

THE MISMATCH REPAIR AND MEIOTIC RECOMBINATION ENDONUCLEASE MLH1-
MLH3 IS DIRECTED BY PROTEIN-PROTEIN INTERACTIONS

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

Najla Abdulaziz Al-Sweel

December 2017

© 2017 Najla Abdulaziz Al-Sweel

THE MISMATCH REPAIR AND MEIOTIC RECOMBINATION ENDONUCLEASE MLH1- MLH3 IS DIRECTED BY PROTEIN-PROTEIN INTERACTIONS

Najla Abdulaziz Al-Sweel, Ph. D.

Cornell University 2017

During meiosis, diploid germ cells that will become eggs or sperm undergo a single round of DNA replication followed by two consecutive chromosomal divisions. The segregation of chromosomes at the first meiotic division is dependent in most organisms on at least one genetic exchange, or crossover event, between chromosome homologs. Homologs that do not receive a crossover frequently undergo nondisjunction at the first meiotic division, yielding aneuploid gametes. Such events have been linked to human disease and infertility. Recent studies suggest that the Mlh1-Mlh3 complex is an endonuclease that resolves recombination intermediates into crossovers. Interestingly, this complex also acts as a matchmaker in DNA mismatch repair (MMR) to remove DNA replication errors.

How does Mlh1-Mlh3 act in two different processes? I investigated this question by performing a mutational analysis of the baker's yeast Mlh3 protein. This resulted in a comprehensive structure-function map that illustrates the effects of 60 new *mlh3* alleles on both functions. From this map, I identified domains critical for one function but not the other. Five alleles (*mlh3-6*, *-23*, *-42*, *-45*, and *-54*) disrupt the MMR function while crossing over remains at nearly wild-type levels. Based on two-hybrid analysis, two of these five alleles (*mlh3-42*, and *-54*) disrupt interaction with Mlh1. This information suggests that the weakened Mlh1-mlh3 heterodimer is stabilized during crossing over by other protein factors. Alternatively, a weak defect in Mlh1-Mlh3 interaction may not be a limiting factor for crossover resolution in meiosis.

In addition, I identified one allele (*mlh3-32*) that disrupted crossing over while maintaining nearly wild-type MMR. 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. Based on these and other observations, my work supports the model that protein-protein interactions with MMR or crossover specific factors are critical in regulating Mlh1-Mlh3's enzymatic activity.

Furthermore, I present five *mlh3* mutations (*mlh3-39*, *-40*, *-57*, *-58*, *59*) predicted to disrupt metal binding based on Pms1 structure, which conferred defects in both functions, suggesting the essential endonuclease active site in Mlh3 folds in a manner analogous to Pms1. Finally, whole genome recombination maps were constructed for wild-type and MMR⁺ crossover⁻, MMR⁻ crossover⁺, endonuclease defective and null *mlh3* mutants in an S288c/YJM789 hybrid background. Compared to wild-type, all of the *mlh3* mutants showed increases in the number of noncrossover events, consistent with recombination intermediates being resolved through alternative recombination pathways. Together, these data illustrate how defective meiotic components can alter the fate of meiotic recombination intermediates, providing new insights into how meiotic recombination pathways are regulated.

BIOGRAPHICAL SKETCH

The author was born in Seattle, Washington where her father pursued a Ph.D. and her mother pursued a B.S. Her father's career next took the family to the Washington, D.C. area where he worked as a diplomat in the Saudi Arabian Embassy. The author's mother specialized in early childhood education and prioritized providing the author and her siblings with an enriching educational experience, in and outside the home. The author spent her childhood in the United States and at the age of thirteen moved with her family to Saudi Arabia. There, she pursued undergraduate studies at the College of Applied Medical Sciences, King Saud University, where she majored in clinical laboratory sciences. She was inspired by her professors and took an interest in genetics and biochemistry. After graduation, she was offered a junior faculty position in the College of Applied Medical Sciences and a full scholarship to pursue M.S. and Ph.D. degrees abroad. She taught a clinical biochemistry lab course to all-female classes for three semesters before pursuing graduate studies. She returned to Washington, D.C. and obtained her M.S. in biochemistry and molecular biology from Georgetown University, where she met her husband Ahmad Cluntun. Both she and her husband then decided to pursue their Ph.D. graduate work at Cornell University. The author joined Eric Alani's lab where she developed a deeper understanding and appreciation for classical genetics. Her son Fouzan was born soon after passing her candidacy exam. She aims to ultimately utilize her education, training and network of collaborators to bridge the gap between science in Saudi Arabia and the United States.

I would like to dedicate this work to my mother, Monirah Al-Fouzan.

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my advisor Dr. Eric Alani. He invited me to join his research group with open arms and was a true mentor in every sense of the word. He balanced giving me guidance and independence as needed. It is an understatement to say I learned a lot from Eric. I am most grateful for his optimism (Eric is always smiling no matter how hard things got) and constant support, in and outside the lab. I feel fortunate to have had the pleasure to train with him, as I was able to witness what true leadership entails. Thank you for everything Eric!

I would also like to thank the various members of the Alani lab, past and present. They were wonderful colleagues and are even better friends. I cherish our friendship and I know we will always stay connected.

Importantly, I would like to thank my committee members, Dr. John Schiemnti and Dr. Mariana Wolfner, for always having an open door. They have kindly taken the time, year after year, to learn about my progress and I am truly appreciative of their guidance and insightful suggestions.

I would also like to thank my collaborators from the Indian Institute of Science Education and Research, Dr. K.T Nishant and his team. Their high-resolution recombination maps were an integral component of this work and it has been a great pleasure to work with them.

I would also like to thank the many Cornell faculty members that helped sculpt my scientific thinking, particularly, Dr. Susan Henry. She has been an inspiration and one of my biggest advocates and for that I will forever be grateful.

I would also like to thank the many friends I made over the years here at Cornell, whether through the Cornell Saudi Club, MECA or BMCB. Your friendships will always be the greatest

gift I obtained from Cornell. Specifically, Auntie Juju, who I consider my Ithacan family. Not to mention the BMCB GFAs Vic, Casey, and Ginger. They always knew what to say to brighten my day.

This work would not have been possible without the generous financial support from King Saud University, the Saudi Arabian Cultural Mission, and the Harry and Samuel Mann Outstanding Graduate Student Award.

Importantly, I would like to thank my parents, Abdulaziz Al-Sweel and Monirah Al-Fouzan. My father, Abdulaziz, is the reason I am here today and my mother, Monirah, is my role model. They instilled in me a sense of pride and confidence to achieve my goals, no matter how hard they may seem. Sadly, my mother passed away during my third year at Cornell, but her strong presence is always with me and remains my constant source of love and motivation. I can never thank my siblings enough, Reema, Farah, Ruba, and Rayed. They are my best friends and each is inspiring in their own way. I am immensely proud of them. I cannot put into words my love for and appreciation of them. What my family has given me is invaluable and unparalleled.

Finally and essentially, I would like to thank my loving husband, Ahmad Cluntun, and my son, Fouzan. Ahmad's faith in me is what made this degree possible. We came to Cornell together and we made it to the finish line together, with the addition of our beautiful baby. They give my life meaning and purpose. Ahmad said it best, his "support, encouragement, quiet patience and unwavering love were undeniably the bedrock upon which the past seven years of my life have been built on". Thank you.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH	v
LIST OF FIGURES	xi
LIST OF TABLES	xiii
CHAPTER 1	1
Introduction.....	1
Mlh1-Mlh3: one endonuclease, two pathways	1
Overview of DNA mismatch repair	1
Overview of meiotic crossing over	2
Roles for Mlh1-Mlh3 in mismatch repair	9
Roles for Mlh1-Mlh3 in crossing over.....	11
Mlh1-Mlh3 is an MMR and meiotic crossing over endonuclease	13
Mlh1-Mlh3 has unique <i>in vitro</i> DNA binding and endonuclease activities	15
How does one endonuclease nick two different DNA substrates?	17
Unanticipated observations infer diverse Mlh1-Mlh3 functions in meiotic recombination	28
Conclusions.....	30
References.....	31
CHAPTER 2	39
<i>mlh3</i> mutations in baker’s yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide	39
Abstract.....	40
Author Summary.....	40
Introduction.....	41
Results.....	47
Discussion	100
Closing thoughts.	106
Methods.....	106
Acknowledgements.....	114
References.....	115
CHAPTER 3	127

Future Directions	127
Working model of Mlh1-Mlh3 function	127
1- Remaining questions and immediate future plans	129
2- Future projects	130
References	135
APPENDIX A	139
Segregation of SNPs in all 44 tetrads as labeled in Chapter 2, Table 2.9	139
APPENDIX B	184
Mismatch repair incompatibilities in diverse yeast populations	184

LIST OF FIGURES

Figure 1.1. Post replicative Mismatch repair (MMR).....	3
Figure 1.2. Meiotic double strand break (DSB) repair pathways.	6
Figure 1.3. Functional interactions of the MutL homologs (MLHs) and the MutS homologs (MSHs).....	8
Figure 1.4. Mlh1-Mlh3 endonuclease can nick two different DNA substrates that form during different genomic stability pathways.	19
Figure 1.5. Distinct protein-protein interactions at the two junctions facilitate asymmetric nicking of the double Holliday junction (dHJ) by Mlh1-Mlh3.....	21
Figure 1.6. Mlh1-Mlh3 conformational changes mediated by an ATPase cycle.....	24
Figure 2.1. DSB repair pathways in meiosis.....	43
Figure 2.2. Site directed mutagenesis of <i>MLH3</i>	48
Figure 2.3. Structure-function map of <i>Saccharomyces cerevisiae</i> Mlh3.....	57
Figure 2.4. Identification and characterization of <i>mlh3</i> separation of function alleles.....	58
Figure 2.5. Cumulative genetic distance and spore viability of <i>mlh3</i> separation of function mutants.	66
Figure 2.6. Genome-wide increase in simple noncrossover events (E1) compared to wild-type in <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> and <i>mlh3Δ</i> mutants.	71
Figure 2.7. Total interhomolog events and distribution of gene conversion tract lengths associated with NCO and CO events in wild-type, <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> and <i>mlh3Δ</i> mutants.	84
Figure 2.8. Distribution of gene conversion tract lengths associated with crossovers and noncrossovers (NCOs) for wild-type, <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> , and <i>mlh3Δ</i>	86
Figure 2.9. Crossover and noncrossover distribution on chromosomes for wild-type, <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> and <i>mlh3Δ</i>	89
Figure 2.10. <i>mlh3</i> mutants display normal meiotic prophase progression as measured by the completion of the first meiotic division.	94
Figure 2.11. Mlh1- <i>mlh3-32</i> and Mlh1- <i>mlh3-45</i> display wild-type endonuclease activities that are differentially stimulated by Msh2-Msh3.	95
Figure 2.12. Mlh1- <i>mlh3-6</i> exhibits wild-type endonuclease and ATPase activity.....	96

Figure 2.13. Sgs1 but not sgs1-hd overexpression differentially affects spore viability in *mlh3Δ*
vs. *mlh3-32*..... 99

LIST OF TABLES

Table 1.1. Mismatch repair (MMR) and meiotic factors described in this chapter.	4
Table 2.1. Yeast strains used in this study.	49
Table 2.2. Diploid strains used to measure % tetatype, spore viability, meiotic progression, genetic map distances and for whole genome recombination mapping.	52
Table 2.3. Plasmids used in this study.	53
Table 2.4. Mismatch repair (MMR) and crossover (CO) phenotypes of the <i>mlh3</i> variants as measured in <i>lys2-A₁₄</i> reversion and spore autonomous fluorescent assays.	59
Table 2.5. Genetic map distances for <i>mlh3</i> separation of function mutants on chromosome XV from single spores and tetrads.	67
Table 2.6. Summary of <i>mlh3</i> separation of function phenotypes.	69
Table 2.7. Spore viability (SV), crossover (CO) and noncrossover (NCO) values for <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> , and <i>mlh3Δ</i> mutants in the S288c/YJM789 hybrid.	73
Table 2.8. Sequencing statistics for spores derived from wild type, <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3- D523N</i> and <i>mlh3Δ</i> in S288c/YJM789 hybrid bearing SK1- <i>MLH1/MLH3</i> alleles.	75
Table 2.9. Crossovers (CO) and noncrossovers (NCO) in tetrads of wild-type, <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> and <i>mlh3Δ</i>	80
Table 2.10. Average crossovers (CO) and noncrossovers (NCO) per chromosome for WT (wild type), <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> and <i>mlh3Δ</i> mutants.	82
Table 2.11. Summary of all event categories, as defined in Oke et al. [34], in wild-type (WT), <i>mlh3-23</i> , <i>mlh3-32</i> , <i>mlh3-D523N</i> and <i>mlh3Δ</i>	87
Table 2.12. Gene conversion tract lengths in <i>mlh3Δ</i> , wild-type, <i>mlh3-23</i> , <i>mlh3-32</i> and <i>mlh3- D523N</i> mutants.	90

CHAPTER 1

Introduction

Mlh1-Mlh3: one endonuclease, two pathways

The MutL homolog heterodimer Mlh1-Mlh3 plays a minor role in mismatch repair (MMR) and a major role in meiotic crossing over. The endonuclease activity of Mlh1-Mlh3 is required for both functions. An area of active study is how one endonuclease accurately nicks two different DNA structures as required in MMR and crossing over, which are two diverse and critical genomic integrity pathways. It is likely that Mlh1-Mlh3 participates in these processes through distinct mechanisms. I will review the studies that support a backup role for Mlh1-Mlh3 in MMR and a central role in meiosis. Furthermore, I will present models that explain how one protein can be active in two seemingly unrelated processes. Finally, I will discuss recent unanticipated observations that reveal an *in vitro* endonuclease and DNA binding activity that is distinct from other meiotic nucleases and observations that imply more diverse functions for Mlh1-Mlh3 than initially characterized.

Overview of DNA mismatch repair

The eukaryotic MMR system is an evolutionarily conserved process that acts to correct mismatches that were generated during replication but escaped polymerase proofreading (reviewed in [1]–[4]). DNA polymerase misincorporation events can result in base-base mismatches or insertion-deletion loops. In eukaryotes, mismatches are recognized by MutS homologs (MSH), which in turn recruit the MutL homologs (MLH) 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 [1]–[4]. In *Saccharomyces cerevisiae* MMR, the heterodimers Msh2-Msh6 and Msh2-Msh3 recognize mismatches and recruit the MLH heterodimer, primarily Mlh1-Pms1, which acts as the major endonuclease in this process [5], [6]. The endonuclease activity of Mlh1-Pms1 is activated by the processivity clamp PCNA to introduce a nick in the newly synthesized strand, in a strand discrimination step that has yet to be determined but is hypothesized to involve pre-existing discontinuities [7]. This nick, 3' or 5' to the mismatch, serves as an entry point for the Exo1 exonuclease to excise the newly synthesized strand and remove the mismatch thus permitting re-synthesis by DNA polymerases followed by ligation [1]–[4] (Figure 1.1; Table 1.1). In addition to forming a heterodimeric complex with Pms1, Mlh1 can interact with either Mlh2 or Mlh3. A minor MMR pathway that involves mismatch recognition by Msh2-Msh3 followed by interaction with Mlh1-Mlh3 has also been observed [8]–[10]. Curiously, beyond its minimal participation in MMR, Mlh1-Mlh3 has been shown to play a central role in meiotic crossing over.

