peter, J., De Chiara, M., Friedrich, A., Yue, J.X., Pflieger, D. *et al.* (2018) Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature 556, 339-344 DOI: 10.1038/s41586-018-0030-5.

37. Richard, M. and Yvert, G. (2014) How does evolution tune biological noise? Front Genet 5, 374 DOI: 10.3389/fgene.2014.00374.

38. Rose, M., Winston, F. and Hieter, P. (1990) Methods in yeast genetics: A Laboratory Course Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

39. Gietz, R.D. and Schiestl, R.H. (2007) Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc 2, 38-41 DOI: 10.1038/nprot.2007.15.

40. Argueso, J.L., Kijas, A.W., Sarin, S., Heck, J., Waase, M. *et al.* (2003) Systematic mutagenesis of the *Saccharomyces cerevisiae* MLH1 gene reveals distinct roles for Mlh1p in meiotic crossing over and in vegetative and meiotic mismatch repair. Mol Cell Biol 23, 873-86.

41. Tran, H.T., Keen, J.D., Kricker, M., Resnick, M.A. and Gordenin, D.A. (1997) Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. Mol Cell Biol 17, 2859-65.

42. Hoffman, C.S. and Winston, F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene 57, 267-72.

43. Demogines, A., Wong, A., Aquadro, C. and Alani, E. (2008) Incompatibilities involving yeast mismatch repair genes: a role for genetic modifiers and implications for disease penetrance and variation in genomic mutation rates. PLoS Genet 4, e1000103 DOI: 10.1371/journal.pgen.1000103.

44. Zubko, E.I. and Zubko, M.K. (2014) Deficiencies in mitochondrial DNA compromise the survival of yeast cells at critically high temperatures. Microbiol Res 169, 185-95 DOI: 10.1016/j.micres.2013.06.011.

45. Dixon, W.J. and Massey, F.J. (1969) Introduction to Statistical Analysis. McGraw-Hill, New York.

46. Wilcoxon, F. (1945) Individual comparisons by ranking methods. Biometrics Bulletin 80-83 DOI: https://doi.org/10.2307/3001968.

47. Mann, H.B. and Whitney, D.R. (1947) On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 50-60 DOI: https://doi.org/10.1214/aoms/1177730491.

48. Lang, G.I. and Murray, A.W. (2008) Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. Genetics 178, 67-82 DOI: 10.1534/genetics.107.071506.

49. Rosebrock, A.P. (2017) Analysis of the Budding Yeast Cell Cycle by Flow Cytometry. Cold Spring Harb Protoc 2017, pdb prot088740 DOI: 10.1101/pdb.prot088740.

50. Chan, C.S. and Botstein, D. (1993) Isolation and characterization of chromosome-gain and increase-in-ploidy mutants in yeast. Genetics 135, 677-91.

51. Clemons, K.V., Park, P., McCusker, J.H., McCullough, M.J., Davis, R.W. *et al.* (1997) Application of DNA typing methods and genetic analysis to epidemiology and taxonomy of *Saccharomyces* isolates. J Clin Microbiol 35, 1822-8.

52. Wickert, S., Finck, M., Herz, B. and Ernst, J.F. (1998) A small protein (Ags1p) and the Pho80p-Pho85p kinase complex contribute to aminoglycoside antibiotic resistance of the yeast *Saccharomyces cerevisiae*. J Bacteriol 180, 1887-94.

53. Ernst, J.F. and Chan, R.K. (1985) Characterization of *Saccharomyces cerevisiae* mutants supersensitive to aminoglycoside antibiotics. J Bacteriol 163, 8-14.

54. Zeyl, C. and DeVisser, J.A. (2001) Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. Genetics 157, 53-61.

55. Mortimer, R.K. (2000) Evolution and variation of the yeast (*Saccharomyces*) genome. Genome Res 10, 403-9.

56. Butler, G., Kenny, C., Fagan, A., Kurischko, C., Gaillardin, C. *et al.* (2004) Evolution of the MAT locus and its Ho endonuclease in yeast species. Proc Natl Acad Sci U S A 101, 1632-7 DOI: 10.1073/pnas.0304170101.

57. Nasmyth, K. (1993) Regulating the HO endonuclease in yeast. Curr. Opin. Genet. Dev. 286–294.

58. Gorsich, S.W. and Shaw, J.M. (2004) Importance of mitochondrial dynamics during meiosis and sporulation. Mol Biol Cell 15, 4369-81 DOI: 10.1091/mbc.e03-12-0875.

59. Yupsanis, T., Eleftheriou, P. and Kelepiri, Z. (1996) Separation and purification of both acid and neutral nucleases from germinated alfalfa seeds. Journal of Plant Physiology 149, 641-649.

60. Xu, K., Lu, T., Zhou, H., Bai, L. and Xiang, Y. (2010) The role of MSH5 C85T and MLH3 C2531T polymorphisms in the risk of male infertility with azoospermia or severe oligozoospermia. Clin Chim Acta 411, 49-52 DOI: 10.1016/j.cca.2009.09.038.

61. Lynch, M., Ackerman, M.S., Gout, J.F., Long, H., Sung, W. *et al.* (2016) Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet 17, 704-714 DOI: 10.1038/nrg.2016.104.

62. Magwene, P.M., Kayikci, O., Granek, J.A., Reininga, J.M., Scholl, Z. *et al.* (2011) Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 108, 1987-92 DOI: 10.1073/pnas.1012544108.

63. Marsit, S. and Dequin, S. (2015) Diversity and adaptive evolution of *Saccharomyces* wine yeast: a review. FEMS Yeast Res 15, pii: fov067 DOI: 10.1093/femsyr/fov067.

64. Wanat, J.J., Singh, N. and Alani, E. (2007) The effect of genetic background on the function of *Saccharomyces cerevisiae* mlh1 alleles that correspond to HNPCC missense mutations. Hum Mol Genet 16, 445-52 DOI: 10.1093/hmg/ddl479.

65. Pang, Q., Prolla, T.A. and Liskay, R.M. (1997) Functional domains of the *Saccharomyces cerevisiae* Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations. Mol Cell Biol 17, 4465-73.

66. Romanova, N.V. and Crouse, G.F. (2013) Different roles of eukaryotic MutS and MutL complexes in repair of small insertion and deletion loops in yeast. PLoS Genet 9, e1003920 DOI: 10.1371/journal.pgen.1003920.

67. Shcherbakova, P.V. and Kunkel, T.A. (1999) Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. Mol Cell Biol 19, 3177-83.

68. Smith, C.E., Bowen, N., Graham, W.J.t., Goellner, E.M., Srivatsan, A. *et al.* (2015) Activation of *Saccharomyces cerevisiae* Mlh1-Pms1 Endonuclease in a Reconstituted Mismatch Repair System. J Biol Chem 290, 21580-90 DOI: 10.1074/jbc.M115.662189.

69. Smith, C.E., Mendillo, M.L., Bowen, N., Hombauer, H., Campbell, C.S. *et al.* (2013) Dominant mutations in *S. cerevisiae* PMS1 identify the Mlh1-Pms1 endonuclease active site and an exonuclease 1-independent mismatch repair pathway. PLoS Genet 9, e1003869 DOI: 10.1371/journal.pgen.1003869.

70. Takahashi, M., Shimodaira, H., Andreutti-Zaugg, C., Iggo, R., Kolodner, R.D. *et al.* (2007) Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. Cancer Res 67, 4595-604 DOI: 10.1158/0008-5472.CAN-06-3509.

71. Tran, P.T. and Liskay, R.M. (2000) Functional studies on the candidate ATPase domains of *Saccharomyces cerevisiae* MutLalpha. Mol Cell Biol 20, 6390-8.

72. Welz-Voegele, C., Stone, J.E., Tran, P.T., Kearney, H.M., Liskay, R.M. *et al.* (2002) Alleles of the yeast Pms1 mismatch-repair gene that differentially affect recombination- and replication-related processes. Genetics 162, 1131-45.

73. Gueneau, E., Dherin, C., Legrand, P., Tellier-Lebegue, C., Gilquin, B. *et al.* (2013) Structure of the MutLα C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. Nat Struct Mol Biol 20, 461-8 DOI: 10.1038/nsmb.2511.

74. Steinmetz, L.M., Sinha, H., Richards, D.R., Spiegelman, J.I., Oefner, P.J. *et al.* (2002) Dissecting the architecture of a quantitative trait locus in yeast. Nature 416, 326-30 DOI: 10.1038/416326a.

75. Sinha, H., David, L., Pascon, R.C., Clauder-Munster, S., Krishnakumar, S. *et al.* (2008) Sequential elimination of major-effect contributors identifies additional quantitative trait loci conditioning high-temperature growth in yeast. Genetics 180, 1661-70 DOI: 10.1534/genetics.108.092932.

76. Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L. *et al.* (2009) Population genomics of domestic and wild yeasts. Nature 458, 337-41 DOI: 10.1038/nature07743.

77. Gerstein, A.C. and Berman, J. (2015) Shift and adapt: the costs and benefits of karyotype variations. Curr Opin Microbiol 26, 130-6 DOI: 10.1016/j.mib.2015.06.010.

78. Scott, A.L., Richmond, P.A., Dowell, R.D. and Selmecki, A.M. (2017) The Influence of Polyploidy on the Evolution of Yeast Grown in a Sub-Optimal Carbon Source. Mol Biol Evol 34, 2690-2703 DOI: 10.1093/molbev/msx205.

CHAPTER 3

Purification of yeast and mouse Mlh1-Mlh3 complexes; biochemical analysis of yeast Mlh1-mlh3 separation of function complexes and initial purification of mouse Mlh1-Mlh3.

This work is a part of a study that was published in PLoS Genetics [1] and bioRxiv [2].

Appendix A represents a new protocol for purification of yeast Mlh1-Mlh3

Abstract

Mlh1-Mlh3 is an endonuclease hypothesized to act in meiosis to resolve double Holliday junctions into crossovers. It also plays a minor role in eukaryotic DNA mismatch repair (MMR). To understand how Mlh1-Mlh3 functions in both meiosis and MMR, we analyzed in baker's yeast 60 new *mlh3* alleles. Five alleles specifically disrupted MMR, whereas one (*mlh3-32*) specifically disrupted meiotic crossing over. Mlh1-mlh3 representatives for each class were purified and characterized. Both Mlh1-mlh3-32 (MMR^+, $crossover^-$) and Mlh1- mlh3-45 (MMR^-, $crossover^+$) displayed wild-type endonuclease activities *in vitro*. Msh2- Msh3, an MSH complex that acts with Mlh1-Mlh3 in MMR, stimulated the endonuclease activity of Mlh1-mlh3-32 but not Mlh1-mlh3-45, suggesting that Mlh1-mlh3-45 is defective in MSH interactions. My observations reveal the importance of protein-protein interactions in regulating the enzymatic activity of Mlh1-Mlh3 protein complex. As part of this work, I also developed a strategy to purify the mouse Mlh1-Mlh3 complex and demonstrate that both the wild type and endonuclease active site mutant complexes can be purified as heterodimers.

Introduction

During mismatch repair (MMR), insertion/deletion and base-base mismatches that form as the result of DNA replication errors are recognized by MutS homolog (MSH) proteins, which in turn recruit MutL homolog (MLH) proteins to form ternary complexes containing mismatched DNA, MSH factors, and MLH factors. These interactions result in the recruitment of downstream excision and resynthesis proteins to remove the error [3]. In *S. cerevisiae* repair of insertion deletion loops greater than one nucleotide in size primarily involves the MSH heterodimer Msh2-Msh3 and the MLH heterodimer Mlh1-Pms1 [3]. The MLH heterodimer Mlh1-Mlh3 has been shown to play a minor role in this process and can partially substitute for Mlh1-Pms1 in Msh2-Msh3-dependent MMR [4-6]. However, Mlh1-Mlh3 has been shown to play a major role in meiotic crossing over [7-10]. Accurate chromosome segregation in Meiosis I in most eukaryotes requires reciprocal exchange of genetic information (crossing over) between homologs [11-14]. Failure to achieve at least one crossover (CO) per homolog pair results in homolog nondisjunction and the formation of aneuploid gametes. Errors in meiotic chromosome segregation are a leading cause of spontaneous miscarriages and birth defects [15].

Mlh1-Mlh3 is an endonuclease that nicks circular duplex DNA *in vitro*, and Mlh1-mlh3-D523N is defective in endonuclease activity [16, 17]. This activity is dependent on the integrity of a highly conserved (DQHA(X)$_2$E(X)$_4$E) metal binding motif found in Mlh3. Previous work demonstrated that a point mutation within this motif (*mlh3-D523N*) conferred *mlh3*Δ- like defects in MMR and crossing over. These included a mutator phenotype, a decrease in spore viability to 70% (from 97% in wild-type), and a two-fold reduction in genetic map distances [8].

Approximately 200 double strand breaks (DSBs) are induced throughout the genome in a *S. cerevisiae* cell in meiotic prophase, of which ~90 are repaired to form COs between

homologous chromosomes, with the rest repaired to form noncrossovers (NCOs; [18-20]). In this pathway a DSB is resected, resulting in the formation of 3' single-strand tails. One of these tails invades the other homolog and creates a single-end invasion intermediate (SEI). A second invasion event initiating from the SEI, known as second-end capture, can re-anneal and ligate to the other side of the DSB resulting in the formation of a double Holliday junction (dHJ). The dHJ can be acted upon by Holliday junction (HJ) resolvases to form CO and NCO products. In baker's yeast the majority of COs are formed through an interference-dependent CO pathway (class I COs) in which the vast majority of dHJs are resolved to form COs in steps requiring the ZMM proteins Zip1-4, Mer3, and Msh4-Msh5 as well as the Sgs1-Top3-Rmi1 (STR) helicase/topoisomerase complex, Mlh1-Mlh3, and Exo1 [10, 21-28]. These steps are biased to resolve the two junctions present in the dHJ so that the resulting product is exclusively a CO. Interestingly, Exo1's role in maintaining wild-type levels of crossing over is independent of its catalytic activity, suggesting a structural role for this pro- CO factor [28]. Consistent with the above observations, Msh4-Msh5, STR, Exo1 and Zip3 have all been shown to interact with one another and/or with Mlh1-Mlh3 [29].