Overview of meiotic crossing over

Crossing over during Meiosis I creates physical connections between homologous chromosomes that are critical for their accurate segregation. Failure to achieve at least one crossover (CO) per homolog pair results in nondisjunction and the formation of aneuploid gametes [11]. Errors in meiotic chromosome segregation are a leading cause of spontaneous miscarriages and birth defects [12]. In baker's yeast, the recombination process leading to CO formation begins in meiotic prophase, after DNA replication, with the programmed induction of ~150-200 double strand breaks (DSBs) that appear throughout the genome. DSBs are resected to

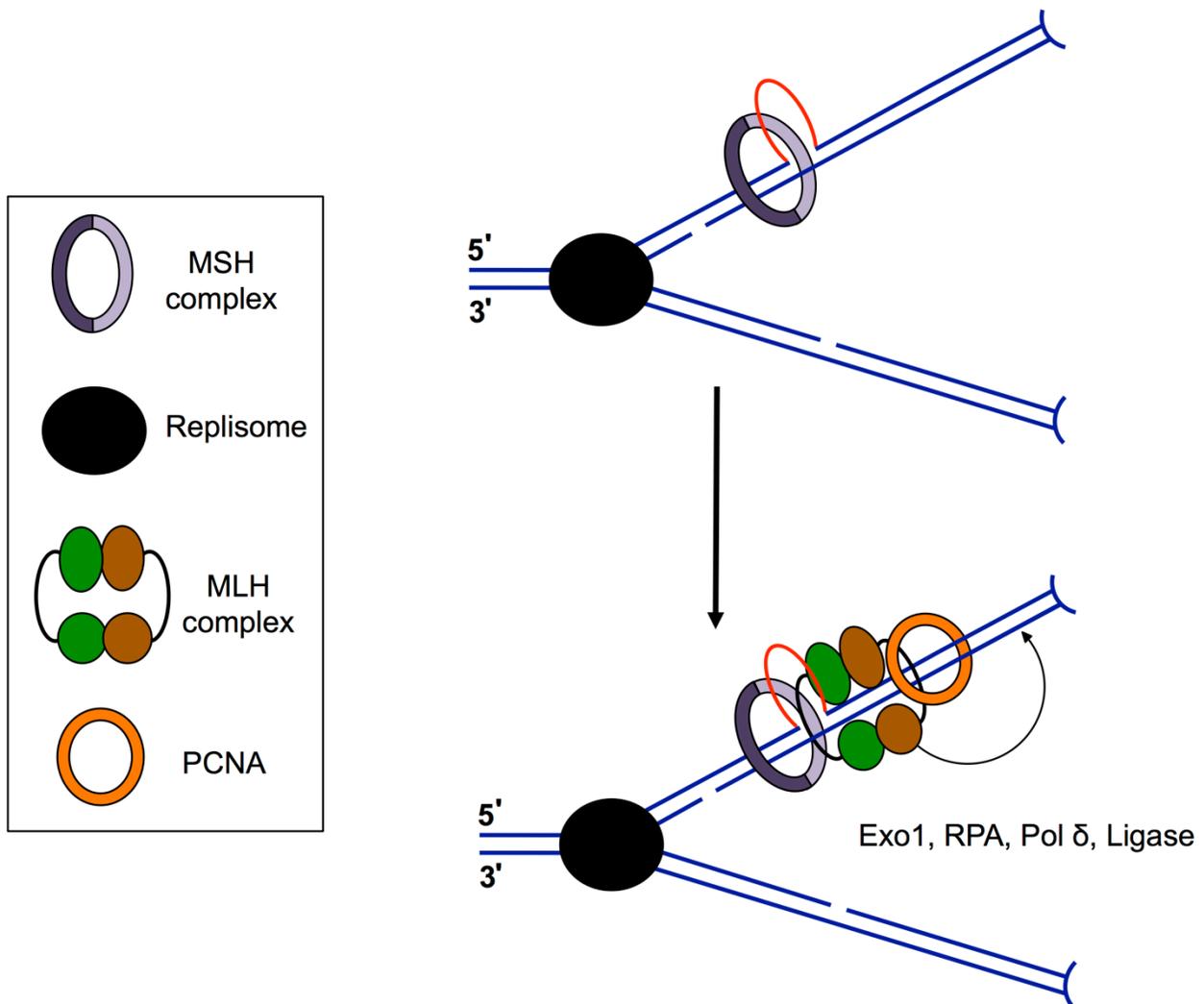


Figure 1.1. Post replicative Mismatch repair (MMR).

Overview of the eukaryotic MMR process. Insertion/deletion loops (deletion loop shown in red) are a result of polymerase slippage events, which can occur during the replication of repetitive DNA sequences. In eukaryotes, MMR proteins are hypothesized to travel with the replication fork. Recognition of a mismatch will trigger the assembly of a “repirosome” consisting of an MSH complex, an MLH complex, and PCNA (shown), in addition to RFC and RPA (not shown). This complex will identify the newly synthesized strand by an as yet undetermined mechanism involving a pre-existing nick. The MLH complex (primarily Mlh1-Pms1) will then introduce a nick at distal sites on the same strand (black arrow). Loading of ExoI will activate its 5'-to-3' exonuclease activity until the mismatch is removed. DNA polymerase δ will then fill in the resulting single-stranded gap, and DNA ligase will seal the remaining nick.

Table 1.1. Mismatch repair (MMR) and meiotic factors described in this chapter.

Protein in <i>Saccharomyces cerevisiae</i>	Function
Msh2-Msh6 Msh2-Msh3	mismatch recognition complexes
PCNA	DNA polymerase processivity factor
RFC	PCNA loader
RPA	ssDNA binding protein
Exo1	5'-3' dsDNA exonuclease
STR (Sgs1-Top3-Rmi1 complex)	DNA helicase/topoisomerase
ZMM proteins: Zip3 Mer3 Msh4-Msh5	SUMO E3 ligase meiosis specific DNA helicase meiosis specific MSH dimer important for joint molecule intermediate stability
Rad52	contains single-strand annealing activity
Rad54	DNA dependent ATPase that stimulates strand exchange

create 3' single stranded tails that can invade the homologous template to create a strand invasion intermediate known as a displacement-loop (D-loop). The D-loop can be extended to form a stable heteroduplex intermediate known as the single end invasion (SEI) [13]–[15]. Working models propose that the SEI intermediate can be processed by a number of pathways facilitated by the Sgs1-Top3-Rmi1 (STR) helicase/topoisomerase complex, which has recently been labeled as the master regulator of meiotic DSB repair [16]–[18] (Figure 1.2).

Initially the STR complex was characterized as playing an anti-crossover role due to its role in promoting noncrossover (NCO) formation via unwinding of the SEI intermediate in a pathway known as synthesis dependent strand annealing (SDSA) [19]. However, recent studies have uncovered a pro-crossover function for the STR complex [16]–[18]. Taken together, these findings led to the “recombination intermediate chaperone model” in which the STR complex acts as a chaperone designating different resolution pathways for early strand invasion intermediates [16] (Figure 1.2). In this model the SEI intermediate is metabolized by the STR complex, to form early NCOs (via SDSA), or to facilitate the return of events to the original DSB state thus promoting capture and protection by the ZMM proteins (Zip1-4, Mer3, and Msh4-Msh5). In the ZMM pathway, SEI intermediates are stabilized by Msh4-Msh5 and subject to further processing by a second-end capture mechanism catalyzed by RPA, Rad52, and Rad54 (Table 1.1) [20]. During second-end capture, the newly synthesized invading strand can re-anneal and ligate to the other side of the DSB creating a joint molecule intermediate known as a double Holliday junction (dHJ).

dHJs formed in the ZMM pathway are biased to resolve the two junctions asymmetrically, such that the resulting product is exclusively a CO, in steps requiring the endonuclease activity of Mlh1-Mlh3 and exonuclease-independent activities of Exo1 [21]. In this

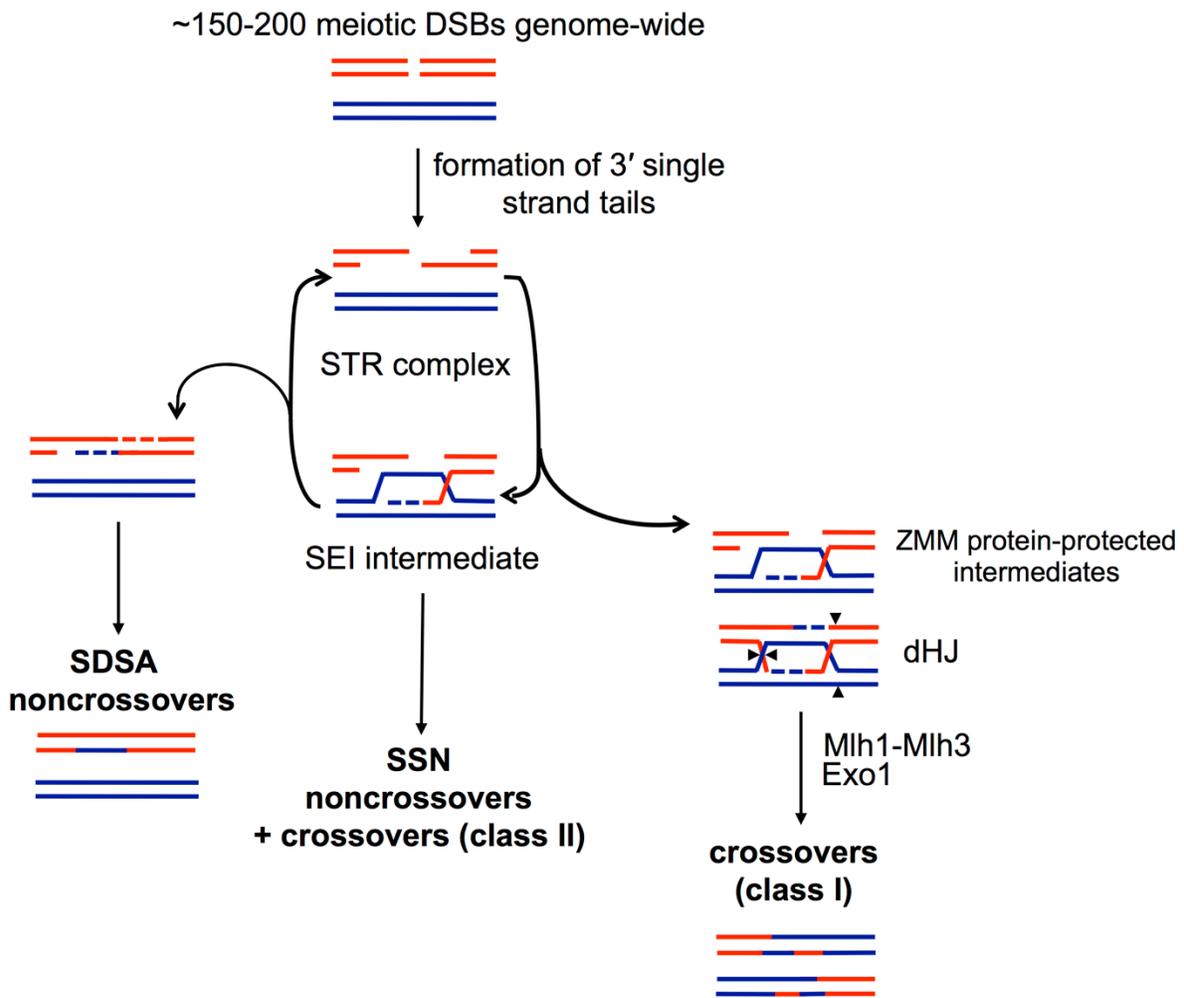


Figure 1.2. Meiotic double strand break (DSB) repair pathways.

DSBs, induced early in Meiosis I, are resected to create 3' single stranded tails that can invade the homologous template. Early strand invasion intermediates can be extended to create a single end invasion (SEI) intermediate. The Sgs1-Top3-Rmi1 (STR) helicase/topoisomerase complex can disassemble the SEI intermediate to facilitate noncrossover formation via synthesis dependent strand annealing (SDSA). Alternatively, return of early strand invasion events to the original DSB state by STR activity facilitates capture by the ZMM proteins (Zip1-4, Mer3, and Msh4-Msh5). SEI stabilization by the ZMM proteins and subsequent second-end capture gives rise to the double Holliday junction (dHJ). dHJs are nicked in a biased manner (black arrows depicted in opposite orientations) to result exclusively in crossovers (class I) in a process that involves Mlh1-Mlh3 and Exo1. Strand invasion events that escape STR disassembly form unregulated joint molecules in addition to dHJs that are susceptible to the unbiased resolution of structure selective nucleases (SSNs) giving rise to both crossovers (class II) and noncrossovers. Model adapted from Kaur et al. [17].

pathway COs are interference-dependent and known as class I COs. Interference is a mechanism that regulates CO placement such that the presence of one CO reduces the probability of a second CO in a nearby region [22]. Class I COs account for the majority of COs in baker's yeast. Furthermore, in this model, it is proposed that strand invasion events which escape STR disassembly form unregulated joint molecules that can be resolved by the structure selective nucleases (SSNs: Mus81-Mms4, Slx1-Slx4, and Yen1) [23]. In this pathway the two junctions in a dHJ are resolved independently into both COs and NCOs. COs formed by SSNs are interference-independent and known as class II COs. It is important to note that there is redundancy in the joint molecule resolution pathways described above. Functional links exist between the different resolvases such that loss of one pathway results in a compensatory rise in the other [16], [23], [24].

In addition to Mlh1-Mlh3's involvement in the formation of COs, the MMR machinery has been shown to play other roles during meiotic DSB repair. When the donor DNA sequence differs from the acceptor, mismatches are formed in heteroduplex strand invasion intermediates. Mlh1-Pms1 has been shown to initiate repair of mismatches present in one chromosome by using sequences present on the other as a template. This process is known as gene conversion (GC) [25], [26]. Furthermore, MMR proteins prevent recombination between divergent DNA sequences in a process known as heteroduplex rejection [27], [28]. These meiotic roles are similar to gene conversion and anti-recombination roles observed for MMR proteins during DSB repair in non-meiotic cells. It is important to note that Mlh1-Mlh2, also dubbed a minor MMR factor [9], has recently been shown to play an essential role in limiting GC tract lengths without affecting CO formation [29] (Figure 1.3).