Genetic and physical studies summarized below support a major role for Mlh1-Mlh3 in promoting meiotic CO formation in the interference-dependent CO pathway. 1. Genetic studies performed in yeast showed that *mlh1* and *mlh3* mutants display approximately two-fold reductions in crossing over [7, 30, 31]. 2. There are several factors that can resolve dHJs into COs in yeast using different pathways. This involves the endonucleases Mlh1-Mlh3, Mus81-Mms4, Yen1, and Slx1-Slx4 [7, 8, 10], with Yen1 and Slx1-Slx4 acting in cryptic or backup roles. When all four factors were removed, crossing over was reduced to very low levels; however, in an *mms4 slx4 yen1* triple mutant, in which Mlh1-Mlh3 is maintained, relatively high

CO levels (~70% of wild-type levels) were observed, suggesting that Mlh1-Mlh3 is the primary factor required for CO resolution in the interference-dependent CO pathway [10]. 3. MLH1 and MLH3 play critical roles in mammalian meiosis [32, 33]. For example, *mlh3*$^{-/-}$ mice are sterile with an 85–94% reduction in the number of COs; germ cells in these mice fail to maintain homologous pairing at metaphase and undergo apoptosis [32, 34].

Much remains to be understood on how biased resolution of dHJs in the interference-dependent pathway is achieved. A working model, supported by genetic and molecular studies, is that the STR complex and a subset of ZMM proteins process and interact with DSB repair and SEI intermediates to create a dHJ substrate that can be resolved by the Mlh1-Mlh3 endonuclease and Exo1 to form primarily COs [7-10, 17, 25, 27, 28, 31, 35-41]. In this model, the biased cleavage of a dHJ suggests coordination between the two junctions that would likely require asymmetric loading of meiotic protein complexes at each junction. However, little is known at the mechanistic level about how such coordination could be accomplished. The Fung group proposed that Msh4-Msh5 is required at the invading end of the DSB to stabilize recombination intermediates such as SEIs, while Zip3 acts to promote second-end capture steps at the ligating end of the DSB [42]. In support of this model, the ZMM heterodimer Msh4-Msh5 has been shown to promote COs in the same pathway as Mlh1-Mlh3, and human MSH4-MSH5 was shown to bind to SEI and Holliday junction substrates *in vitro* [7, 40]. Furthermore, cytological observations in mouse have shown that MSH4-MSH5 foci appear prior to MLH1-MLH3 [37, 43-45]. Consistent with these observations, MLH1 and MLH3 foci formation requires MSH4-MSH5 [44].

In this study, we created a structure-function map of Mlh3 by analyzing 60 new *mlh3* alleles in *S. cerevisiae*. Five alleles predicted to disrupt the Mlh1-Mlh3 endonuclease motif

conferred defects in both MMR and crossing over, providing further support that endonuclease activity is required for both functions. Importantly, we identified five *mlh3* mutations that specifically disrupted MMR, and one mutation that specifically disrupted crossing over. By performing biochemical and genetic analyses of the separation of function Mlh1-mlh3 complexes we suggest that the defects seen in our mutants can be explained by a weakening of protein-protein interactions, which can be tolerated in meiosis, but not MMR.

Materials and Methods

Media

S. cerevisiae SK1, S288c, and YJM789 strains were grown on either yeast extract-peptone- dextrose (YPD) or minimal complete media at 30℃ [46]. For selection purposes, minimal dropout media lacking uracil was used when needed. Geneticin (Invitrogen, San Diego), Nourseothricin (Werner BioAgents, Germany) and Hygromycin (HiMedia) were added to media when required at recommended concentrations [47, 48]. Cells were sporulated as described by Argueso et al. [7].

Site-directed mutagenesis of *MLH3*

60 *mlh3* alleles were constructed, resulting in the mutagenesis of 139 amino acids in the 715 amino acid Mlh3 polypeptide. The single-step integration vector (pEAI254), containing the SK1 *MLH3* gene with a *KANMX4* selectable marker inserted 40 bp downstream of the stop codon [8], was used as a template to create plasmids bearing the *mlh3* mutant alleles via QuickChange site directed mutagenesis (Stratagene, La Jolla, CA). *mlh3-60*, in which the last 11 residues of Pms1 (DWSSFSKDYEI) were inserted before the *MLH3* stop codon, was also made

by QuickChange. Mutations were confirmed by sequencing the entire open reading frame (Sanger method), as well as 70 bp upstream and 150 bp downstream. Primer sequences used to make and sequence these variants are available upon request.

Construction of strains to measure meiotic crossing over and MMR

The SK1 strain EAY3255 was constructed to allow for the simultaneous analysis of *mlh3* MMR and meiotic crossing over phenotypes (Figure 3.1 and 3.2). It carries a spore autonomous fluorescent protein marker (RFP) on chromosome VIII to monitor chromosome behavior (crossing over and nondisjunction; [49]) as well as the *lys2*::*InsE-A14* cassette to measure reversion to Lys^+ [50]. pEAI254 and mutant derivatives were digested with *Bam*HI and *Sal*I and introduced into EAY3255 by gene replacement using the lithium acetate transformation method as described in Gietz *et al.* [51]. At least two independent transformants for each genotype (verified by sequencing) were made resulting in a total of 120 haploid strains bearing the *mlh3* variants described in this study. These haploid strains were used to measure the effect of *mlh3* mutations on reversion rate and were mated to EAY3486, an *mlh3*Δ strain containing the CFP marker, resulting in diploid strains suitable for analysis of crossing over. Diploids were selected on media lacking the appropriate nutrients and maintained as stable strains. Meiosis was induced upon growing the diploid strains on sporulation media as described in Argueso et al. [7]. Wild-type strains carrying the fluorescent protein markers used to make the above test strains were a gift from the Keeney lab.

Lys^+ reversion assays

The haploid strains described above were analyzed for reversion to Lys^+ as described in Tran *et al.* (Figure 3.2, [50]). At least 10 independent cultures were analyzed for each mutant allele alongside wild-type or *mlh3*Δ controls. Analyses were performed for two independent transformants per allele. Reversion rates were measured as described [52, 53], and each median rate was normalized to the wild-type median rate (1X) to calculate fold increase. Alleles were classified into a wild-type, intermediate, or null phenotype based on the 95% confidence intervals which were determined as described [54].

Spore autonomous fluorescent protein expression to measure percent tetratype

Diploids in the EAY3255/EAY3486 background described above were sporulated on media described in Argueso et al. [7]. Spores were treated with 0.5% NP40 and briefly sonicated before analysis using the Zeiss AxioImager.M2 [49]. At least 250 tetrads for each *mlh3* allele were counted to determine the % tetratype. Two independent transformants were measured per allele. A statistically significant difference ($p<0.01$) from wild-type and *mlh3*Δ controls based on χ^2 analysis was used to classify each allele as exhibiting a wild-type, intermediate, or null phenotype.

Purification of Mlh1-Mlh3 and mutant complexes from baculovirus- infected Sf9 cells

For the results shown in Figure 3.3 and Figure 3.4, Mlh1-Mlh3 was purified using the following protocol: Mlh1-Mlh3 and Mlh1-mlh3 mutant derivatives were purified from Sf9 cells infected with Bac- to-Bac baculovirus expression system using pFastBacDual constructs [17]. Mutant Mlh1-mlh3 complexes were purified using the same protocol developed to purify wild-type Mlh1-Mlh3. This involved the use of successive nickel-nitroloacetic acid-agarose [55] and

heparin sepharose (GE Healthcare) column purifications. Mlh1-Mlh3 and mutant derivative yields were ~150 μg per 5 x 10^8 cells; aliquots from the final heparin purification were frozen in liquid N_2 and stored at -80℃. Protein concentrations were determined by densitometry on SDS-PAGE using BSA standard. The *mlh3-6*, *mlh3-32* and *mlh3-45* mutations were introduced into pEAE358 (pPH-His_{10}-MLH3-HA pFastBacDual construct; [17]) by Quick Change (Stratagene). *His_{10}-mlh3-HA* fragments were individually subcloned by restriction digestion into pEAE348 to form pFastBacDual constructs pEAE382 (Mlh1-mlh3-6), pEAE383 (Mlh1-mlh3-32) and pEAE384 (Mlh1-mlh3-45), in which the *MLH1-FLAG* gene is downstream of the p10 promoter and the *His_{10} -mlh3-HA* gene is downstream of the pPH promoter. The sequence of the restriction fragments inserted into pEAE348 were confirmed by DNA sequencing (Cornell Biotechnology Resource Center). Msh2-Msh3 was purified as described previously [56].

Endonuclease assay on supercoiled plasmid DNA and ATPase assay

Mlh1-Mlh3 nicking activity was assayed on supercoiled pBR322 (Thermo Scientific) (Figure 3.3 and 3.4). DNA (2.2 nM) was incubated in 20 μl reactions containing indicated amounts of Mlh1-Mlh3 and Msh2-Msh3 [56] in 20 mM HEPES-KOH pH 7.5, 20 mM KCl, 0.2 mg/ml BSA, 1% glycerol, and 1 mM $MgCl_2$ for 1 h at 37℃. Reactions were quenched by incubation for 20 min at 37℃ with 0.1% SDS, 14 mM EDTA, and 0.1 mg/ml proteinase K (New England Biolabs) (final concentrations). Samples were resolved by 1% agarose gel with 0.1 μg/ml ethidium bromide run in 1X TAE buffer for 50 min at 95 V. All quantifications were performed using GelEval (FrogDance Software, v1.37). The amount of nicked product was quantified as a fraction of the total starting substrate in independent experiments. *bkg* indicates

that amount of nicked product was not above background levels established by negative controls. ATPase assays were performed as described [17].

Cloning mouse *Mlh1* and *Mlh3*

cDNA was synthesized from total testis RNA from wildtype C57B/6J adult males using the SuperScript III Reverse Transcriptase Kit from ThermoFisher. *Mlh1* and *Mlh3* open reading frames were PCR amplified from cDNA using Expand High Fidelity DNA polymerase using primer pairs AO3365 (5'GCTAGCAGCTGATGCATATGGCGTTTGTAGCAGGAG) and AO3366 (5'TACCGCATGCTATGCATTAACACCGCTCAAAGACTTTG) for *Mlh1*, and AO3367 (5'ACGTCGACGAGCTCATATGCATCACCATCACCATCACCATCACCATCACATCAGG TGTCTATCAGATGAC) and AO3368 (5'CGAAAGCGGCCGCGATCATGGAGGCTCACAAGG) for *His$_{10}$-Mlh3*. Each fragment was cloned into the *Spe*1 site of pFastBac1 (ThermoFisher) using Gibson assembly PCR (NEB) to create pEAE393 (*Mlh1*) and pEAE397 (*His$_{10}$-Mlh3*). Constructs were verified by DNA sequencing with NCBI reference sequences NM_026810.2 and NM_175337.2 for *Mlh1* and *Mlh3,* respectively. These constructs were then modified as follows:

1. The MBP –TEV sequence was inserted at the N-terminus of *Mlh1* in pEAE393 to create pEAE395.
2. The *mlh3-D1185N* mutation was introduced into pEAE397 by Quick Change [57] to create pEAE413.

Chromatography analysis of the mouse MLH1-MLH3-D1185N heterodimer from Baculovirus-infected Sf9 cells

Sf9 cells were transfected with pEAE397 (*His$_{10}$-Mlh3*), pEAE413 (*His$_{10}$-Mlh3-D1185N*) and pEAE395 (*MBP-Mlh1*) using the Bac-to-Bac baculovirus infection system. Fresh *Sf9* cells were co-infected with both viruses (containing *Mlh1* and *Mlh3* or *Mlh3-D1185N*). Cells were harvested 60 hours post infection, washed with phosphate buffered saline, and kept at -80°C until use.

Cell pellets from 250 ml of cells were thawed, resuspended in 60 ml hypotonic lysis buffer (20 mM HEPES-KOH pH 7.5, 5 mM NaCl, 1 mM $MgCl_2$, 1 mM PMSF and EDTA free protease inhibitor mixture from Roche and Thermo Scientific) and incubated for 15 min on ice. The suspension was adjusted to 250 mM NaCl, 15 mM imidazole, 10% glycerol, 2 mM ß-mercaptoethanol (BME), and clarified by centrifugation at 17,000 g for 20 min at 4°C. The supernatant was mixed with 6 ml of 50% nickel-nitrolotriaceticacid-agarose (Ni-NTA) resin and allowed to bind for 2 hours or overnight followed by centrifugation to remove the unbound fraction. The resin was packed onto a column and washed with 7-10 column volumes of wash buffer (50 mM HEPES-KOH pH 7.5, 250 mM NaCl, 40 mM imidazole, 10% glycerol, 2 mM BME, 1 mM PMSF). Protein was eluted with 15 ml of 300 mM imidazole in 50 mM HEPES-KOH pH 7.5, 250 mM NaCl, 40 mM imidazole, 10% glycerol, 2 mM BME and 1 mM PMSF. Elution fractions containing MLH1-MLH3, determined by SDS-PAGE, were pooled and loaded onto 1 ml 100% amylose resin (NEB). The resin was washed with 10 column volumes of wash buffer (50 mM HEPES-KOH pH 7.5, 250 mM NaCl, 10% glycerol, 2 mM BME, 1 mM PMSF) and eluted with 6 ml wash buffer containing 10 mM maltose. Fractions containing MLH1-MLH3 were pooled and aliquots were flash frozen and stored in -80°C. The protein yield, following amylose chromatography, was similar for wild-type and mutant complexes (Figure 3.5, approximately 120-150 µg per 250 ml cells).