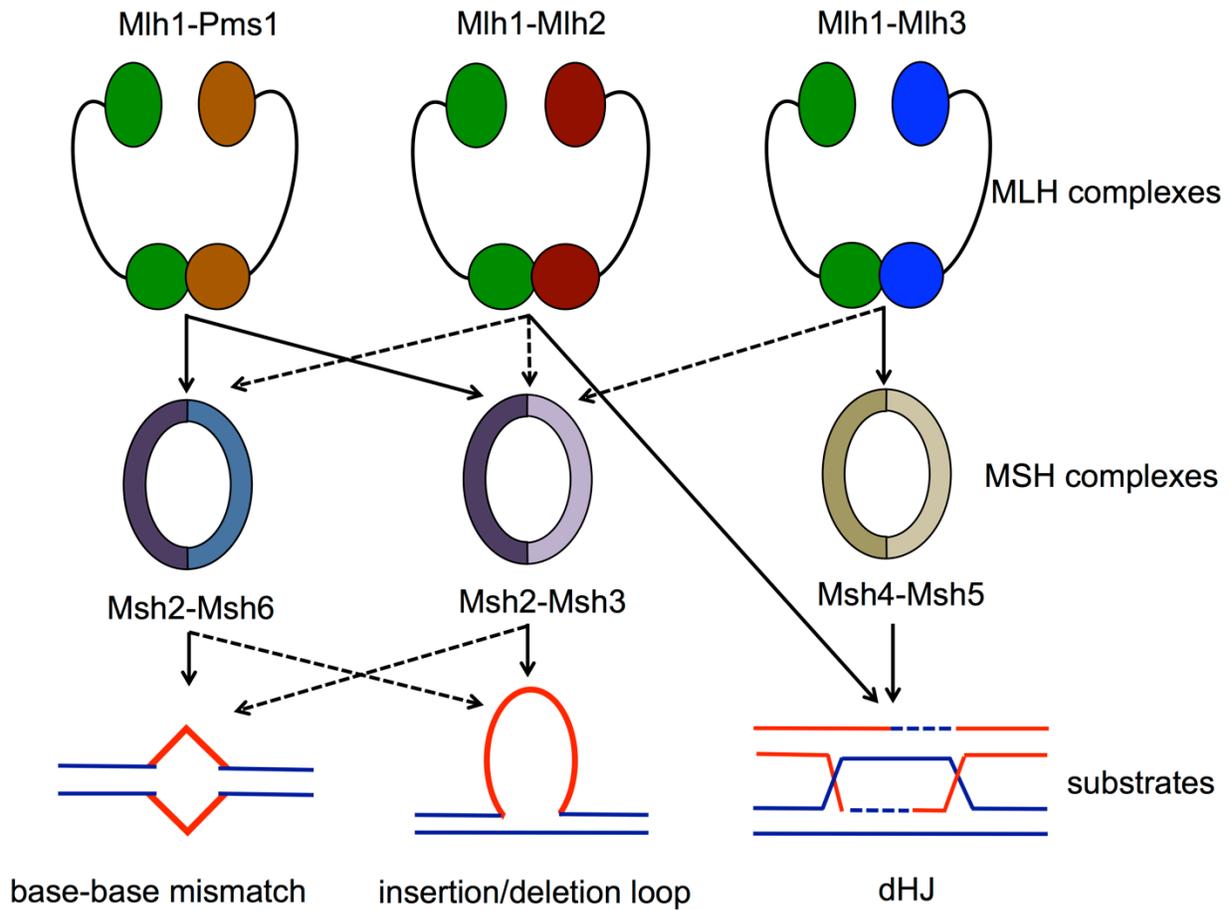


Figure 1.3. Functional interactions of the MutL homologs (MLHs) and the MutS homologs (MSHs).

MLHs are composed of globular N- and C-terminal domains (depicted as ovals and circles respectively) connected by a linker arm. Arrows indicate functional interactions between the MSH and MLH heterodimers and their DNA substrates. Dashed arrows (in black) denote minor roles. Mlh1-Mlh2 interacts with Mer3 (not shown) to limit gene conversion tract lengths in meiotic recombination, in addition to its minor MMR role. Mlh1-Mlh2 interacts with the single end invasion intermediate (not shown) that forms prior to the double Holliday junction (dHJ). Figure adapted from Reyes et al. [2].

Roles for Mlh1-Mlh3 in mismatch repair

Eukaryotic MMR entails mismatch recognition by MSH proteins, followed by processing and repair events that involve MLH proteins. MSH and MLH proteins perform their functions as heterodimers (reviewed in [1]–[4]). In *Saccharomyces cerevisiae* MMR, there are two MSH complexes: MutS α (Msh2-Msh6) and MutS β (Msh2-Msh3), and three MLH complexes: MutL α (Mlh1-Pms1), MutL β (Mlh1-Mlh2) and MutL γ (Mlh1-Mlh3). Msh2-Msh6 and Msh2-Msh3 recognize mismatches with partial redundancy (Figure 1.3). Msh2-Msh6 primarily recognizes base-base mismatches and small insertion/deletion loops, while Msh2-Msh3 primarily recognizes larger insertion deletion/loops. The major MMR pathway involves mismatch recognition by Msh2-Msh6 and the endonuclease activity of Mlh1-Pms1. Mlh1-Pms1 nicks the newly synthesized strand distal to the mismatch providing an entry point for downstream repair factors [5], [6]. Although Mlh1-Pms1 is central to the MMR process, both Mlh1-Mlh2 and Mlh1-Mlh3 have been shown to minimally contribute to the repair of mutational intermediates [9].

Mlh1-Mlh3 can partially substitute for Mlh1-Pms1 in MMR [8]–[10]. Flores-Rozas and Kolodner [8] showed that yeast strains deficient in Mlh3 display a 2-3-fold increase in mutation rates relative to wild-type, a small but significant effect in assays where a *pms1* deletion displays a 50 to 300-fold increase. Likewise Harfe et al. [9] observed an increase in mutation rates in *mlh3 Δ* strains of up to 21-fold relative to wild-type in assays where *msh2 Δ* resulted in an approximately 10,000-fold increase. Furthermore, epistasis analysis of *msh3 mlh3* double mutants suggest the two proteins function in the same repair pathway. Flores-Rozas and Kolodner [8] examined mutator phenotypes by determining the rate of reversion of the *hom3-10* and *lys2-Bgl* alleles, two assays that are particularly sensitive to defects in MMR. They observed that while *msh6 mlh3* double mutants display synergistic increases in mutation rates as compared

to each single mutant, *msh3 mlh3* double mutants had essentially the same mutation rates as each single mutant [8]. Thus, these data demonstrate that Mlh1-Mlh3 likely participates in the repair of some insertions/deletions in the Msh2-Msh3 but not the Msh2-Msh6 pathway. Although studies reported to date do not provide evidence for an interaction between Msh6 and Mlh3, it is important to note that other assays, utilizing different substrates, might detect such an interaction.

Romanova and Crouse propose a model in which Mlh1-Mlh3 mediated repair requires Mlh1-Pms1 [10]. In this study, the authors aimed to investigate the different contributions of the MMR machinery in the repair of insertion versus deletion loops. Insertion and deletion loops are a result of polymerase slippage events, which can occur during the replication of repetitive DNA sequences [30]. Loops that form on the primer strand lead to insertions, whereas loops on the template strand lead to deletions. In order to compare the effects of insertion and deletion mismatches of the same size, the authors utilized two complementary reversion assays: the -1 *lys2ΔA746* and the +1 *lys2ΔBgl* frameshift alleles in the same *LYS2* gene region. Loop mutations were generated *in vivo* via single-strand oligonucleotide (oligo) transformations. Oligos transform by serving as primers on either the leading or lagging strands. Additional nucleotides in the oligo will create a primer-strand loop resulting in an insertion mutation, while missing nucleotides in the oligo will create template-strand loop resulting in a deletion mutation. If the mismatch created by the oligo is not removed, a reverting frameshift will result in the next round of replication. From this analysis, Romanova and Crouse observed that deletion loop repair was more dependent on Msh2-Msh3 while insertion loops were more biased towards repair by Msh2-Msh6 [10]. In accordance with these observations, Mlh1-Mlh3 was found to function in the repair of deletion loops, presumably mediated by Msh2-Msh3, and not insertion loops. *mlh3Δ* resulted in a 4-fold decrease in deletion loop repair in an assay where *pms1* mutations exhibit a

92-fold decrease as Mlh1-Pms1 is needed for most repair. Interestingly, the authors observed that a *PMS1* point mutant (*pms1-H888R*), which shows a deficiency in insertion repair, displayed a modest 2-fold decrease in deletion repair. Strikingly, given the small effect of each individual mutation, *pms1-H888R mlh3Δ* double mutants exhibited a synergistic 13-fold reduction in deletion repair. Thus, the authors conclude that deletion repair in *pms1-H888R* strains has a major dependence on Mlh1-Mlh3. Furthermore, in *pms1-H888R mlh3Δ msh6Δ* strains, where only Msh2-Msh3 is present, no further effect was observed. Taken together, these data suggest a model in which deletion loops are recognized by Msh2-Msh3 that acts to recruit Mlh1-Mlh3, which in turn facilitates interactions with Mlh1-pms1-H888R.

Roles for Mlh1-Mlh3 in crossing over

Accurate chromosome segregation in Meiosis I requires reciprocal exchange of genetic information between homologous chromosomes known as crossing over. In most eukaryotes, the majority of COs are formed through a DSB repair pathway that involves stabilization of joint molecule intermediates by the ZMM proteins, leading to their ultimate resolution by Mlh1-Mlh3 and the exonuclease-independent functions of Exo1 [13]–[15], [31]. As outlined earlier in this chapter, current models propose a central role for the STR complex in this process [16]–[18]. A critical role for yeast MLH proteins in crossing over was established by Hunter and Borts [32]. The authors analyzed crossing over frequencies in many MMR mutants including *mlh1*, *pms1*, *msh2*, *msh3*, and *msh6*. Interestingly, they observed that the ability to promote crossing over is unique to Mlh1 among the MMR proteins tested. Furthermore, the authors observed that the *mlh1 msh4* double mutant is indistinguishable from the *msh4* single mutant. Hence, they concluded that Mlh1 acts in the same pathway as Msh4-Msh5 to generate COs. Msh4-Msh5 is a

meiosis-specific MSH complex with no role in MMR [33], [34] (Figure 1.3).

Wang et al. [35] first identified a major role for Mlh3 in crossing over in a study that aimed to functionally characterize the three Mlh1-based heterodimers in meiosis. The authors investigated crossing over at a recombination hot spot on chromosome III in *Saccharomyces cerevisiae* and observed a 21-33% reduction in crossing over in *mlh3* mutants, indistinguishable to reductions observed in *mlh1*. In addition, homolog nondisjunction rates were similarly elevated in *mlh1* and *mlh3* mutants relative to wild-type strains.

Argueso et al. [24] proposed that competing pathways exist to promote COs during meiosis. These authors aimed to investigate the relationship between different complexes known to promote COs in yeast. Mus81-Mms4 is an endonuclease complex shown to be important for wild-type levels of crossing over. Interestingly, they observed the most significant decrease in crossing over (13 to 15-fold relative to wild-type) in *mlh1Δ mms4Δ* double mutants. Similar significant reductions in CO levels (6 to 17-fold relative to wild-type) were observed in *mlh3Δ mms4Δ* by Sonntag Brown et al. [36]. Furthermore, Zakharyevich et al. [23] performed a comprehensive analysis of joint molecule resolution during meiosis in budding yeast and concluded that Mlh1-Mlh3 together with the SSNs (Mus81-Mms4, Slx1-Slx4, and Yen1) ultimately account for all joint molecule resolution, and that they can sometimes compensate for each other.

In their study, Zakharyevich et al. [23] used a DNA physical assay system to monitor recombination at a recombination hot spot. Interestingly, the authors observed efficient joint molecule formation and resolution in strains lacking the SSNs (*mms4 slx4 yen1* triple mutants), with NCOs at wild-type levels and COs at about 70% of wild-type levels. The authors conclude that the SSNs are dispensable for the bulk of meiotic joint molecule resolution. More

importantly, the authors found that deleting either Sgs1 or Mlh3 in this background (*mms4 slx4 yen1*) resulted in a similar reduction of CO levels, to just 10% of wild-type levels, suggesting a pathway where Sgs1-dependent COs require Mlh1-Mlh3. Similar results were observed in *mms4 slx4 yen1* strains deficient in Top3 or Rmi1 [17], [18]. Thus, these data indicate that the STR complex promotes the majority of COs in conjunction with a resolvase that is not Mus81-Mms4, Slx1-Slx4 or Yen1. However, it is important to note that while *mms4 slx4 yen1 sgs1* quadruple mutants exhibit a severe defect in NCO levels, the *mms4 slx4 yen1 mlh3* mutants do not, as Sgs1 plays an integral role in both CO and NCO production. Similar conclusions were made by De Muyt et al. [16].

The studies described above indicate a critical role for Mlh1-Mlh3 in yeast meiosis. To examine Mlh1-Mlh3 in mammalian meiosis, Lipkin et al. generated *Mlh3^{-/-}* mice [37]. The authors found that *Mlh3^{-/-}* mice were viable but infertile. They observed smaller male testes in *Mlh3^{-/-}* mice and severe depletion of spermatocytes due to aneuploidy, which resulted in apoptosis. Furthermore, chromosome preparations showed a significant defect in chiasma formation, the physical manifestation of CO sites. The authors did not observe a role for Mlh3 in mammalian MMR.

Mlh1-Mlh3 is an MMR and meiotic crossing over endonuclease

Based on homology with Pms1, Mlh3 contains a highly conserved ATP binding motif in the N-terminal domain (NTD), which is connected to the C-terminal domain (CTD) by a dynamic and unstructured linker arm. As in other MLH dimers, the Mlh3 CTD dimerizes with the Mlh1 CTD [38] (Figure 1.3). Kadyrov et al. categorized Mlh1-Pms1 in yeast (and Mlh1-Pms2 in humans) as a latent endonuclease whose catalytic activity is dependent on the integrity

of a highly conserved metal-binding motif (DQHA(X)₂E(X)₄E) in the CTD of the Pms1 subunit [5], [6]. A study performed by Nishant et al. [39] investigated this motif in yeast Mlh3. They found a single point mutation within this motif (*mlh3-D523N*) conferred *mlh3Δ* like defects in both MMR and meiotic crossing over. A 7.9-fold increase in the mutation rate of *mlh3-D523N* strains was observed, similar to a 6.3-fold increase observed in *mlh3Δ*. CO frequency was measured in four genetic intervals on chromosome XV and about a 2-fold reduction in crossing over was observed in both *mlh3-D523N* and *mlh3Δ*. Similar reductions in spore viabilities were also detected in *mlh3-D523N* and *mlh3Δ*. Importantly, the *mlh3-D523N* mutation did not affect protein stability or interaction with Mlh1 as measured by yeast two-hybrid assays, chromatography and western blots. These data suggest that Mlh1-Mlh3 possesses a catalytic activity critical for its functions.

The DQHA(X)₂E(X)₄E motif is hypothesized to be part of the Mlh1-Pms1 endonuclease active site. Indeed, in the crystal structures of the Mlh1-Pms1 CTDs solved from *Saccharomyces cerevisiae*, Gueneau et al. [38] found residues within this motif implicated in metal binding in addition to two other highly conserved motifs (ACR and C(P/N)HGRP). In particular, the authors identified five Pms1 residues, located within these motifs (H703, E707, C817, C848, H850), which form the metal binding site through folding of the Pms1 CTD. Based on this and on the study performed by Nishant et al. [39] described above, Al-Sweel et al. [40] identified the putative Mlh1-Mlh3 endonuclease active site. By sequence alignment and homology modeling of Mlh3 based on Pms1, the authors identified and mutated the following residues predicted to form the metal binding site in the Mlh3 CTD: H525, E529, C670, C701, and H703. They observed null-like phenotypes in MMR and crossing over, similar to defects observed in *mlh3-D523N*, indicating that these residues are essential for Mlh1-Mlh3 functions. Thus, the authors conclude

that the catalytic sites in yeast Pms1 and Mlh3 likely fold in a similar manner.