It is important to note that we were unable to detect a specific endonuclease activity for the mouse MBP-MLH1-MLH3 complex, suggesting that the MBP tag interferes with MLH1-MLH3 functions. We were unable to test this directly because, despite numerous attempts, we were unable to efficiently remove the MBP tag from MLH1 by treating MBP-MLH1-MLH3 with TEV protease.

Mass-spectrometry of MLH1 and MLH3 bands from SDS-PAGE

SDS-PAGE bands following amylose chromatography predicted to contain MBP-MLH1 and His_{10}-MLH3 were excised and analyzed by the Cornell University Proteomics facility using a Thermo LTQ Orbitrap Velos Mass Spectrometer (Figure 3.5C).

Results

Rationale for site-directed mutagenesis of *MLH3*

Mlh3 contains a highly conserved N-terminal ATP binding motif, a dynamic and unstructured motif known as the linker arm, and an endonuclease active site that overlaps with a C- terminal Mlh1 interaction domain [58]. We performed a clustered charged-to-alanine scanning mutagenesis of the *S. cerevisiae MLH3* gene to create 60 *mlh3* variants (Figure 3.1). Charged residues were considered "clustered" if there were at least two charged residues, consecutive or separated by at most one amino acid, within the primary sequence of Mlh3. Such a directed approach, in the absence of a complete crystal structure, is aimed at targeting the surface of a protein where clusters of charged residues likely reside, while minimizing changes within the interior. In this model, replacement of a charged patch from Mlh3's surface with alanine residues would disrupt protein-protein or protein-DNA interactions without affecting

Mlh3 structure. As shown below, we identified a subset of mutations that caused strong defects in either MMR or crossing over, but not both, likely through disrupted interactions with Mlh1 and other MMR and meiotic CO factors.

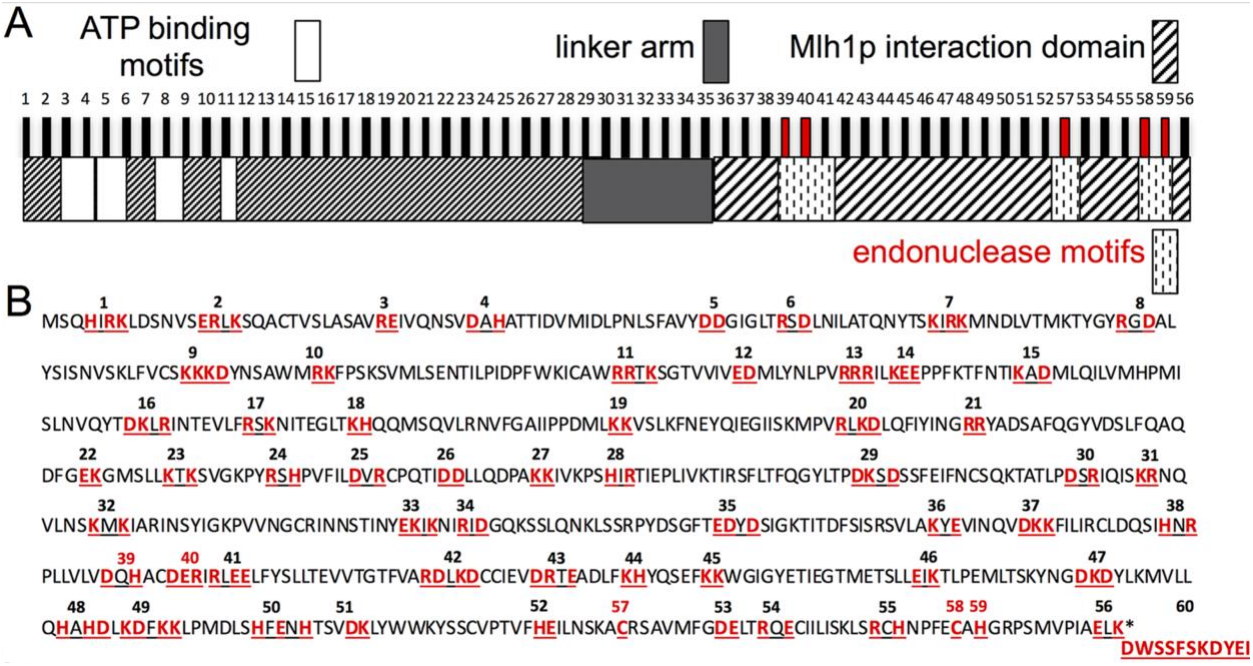


Figure 3.1 Site directed mutagenesis of *MLH3.* A. Functional organization of Mlh3 based on sequence homology and secondary structure prediction [58]. The vertical bars indicate the approximate position of the *mlh3* mutations (except *mlh3-60*) analyzed in this study and described in panel B. *mlh3-39*, *-40*, *-57*, *-58*, and *-59* colored in red are based on highly conserved residues in the endonuclease motifs of Pms1 which were shown in the crystal structure of Mlh1-Pms1 to form a single metal binding site [58]. B. Amino acid positions of charged-to-alanine substitutions presented in red on the primary sequence of *Saccharomyces cerevisiae* Mlh3. Each cluster of underlined residues represents one allele corresponding to the vertical bars in panel A. *mlh3-39*, *-40*, *-57*, *-58*, and *-59* are colored in red as in panel A. *mlh3-60* represents the last 11 residues of Pms1 which constitute patch II of the heterodimerization interface of Mlh1-Pms1 [58].

Comparison of the MMR and CO assay results for each individual allele led to the identification of six separation of function mutations, defined as showing strong defects in one function (e.g. MMR) relative to another (e.g. CO), in the Mlh3 ATP-binding motifs, N-terminal domain beyond the ATP-binding motifs, linker arm, and the interaction domain (Figure 3.2, indicated by stars). One of these alleles (*mlh3-32*) conferred a nearly wild-type phenotype for MMR and a null phenotype for crossing over on chromosome VIII (hereafter referred to as MMR^{+}, CO^{-}). The remaining five mutations (*mlh3-6*, *mlh3-23*, *mlh3-42*, *mlh3-45*, and *mlh3-54*) conferred null MMR phenotypes and nearly wild-type levels of crossing over (hereafter referred to as MMR^{-}, CO^{+}).

The phenotypes observed in the separation of function mutants may result from a defect in DNA binding/substrate specificity, endonuclease activity, interactions with specific MMR and meiotic CO factors, or changes in protein conformation. It is important to note that a co-crystal structure of the N-terminal domain of *E. coli* MutL (LN40) and *E. coli* MutS was recently solved. This work showed that conformational changes license MutS-MutL interaction and are essential for MMR [59, 60].

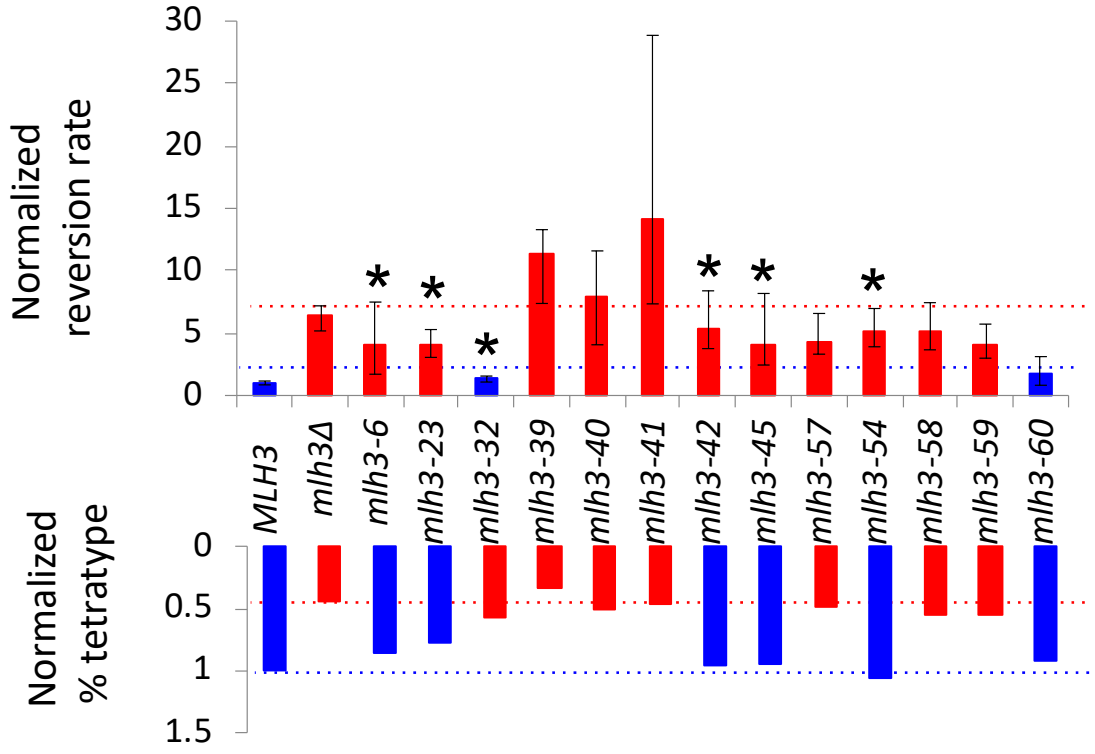


Figure 3.2 Separation of function mutants of *mlh3*. MMR (top) and CO (bottom) phenotypes for *MLH3* and *mlh3* null (*mlh3*Δ), separation of function, endonuclease, and C-terminal tail (*mlh3-60*) mutants. Mismatch repair was measured using the *lys2-A14* reversion assay [50][and crossing over was measured using the assay depicted in panel A. Bars represent the median reversion rates (error bars based on 95% confidence intervals) and percent tetratype normalized to *MLH3* (1X). For mismatch repair (top), bars represent reversion rates of at least 10 independently tested cultures from two independently constructed strains presented here normalized to *MLH3* median rate of 1X = 1.43×10^{-6} (n = 140). For crossing over (bottom), bars represent percent tetratype of at least 250 tetrads from two independently constructed strains presented here normalized to *MLH3* percent tetratype 1X = 36.7% (n = 226). Blue and red dotted lines represent *MLH3* and *mlh3*Δ respectively. * indicate separation of function mutants.

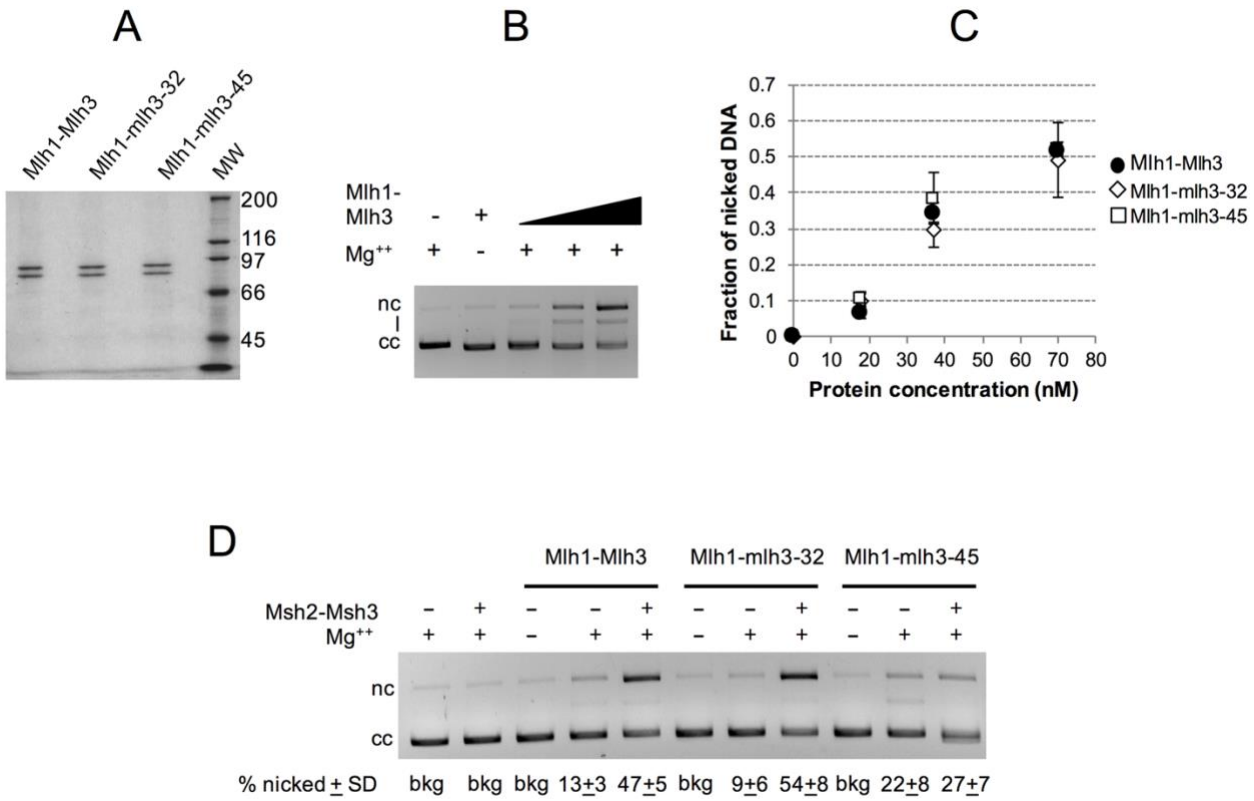


Fig 3.3 Mlh1-mlh3-32 and Mlh1-mlh3-45 display wild-type endonuclease activities that are differentially stimulated by Msh2-Msh3. A. SDS-PAGE analysis of purified Mlh1-Mlh3, Mlh1-mlh3-32 and Mlh1-mlh3-45. Coomassie Blue R250-stained 8% Tris-glycine gel. 0.5 μg of each protein is shown. MW = Molecular Weight Standards from top to bottom- 200, 116, 97, 66, 45 kD). B, C. Mlh1-Mlh3, Mlh1-mlh3-32 and Mlh1-mlh3-45 (18, 37, 70 nM) were incubated with 2.2 nM supercoiled pBR322 DNA, and analyzed in agarose gel electrophoresis (C) and the endonuclease activity was quantified (average of 6 independent experiments presented +/-SD) as described in the Methods. Ladder: 1 kb DNA ladder (New England BioLabs). Migration of closed circular (cc), nicked (nc) and linear (l) pBR322 DNA is indicated. D. Endonuclease assays were performed as in B., but contained 20 nM of the indicated wild-type or mutant Mlh1-Mlh3 complex and 40 nM Msh2-Msh3 when indicated. Reactions were performed in triplicate, samples were resolved on agarose gels, and the fraction of nicked DNA was quantified, averaged, and the standard deviation between experiments was calculated. The average fraction of supercoiled substrate cleaved is presented +/-S.D. below the gel. (bkg) background, (cc) closed circular DNA, (nc) nicked DNA.