Rogacheva et al. [41] and Ranjha et al. [42] purified yeast Mlh1-Mlh3 that was active *in vitro*. Rogacheva et al. observed that, like Mlh1-Pms1, Mlh1-Mlh3 has an intrinsic endonuclease activity that is dependent on the integrity of its endonuclease motif; Mlh1-Mlh3, but not the endonuclease mutant Mlh1-mlh3-D523N, was shown to nick supercoiled plasmid DNA *in vitro*. Surprisingly, unlike Mlh1-Pms1, the catalytic activity of Mlh1-Mlh3 was not stimulated by ATP or the replication factors RFC (clamp loader) and PCNA (processivity clamp). Rather, Rogacheva et al. observed stimulation by Msh2-Msh3. The authors conclude that physical interactions between Msh2-Msh3 and Mlh1-Mlh3 act to coordinate substrate nicking during MMR. Ranjha et al. [42] observed similar Mlh1-Mlh3 endonuclease activity on supercoiled plasmid DNA *in vitro* that was not stimulated by ATP or RFC and PCNA. Furthermore, the authors tested the binding selectivity of Mlh1-Mlh3 on different oligo-based DNA substrates and demonstrated a preference for binding the un-stacked or “open” form of a Holliday junction (HJ). However, in the following section, I will present arguments suggesting that, although Mlh1-Mlh3 can bind HJs, it is not the substrate for activation of the Mlh1-Mlh3 endonuclease.

Mlh1-Mlh3 has unique *in vitro* DNA binding and endonuclease activities

Manhart et al. [43] propose a novel model in which Mlh1-Mlh3 is loaded onto DNA to form an activated polymer that, directed by other recombination proteins, cleaves recombination intermediates to form COs. The authors observe inhibition of Mlh1-Mlh3 endonuclease activity upon binding synthetic mismatched and HJ DNA substrates. Thus, they argue that Mlh1-Mlh3 does not act according to archetypes set by canonical HJ resolvases that recognize and cleave branched DNA structures. Interestingly, they found Mlh1-Mlh3’s endonuclease activity to be

more active on circular DNA than on linear DNA and that this activity increases as substrate size increases, but is inhibited by disruptions in DNA, such as a loop mismatch or biotin-streptavidin linkages. Based on these and other findings, the authors conclude that multiple Mlh1-Mlh3 molecules interact on the DNA substrate to form a polymer, which in turn licenses its nicking activity. In support of this model, the authors visualized, via electron microscopy, tightly packed protein-DNA clusters that form when circular DNA and Mlh1-Mlh3 are mixed together under conditions that result in optimal endonuclease activity. Similar higher-order protein-DNA structures were visualized by Claeys Bouuaert and Keeney [44] using atomic force microscopy (AFM).

Claeys Bouuaert and Keeney [44] performed a detailed biochemical characterization of the DNA-binding properties of yeast Mlh1-Mlh3. When Mlh1-Mlh3 was pulled down using beads coated with different DNA substrates, efficient pull down was observed with HJs, as reported previously[42], and, surprisingly, with ssDNA. Moreover, they observed that Mlh1-Mlh3's DNA-binding activities to these two substrates are largely non-competing. This led the authors to hypothesize that the Mlh1-Mlh3-DNA binding sites are likely different for HJs and ssDNA, therefore raising the possibility that the complex can potentially bind two substrates at once. To test these ideas, the authors used hydroxyl radical footprinting to map the HJ and ssDNA-binding interfaces of Mlh1-Mlh3. Indeed, they observed that the DNA-binding interfaces for the two substrates only partially overlap, with HJ binding mapping predominantly to the linkers of Mlh1 and Mlh3, and ssDNA binding mapping to both the N-terminal domains and the linkers of Mlh1 and Mlh3. Furthermore, Claeys Bouuaert and Keeney incubated Mlh1-Mlh3 with equimolar concentrations of biotin-labeled and radiolabeled substrates (HJ or ssDNA) and, after complex assembly, observed pull-down of radiolabeled substrates with streptavidin beads.

Thus the authors conclude that Mlh1-Mlh3 can bind multiple substrates simultaneously. In agreement with this model, Manhart et al. [43] showed that an active Mlh1-Mlh3-DNA nucleoprotein complex can interact with another DNA substrate in trans. They observed Mlh1-Mlh3 nicking on small linear substrates only when incubated with large closed circular substrates.

These studies reveal distinct DNA binding and endonuclease activities unique to Mlh1-Mlh3 that do not conform to paradigms set by previously characterized dHJ resolvases. Rather, Mlh1-Mlh3 displays DNA binding properties similar to Mlh1-Pms1 that are critical for its MMR function. Thus, the authors hypothesized that such properties are utilized by Mlh1-Mlh3 in meiosis to resolve dHJs into COs.

How does one endonuclease nick two different DNA substrates?

Recent studies by Al-Sweel et al. [40] and Claeys Bouuaert and Keeney [44] have shown that the MMR and meiotic functions of *Saccharomyces cerevisiae* Mlh3 are genetically separable. Al-Sweel et al. created a structure-function map of Mlh3 by analyzing 60 new *mlh3* alleles. The authors applied a clustered charged-to-alanine scanning mutagenesis approach. They hypothesized that replacement of a charged patch from Mlh3's surface with alanine residues would disrupt protein-protein or protein-DNA interactions with minimal affect on Mlh3 folding. Claeys Bouuaert and Keeney, on the other hand, mutated conserved residues that appear to be in proximity to DNA as judged from hydroxyl radical footprinting. In both studies the authors tested the effects of their mutations on MMR and meiotic functions *in vivo*. Although the two independent studies utilized different mutagenesis strategies, many of the same residues were targeted and comparable overall phenotypes were observed. N-terminal mutants were more

uniformly defective in MMR, suggesting the ATP-binding domain may be more critical for this function, while more severe defects in crossing over were observed in the linker mutants.

Importantly, both studies identified *mlh3* alleles with separable functions in MMR and crossing over. Similar observations regarding the importance of the Mlh3 ATP-binding domain in MMR were made by Sonntag Brown et al. [36] in a study that characterized eight *mlh3* ATPase mutants in MMR and crossing over.

In the following section, I will present models, which are not mutually exclusive, that can explain how one protein is regulated to act in two different genomic stability pathways (Figure 1.4).

a. Protein-protein interactions with MMR or pro-crossover factors direct endonuclease activity

Al-Sweel et al. [40] propose a model in which the Mlh1-Mlh3 endonuclease activity on insertion/deletion loops vs. dHJs is activated and directed by protein-protein interactions with MMR or CO specific factors respectively. They hypothesize that the positioning of other protein factors on DNA provides a unique substrate for Mlh1-Mlh3 to act on such that these factors spatially and temporally coordinate its nicking activity to result in strand specific repair or CO products. The authors characterized two variants with opposing separation of function phenotypes. *mlh3-32* conferred null-like defects in crossing over with MMR at levels nearly indistinguishable from wild-type. On the other hand, *mlh3-45* conferred null-like defects in MMR with CO frequencies nearly indistinguishable from wild-type. Interestingly, they did not observe defects in the *in vitro* endonuclease, and by extension, DNA-binding activities of either Mlh1-mlh3 mutant complex tested. The authors found that, in agreement with the *in vivo*

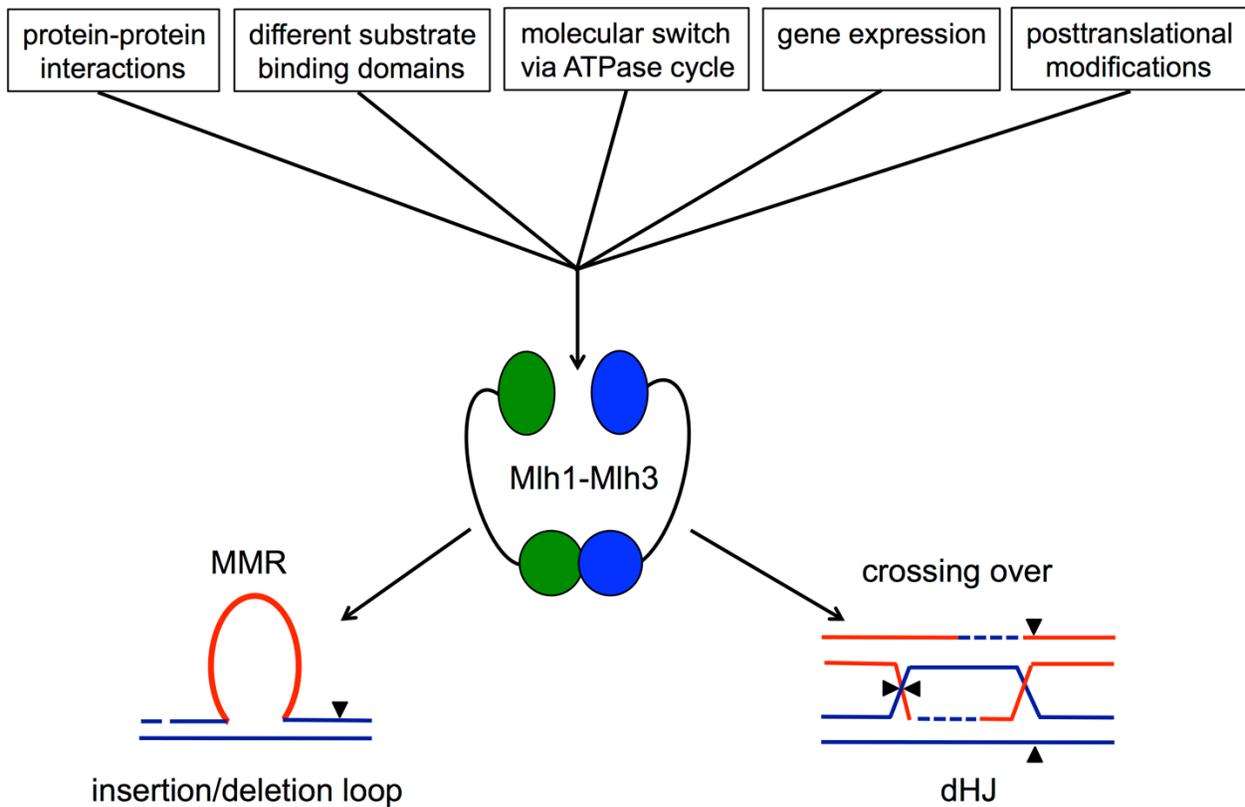


Figure 1.4. Mlh1-Mlh3 endonuclease can nick two different DNA substrates that form during different genomic stability pathways.

Factors that are hypothesized to regulate Mlh1-Mlh3 activity and subsequently direct and activate nicking of the MMR insertion/deletion substrate or the double Holliday junction (dHJ) substrate formed during meiotic recombination. Black arrowheads indicate nicking positions that would yield strand specific repair or crossover products.

phenotypes, the endonuclease activity of the MMR defective Mlh1-mlh3-45 complex was not stimulated by the MMR factor Msh2-Msh3, unlike the endonuclease activity of Mlh1-mlh3-32.

Al-Sweel et al. [40] suggest 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. Interestingly, *mlh3-45* maintained COs at nearly wild-type levels despite a possibly disrupted interaction with Msh4-Msh5. In addition, they identified two *mlh3* alleles that weakened dimerization with Mlh1. While this affected the MMR function of these two alleles, no significant defects in crossing over were observed. Thus, the authors observed that destabilized protein-protein interactions were tolerated in meiosis but not MMR. Taken together, these findings support a model in which pro-crossover factors act in concert to stabilize weakened interactions during meiotic CO resolution *in vivo*. The authors propose that Mlh1-Mlh3 is part of a resolvase complex at dHJs that regulates the endonuclease activity of Mlh1-Mlh3 to generate COs. In agreement with this, Ranjha et al. [42], based on their *in vitro* observations that Mlh1-Mlh3 prefers to bind open HJs, propose that Mlh1-Mlh3 *in vivo* acts as a complex with other recombination factors that facilitate its access to the junction.

Much remains to be understood about how biased resolution of dHJs in the ZMM pathway is implemented. A working model suggests that, in the steps leading to dHJ formation, the coordinated assembly of diverse protein factors at the two junctions imposes biased nicking (Figure 1.5). Consistent with this, a recent bioinformatics study identified a possible role for Zip3, a SUMO E3 ligase, in implementing asymmetric resolution [45]. Whole genome sequencing of recombination events in yeast spores derived from *zip3* mutants showed biased cutting was lost while surprisingly maintained in *msh4* mutants. Based on this and other findings, the authors propose that Msh4-Msh5 is required at the invading end of the DSB, to stabilize SEI

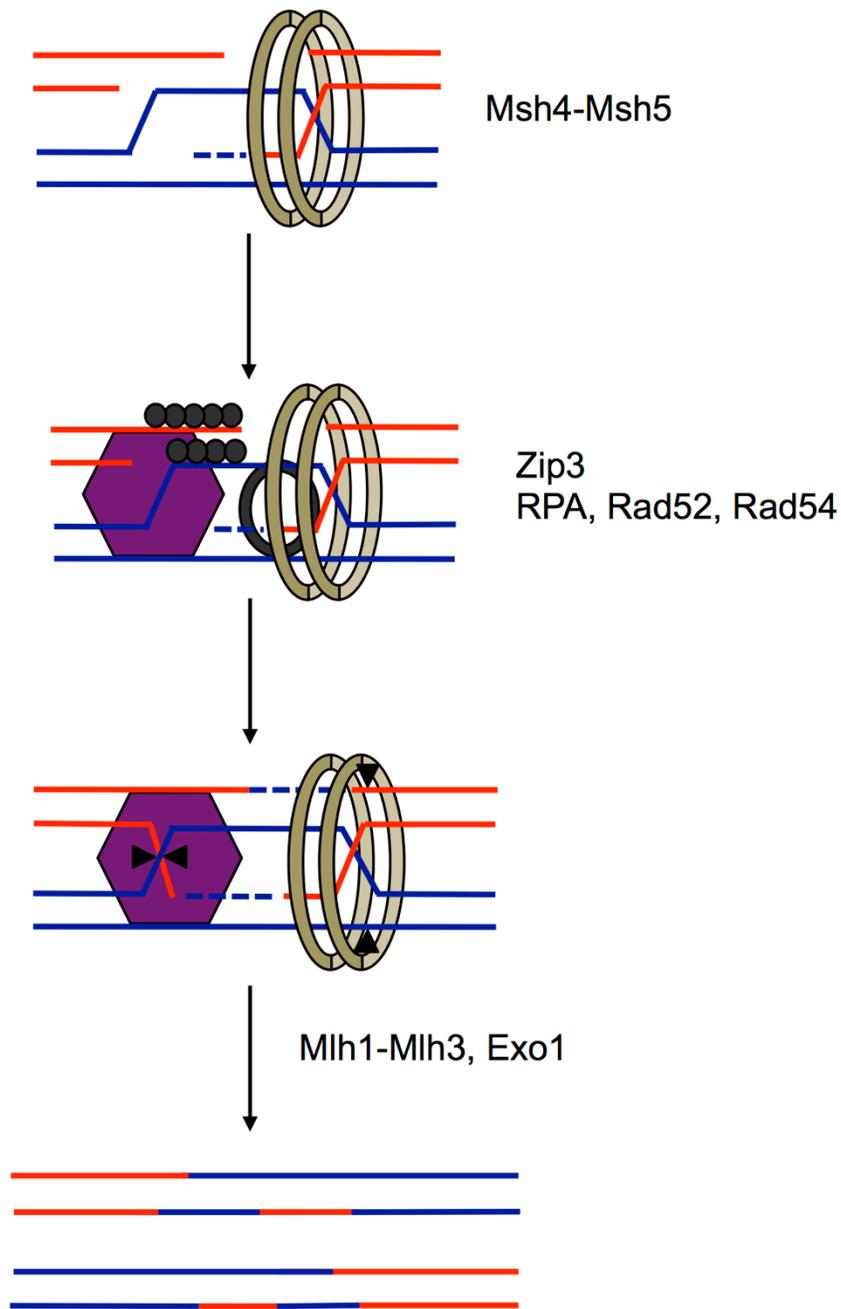


Figure 1.5. Distinct protein-protein interactions at the two junctions facilitate asymmetric nicking of the double Holliday junction (dHJ) by Mlh1-Mlh3.