Mlh1-mlh3-32 and Mlh1-mlh3-45 display wild-type endonuclease activities but only Mlh1-mlh3-32 endonuclease is stimulated by Msh2-Msh3

I examined Mlh1-mlh3 mutant complexes for endonuclease activity [16, 17], focusing on opposite separation of function mutants Mlh1-mlh3-32 (MMR^+, CO^-), Mlh1-mlh3-6, and Mlh1-mlh3-45 (MMR^-, CO^+). Mlh1-mlh3-45, located in the C-terminal Mlh1 interaction domain, was chosen because it is the only separation of function mutant in that domain that displayed wild-type Mlh1-Mlh3 interactions as measured in the two-hybrid assay. As shown in Figure 3.3 and Figure 3.4, all three mutant complexes purified as heterodimers and display endonuclease activities similar to wild-type. When this work was initiated I thought that separation of function mutant complexes might show endonuclease defects indicating that this activity is more critical for MMR or crossing over, or show no endonuclease defects because mutant complexes were defective in interacting with MMR or CO specific factors. My finding that all three mutants have enzymatic activity comparable to wild-type is consistent with the interaction defect model (see below). The *mlh3-6* mutation maps close to conserved sites in the ATP binding motif. I then tested whether the mutant complex displayed a defect in ATPase activity. As shown in Figure 3.4, Mlh1-Mlh3 and Mlh1-mlh3-6 displayed similar ATPase activities.

Because Mlh1-Mlh3's endonuclease activity is enhanced by Msh2-Msh3 [17], I tested whether the opposite separation of function phenotypes of Mlh1-mlh3-32 and Mlh1-mlh3-45 could be explained by defective interactions with MSH complexes. As shown in Figure 3.3, Mlh1-mlh3-32 endonuclease activity but not Mlh1-mlh3-45 could be stimulated by Msh2-Msh3. These data are consistent with the MMR^-, CO^+ phenotype exhibited by *mlh3-45* mutants resulting from a defect in interacting with the MMR component Msh2-Msh3, and the *mlh3-32* mutant likely being defective in interactions with meiosis-specific factors.

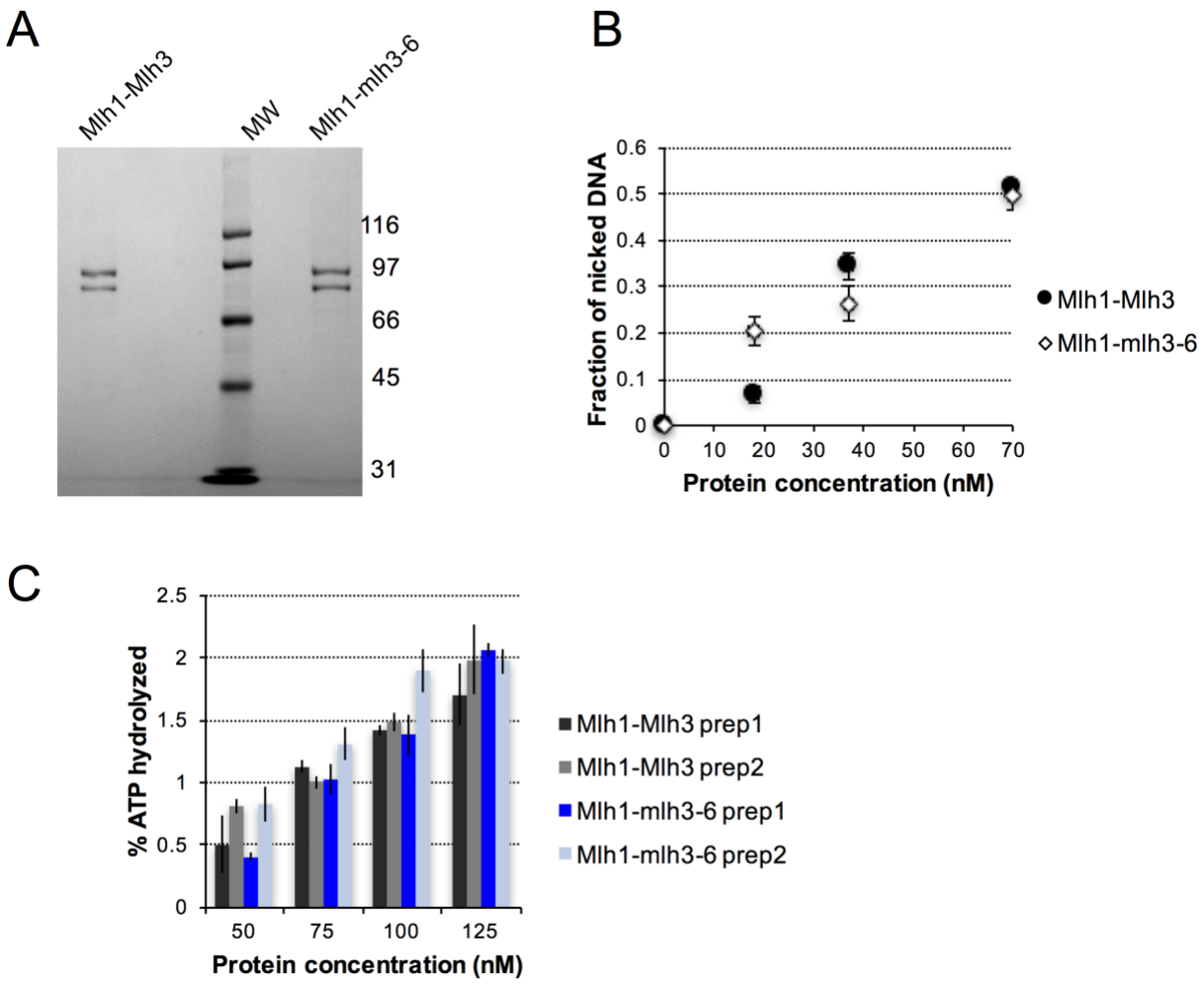
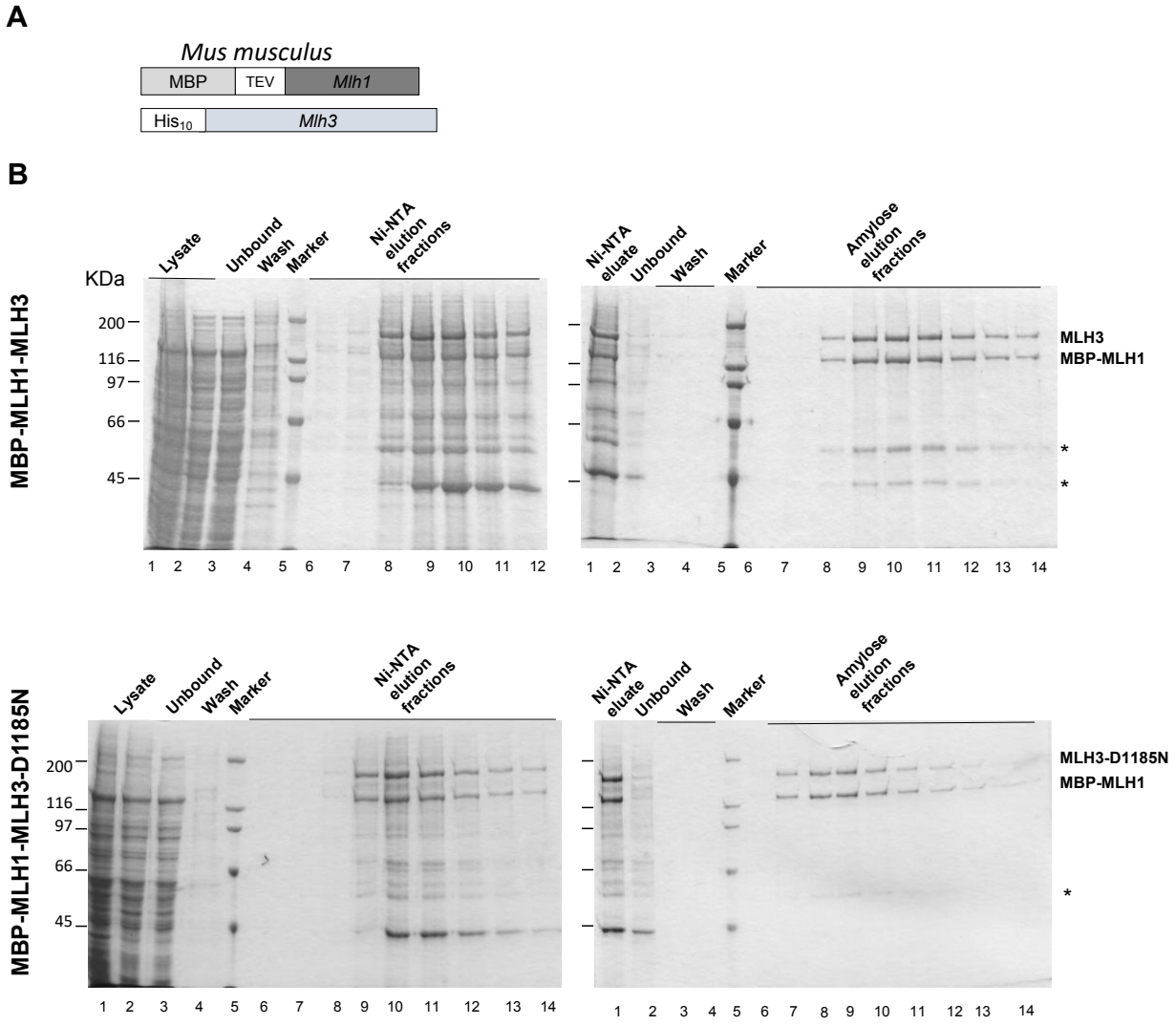


Figure 3.4: Mlh1-mlh3-6 exhibits wild-type endonuclease and ATPase activity.
A. SDS-PAGE analysis of purified Mlh1-Mlh3 and Mlh1-mlh3-6. Coomassie Blue R250-stained 8% Tris-glycine gel. 0.5 µg of each protein is shown. MW = Molecular Weight Standards from top to bottom-116, 97, 66, 45, 31 kD). B, C. Mlh1-Mlh3 and Mlh1-mlh3-6 (18, 37, 70 nM) were incubated with 2.2 nM supercoiled pBR322 DNA, analyzed in agarose gel electrophoresis, and endonuclease activity was quantified (average of 6 independent experiments presented +/-SD) as described in the Methods. C. ATPase assays were performed as described in Rogacheva et al. [17], but contained the indicated amounts of Mlh1-Mlh3 and Mlh1-mlh3-6 incubated with 100 µM ^{32}P-γ-ATP. Reactions were performed in duplicate for two separate purifications of each, and the average values, +/-SD, are presented

Cloning and analysis of the mouse MLH1-MLH3 complex

To determine if the yeast and mouse Mlh1-Mlh3 complexes displayed similar activities, I developed methods to overexpress and purify mouse MLH1-MLH3. Mouse *Mlh1* and *Mlh3* were amplified from cDNA and cloned into pFastBac1 vectors as described in the Methods. The MLH1-MLH3 and MLH1-MLH3-D1185N complexes were expressed from Sf9 cells infected with baculoviruses containing *MBP-Mlh1* and *His_{10}-Mlh3* or *His_{10}-Mlh3-D1185N* constructs (Figure 3.5A). Extracts from these cells were applied to a Ni-NTA column. Fractions containing induced proteins were pooled and then applied to an amylose column. Two major bands of molecular weights predicted for an MBP-MLH1-His_{10}-MLH3 complex were detected on SDS-PAGE after amylose chromatography (Figure 3.5B). These bands were further analyzed by mass spectrometry, and the results from this analysis confirmed their identity (Figure 3.5C). Importantly, MLH1-MLH3 and MLH1-MLH3-D1185N eluted with an apparent 1:1 stoichiometry in both chromatography steps, indicating that the heterodimers were stable, and the protein yields of the two complexes after amylose chromatography were similar (Figure 3.5B).