Early strand invasion intermediates are acted upon by different factors as they mature to form the dHJ. Single end invasion intermediates (SEI) are stabilized by Msh4-Msh5 (brown) giving rise to the first junction. Second-end capture processed via RPA, Rad52, Rad54, (gray) and possibly Zip3 (purple) results in the formation of the second junction. This implies a sequential loading of diverse factors at the two junctions such that a symmetric DNA intermediate is stabilized by asymmetric protein factors. These factors recruit Mlh1-Mlh3 and possibly direct biased resolution of the dHJ to form exclusively crossover products. Model adapted from Manhart and Alani [31].

intermediates, while Zip3 acts to promote second-end capture at the ligating end. It is tempting to propose that pro-crossover factors, present in precise positions on the dHJ substrate, recruit and subsequently orient Mlh1-Mlh3's biased nicking activity. In support of this, cytological observations in mouse have shown that MLH1-MLH3 foci formation requires MSH4-MSH5 [46]. Furthermore, Zakharyevich et al. [21] observed pro-crossover functions for Exo1 independent of its exonuclease activity and dependent on its interaction with Mlh1-Mlh3.

Additional studies suggest that specific protein-protein interactions influence and direct MLH endonuclease activity during MMR. *In vitro* studies performed by Kadyrov et al. [6] showed that RFC-loaded PCNA can activate Mlh1-Pms1, but this was not observed for Mlh1-Mlh3 by Rogacheva et al. [41] or Ranjha et al. [42]. Mlh3 lacks a PCNA binding motif present in Pms1. Romanova and Crouse suggest that Mlh1-Mlh3 acts in conjunction with Mlh1-Pms1 in Msh2-Msh3 dependent MMR [10]. This observation helps address how Mlh1-Mlh3 is involved in MMR in the absence of a PCNA interaction. Mlh1-Mlh3 is recruited and activated by Msh2-Msh3, and also forms a complex with Mlh1-Pms1, which can be directed by PCNA to promote efficient repair [7]. Al-Sweel et al. [40] hypothesize that if dimerization between Mlh1 and Mlh3 is weakened, the ability to be recruited by Msh2-Msh3 and interact with Mlh1-Pms1 is likely inhibited; creating a defect in MMR that is minor because Mlh1-Pms1 is still active. For meiotic crossing over, a relatively slow process compared to MMR at the replication fork, a weakened dimer can be compensated for by interactions with other meiotic factors (e.g. Msh4-Msh5, Exo1 and STR). Alternatively, the defect observed in Mlh1-Mlh3 interaction is weak and may not be a limiting factor for CO resolution in meiosis. Taken together these observations provide further motivation to examine Mlh1-Mlh3 activity on recombination substrates *in vitro* in the presence of pro-crossover factors.

b. Different DNA-binding domains of Mlh1-Mlh3 exert substrate specificity

Claeys Bouuaert and Keeney [44] suggest that distinct DNA-binding surfaces in the ATPase domains and linker arms of Mlh1-Mlh3 determine substrate specificity and exert separable contributions to MMR or CO functions. The authors mapped the DNA binding interface of Mlh1-Mlh3 by hydroxyl radical footprinting, and found that most contacts with HJ substrates were located in the linkers. Whereas ssDNA contacts mapped within the ATPase domains in the N-terminus as well as the linkers. Next, the authors mutated conserved residues within these protein surfaces that are inferred to contact DNA directly and tested the effects of these mutations on substrate binding *in vitro* in addition to their effect on MMR and CO functions *in vivo*. Interestingly, they observed binding defects *in vitro* that globally agreed with the predictions of the footprinting assay. For example, the mutants in the Mlh1 linker had a stronger defect in HJ binding than mutants in the NTD. But this was not as obvious for Mlh3. Nonetheless, the effects observed on *in vivo* function suggested that the Mlh3 linker is more important for crossing over, further supporting their model. Importantly, when the authors plotted the *in vivo* phenotypes of these binding mutants on MMR versus crossing over, an intriguing pattern emerged. The mutants clustered based on their location in the N-terminal domain or linker region, indicating that particular domains within the protein convey differential contributions to MMR and crossing over.

c. Molecular switch modulated by an ATPase cycle

MLH proteins act as dimers and contain long unstructured linkers that connect the N- and C-terminal domains of each subunit. The linkers vary in length and are resistant to amino acid substitutions [47]. Sacho et al. [48] imaged, via AFM, asymmetric global conformational

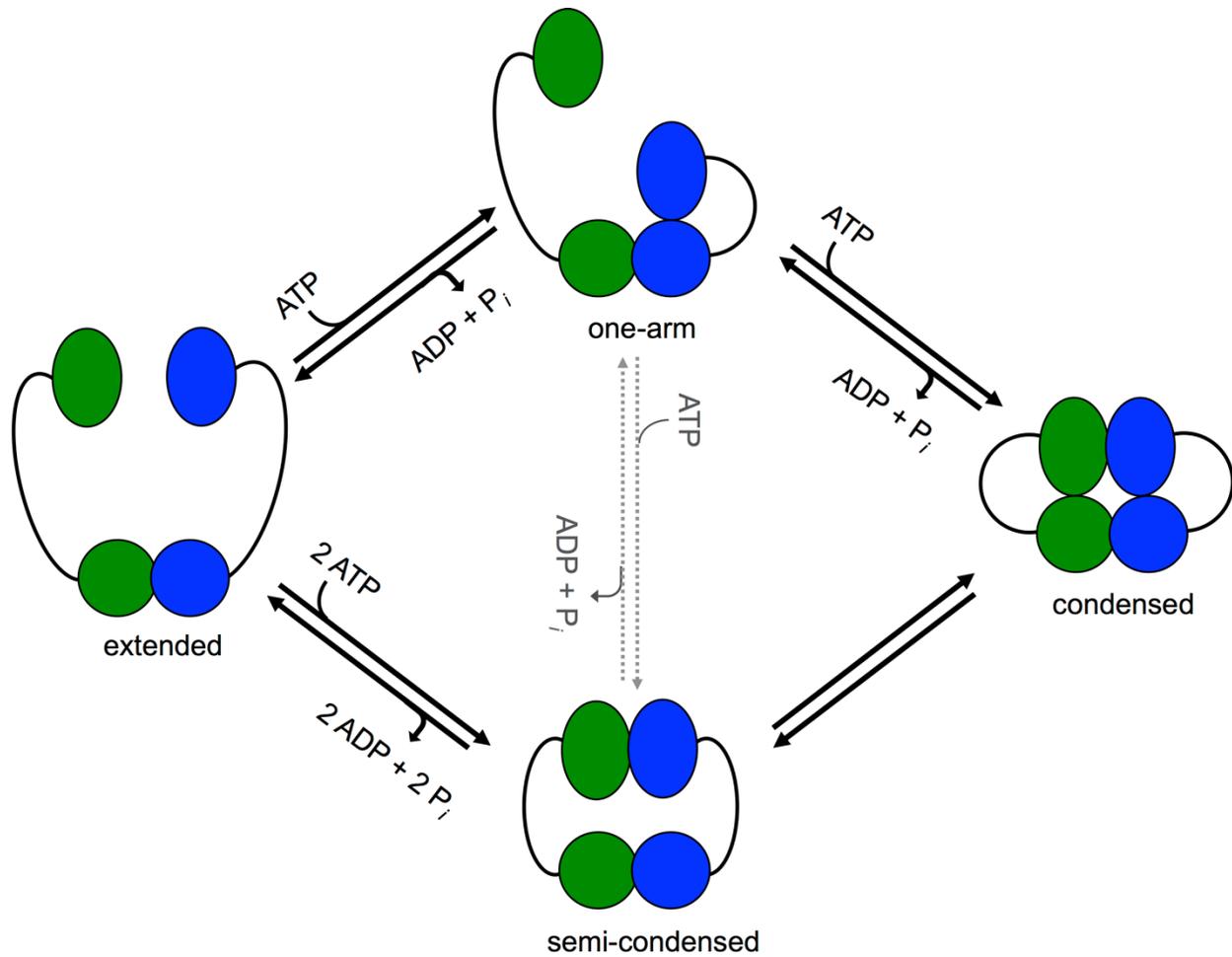


Figure 1.6. Mlh1-Mlh3 conformational changes mediated by an ATPase cycle.

Sequential ATP binding facilitates formation of the one-arm and condensed state respectively. Hydrolysis and release of both nucleotides, either sequentially (dashed arrows) or at the same time returns the heterodimer to the extended conformation. Model adapted from Sacho et al. [48] based on data from Claeys Bouuaert and Keeney [44].

changes in yeast Mlh1-Pms1 regulated by an ATP binding, hydrolysis, and release cycle. The linker arms, which switch between extended and condensed states, facilitate the conformational changes (Figure 1.6). The extended state, in which only the CTDs of the subunits interact, corresponds to no nucleotide bound. Sequential binding of ATP to each subunit forms the one-arm folded and the condensed state respectively. In the condensed state the NTDs of the subunits interact in addition to the CTDs. The authors hypothesize that these conformational changes expose different regions of the heterodimer to modulate diverse protein-protein or protein-DNA interactions in addition to mediating the timing of these interactions.

Extending these ideas to Mlh1-Mlh3, Claeys Bouuaert and Keeney [44] imaged the heterodimer by AFM and indeed observed the different configurations. They also observed an increase in the population of condensed particles upon the addition of ATP, at the expense of extended and one-arm folded particles. They conclude that, like Mlh1-Pms1, Mlh1-Mlh3 undergoes conformational changes, mediated by an ATPase cycle, that are likely critical in dictating diverse protein-protein or protein-DNA interactions (Figure 1.6). It is interesting to note that both Al-Sweel et al. [40] and Claeys Bouuaert and Keeney [44] identified two lysine residues in the Mlh3 linker arm (414K and 416K) that conferred a greater defect in crossing over than in MMR. It is tempting to speculate that mutating these residues obstructs the formation of a configuration that may be more important for facilitating interactions with recombination specific protein factors or substrates.

Cotton et al. [49] constructed and tested *mlh3* mutants predicted to disrupt either ATP binding (*mlh3-N35A*) or ATP hydrolysis (*mlh3-E31A*). They observed significant defects in MMR and crossing over, indistinguishable from *mlh3Δ*, in the binding mutant. In contrast, the hydrolysis mutant affected only the MMR function while COs were maintained at close to wild-

type levels. Thus, they suggest that ATP binding, and the accompanying conformational change, is essential for both MMR and crossing over. Whereas ATP hydrolysis by Mlh3 is important only in MMR and is dispensable for crossing over. Based on the model proposed by Sacho et al. [48], ATP binding is required for formation of the condensed state. ATP hydrolysis was proposed to be required for the subsequent release of ADP and P_i and recycling of the protein complex. Cotton et al. [49] hypothesize that the inability to hydrolyze ATP prevents Mlh1-Mlh3 from being free to undergo another round of ATP binding implying that the availability of Mlh1-Mlh3 is rate limiting only for crossing over. However, in contrast to Cotton et al., Sonntag Brown et al. [36] observed meiotic defects in the *mlh3* hydrolysis mutant. This is consistent with both ATP binding and hydrolysis being important for Mlh1-Mlh3 functions in MMR and crossing over.

It is important to note that Mlh1-Mlh3 exhibited a low ATPase activity *in vitro*, a few fold lower than that observed for Mlh1-Pms1. The ATPase activity was not significantly stimulated by DNA [41],[44]. Furthermore, addition of ATP does not promote the *in vitro* endonuclease activity of Mlh1-Mlh3 in contrast to Mlh1-Pms1 [41],[42].

d. Regulation of gene expression

Wang et al. [35] propose a model in which repair or recombination substrates are channeled into the appropriate pathway through differential expression of Pms1 or Mlh3. Kramer et al. [50] and Chu et al. [51] demonstrate that in *Saccharomyces cerevisiae* the transcript level of *MLH1*, the common component in Mlh complexes, appears to be constant throughout mitotic and meiotic cell cycles, while both *PMS1* and *MLH3* transcript levels are cell cycle regulated. Transcription of *PMS1* is induced around the time of DNA replication in vegetative and meiotic

cells in agreement with a role in MMR. Whereas, transcription of *MLH3* is up-regulated during meiosis, somewhat later than *PMS1*, in agreement with a primary role in recombination required after bulk DNA replication [50]–[52]. Based on these observations, Wang et al. [35] hypothesize that the relative abundance of Mlh complexes is regulated via programmed changes in gene expression such that differential expression of Pms1 and Mlh3 could promote preferential formation of specific Mlh1-containing heterocomplexes.

e. Posttranslational modifications

Matos et al. [53] demonstrate that the endonuclease activities of the SSNs Mus81-Mms4 and Yen1 are controlled by their phosphorylation states. The authors hypothesize that this regulatory system directs joint molecule resolution by timing the actions of resolvases according to cellular needs. They observed that Mus81-Mms4 is hyperactivated by the Polo-like kinase Cdc-5 at the onset of meiosis I, to generate some of the COs necessary for chromosome segregation, while Yen1 is activated by dephosphorylation at the onset of meiosis II to resolve persistent HJs. Manhart et al. [43] did not observe an effect on the *in vitro* endonuclease activity of Mlh1-Mlh3 upon phosphatase or kinase treatment, suggesting that it is unlikely to be regulated by phosphorylation. The authors do not rule out the possibility that other posttranslational modifications, such as SUMOylation, may play an as yet unknown role in controlling Mlh1-Mlh3's activity. It is important to note that Qiao et al. [54] observed antagonistic roles for the ubiquitin ligase HEI10 and the SUMO ligase RNF212, mammalian ortholog of Zip3, that are balanced during meiotic recombination in mice to effect differential stabilization of recombination factors at CO and NCO sites.