C

Band	Protein name	Protein MW (KDa)	Peptide number	Protein score	% Protein Coverage	Molar %
Upper	His_{10}-MLH3	160	90	2077	64	95
Lower	MBP-MLH1	127.5	90	2413	79	94

Figure 3.5: Mouse MLH3-D1185N forms a stable complex with MLH1. A) Schematic of mouse *Mlh1* and *Mlh3* constructs (see Methods for details). (B) Representative purification of MBP-MLH1-MLH3 (top) and MBP-MLH1-mlh3-D1185N (bottom) using Ni-NTA and amylose resin chromatography as described in the Methods. Fractions were analyzed using SDS- PAGE, stained by Coomassie brilliant blue. MLH1-MLH3 and MLH1-MLH3-D1185N were eluted from amylose in the same fractions. The mass of molecular weight standards is indicated on the left and the expected positions of MBP-MLH1 (127.5 KDa) and His10-MLH3 (165 KDa) is indicated in the center. *Likely to be degradation products of MLH1-MLH3. (C) Mass spectrometry analysis of the two major bands in SDS-PAGE detected after amylose chromatography.

Discussion

We performed a structure-function analysis of Mlh3, a factor that acts in both MMR and meiotic crossing over. This work was pursued because little is known about how Mlh1-Mlh3 acts as a meiotic endonuclease. This is due in part to Mlh1-Mlh3 sharing little in common with the well-characterized structure-selective endonucleases (SSNs are structure selective endonucleases that recognize and cleave dHJs or other such structures specifically; Mus81-Mms4, Slx1-Slx4, and Yen1) in terms of homology and intrinsic behavior *in vitro* (reviewed in [29]). Obtaining new mechanistic insights has been complicated by the fact that Mlh1-Mlh3 can bind to model HJ substrates, but cannot cleave them, and by genetic studies suggesting that Mlh1-Mlh3 acts in concert with other pro-CO factors [16, 17, 29, 61]. Recent work has suggested that multiple Mlh1-Mlh3 heterodimers are required to activate the endonuclease and that the complex is at least partially inhibited by incorporation of a DNA secondary structure that is not part of a continuous homoduplex substrate [61]. These data suggest that other protein factors likely recruit and position Mlh1-Mlh3 complexes during meiotic recombination. The identities of these factors are for the most part known, though it is not understood how they directly contribute to Mlh1-Mlh3's ability to nick DNA in the directed manner required to generate COs.

Our analysis of *mlh3-32* (MMR^+, CO^-) and *mlh3-45* (MMR^-, CO^+), support the above hypothesis that protein-protein interactions are critical for directing Mlh1-Mlh3 endonuclease activity (Figure 3.3). Mlh1-Mlh3 has been shown genetically to act downstream of Msh4-Msh5 [37-39, 43]; this order of events is analogous to steps in DNA MMR where MLH acts following MSH recognition [62, 63]. As outlined in the introduction, Msh4-Msh5, STR, Exo1 (independent of its enzymatic activity) and Zip3 have been classified as pro-CO factors, and have all been shown to interact with one another and/or with Mlh1-Mlh3 (reviewed in [29]). Our biochemical

studies are consistent with Mlh1-mlh3-45 having interaction defects that prevent its endonuclease activity from being stimulated by Msh2-Msh3 in MMR. We hypothesize that MLH complexes interact with MSH complexes via a common mechanism, and that defective interactions with Msh2-Msh3 are also indicative of defective interactions with Msh4-Msh5, but during meiotic CO resolution, additional factors such as Exo1, Sgs1-Top3-Rmi1 and Zip3 act in concert to strengthen a possibly weakened Msh4-Msh5 and Mlh1-mlh3-45 interaction. This model also helps explain why we identified several MMR^- CO^+ *mlh3* mutants (*mlh3-42* and *-54*) in which the mutant mlh3 protein fails to interact with Mlh1. For the Mlh1-mlh3-32 complex, the MSH interaction and enhancement is retained, but interaction with other critical meiotic factors is likely lost, possibly resulting in an unstable complex that cannot resolve dHJs.

Appendix A

A new protocol for purifying the yeast Mlh1-Mlh3 complex.

Introduction and Results

A year after the work published in Al-Sweel *et al* [1], we noticed in newer preparations that there was a contaminating endonuclease activity copurifying with Mlh1-mlh3-DN using our standard protocol. Carol Manhart in our lab has shown multiple times using this protocol that endonuclease dead protein could be purified without any nuclease contamination, but something had changed in the reagents or cells of our newer preparations that we were unable to track down. Because of this issue, I developed a new protocol written in detail below to purify yeast Mlh1-Mlh3 and endonuclease dead mutant and demonstrated that the wild type was active and the endonuclease dead version was inactive, indicating that the purification was free of contaminating endonucleases (Figure 3.6). The new protocol for purification of the Mlh1-Mlh3 complex involved affinity chromatography with anti-FLAG resin, followed by affinity chromatography with Ni-NTA and dialysis. Mlh1-Mlh3 and endonuclease active site mutant Mlh1-mlh3-DN were purified as heterodimers with 1:1 stoichiometry of Mlh1 and Mlh3 or mlh3-DN (Figure 3.6 A). The protein yields were greatly increased using this method (1mg versus 150µg per 5 x 10^8 cells). The wild type was endonuclease active with the highest activity seen with 5mM $MnSO_4$ (Figure 3.6 B, C). Lower amounts of endonuclease activity were observed with 1mM $MnSO_4$, 5mM $MgCl_2$ and 1mM $MgCl_2$ in this order (Figure 3.6 B). The endonuclease active site mutant Mlh1-mlh3-DN only had a low endonuclease activity at very high protein concentration of 300nM and did not show any endonuclease activity above

background otherwise (Figure 3.6 B, C). Moreover, similar to the older established preparation, using the new protocol for purification, I observed similar preference for larger sized plasmid substrate as seen before.

Materials and Methods

Purification of yeast Mlh1-Mlh3

Mlh1-Mlh3 and Mlh1-mlh3-DN were expressed in Sf9 cells as described above. The pellets, each containing 300ml cells were frozen as pellets in -80ºC. The purification was divided into four parts:

A. Cell lysis
B. Affinity purification using the FLAG resin
C. Affinity purification using the NiNTA resin
D. Dialysis to remove Imidazole.

1. The cell pellet containing 300ml cells was thawed on ice in the cold room for 1 hour. It was resuspended in 15 mL hypotonic lysis buffer (20mM HEPES-KOH, pH7.5, 5mM KCl, 1mM $MgCl_2$, 0.5% Tween-20, 1mM PMSF, 1X Halt protease inhibitor cocktail EDTA-free (Thermo), 1 tablet complete protease inhibitor tablet EDTA-free [64]), in the cold room. The lysate was incubated on ice for 20 min.
2. The suspension was adjusted to have final concentrations of the components as follows: 100mM KCl, 400mM NaCl, 1mM 2-mercaptoethanol, 10% glycerol. The lysate was centrifuged at 17,000xg for 20 min at 4ºC.

3. The supernatant was applied over 1.5 ml of 50% M2 anti-FLAG agarose beads (Sigma-Aldrich), pre-equilibrated in FLAG equilibration buffer (25mM HEPES-KOH, pH 7.5, 100mM KCl, 400mM NaCl, 1mM EDTA, 10% Glycerol, 0.2% Tween-20, 1.4mM 2-mercaptoethanol, 1mM PMSF) and incubated with rotation for 1 hour.
4. The column was washed with 25 column volumes of FLAG wash buffer (25mM HEPES-KOH, pH 7.5, 100mM KCl, 400mM NaCl, 10% glycerol, 02% Tween-20, 1.4mM 2-mercaptoethanol, 1mM PMSF) and 0.5ml fractions were eluted using FLAG elution buffer (6ml FLAG wash buffer + 1500µg 3x-FLAG peptide (SIGMA)).
5. Fractions containing Mlh1-Mlh3, determined by SDS-PAGE were pooled, imidazole was added to a concentration of 15mM. The protein was loaded over 3ml of 50% Ni-NTA agarose beads [55], pre-equilibrated in Ni-NTA binding buffer (25mM HEPES-KOH, pH 7.5, 100mM KCl, 400mM NaCl, 15mM Imidazole, 10% Glycerol, 0.5% Tween-20, 1.4mM 2-mercaptoethanol, 1mM PMSF) and incubated with rotation overnight at 4ºC.
6. The column was washed with 10 column volumes of Ni-NTA wash buffer (25mM HEPES-KOH, pH 7.5, 100mM KCl, 300mM NaCl, 40mM Imidazole, 10% Glycerol, 0.2% Tween-20, 1.4mM 2-mercaptoethanol, 1mM PMSF) and eluted with 8 ml Ni-NTA elution buffer (25mM HEPES-KOH pH 7.5, 100mM KCl, 250mM NaCl, 300mM Imidazole, 10% Glycerol, 02% Tween-20, 1.4mM 2-mercaptoethanol, 1mM PMSF) and ~0.62 mL fractions were collected.
7. Fractions containing Mlh1-Mlh3 were pooled and dialyzed for 4 hours using a 20 KDa cut-off Slide-A-Lyzer (Thermo-Fisher Scientific) in 1L dialysis buffer. The dialysis buffer was changed after the first 2 hours. Aliquots were frozen in liquid-N_2, followed by storage at -80ºC.

Protein concentrations were estimated by densitometry on SDS-PAGE, by running BSA (80-500ng per lane) and 2, 5 and 10µl pure protein in the same gel. Two such gels were run simultaneously to reduce error. The typical yields for both wild-type and endonuclease mutant derivative was ~1.2mg per 6 x 10^8 cells.

Endonuclease assay

For the Mlh1-Mlh3 purified using the new protocol, shown in Figure 3.6 the same endonuclease assay method as in Chapter 3 was used with the following exceptions: 5mM $MnSO_4$ or 5mM $MgCl_2$ were used in place of 1mM $MgCl_2$.

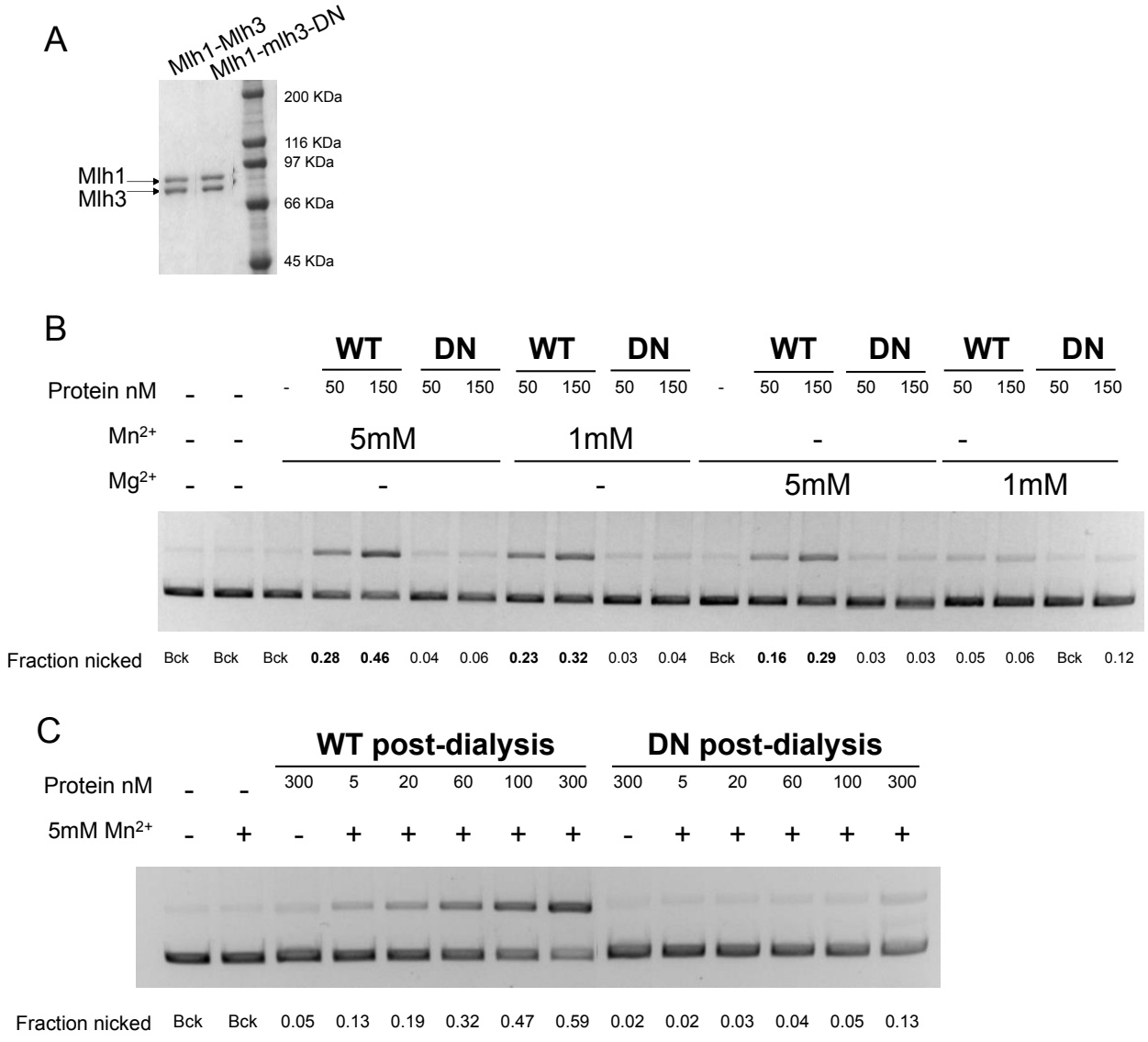


Figure 3.6: Mlh1-Mlh3 purified using a new protocol involving FLAG and Ni-NTA affinity chromatography is an active endonuclease, and Mlh1-mlh3-DN is inactive. A. SDS-PAGE analysis of purified Mlh1-Mlh3 and Mlh1-mlh3-DN. Coomassie Blue R250-stained 8% Tris-glycine gel. 0.5 μg of each protein is shown. MW = Molecular Weight Standards from top to bottom- 200, 116, 97, 66, 45 kD). B Mlh1-Mlh3, Mlh1-mlh3-DN (50 and 150 nM) were incubated with 2.2 nM supercoiled pBR322 DNA, and either 5mM $MnSO_4$/ 5mM $MgCl_2$/ 1mM $MnSO_4$/ 1mM $MgCl_2$ and analyzed in agarose gel electrophoresis. Fraction of nicked DNA is quantified and shown. D. Endonuclease assays were performed as in B., but contained a titration of the wild-type Mlh1-Mlh3 or Mlh1-mlh3-DN complex (5, 20, 60, 100, 300nM) with 5mM $MnSO_4$. Fraction of nicked DNA is quantified and shown. Bck is background nicking, WT is wild-type Mlh1-Mlh3, DN is Mlh1-mlh3-DN.