Unanticipated observations infer diverse Mlh1-Mlh3 functions in meiotic recombination

Al-Sweel et al. [40] propose that *mlh3* mutants can alter the fate of meiotic recombination intermediates. The authors utilized high-resolution recombination maps to characterize wild-type, *mlh3Δ*, and three *mlh3* alleles; the endonuclease defective mutant, *mlh3-D523N*, in addition to two alleles that conferred stronger defects in one function relative to the other (MMR or crossing over). The *mlh3* point and null mutants displayed genome-wide increases in NCO events as compared to wild-type strains. They hypothesize that in mutants defective in Mlh1-Mlh3 functions, dHJs formed in the ZMM pathway become susceptible to the actions of the SSNs to form NCOs and COs indiscriminately. dHJs, in these mutant backgrounds, could also be acted upon by the STR complex through convergent migration of the two junctions until a hemicatenane can be removed by the topoisomerase. The authors propose a model in which Mlh1-Mlh3 acts in concert with the ZMM proteins to protect recombination intermediates from the actions of STR. They speculate that these observations provide new insights for how meiotic recombination pathways are regulated. In agreement with this, Zakharyevich et al. [23] propose that Mlh1-Mlh3 is required for Sgs1-dependent crossing over but is dispensable for efficient joint molecule resolution and formation of NCOs. The authors observed similar reductions in crossing over upon deletion of either Sgs1 or Mlh3 in yeast strains lacking the SSNs (*mms4 slx4 yen1*). However deletion of Mlh3, in contrast to Sgs1, did not affect NCOs or joint molecule resolution.

Claeys Bouuaert and Keeney [44] observed similarly decreased CO levels for *mlh1Δ* and *mlh3Δ* strains, as observed previously. Therefore, given the importance of crossing over in promoting accurate chromosome segregation, the authors anticipated to find similar chromosome segregation defects in *mlh1Δ* as compared to *mlh3Δ*. Surprisingly meiosis I nondisjunction was 2

to 3-fold more frequent in *mlh3Δ*. They speculate that in the absence of Mlh3 joint molecules are taken apart too early to facilitate chromosome biorientation, suggesting a role for Mlh3 in stabilizing joint molecules. Alternatively, in the absence of Mlh3, joint molecules may persist too long thereby interfering with chromosome separation. The authors do not rule out the possibility that the different contributions of Mlh1 and Mlh3 to meiotic recombination reflect the different complexes they inhabit. Mlh1 is a common subunit of three heterodimers, while Mlh3 is known to interact only with Mlh1. For example, Mlh1 interacts with Mlh2 to regulate GC tract lengths. Thus, the authors propose that the more severe *mlh3Δ* nondisjunction phenotype could arise from there no longer being a competition between Mlh3 and other proteins for binding Mlh1. Taken together, the authors conclude that these observations draw a more complex picture of the role of Mlh1-Mlh3 in meiotic recombination. It is important to note that the relative increase in homolog nondisjunction in *mlh3Δ* as compared to *mlh1Δ*, reported in this study, was not observed in studies, described previously, based on spore viability measurements [35]. This discrepancy may be a result of chromosome-specific effects, as spore viability reflects segregation of all chromosomes, whereas in this study missegregation of only chromosome VIII was measured.

Claeys Bouuaert and Keeney [44] also observed an unexpected preference of Mlh1-Mlh3 to bind ssDNA. This is surprising because canonical DSB repair models do not predict that ssDNA will be present at the stage of the dHJ intermediate and Mlh1-Mlh3 involvement. The authors hypothesize that Mlh1-Mlh3 may play an earlier role than typically implicated. They propose that Mlh1-Mlh3 may be recruited to early strand invasion intermediates and initially binds ssDNA and then probes for nearby branched structures. In support of this, immunolocalization studies in mouse spermatocytes performed by Lipkin et al. [37] showed that Mlh3 foci appeared in early to mid pachynema consistently without Mlh1. Moreover, the

absence of Mlh3 results in a failure to load Mlh1. The authors conclude that preloading of Mlh3 is required for binding of Mlh1 to sites of crossing over. Alternatively, the authors speculate that Mlh3 can homodimerize in early pachynema suggesting more diverse Mlh1-Mlh3 functions in meiotic recombination than previously proposed.

Conclusions

Mlh1-Mlh3 is part of the MMR family of proteins yet its endonuclease activity is utilized to provide a major contribution in meiotic recombination. Much remains to be determined regarding Mlh1-Mlh3 function at a mechanistic level. How one endonuclease can recognize and bind two different DNA substrates and introduce precise nicks is intriguing. I utilized a genetic approach to answer this question. In Chapter 2, I investigated the *in vivo* effect of 60 new *mlh3* mutants on both MMR and crossing over in *Saccharomyces cerevisiae*. This provides a structure function map of Mlh3 from which I identified six alleles that conferred stronger defects in one function vs. the other. Thus, I conclude the two functions of Mlh1-Mlh3 are genetically separable. These mutations were validated by different methods that measure frequency of crossing over. In addition, I identified the putative endonuclease active site in Mlh1-Mlh3 based on homology with Mlh1-Pms1. Further characterization of the CTD separation of function mutants by yeast-two hybrid revealed that a weakened interaction between Mlh1-Mlh3 is tolerable in meiosis but not MMR. Moreover, based on *in vitro* biochemical interactions between Msh2-Msh3 and an MMR deficient, CO proficient allele, I hypothesize that weakened MSH-MLH interactions are also permitted in meiosis more than MMR. These observations lend support to the model that Mlh1-Mlh3 acts to promote COs as part of a larger resolvase complex. Finally, whole genome maps of recombination events in tetrads derived from null and point *mlh3*

mutants revealed that defective Mlh1-mlh3 complexes can alter the fate of meiotic recombination intermediates, potentially through the dissolution activity of the STR complex. Thus, I hypothesize that Mlh1-Mlh3 acts to protect recombination intermediates from the actions of STR. In agreement with this, upon overexpression of the Sgs1 helicase, I observed a greater decrease in the spore viability of *mlh3Δ* strains but not in mutant strains where a structurally intact Mlh1-mlh3 complex is thought to be present. In Chapter 3, I discuss immediate future plans, based on these findings, in addition to remaining questions in the field and future projects.

References

- [1] T. A. Kunkel and D. A. Erie, “Eukaryotic Mismatch Repair in Relation to DNA Replication,” pp. 291–313.
- [2] G. X. Reyes, T. T. Schmidt, R. D. Kolodner, and H. Hombauer, “New insights into the mechanism of DNA mismatch repair,” *Chromosoma*, vol. 124, no. 4, pp. 443–462, 2015.
- [3] R. Fishel, “Mismatch repair,” *J. Biol. Chem.*, vol. 290, no. 44, pp. 26395–26403, 2015.
- [4] R. D. Kolodner, “A personal historical view of DNA mismatch repair with an emphasis on eukaryotic DNA mismatch repair,” *DNA Repair (Amst)*., vol. 38, pp. 3–13, 2016.
- [5] F. A Kadyrov, L. Dzantiev, N. Constantin, and P. Modrich, “Endonucleolytic function of MutLalpha in human mismatch repair,” *Cell*, vol. 126, no. 2, pp. 297–308, Jul. 2006.
- [6] F. A Kadyrov, S. F. Holmes, M. E. Arana, O. A Lukianova, M. O’Donnell, T. a Kunkel, and P. Modrich, “*Saccharomyces cerevisiae* MutLalpha is a mismatch repair endonuclease,” *J. Biol. Chem.*, vol. 282, no. 51, pp. 37181–90, Dec. 2007.
- [7] A. Pluciennik, L. Dzantiev, R. R. Iyer, N. Constantin, F. A. Kadyrov, and P. Modrich, “PCNA function in the activation and strand direction of MutL α endonuclease in

- mismatch repair.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 37, pp. 16066–71, 2010.
- [8] H. Flores-Rozas and R. D. Kolodner, “The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 21, pp. 12404–9, Oct. 1998.
- [9] B. D. Harfe, B. K. Minesinger, and S. Jinks-Robertson, “Discrete in vivo roles for the MutL homologs Mlh2p and Mlh3p in the removal of frameshift intermediates in budding yeast.,” *Curr. Biol.*, vol. 10, no. 3, pp. 145–8, Mar. 2000.
- [10] N. V Romanova and G. F. Crouse, “Different roles of eukaryotic MutS and MutL complexes in repair of small insertion and deletion loops in yeast.,” *PLoS Genet.*, vol. 9, no. 10, p. e1003920, Oct. 2013.
- [11] M. D. Champion and R. S. Hawley, “Playing for half the deck: the molecular biology of meiosis.,” *Nat. Cell Biol.*, vol. 4 Suppl, pp. s50-6, Oct. 2002.
- [12] T. Hassold and P. Hunt, “To err (meiotically) is human: the genesis of human aneuploidy.,” *Nat. Rev. Genet.*, vol. 2, no. 4, pp. 280–91, Apr. 2001.
- [13] J. W. Szostak, T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl, “The double-strand-break repair model for recombination.,” *Cell*, vol. 33, no. 1, pp. 25–35, May 1983.
- [14] N. Hunter, “Meiotic recombination,” *Top. Curr. Genet.*, vol. 17, no. 2, pp. 381–442, 2007.
- [15] L. A. Gilbertson and F. W. Stahl, “A Test of the Double-Strand Break Repair Model for Meiotic Recombination in,” 1996.
- [16] A. De Muyt, L. Jessop, E. Kolar, A. Sourirajan, J. Chen, Y. Dayani, and M. Lichten, “BLM helicase ortholog Sgs1 is a central regulator of meiotic recombination intermediate metabolism.,” *Mol. Cell*, vol. 46, no. 1, pp. 43–53, Apr. 2012.
- [17] H. Kaur, A. De Muyt, and M. Lichten, “Top3-Rmi1 DNA Single-Strand Decatenase Is

- Integral to the Formation and Resolution of Meiotic Recombination Intermediates,” *Mol. Cell*, vol. 57, no. 4, pp. 583–594, 2015.
- [18] S. Tang, M. K. Y. Wu, R. Zhang, and N. Hunter, “Pervasive and Essential Roles of the Top3-Rmi1 Decatenase Orchestrate Recombination and Facilitate Chromosome Segregation in Meiosis,” *Mol. Cell*, vol. 57, no. 4, pp. 607–621, 2015.
- [19] L. Wu and I. D. Hickson, “The Bloom’s syndrome helicase suppresses crossing over during homologous recombination.,” *Nature*, vol. 426, no. 6968, pp. 870–874, 2003.
- [20] T. Snowden, S. Acharya, C. Butz, M. Berardini, and R. Fishel, “hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes.,” *Mol. Cell*, vol. 15, no. 3, pp. 437–51, Aug. 2004.
- [21] K. Zakharyevich, Y. Ma, S. Tang, P. Y. H. Hwang, S. Boiteux, and N. Hunter, “Temporally and Biochemically Distinct Activities of Exo1 during Meiosis: Double-Strand Break Resection and Resolution of Double Holliday Junctions,” *Mol. Cell*, vol. 40, no. 6, pp. 1001–1015, 2010.
- [22] S. Gray and P. E. Cohen, “Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation,” *Annu. Rev. Genet.*, vol. 50, no. 1, p. annurev-genet-120215-035111, 2016.
- [23] K. Zakharyevich, S. Tang, Y. Ma, and N. Hunter, “Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase.,” *Cell*, vol. 149, no. 2, pp. 334–47, Apr. 2012.
- [24] J. L. Argueso, J. Wanat, Z. Gemici, and E. Alani, “Competing crossover pathways act during meiosis in *Saccharomyces cerevisiae*.,” *Genetics*, vol. 168, no. 4, pp. 1805–16, Dec. 2004.

- [25] J. A. Surtees, J. L. Argueso, and E. Alani, “Mismatch repair proteins: Key regulators of genetic recombination,” *Cytogenet. Genome Res.*, vol. 107, no. 3–4, pp. 146–159, 2004.
- [26] J.-M. Chen, D. Cooper, N. Chuzhanova, C. Férec, and G. P. Patrinos, “Gene conversion: mechanisms, evolution and human disease.,” *Nat. Rev. Genet.*, vol. 8, no. 10, pp. 762–75, 2007.
- [27] U. Chakraborty and E. Alani, “Understanding how mismatch repair proteins participate in the repair/anti-recombination decision,” *FEMS Yeast Res.*, vol. 16, no. 6, pp. 1–12, 2016.
- [28] K. C. Tham, R. Kanaar, and J. H. G. Lebbink, “Mismatch repair and homeologous recombination,” *DNA Repair (Amst.)*, vol. 38, pp. 75–83, 2016.
- [29] Y. Duroc, R. Kumar, L. Ranjha, C. Adam, R. Guérois, K. Md Muntaz, M.-C. Marsolier-Kergoat, F. Dingli, R. Laureau, D. Loew, B. Llorente, J.-B. Charbonnier, P. Cejka, and V. Borde, “Concerted action of the MutL β heterodimer and Mer3 helicase regulates the global extent of meiotic gene conversion,” *Elife*, vol. 6, pp. 1–24, 2017.
- [30] L. E. Jensen, P. A. Jauert, and D. T. Kirkpatrick, “The large loop repair and mismatch repair pathways of *Saccharomyces cerevisiae* act on distinct substrates during meiosis,” *Genetics*, vol. 170, no. 3, pp. 1033–1043, 2005.
- [31] C. M. Manhart and E. Alani, “Roles for mismatch repair family proteins in promoting meiotic crossing over,” *DNA Repair (Amst.)*, vol. 38, pp. 84–93, 2015.
- [32] N. Hunter and R. H. Borts, “Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis.,” *Genes Dev.*, vol. 11, no. 12, pp. 1573–1582, Jun. 1997.
- [33] G. S. Roeder, “Mutation of a Meiosis-Specific MutS decreases Crossing Over Not Mismatch Correction,” vol. 79, pp. 1069–1080, 1994.