References

1. Al-Sweel, N., Raghavan, V., Dutta, A., Ajith, V.P., Di Vietro, L. *et al.* (2017) mlh3 mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide. PLoS Genet 13, e1006974 DOI: 10.1371/journal.pgen.1006974.

2. Toledo, M., Sun, X., Brieno-Enriquez, M.A., Raghavan, V., Gray, S. *et al.* (2019) A mutation in the endonuclease domain of mouse MLH3 reveals novel roles for MutLγ during crossover formation in meiotic prophase I. bioRxiv DOI: https://doi.org/10.1101/517748.

3. Kunkel, T.A. and Erie, D.A. (2015) Eukaryotic Mismatch Repair in Relation to DNA Replication. Annu Rev Genet 49, 291-313 DOI: 10.1146/annurev-genet-112414-054722.

4. Flores-Rozas, H. and Kolodner, R.D. (1998) The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. Proc Natl Acad Sci U S A 95, 12404-9.

5. Harfe, B.D., Minesinger, B.K. and Jinks-Robertson, S. (2000) Discrete *In vivo* roles for the MutL homologs Mlh2p and Mlh3p in the removal of frameshift intermediates in budding yeast. Curr Biol 10, 145-8.

6. Romanova, N.V. and Crouse, G.F. (2013) Different roles of eukaryotic MutS and MutL complexes in repair of small insertion and deletion loops in yeast. PLoS Genet 9, e1003920 DOI: 10.1371/journal.pgen.1003920.

7. Argueso, J.L., Wanat, J., Gemici, Z. and Alani, E. (2004) Competing crossover pathways act during meiosis in *Saccharomyces cerevisiae*. Genetics 168, 1805-16 DOI: 10.1534/genetics.104.032912.

8. Nishant, K.T., Plys, A.J. and Alani, E. (2008) A mutation in the putative MLH3 endonuclease domain confers a defect in both mismatch repair and meiosis in *Saccharomyces cerevisiae*. Genetics 179, 747-55 DOI: 10.1534/genetics.108.086645.

9. Sonntag Brown, M., Lim, E., Chen, C., Nishant, K.T. and Alani, E. (2013) Genetic analysis of mlh3 mutations reveals interactions between crossover promoting factors during meiosis in baker's yeast. G3 (Bethesda) 3, 9-22 DOI: 10.1534/g3.112.004622.

10. Zakharyevich, K., Tang, S., Ma, Y. and Hunter, N. (2012) Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. Cell 149, 334-47 DOI: 10.1016/j.cell.2012.03.023.

11. Champion, M.D. and Hawley, R.S. (2002) Playing for half the deck: the molecular biology of meiosis. Nat Cell Biol 4 Suppl, s50-6 DOI: 10.1038/ncb-nm-fertilityS50.

12. Hunter, N. (2007) Meiotic recombination. Molecular Genetics of Recombination chapter.

13. Maguire, M.P. (1974) Letter: The need for a chiasma binder. J Theor Biol 48, 485-7.

14. Zickler, D. (2006) From early homologue recognition to synaptonemal complex formation. Chromosoma 115, 158-74 DOI: 10.1007/s00412-006-0048-6.

15. Hassold, T. and Hunt, P. (2001) To err (meiotically) is human: the genesis of human aneuploidy. Nat Rev Genet 2, 280-91 DOI: 10.1038/35066065.

16. Ranjha, L., Anand, R. and Cejka, P. (2014) The *Saccharomyces cerevisiae* Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. J Biol Chem 289, 5674-86 DOI: 10.1074/jbc.M113.533810.

17. Rogacheva, M.V., Manhart, C.M., Chen, C., Guarne, A., Surtees, J. *et al.* (2014) Mlh1-Mlh3, a meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated endonuclease. J Biol Chem 289, 5664-73 DOI: 10.1074/jbc.M113.534644.

18. Chen, S.Y., Tsubouchi, T., Rockmill, B., Sandler, J.S., Richards, D.R. *et al.* (2008) Global analysis of the meiotic crossover landscape. Dev Cell 15, 401-15 DOI: 10.1016/j.devcel.2008.07.006.

19. Gong, X., Tao, R. and Li, Z. (2006) Quantification of RNA damage by reverse transcription polymerase chain reactions. Anal Biochem 357, 58-67 DOI: S0003-2697(06)00449-0 [pii] 10.1016/j.ab.2006.06.025.

20. Keeney, S., Giroux, C.N. and Kleckner, N. (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88, 375-84.

21. Allers, T. and Lichten, M. (2001) Intermediates of yeast meiotic recombination contain heteroduplex DNA. Mol Cell 8, 225-31.

22. Allers, T. and Lichten, M. (2001) Differential timing and control of noncrossover and crossover recombination during meiosis. Cell 106, 47-57.

23. Borner, G.V., Kleckner, N. and Hunter, N. (2004) Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. Cell 117, 29-45.

24. Hunter, N., Borner, G.V., Lichten, M. and Kleckner, N. (2001) Gamma-H2AX illuminates meiosis. Nat Genet 27, 236-8 DOI: 10.1038/85781.

25. Kaur, H., De Muyt, A. and Lichten, M. (2015) Top3-Rmi1 DNA single-strand decatenase is integral to the formation and resolution of meiotic recombination intermediates. Mol Cell 57, 583-94 DOI: 10.1016/j.molcel.2015.01.020.

26. Schwacha, A. and Kleckner, N. (1995) Identification of double Holliday junctions as intermediates in meiotic recombination. Cell 83, 783-91.

27. Tang, S., Wu, M.K., Zhang, R. and Hunter, N. (2015) Pervasive and essential roles of the Top3-Rmi1 decatenase orchestrate recombination and facilitate chromosome segregation in meiosis. Mol Cell 57, 607-21 DOI: 10.1016/j.molcel.2015.01.021.

28. Zakharyevich, K., Ma, Y., Tang, S., Hwang, P.Y., Boiteux, S. *et al.* (2010) Temporally and biochemically distinct activities of Exo1 during meiosis: double-strand break resection and resolution of double Holliday junctions. Mol Cell 40, 1001-15 DOI: 10.1016/j.molcel.2010.11.032.

29. Manhart, C.M. and Alani, E. (2016) Roles for mismatch repair family proteins in promoting meiotic crossing over. DNA Repair (Amst) 38, 84-93 DOI: 10.1016/j.dnarep.2015.11.024.

30. Hunter, N. and Borts, R.H. (1997) Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. Genes Dev 11, 1573-82.

31. Wang, T.F., Kleckner, N. and Hunter, N. (1999) Functional specificity of MutL homologs in yeast: evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. Proc Natl Acad Sci U S A 96, 13914-9.

32. Lipkin, S.M., Moens, P.B., Wang, V., Lenzi, M., Shanmugarajah, D. *et al.* (2002) Meiotic arrest and aneuploidy in MLH3-deficient mice. Nat Genet 31, 385-90 DOI: 10.1038/ng931.

33. Woods, L.M., Hodges, C.A., Baart, E., Baker, S.M., Liskay, M. *et al.* (1999) Chromosomal influence on meiotic spindle assembly: abnormal meiosis I in female Mlh1 mutant mice. J Cell Biol 145, 1395-406.

34. Svetlanov, A., Baudat, F., Cohen, P.E. and de Massy, B. (2008) Distinct functions of MLH3 at recombination hot spots in the mouse. Genetics 178, 1937-45 DOI: 10.1534/genetics.107.084798.

35. Fasching, C.L., Cejka, P., Kowalczykowski, S.C. and Heyer, W.D. (2015) Top3-Rmi1 dissolve Rad51-mediated D loops by a topoisomerase-based mechanism. Mol Cell 57, 595-606 DOI: 10.1016/j.molcel.2015.01.022.

36. Hoffmann, E.R. and Borts, R.H. (2004) Meiotic recombination intermediates and mismatch repair proteins. Cytogenet Genome Res 107, 232-48 DOI: 10.1159/000080601.

37. Kneitz, B., Cohen, P.E., Avdievich, E., Zhu, L., Kane, M.F. *et al.* (2000) MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. Genes Dev 14, 1085-97.

38. Santucci-Darmanin, S., Neyton, S., Lespinasse, F., Saunieres, A., Gaudray, P. *et al.* (2002) The DNA mismatch-repair MLH3 protein interacts with MSH4 in meiotic cells, supporting a role for this MutL homolog in mammalian meiotic recombination. Hum Mol Genet 11, 1697-706.

39. Santucci-Darmanin, S., Walpita, D., Lespinasse, F., Desnuelle, C., Ashley, T. *et al.* (2000) MSH4 acts in conjunction with MLH1 during mammalian meiosis. FASEB J 14, 1539-47.

40. Snowden, T., Acharya, S., Butz, C., Berardini, M. and Fishel, R. (2004) hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. Mol Cell 15, 437-51 DOI: 10.1016/j.molcel.2004.06.040.

41. Wang, T.F. and Kung, W.M. (2002) Supercomplex formation between Mlh1-Mlh3 and Sgs1-Top3 heterocomplexes in meiotic yeast cells. Biochem Biophys Res Commun 296, 949-53.

42. Oke, A., Anderson, C.M., Yam, P. and Fung, J.C. (2014) Controlling meiotic recombinational repair - specifying the roles of ZMMs, Sgs1 and Mus81/Mms4 in crossover formation. PLoS Genet 10, e1004690 DOI: 10.1371/journal.pgen.1004690.

43. Kolas, N.K. and Cohen, P.E. (2004) Novel and diverse functions of the DNA mismatch repair family in mammalian meiosis and recombination. Cytogenet Genome Res 107, 216-31 DOI: 10.1159/000080600.

44. Kolas, N.K., Svetlanov, A., Lenzi, M.L., Macaluso, F.P., Lipkin, S.M. *et al.* (2005) Localization of MMR proteins on meiotic chromosomes in mice indicates distinct functions during prophase I. J Cell Biol 171, 447-58 DOI: 10.1083/jcb.200506170.

45. Lipkin, S.M., Wang, V., Jacoby, R., Banerjee-Basu, S., Baxevanis, A.D. *et al.* (2000) MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. Nat Genet 24, 27-35 DOI: 10.1038/71643.

46. Rose, M., Winston, F. and Hieter, P. (1990) Methods in Yeast Genetics -- A Laboratory Course Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

47. Goldstein, A.L. and McCusker, J.H. (1999) Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast 15, 1541-53 DOI: 10.1002/(SICI)1097-0061(199910)15:14<1541::AID-YEA476>3.0.CO;2-K.

48. Wach, A., Brachat, A., Pohlmann, R. and Philippsen, P. (1994) New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. Yeast 10, 1793-808.

49. Thacker, D., Lam, I., Knop, M. and Keeney, S. (2011) Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in *Saccharomyces cerevisiae*. Genetics 189, 423-39 DOI: 10.1534/genetics.111.131326.

50. Tran, H.T., Keen, J.D., Kricker, M., Resnick, M.A. and Gordenin, D.A. (1997) Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. Mol Cell Biol 17, 2859-65.

51. Gietz, R.D., Schiestl, R.H., Willems, A.R. and Woods, R.A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast 11, 355-60 DOI: 10.1002/yea.320110408.

52. Drake, J.W. (1991) A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A 88, 7160-4.

53. Heck, J.A., Argueso, J.L., Gemici, Z., Reeves, R.G., Bernard, A. *et al.* (2006) Negative epistasis between natural variants of the *Saccharomyces cerevisiae* MLH1 and PMS1 genes results in a defect in mismatch repair. Proc Natl Acad Sci U S A 103, 3256-61 DOI: 10.1073/pnas.0510998103.

54. Gibbons, J.D. and Chakraborti, S. (2014) Nonparametric Statistical Inference, Fourth Edition. Marcel Dekker, Inc., New York.

55. Qiagen, Plasmid purification.

56. Surtees, J.A. and Alani, E. (2006) Mismatch repair factor MSH2-MSH3 binds and alters the conformation of branched DNA structures predicted to form during genetic recombination. J Mol Biol 360, 523-36 DOI: 10.1016/j.jmb.2006.05.032.

57. Strategene, A., pBluescript II phagemid vectors.

58. Gueneau, E., Dherin, C., Legrand, P., Tellier-Lebegue, C., Gilquin, B. *et al.* (2013) Structure of the MutLα C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. Nat Struct Mol Biol 20, 461-8 DOI: 10.1038/nsmb.2511.

59. Groothuizen, F.S., Winkler, I., Cristovao, M., Fish, A., Winterwerp, H.H. *et al.* (2015) MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. Elife 4, e06744 DOI: 10.7554/eLife.06744.