- [34] N. M. Hollingsworth, L. Ponte, and C. Halsey, “MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair.,” *Genes Dev.*, vol. 9, no. 14, pp. 1728–1739, Jul. 1995.
- [35] T. F. Wang, N. Kleckner, and N. Hunter, “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.*, vol. 96, no. 24, pp. 13914–9, Nov. 1999.
- [36] M. S. Brown, E. Lim, C. Chen, K. T. Nishant, and E. Alani, “Genetic Analysis of *mlh3* Mutations Reveals Interactions Between Crossover Promoting Factors During Meiosis in Baker’s Yeast.,” *G3 (Bethesda)*, vol. 3, no. 1, pp. 9–22, Jan. 2013.
- [37] S. M. Lipkin, P. B. Moens, V. Wang, M. Lenzi, D. Shanmugarajah, A. Gilgeous, J. Thomas, J. Cheng, J. W. Touchman, E. D. Green, P. Schwartzberg, F. S. Collins, and P. E. Cohen, “Meiotic arrest and aneuploidy in MLH3-deficient mice.,” *Nat. Genet.*, vol. 31, no. 4, pp. 385–90, Aug. 2002.
- [38] E. Gueneau, C. Dherin, P. Legrand, C. Tellier-Lebegue, B. Gilquin, P. Bonnesoeur, F. Londino, C. Quemener, M.-H. Le Du, J. a Márquez, M. Moutiez, M. Gondry, S. Boiteux, and J.-B. Charbonnier, “Structure of the MutL α C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site.,” *Nat. Struct. Mol. Biol.*, vol. 20, no. 4, pp. 461–8, 2013.
- [39] K. T. Nishant, A. J. Plys, and E. Alani, “A mutation in the putative MLH3 endonuclease domain confers a defect in both mismatch repair and meiosis in *Saccharomyces cerevisiae*.,” *Genetics*, vol. 179, no. 2, pp. 747–55, Jun. 2008.
- [40] N. Al-Sweel, V. Raghavan, A. Dutta, V. P. Ajith, L. Di Vietro, K. Nabila, C. M. Manhart,

- J. A. Surtees, K. T. Nishant, and E. Alani, “mlh3 mutations in baker’s yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide,” *PLoS Genet.*, p. In press, 2017.
- [41] M. V Rogacheva, C. M. Manhart, C. Chen, A. Guarne, J. Surtees, and E. Alani, “Mlh1-Mlh3, a meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated endonuclease.,” *J. Biol. Chem.*, vol. 289, no. 9, pp. 5664–73, Feb. 2014.
- [42] L. Ranjha, R. Anand, and P. Cejka, “The *Saccharomyces cerevisiae* Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions.,” *J. Biol. Chem.*, vol. 289, no. 9, pp. 5674–86, Mar. 2014.
- [43] C. M. Manhart, X. Ni, M. A. White, J. Ortega, J. A. Surtees, and E. Alani, “The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans,” *PLoS Biol.*, p. In press, 2017.
- [44] C. C. Bouuaert and S. Keeney, “Distinct DNA-binding surfaces in the ATPase and linker domains of MutL γ determine its substrate specificities and exert separable functions in meiotic recombination and mismatch repair,” *PLoS Genet.*, pp. 1–32, 2017.
- [45] A. Oke, C. M. Anderson, P. Yam, and J. C. Fung, “Controlling meiotic recombinational repair - specifying the roles of ZMMs, Sgs1 and Mus81/Mms4 in crossover formation.,” *PLoS Genet.*, vol. 10, no. 10, p. e1004690, Oct. 2014.
- [46] N. K. Kolas, A. Svetlanov, M. L. Lenzi, F. P. Macaluso, S. M. Lipkin, R. M. Liskay, J. Greally, W. Edelmann, and P. E. Cohen, “Localization of MMR proteins on meiotic chromosomes in mice indicates distinct functions during prophase I.,” *J. Cell Biol.*, vol. 171, no. 3, pp. 447–58, Nov. 2005.
- [47] J. L. Argueso, A. W. Kijas, S. Sarin, J. Heck, M. Waase, and E. Alani, “Systematic

- Mutagenesis of the *Saccharomyces cerevisiae* MLH1 Gene Reveals Distinct Roles for Mlh1p in Meiotic Crossing Over and in Vegetative and Meiotic Mismatch Repair,” vol. 23, no. 3, pp. 873–886, 2003.
- [48] E. J. Sacho, F. A. Kadyrov, P. Modrich, T. A. Kunkel, and D. A. Erie, “Direct visualization of asymmetric adenine-nucleotide-induced conformational changes in MutL alpha.,” *Mol. Cell*, vol. 29, no. 1, pp. 112–21, Jan. 2008.
- [49] V. E. Cotton, E. R. Hoffmann, and R. H. Borts, “Distinct regulation of Mlh1p heterodimers in meiosis and mitosis in *Saccharomyces cerevisiae*.,” *Genetics*, vol. 185, no. 2, pp. 459–67, Jul. 2010.
- [50] W. Kramer, B. Fartmann, and E. C. Ringbeck, “Transcription of *mutS* and *mutL*-homologous genes in *Saccharomyces cerevisiae* during the cell cycle.,” *Mol. Gen. Genet.*, vol. 252, no. 3, pp. 275–83, 1996.
- [51] S. Chu, J. DeRisi, M. Eisen, J. Mulholland, D. Botstein, P. O. Brown, and I. Herskowitz, “The transcriptional program of sporulation in budding yeast.,” *Science*, vol. 282, no. 5389, pp. 699–705, Oct. 1998.
- [52] A. Morrison, A. L. Johnson, L. H. Johnston, and A. Sugino, “Pathway correcting DNA replication errors in *Saccharomyces cerevisiae*.,” *EMBO J.*, vol. 12, no. 4, pp. 1467–1473, 1993.
- [53] J. Matos, M. G. Blanco, S. Maslen, J. M. Skehel, and S. C. West, “Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis.,” *Cell*, vol. 147, no. 1, pp. 158–72, Sep. 2011.
- [54] H. Qiao, H. B. D. P. Rao, Y. Yang, J. H. Fong, J. M. Cloutier, D. C. Deacon, K. E. Nagel, R. K. Swartz, E. Strong, J. K. Holloway, P. E. Cohen, J. Schimenti, J. Ward, and N.

Hunter, “Antagonistic roles of ubiquitin ligase HEI10 and SUMO ligase RNF212 regulate meiotic recombination,” *Nat. Publ. Gr.*, vol. 46, no. 2, pp. 194–199, 2014.

CHAPTER 2

***mlh3* mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide**

Najla Al-Sweel¹, Vandana Raghavan¹, Abhishek Dutta², V. P. Ajith², Luigi Di Vietro^{3,5}, Nabila Khondakar¹, Carol M. Manhart¹, Jennifer A. Surtees⁴, K. T. Nishant^{2,6}, and Eric Alani¹

¹ Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA

² School of Biology, and ⁶Center for Computation Modelling and Simulation, Indian Institute of Science Education and Research Thiruvananthapuram, Trivandrum 695016, India

³ Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy.

⁴ Department of Biochemistry, University at Buffalo, State University of New York, Buffalo, NY, USA

This chapter was published in PLOS Genetics on 8/21/17:

Al-Sweel N, Raghavan V, Dutta A, Ajith VP, Di Vietro L, Khondakar N, et al. (2017) *mlh3* mutations in baker's yeast alter meiotic recombination outcomes by increasing noncrossover events genome-wide. PLoS Genet 13(8): e1006974.

Contributions: V.R. performed the *in vitro* biochemical assays. A.D., V.P.A., and K.T.N. provided the whole genome recombination maps. L.D.V. and N.K. assisted with the *lys2-A₁₄* reversion assays. J.A.S. provided purified Msh2-Msh3. E.A. performed the yeast two-hybrid experiment. C.M.M., J.A.S., K.T.N., and E.A. contributed to interpretation and editing. All other work performed by N.A.S.

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. Whole genome recombination maps were constructed for wild-type and MMR⁺ crossover⁻, MMR⁻ crossover⁺, endonuclease defective and null *mlh3* mutants in an S288c/YJM789 hybrid background. Compared to wild-type, all of the *mlh3* mutants showed increases in the number of noncrossover events, consistent with recombination intermediates being resolved through alternative recombination pathways. Our observations provide a structure-function map for Mlh3 that reveals the importance of protein-protein interactions in regulating Mlh1-Mlh3's enzymatic activity. They also illustrate how defective meiotic components can alter the fate of meiotic recombination intermediates, providing new insights for how meiotic recombination pathways are regulated.

Author Summary

During meiosis, diploid germ cells that will become eggs or sperm undergo a single round of DNA replication followed by two consecutive chromosomal divisions. The segregation

of chromosomes at the first meiotic division is dependent in most organisms on at least one genetic exchange, or crossover event, between chromosome homologs. Homologs that do not receive a crossover frequently undergo nondisjunction at the first meiotic division, yielding gametes that lack chromosomes or contain additional copies. Such events have been linked to human disease and infertility. Recent studies suggest that the Mlh1-Mlh3 complex is an endonuclease that resolves recombination intermediates into crossovers. Interestingly, this complex also acts as a matchmaker in DNA mismatch repair (MMR) to remove DNA replication errors. How does one complex act in two different processes? We investigated this question by performing a mutational analysis of the baker's yeast Mlh3 protein. Five mutations were identified that disrupted MMR but not crossing over, and one mutation disrupted crossing over while maintaining MMR. Using a combination of biochemical and genetic analyses to further characterize these mutants we illustrate the importance of protein-protein interactions for Mlh1-Mlh3's activity. Importantly, our data illustrate how defective meiotic components can alter the outcome of meiotic recombination events. They also provide new insights for the basis of infertility syndromes.

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 [1]. 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 [1]. 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 [2-4]. However, Mlh1-Mlh3 has been shown to play a major role in meiotic crossing over [5-8]. Accurate chromosome segregation in Meiosis I in most eukaryotes requires reciprocal exchange of genetic information (crossing over) between homologs [9-12]. 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 [13].

Yeast Mlh1-Pms1 and its human ortholog MLH1-PMS2 both exhibit an endonuclease activity that is essential for MMR [14-15]. This activity is dependent on the integrity of a highly conserved (DQHA(X)₂E(X)₄E) metal binding motif also found in Mlh3. Previous work demonstrated that a point mutation within this motif (*mlh3-D523N*) conferred *mlh3Δ*-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 [5]. Consistent with these observations, Mlh1-Mlh3 is an endonuclease that nicks circular duplex DNA *in vitro*, and Mlh1-*mlh3-D523N* is defective in endonuclease activity [16-17].

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; [18-23]). In this pathway a DSB, which forms on one chromatid of a homologous pair, is resected by 5' to 3' exonucleases, resulting in the formation of 3' single-strand tails on both sides of the DSB (Figure 2.1). One of these tails invades the other unbroken homolog and is extended and stabilized to create a single-end invasion intermediate (SEI). A second invasion event initiating from the SEI,

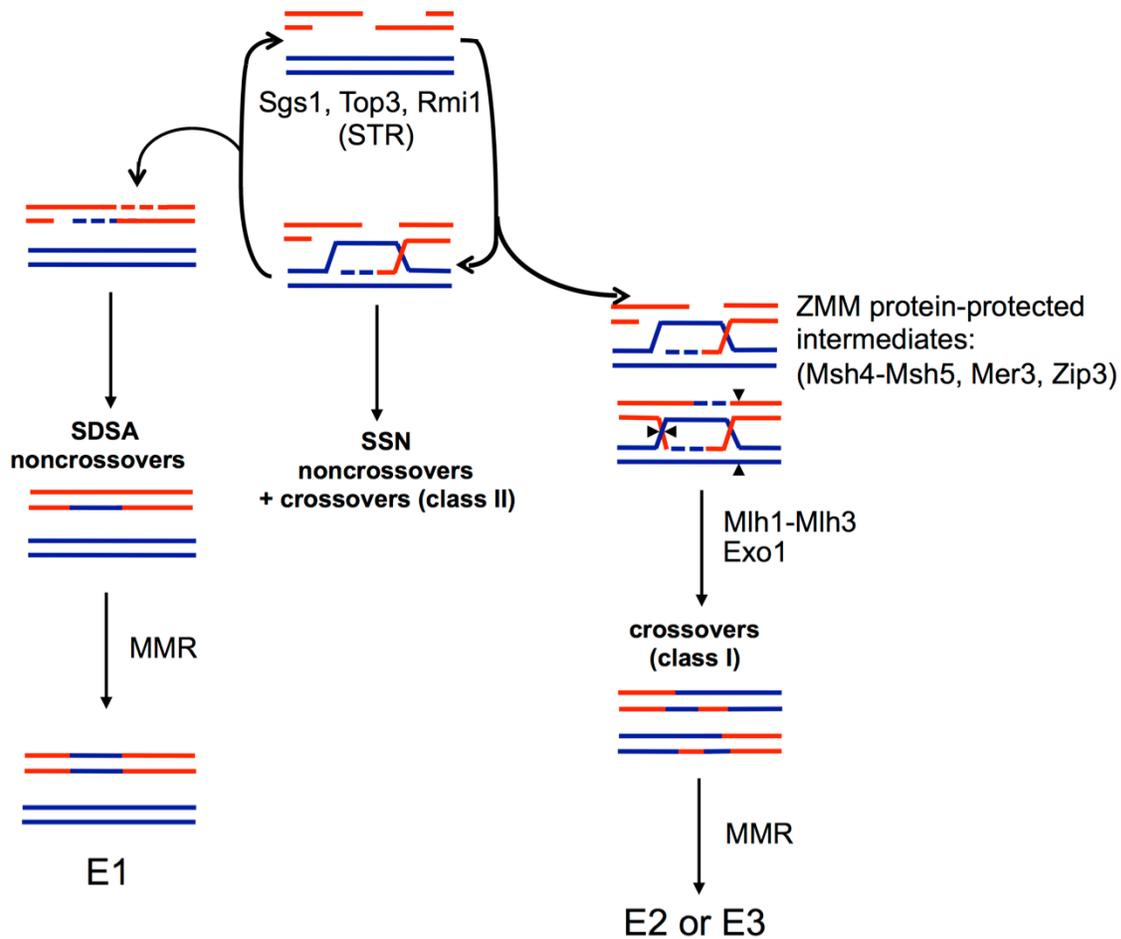


Figure 2.1. DSB repair pathways in meiosis.

Model adapted from Kaur et al. [30] depicting wild-type meiosis and the central role of the STR complex (Sgs1-Top3-Rmi1 helicase/topoisomerase) in disassembling strand invasion intermediates to facilitate synthesis dependent strand annealing (SDSA) or return of events to the original DSB state to allow capture and protection by the ZMM proteins and dHJ formation for ultimate resolution as class I crossovers by Mlh1-Mlh3 and Exo1. Events that escape STR disassembly form unregulated joint molecules that are resolved by the structure selective nucleases (SSNs-Mus81-Mms4, Yen1, Slx-Slx4) as noncrossovers or class II crossovers. The “E” classification of recombination classes was described in Oke et al. [34]. The majority event classes are presented here and result from MMR of heteroduplex DNA intermediates. E1 events are simple noncrossovers (NCO), E2 are simple crossovers (CO) with or without continuous gene conversion, and E3 are COs with discontinuous gene conversion. A set of definitions for these classes can also be found in Figure 2.6.

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 evenly spaced 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 [8, 24-31]. These steps are biased to resolve the two junctions present in the dHJ so that the resulting product is exclusively a CO. A second interference-independent pathway was identified that accounts for a small (~10%) number of CO events (class II COs). In this pathway, which does not involve the ZMM proteins, the two junctions are resolved independently by the Mus81-Mms4 endonuclease, leading to a mixture of CO and NCO products [7, 8, 32, 33].

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, 35, 36]. 2. There is significant redundancy of factors required to resolve dHJs into COs. This redundancy involves the endonucleases Mlh1-Mlh3, Mus81-Mms4, Yen1, and Slx1-Slx4 [5, 7, 8], 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 [8]. 3. MLH1 and MLH3 play critical roles in mammalian meiosis [37, 38]. 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 [37, 39].

Much remains to be understood about how biased resolution of dHJs in the interference-dependent pathway is achieved. A working model, supported by genetic and molecular studies outlined below, 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 [5-8, 16, 29, 30, 31, 36, 40-46]. 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. A recent bioinformatics study by the Fung group, which involved the analysis of CO-associated gene conversion patterns in yeast tetrads, suggested that Zip3, a SUMO E3 ligase, is required for biased cleavage [34]. Curiously, they found that biased resolution of dHJs was maintained in *msh4* mutants. Based on these findings and other observations they propose 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 [34]. 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, 43]. Furthermore, cytological observations in mouse have shown that MSH4-MSH5 foci appear prior to MLH1-MLH3 [37, 44, 47, 48]. Consistent with these observations, MLH1 and MLH3 foci formation requires MSH4-MSH5 [48].