60. Hargreaves, V.V., Putnam, C.D. and Kolodner, R.D. (2012) Engineered disulfide-forming amino acid substitutions interfere with a conformational change in the mismatch recognition complex Msh2-Msh6 required for mismatch repair. J Biol Chem 287, 41232-44 DOI: 10.1074/jbc.M112.402495.

61. Manhart, C.M., Ni, X., White, M.A., Ortega, J., Surtees, J.A. *et al.* (2017) The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans. PLoS Biol 15, e2001164 DOI: 10.1371/journal.pbio.2001164.

62. Hombauer, H., Campbell, C.S., Smith, C.E., Desai, A. and Kolodner, R.D. (2011) Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. Cell 147, 1040-53 DOI: 10.1016/j.cell.2011.10.025.

63. Kadyrov, F.A., Dzantiev, L., Constantin, N. and Modrich, P. (2006) Endonucleolytic function of MutLalpha in human mismatch repair. Cell 126, 297-308 DOI: 10.1016/j.cell.2006.05.039.

64. Perez-Amador, M.A., Abler, M.L., De Rocher, E.J., Thompson, D.M., van Hoof, A. *et al.* (2000) Identification of BFN1, a bifunctional nuclease induced during leaf and stem senescence in Arabidopsis. Plant Physiol 122, 169-80.

CHAPTER 4

Future directions: An analysis of the role of Exo1 in meiotic crossing over

In this chapter, I suggest future directions for the Mlh1-Mlh3 meiosis work described in Chapter 3. The future directions for the Mlh1-Pms1 incompatibility work with clinical isolates is written as a part of Chapter 1.

Introduction

The crossover resolution step involves ZMM proteins Zip1-4, Mer3, and Msh4-Msh5 as well as the Sgs1-Top3-Rmi1 (STR) helicase/topoisomerase complex, Mlh1-Mlh3, and Exo1. Biochemical analyses of Mlh1-Mlh3 protein complex indicate that it does not function as a structure specific nuclease to nick Holliday junctions and that Mlh1-Mlh3 is being directed by other protein complexes to nick a dHJ. Hence, it is important to study Mlh1-Mlh3 in the context of other recombination proteins such as Msh4-Msh5, Exo1, Sgs1, Top3, Rmi1, as also to develop substrates that will resemble the *in-vivo* Holliday junction more accurately. Towards this goal I started to tease apart the role of Exo1 in recombination, and to understand how it interacts with Mlh1-Mlh3 to resolve dHJs.

Exo1 is a nuclease of Rad2/XPG family and has both 5'-3' exonuclease as well as 5' flap endonuclease activities [1]. Interestingly, Exo1 is thought to have a structural or a non-catalytic role in resolving crossovers and data from the Hunter and Borts laboratory suggests that the nuclease function of Exo1 is not necessary for crossing over [2, 3]. But it is not completely understood how it interacts with Mlh1-Mlh3 and whether it has any role in directing Mlh1-Mlh3 to nick at the double Holliday junction to generate crossover products.

My hypothesis is that Exo1 acts as a scaffold at the crossover resolution step, interacting with multiple protein factors such as Mlh1-Mlh3, Sgs1, with which it is known to interact as well as mediating interactions with other factors such as Zip3, Msh4-Msh5. Exo1 in addition to Msh4-Msh5 may be recruiting and directing Mlh1-Mlh3 to create the nicks in an asymmetric manner at the dHJ site to facilitate crossover resolution. Hence, I wanted to test if Exo1 can stimulate the enzymatic activity of Mlh1-Mlh3 *in-vitro*. To do this we need an Exo1 that is nuclease deficient, to eliminate its contribution as a nuclease and focus on its structural function.

Others in the Alani laboratory observed that Exo-D173A, the nuclease active site mutant reported in the literature had residual endonuclease activity *in vitro*. So, I set out to identify a mutant that was active in crossing over but inactive as a nuclease. Towards this goal, I generated a system to study the crossover phenotype in *exo1* mutants as described below.

Materials and Methods

Cloning yeast *EXO1*

HiFi DNA Assembly (NEB) was used to construct *EXO1* (pEAI423) and *exo1Δ* (pEAI422) constructs (Figure 4.1). *EXO1* was PCR amplified from genomic DNA isolated from SKY3576 (SK1 strain) using AO4030 (gcatgcctgcaggtcgactctagagGCGTCTTTAGCAAAGGCG) and AO4031 (gggcctccatgtcTAGCGGCTTGATTAGATAAATAGAG), *KANMX* was PCR amplified from pFA6 plasmid using AO4032 (aatcaagccgctaGACATGGAGGCCCAGAATAC) and AO 4033 (gaatgtagaccgcCAGTATAGCGACCAGCATTC), downstream region of *EXO1* was PCR amplified from genomic DNA using AO4034 (ggtcgctatactgGCGGTCTACATTCGCTATC) and AO4035 (tacgaattcgagctcggtacccgggGGCATAAGAAGGCACGTC). The three inserts were cloned into pUC18 plasmid linearized with *BamH1*.

To construct pEAI422 (*exo1Δ*), *KANMX* was amplified from pFA6 plasmid using AO4037 (tcttcctcagttaGACATGGAGGCCCAGAATAC) and AO4038 (attgggaaagcaaCAGTATAGCGACCAGCATTC), *EXO1* 5' flanking region (280bp) PCR amplified from SK1 genomic DNA using AO4030 and AO4036 (gggcctccatgtcTAACTGAGGAAGAAGACCTTG), *EXO1* 3' flank (340bp) PCR amplified from SK1 genomic DNA using AO4039 (ggtcgctatactgTTGCTTTCCCAATTTGTTTATAAAG) and

AO4040 (tacgaattcgagctcggtacccgggGGAAAAAAAAATGTGAATTGC). The three inserts were cloned into pUC18 linearized with *BamH1*.

pEAI424-427, pEAI436, pEAI437 were constructed with Q5 mutagenesis kit (NEB) using pEAI423 as template.

Mutational analysis of Exo1

pEAI423 and mutant derivatives (generated by Q5 mutagenesis kit-NEB) were digested with *Spe1* and *Nhe1* and the null strain was generated by digesting pEAI422 with *Spe1* and *Sma1* and they were all introduced into SKY3576 and SKY3575 by gene replacement using the lithium acetate transformation method as described in Gietz et al. [4]. At least two independent transformants for each genotype (verified by sequencing) were made resulting in a total of 39 haploid strains bearing either the *exo1* variants, wild type or null. The haploid of each variant in SKY3576 was mated with the haploid of the same variant in SK3575 and the diploid was selected on media lacking leucine and tryptophan and maintained as stable strains. Meiosis was induced upon growing the diploid strains on sporulation media as described in Argueso et al. [5]. Wild-type strains carrying the fluorescent protein markers used to make the above test strains were a gift from the Keeney lab. Strains and plasmids used in this work are shown in Table 4.1, Table 4.2 and Figure 4.1.

Exo1 homology model

The amino acid sequence of yeast Exo1 (amino acids 2-356) was used to construct a homology model (Figure 4.2) using the Phyre2 software (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index). This predicted structure was then aligned to human Exo1 (PDB ID: 3QEB) using Pymol (https://pymol.org/2/).

Spore autonomous fluorescent protein expression to measure percent tetratype

Diploids in the SKY3576 (*MATa; ho::LYS2; lys2; ura3; leu2::hisG; trp1::hisG; THR1::m-Cerulean-TRP1*) /SKY3575(*MATα; ho::LYS2; lys2; ura3; leu2::hisG; trp1::hisG; CEN8::tdTomato-LEU2*) background described above were sporulated on media described in Argueso *et al.* [5]. Spores were treated with 0.5% NP40 and briefly sonicated before analysis using the Zeiss AxioImager.M2 [6]. At least 500 tetrads for each *exo1* allele were counted to determine the % tetratype. Two independent transformants were measured per allele (Table 4.3). A statistically significant difference ($p<0.01$) from wild-type and *mlh3*Δ controls based on χ^2 analysis was used to classify each allele as exhibiting a wild-type, intermediate, or null phenotype.

Table 4.1: Yeast strains used to study *exo1* mutants

Strain	**Genotype in the SK1 strain background**	
EAY4154-4156	*MATa; ho::LYS2; lys2; ura3; leu2::hisG; trp1::hisG; THR1::m-Cerulean-TRP1*, *EXO1::KANMX*	
EAY4157-4159	*MATα; ho::LYS2; lys2; ura3; leu2::hisG; trp1::hisG; CEN8::tdTomato-LEU2*, *EXO1::KANMX*	
EAY4149-4150	*MATa; ho::LYS2; lys2; ura3; leu2::hisG; trp1::hisG; THR1::m-Cerulean-TRP1*, *exo1Δ::KANMX*	
EAY4151-4153	*MATα; ho::LYS2; lys2; ura3; leu2::hisG; trp1::hisG; CEN8::tdTomato-LEU2*, *exo1Δ::KANMX*	
		Amino acid substitution
EAY4160-4162	Same as EAY4154, but *exo1- D78A::KANMX*	D78A
EAY4163-4164	Same as EAY4157, but *exo1*- D78A*::KANMX*	D78A
EAY4165-4167	Same as EAY4154, but *exo1*- D173A*::KANMX*	D173A
EAY4168-4170	Same as EAY4157, but *exo1*- D173A*::KANMX*	D173A
EAY4171-4172	Same as EAY4154, but *exo1- G236D::KANMX*	G236D
EAY4173-4174	Same as EAY4157, but *exo1- G236D::KANMX*	G236D
EAY4175-4177	Same as EAY4154, but *exo1- D78A, D173A::KANMX*	D78A, D173A
EAY4178	Same as EAY4157, but *exo1- D78A, D173A::KANMX*	D78A, D173A
EAY4179	Same as EAY4154, but *exo1*- G236D, D173A*::KANMX*	G236D, D173A
EAY4180-4181	Same as EAY4157, but *exo1- G236D, D173A::KANMX*	G236D, D173A
EAY4182-4184	Same as EAY4154, but *exo1- F447A, F448A::KANMX*	F447A, F448A (MIP)
EAY4185-4187	Same as EAY4157, but *exo1- F447A, F448A::KANMX*	F447A, F448A (MIP)

Table 4.2: Plasmids containing *EXO1* and *exo1* alleles

Plasmid	Allele
pEAI423	*EXO1::KANMX*
pEAI422	*exo1Δ::KANMX*
pEAI424	*exo1-D78A::KANMX*
pEAI425	*exo1-D173A::KANMX*
pEAI426	*exo1-G236D::KANMX*
pEAI427	*exo1- D78A, D173A::KANMX*
pEAI436	*exo1- G236D, D173A::KANMX*
pEAI437	*exo1- F447A, F448A::KANMX*

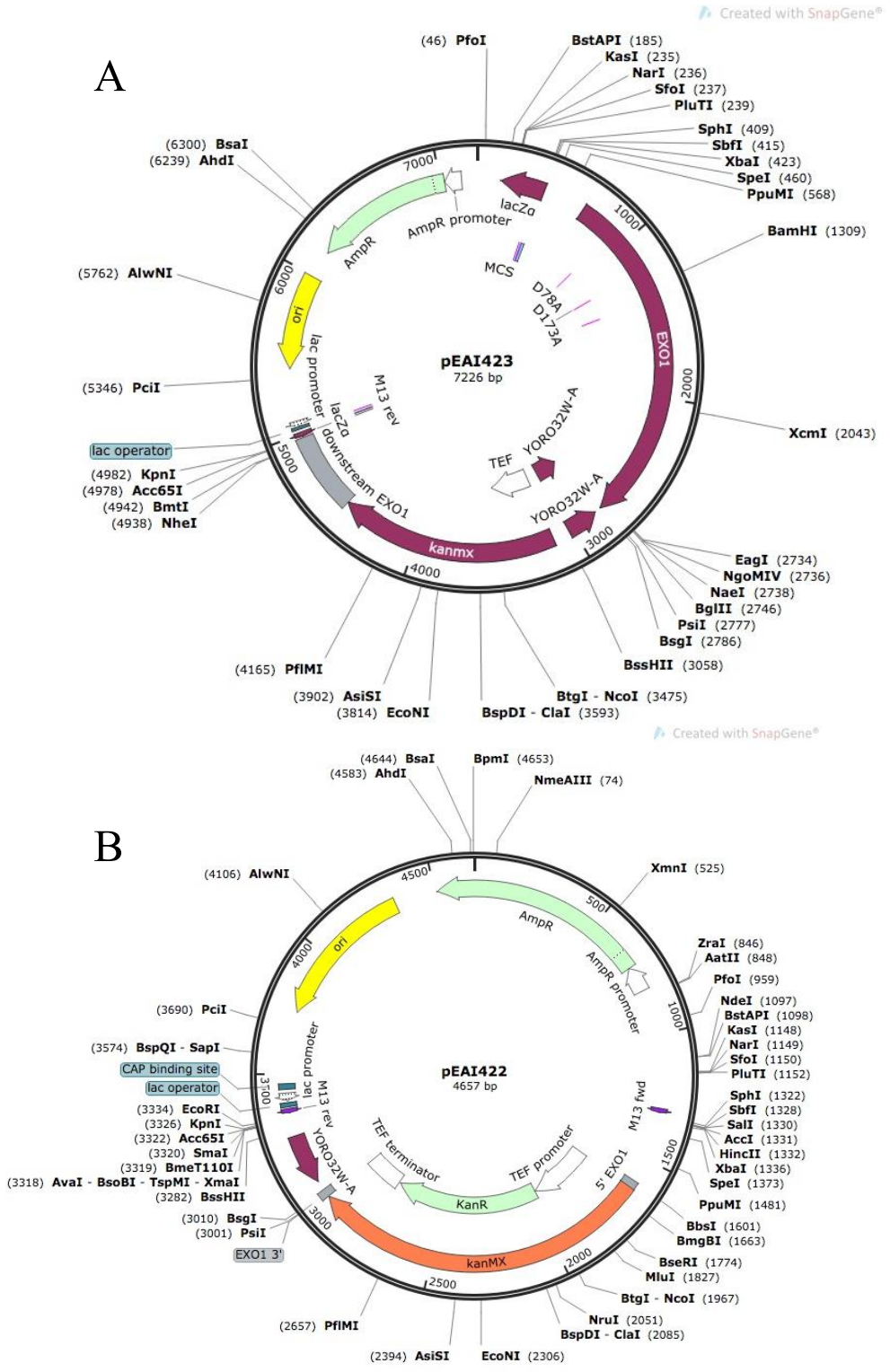


Figure 4.1: *EXO1* Plasmid maps. A. Construct containing wild-type *EXO1::KANMX* or pEAI423. pEAI424-427 and pEAI436, pEAI437 were constructed by Q5 mutagenesis of pEAI423 using appropriate primers. B. Construct containing upstream and downstream sequences of *EXO1* interrupted by a *KANMX* sequence for making *exo1Δ* yeast strains by integration.

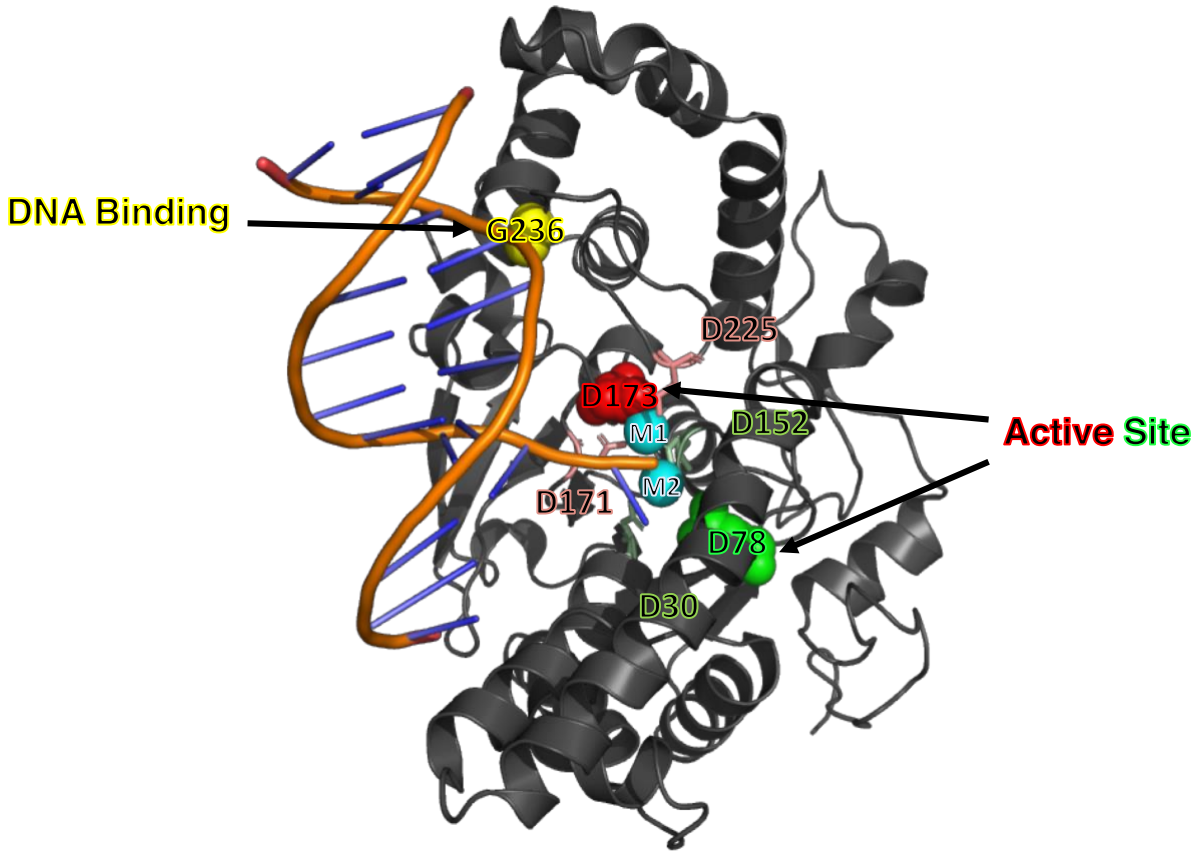


Figure 4.2: Yeast Exo1 modelled on a DNA substrate based on sequence homology to human Exo1. The amino acids predicted to be involved in DNA binding: G236, and the active site: D173, D78, D152, D171 and other charged residues in the vicinity of the active site: D225, D30 are shown. M1 and M2 are the metals at the active site that aid in the catalytic activity of this enzyme.

Results

I generated a system to study the crossover phenotype in *exo1* mutants using the Keeney laboratory's spore autonomous fluorescent assay (Table 4.1, 4.2). I compared *EXO1*, *exo1Δ*, *exo1-F447A, 448A exo1-D78A*, *exo1-D173A*, *exo1-D78A, D173A, exo1-G236D* and *exo1-G236D, D173A*. The crossing over phenotypes for the wild type and null were as expected and the *exo1-F447A, F448A* mutant, mutated in the Mlh1 interaction domain (MIP domain), had an intermediate phenotype, as observed previously (Table 4.3, [3]). The mutants selected to disrupt the nuclease active site: D78A, D173A and the double mutant, were also functional in crossing over (Table 4.3). These results further support the idea that Exo1 has a structural role in crossover resolution. The predicted DNA binding mutant, *exo1-G236D* has an intermediate phenotype (Table 4.3). Interestingly the double mutant of *exo1-G236D, D173A* was wild-type, indicating that the DNA binding mutant may have some non-specific interactions with DNA, but removal of the exonuclease function by D173A mutation in this background, brings it back to wild type levels (Table 4.3). This system allows us to test more mutants of Exo1 to further understand its role in crossing over (Table 4.4).

Table 4.3: Crossover (CO) phenotypes of the *exo1* variants as measured in spore autonomous fluorescent assays.

Strain	% Tetratype	Number of tetrads counted	CO phenotype
EXO1-KANMX x *EXO1-KANMX*	37.5	550	+
exo1Δ x *exo1Δ*	22.0	549	-
exo1-D78A x *exo1-D78A*	39.7	531	+
exo1-D173A x *exo1-D173A*	37.4	519	+
exo1-D78A, D173A x exo1-D78A, D173A	36.4	544	+
exo1-G236D x *exo1-G236D*	29.9	521	INT
exo1-G236D, D173A x *exo1-G236D, D173A*	35.7	532	+
exo1-F447A, F448A x *exo1-F447A, F448A*	31.1	547	INT

Two independently constructed strains with *exo1* variants in the EAY3339 (containing cyan fluorescent protein) and EAY3341 (containing red fluorescent protein) background were analyzed. For each variant, at least two independent transformants were analyzed for each background. The variant in the EAY3339 background was crossed with the variant in the EAY3341 background so that each mutation was homozygous. Diploid strains were induced for meiosis and % tetratype was measured. At least 500 tetrads were counted for each allele. WT is wild-type. +, indistinguishable from WT as measured by $\chi 2$ ($p<0.01$, for % tetratype). -, indistinguishable from null as measured by $\chi 2$ ($p<0.01$). INT is intermediate, distinguishable from both wild-type and null as measured by $\chi 2$ ($p<0.01$).

Open questions and future plans

How Exo1 functions in crossing over and how it coordinates with Mlh1-Mlh3 complex remains a mystery. A targeted mutation strategy to identify DNA binding residues and catalytic domain of Exo1 would further the understanding of the requirements of Exo1 in its crossing over role. It will also aid in identification of mutants of Exo1 that can be used biochemically to test the combined action of Exo1 and Mlh1-Mlh3. I have compiled a list of residues in Exo1 that can be mutated to disrupt the DNA binding (based on human EXO1 data Orans, J. *et al* [7]) or catalytic activities (based on homology model Figure 4.2) as well as a combined mutation in the Mlh1-interaction domain and a DNA binding mutant (Table 4.4). We still need to understand if DNA binding by Exo1 is required for its role in crossing over. My hypothesis for the intermediate phenotype of the DNA binding mutant *exo1-G236D* (Table 4.3) is that it interacts with DNA non-specifically initiating nuclease cleavage at undesired sites. Mutating other DNA binding residues will help identify residues that may obliterate DNA-binding and some that may result in non-specific binding. One could test a combination of different DNA binding mutants with the catalytic D173A mutation to see if non-specific binding is indeed destructive to the function of Exo1 in crossing over, as well as testing to see if phenotype of a completely DNA binding deficient mutant would not change in combination with D173A. Additionally, it would be useful to test a combined mutation of the DNA binding residues and Mlh1- interaction domain residues. If DNA binding of Exo1 is not necessary for crossing over, the combined mutant of *exo1-DNA binding* and *exo1-MIP* should not be worse than the *exo1-MIP* (intermediate in crossing over phenotype).

Table 4.4: Suggested *exo1* mutations to disrupt catalytic activity and DNA binding of Exo1.

Catalytic	**DNA-binding**
D171A	K61A/D
D171A+D173A	R92A/E
K85A	R96A/E
E150A	K185A/D
D152A	G234A/D
	L235A
	FF447,448AA + G236D

Identification of more catalytic residues of Exo1 that do not disrupt the crossing over function will allow purification of these mutant proteins to verify their lack of nuclease activity. The goal would be to test their interaction with Mlh1-Mlh3 as well as stimulation of the endonuclease activity of Mlh1-Mlh3 on plasmid substrates. Additionally, it would be interesting to test the nicking of Holliday junction-like structures using a combination of Mlh1-Mlh3 and Exo1 mutants to see if Mlh1-Mlh3 can now be directed to nick a Holliday junction [8, 9].

Lastly, Exo1 may be playing an important role in interacting with multiple protein complexes involved in resolving dHJs. Disrupting interactions with other proteins, such as Sgs1, similar to disrupting interaction with Mlh1, may affect the crossing over function of Exo1. An alanine scanning mutagenesis of charged residues on Exo1 (the strategy used for mlh3 mutagenesis in Chapter 3, [10]) to pin-point amino acid residues that are crucial for function may help in identifying regions of the Exo1 that are important in interactions with other protein complexes. This mutagenesis may be useful to prove the 'supercomplex hypothesis' for crossover resolution in which we think that multiple protein factors including Mlh1-Mlh3, Sgs1-Top3-Rmi1, Exo1, Zip3, Msh4-Msh5 are stabilize on the double Holliday junction by protein-protein interactions. In this model disrupting interaction with one factor may not affect crossing over since interaction with other factors may compensate for the lack of one interaction. It would be interesting to test an Exo1 mutant that is mildly defective in crossing over, together with mlh3 mutants: mlh3-42 or mlh3-54 (identified in Al-Sweel et al [10] as not being able to interact with Mlh1, but maintains crossing over function), to see if it worsens the phenotype or behaves like null since more parts of the supercomplex are now affected.

References

1. Tran, P.T., Erdeniz, N., Dudley, S. and Liskay, R.M. (2002) Characterization of nuclease-dependent functions of Exo1p in *Saccharomyces cerevisiae*. DNA Repair (Amst) 1, 895-912.

2. Keelagher, R.E., Cotton, V.E., Goldman, A.S. and Borts, R.H. (2011) Separable roles for Exonuclease I in meiotic DNA double-strand break repair. DNA Repair (Amst) 10, 126-37 DOI: 10.1016/j.dnarep.2010.09.024.

3. Zakharyevich, K., Ma, Y., Tang, S., Hwang, P.Y., Boiteux, S. *et al.* (2010) Temporally and biochemically distinct activities of Exo1 during meiosis: double-strand break resection and resolution of double Holliday junctions. Mol Cell 40, 1001-15 DOI: 10.1016/j.molcel.2010.11.032.

4. Gietz, R.D., Schiestl, R.H., Willems, A.R. and Woods, R.A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast 11, 355-60 DOI: 10.1002/yea.320110408.

5. Argueso, J.L., Wanat, J., Gemici, Z. and Alani, E. (2004) Competing crossover pathways act during meiosis in *Saccharomyces cerevisiae*. Genetics 168, 1805-16 DOI: 10.1534/genetics.104.032912.

6. Thacker, D., Lam, I., Knop, M. and Keeney, S. (2011) Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in *Saccharomyces cerevisiae*. Genetics 189, 423-39 DOI: 10.1534/genetics.111.131326.

7. Orans, J., McSweeney, E.A., Iyer, R.R., Hast, M.A., Hellinga, H.W. *et al.* (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. Cell 145, 212-23 DOI: 10.1016/j.cell.2011.03.005.

8. Johnson, R.D. and Symington, L.S. (1993) Crossed-stranded DNA structures for investigating the molecular dynamics of the Holliday junction. J Mol Biol 229, 812-20 DOI: 10.1006/jmbi.1993.1087.

9. Plank, J.L. and Hsieh, T.S. (2006) A novel, topologically constrained DNA molecule containing a double Holliday junction: design, synthesis, and initial biochemical characterization. J Biol Chem 281, 17510-6 DOI: 10.1074/jbc.M602933200.

10. Al-Sweel, N., Raghavan, V., Dutta, A., Ajith, V.P., Di Vietro, L. *et al.* (2017) mlh3 mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide. PLoS Genet 13, e1006974 DOI: 10.1371/journal.pgen.1006974.