Additional support for the above model was obtained from analysis of the STR complex [8, 30, 31, 46, 49]. The STR complex has recently been labeled the master regulator of meiotic

DSB repair, acting as both a positive and negative CO coordinator (Figure 2.1; [30, 49]). Initially, the Sgs1 helicase was characterized as anti-CO because it facilitates unwinding of DSB repair intermediates to promote NCOs via synthesis-dependent strand annealing (SDSA). However, deleting either Sgs1 or Mlh3 in yeast strains that lack all other meiotic resolvases (*mms4*, *slx4*, *yen1*) results in a similar reduction of CO levels (~10% of wild-type levels) suggesting a pathway where Sgs1-dependent COs require Mlh1-Mlh3 [8]. Similar results were observed in *mms4*, *slx4*, *yen1* strains deficient in Top3 or Rmi1 [30, 31]. These data indicate that the STR complex promotes the majority of COs in conjunction with a resolvase that is not Mus81-Mms4, Slx1-Slx4 or Yen1.

A role for Exo1 in crossing over is supported by genetic studies that show Exo1 and Mlh3 acting in the same CO pathway [29]. 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 [29]. 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 [50].

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. Importantly, our high-resolution recombination mapping of these mutants revealed ways in which defective

meiotic components can alter the fate of meiotic recombination intermediates.

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 [51]. We performed a clustered charged-to-alanine scanning mutagenesis of the *S. cerevisiae MLH3* gene to create 60 *mlh3* variants (Figure 2.2; Tables 2.1-2.3). 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. This unbiased mutagenesis has been successfully applied to study the functional domains of several proteins [52, 53], and has provided a comprehensive view of the functional organization of *MLH1* [54]. 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.

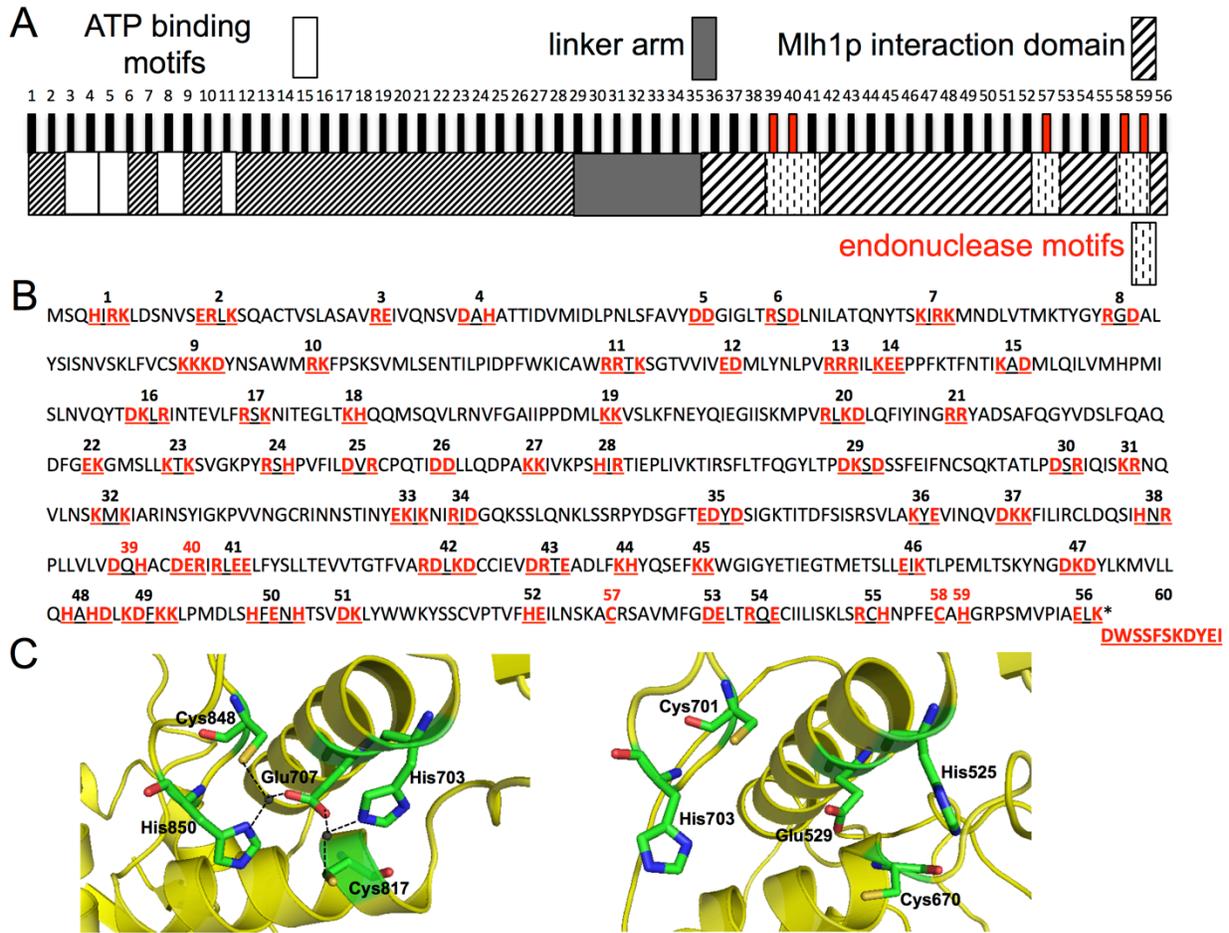


Figure 2.2. Site directed mutagenesis of *MLH3*.

A. Functional organization of Mlh3 based on sequence homology and secondary structure prediction [51]. 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 [51] described in panel C. 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 [51]. C. Metal binding site of Pms1 (left panel) from [51] comprised of the five highlighted residues (H703, E707, C817, C848, and H850) were found to be highly conserved in Mlh3 (right panel) based on sequence alignment and structural modeling (H525, E529, C670, C701, and H703) and were targeted in the mutagenesis described in this study (alleles represented by red bars in A and red numbers in B).

Table 2.1. Yeast strains used in this study.

Strain	Genotype	
EAY3252	<i>MATa ho::hisG, ura3, leu2::hisG, trp1::hisG, ADE2, HIS4, CEN8Tomato::LEU2, MLH3, lys2::insE-A14</i>	
EAY3255	<i>MATa ho::hisG, ura3, leu2::hisG, trp1::hisG, ADE2, his4xB, CEN8Tomato::LEU2, mlh3Δ::NATMX, lys2::insE-A14</i>	
		amino acid substitution
EAY3490-91	Same as EAY3255, but <i>mlh3-1::KANMX</i>	H4A,R6A,K7A
EAY3492-93	Same as EAY3255, but <i>mlh3-2::KANMX</i>	E14A,R15A,K17A
EAY3494-95	Same as EAY3255, but <i>mlh3-3::KANMX</i>	R30A,E31A
EAY3496-97	Same as EAY3255, but <i>mlh3-4::KANMX</i>	D38A,H40A
EAY3498-99	Same as EAY3255, but <i>mlh3-5::KANMX</i>	D59A,D60A
EAY3500-01	Same as EAY3255, but <i>mlh3-6::KANMX</i>	R66A,D68A
EAY3502-03	Same as EAY3255, but <i>mlh3-7::KANMX</i>	K80A,R82A,K83A
EAY3504-05	Same as EAY3255, but <i>mlh3-8::KANMX</i>	R96A,D98A
EAY3506-07	Same as EAY3255, but <i>mlh3-9::KANMX</i>	K114A,K115A,K166A,D117A
EAY3508-09	Same as EAY3255, but <i>mlh3-10::KANMX</i>	R124A,K125A
EAY3510-11	Same as EAY3255, but <i>mlh3-11::KANMX</i>	R151A,R152A,K154A
EAY3512-13	Same as EAY3255, but <i>mlh3-12::KANMX</i>	E162A,D163A
EAY3514-15	Same as EAY3255, but <i>mlh3-13::KANMX</i>	R171A,R172A,R173A
EAY3516-17	Same as EAY3255, but <i>mlh3-14::KANMX</i>	K176A,E177A,E178A
EAY3518-19	Same as EAY3255, but <i>mlh3-15::KANMX</i>	K188A,D190A
EAY3520-21	Same as EAY3255, but <i>mlh3-16::KANMX</i>	D209A,K210A,R212A
EAY3522-23	Same as EAY3255, but <i>mlh3-17::KANMX</i>	R220A,K222A
EAY3524-25	Same as EAY3255, but <i>mlh3-18::KANMX</i>	K230A, H231A
EAY3526-27	Same as EAY3255, but <i>mlh3-19::KANMX</i>	K252A, K253A
EAY3528-29	Same as EAY3255, but <i>mlh3-20::KANMX</i>	R273A, K275A, D276A
EAY3530-31	Same as EAY3255, but <i>mlh3-21::KANMX</i>	R285A, R286A
EAY3532-33	Same as EAY3255, but <i>mlh3-22::KANMX</i>	E307A, K308A
EAY3534-35	Same as EAY3255, but <i>mlh3-23::KANMX</i>	K314A, K316A
EAY3536-37	Same as EAY3255, but <i>mlh3-24::KANMX</i>	R323A, H325A
EAY3538-39	Same as EAY3255, but <i>mlh3-25::KANMX</i>	D331A, R333A
EAY3540-41	Same as EAY3255, but <i>mlh3-26::KANMX</i>	D339A, D340A
EAY3542-43	Same as EAY3255, but <i>mlh3-27::KANMX</i>	K347A, K348A
EAY3544-45	Same as EAY3255, but <i>mlh3-28::KANMX</i>	H354A, R356A
EAY3546-47	Same as EAY3255, but <i>mlh3-29::KANMX</i>	D379A, K380A, D382A
EAY3548-49	Same as EAY3255, but <i>mlh3-30::KANMX</i>	D399A, R401A
EAY3550-51	Same as EAY3255, but <i>mlh3-31::KANMX</i>	K406A, R407A

Table 2.1 (continued)

EAY3552-53	Same as EAY3255, but <i>mlh3-32::KANMX</i>	K414A, K416A
EAY3554-55	Same as EAY3255, but <i>mlh3-33::KANMX</i>	E442A, K443A, K445A
EAY3556-57	Same as EAY3255, but <i>mlh3-34::KANMX</i>	R448A, D450A
EAY3558-59	Same as EAY3255, but <i>mlh3-35::KANMX</i>	E471D, D472A, D474A
EAY3560-61	Same as EAY3255, but <i>mlh3-36::KANMX</i>	K492A, E494A
EAY3562-63	Same as EAY3255, but <i>mlh3-37::KANMX</i>	D500A, K501A, K502A
EAY3564-65	Same as EAY3255, but <i>mlh3-38::KANMX</i>	H514A, R516A
EAY3566-67	Same as EAY3255, but <i>mlh3-39::KANMX</i>	D523A, H525A
EAY3568-69	Same as EAY3255, but <i>mlh3-40::KANMX</i>	D528A, E529A, R530A
EAY3570-71	Same as EAY3255, but <i>mlh3-41::KANMX</i>	R532A, E534A, E535A
EAY3572-73	Same as EAY3255, but <i>mlh3-42::KANMX</i>	R552A, D553A, K555A, D556A
EAY3574-75	Same as EAY3255, but <i>mlh3-43::KANMX</i>	D562A, R563A, E565A
EAY3576-77	Same as EAY3255, but <i>mlh3-44::KANMX</i>	K570A, H571A
EAY3578-79	Same as EAY3255, but <i>mlh3-45::KANMX</i>	K577A, K578A
EAY3580-81	Same as EAY3255, but <i>mlh3-46::KANMX</i>	E596A, K598A
EAY3582-83	Same as EAY3255, but <i>mlh3-47::KANMX</i>	D611A, K612A, D613A
EAY3584-85	Same as EAY3255, but <i>mlh3-48::KANMX</i>	H622A, H624A, D625A
EAY3586-87	Same as EAY3255, but <i>mlh3-49::KANMX</i>	K627A, D628A, K630A, K631A
EAY3588-89	Same as EAY3255, but <i>mlh3-50::KANMX</i>	H638A, E640A, H642A
EAY3590-91	Same as EAY3255, but <i>mlh3-51::KANMX</i>	D646A, K647A
EAY3592-93	Same as EAY3255, but <i>mlh3-52::KANMX</i>	H662A, E663A
EAY3594-95	Same as EAY3255, but <i>mlh3-53::KANMX</i>	D678A, E679A
EAY3596-97	Same as EAY3255, but <i>mlh3-54::KANMX</i>	R682A, E684A
EAY3598-99	Same as EAY3255, but <i>mlh3-55::KANMX</i>	R694A, H696A
EAY3600-01	Same as EAY3255, but <i>mlh3-56::KANMX</i>	E713A, K715A
EAY3701-02	Same as EAY3255, but <i>mlh3-57::KANMX</i>	C670A
EAY3703-04	Same as EAY3255, but <i>mlh3-58::KANMX</i>	C701A
EAY3705-06	Same as EAY3255, but <i>mlh3-59::KANMX</i>	H703A
EAY3707-08	Same as EAY3255, but <i>mlh3-60::KANMX</i>	716-DWSSFSKDYEI
EAY3819-20	Same as EAY3255, but <i>mlh3-D523N::KANMX</i>	
EAY3339	<i>MATa ho::LYS2, lys2, ura3, leu2::hisG, trp1::hisG, THR1::m-Cerulean-TRP1</i>	
EAY3486	Same as EAY3339, but <i>mlh3Δ::NATMX</i>	
EAY1112	<i>MATa ho::hisG, trp1::hisG, leu2::hisG, lys2, ura3, ade2::hisG, his3Δ::hisG, TRP1insertion@CENXV</i>	
EAY1848	Same as EAY1112, but <i>mlh3Δ::KANMX</i>	

Table 2.1 (continued)

EAY3712	Same as EAY1112, but <i>mlh3Δ::URA3</i>
EAY3713-14	Same as EAY1112, but <i>mlh3-6::KANMX</i>
EAY3715-16	Same as EAY1112, but <i>mlh3-23::KANMX</i>
EAY3717-18	Same as EAY1112, but <i>mlh3-32::KANMX</i>
EAY3719-20	Same as EAY1112, but <i>mlh3-42::KANMX</i>
EAY3721-22	Same as EAY1112, but <i>mlh3-45::KANMX</i>
EAY3723-24	Same as EAY1112, but <i>mlh3-54::KANMX</i>
EAY2413	<i>MATa trp1::hisG, leu2::hisG, ho::hisG, ura3, lys2, URA3insertion@CENXV, LEU2insertion@chromXV, LYS2 insertion at position 505193 on chromosome XV, mlh3Δ::NATMX</i>
S288c	<i>MATa ho lys5</i>
YJM789	<i>MATa ho::hisG lys2 cyh</i>
KTY618	Same as YJM789, but <i>SK1-MLH1::natMX4, mlh3Δ::kanMX4</i>
KTY610	Same as S288c, but <i>SK1-MLH3::kanMX4, mlh1Δ::hphMX4</i>
KTY616	Same as S288c, but <i>SK1-mlh3-23::kanMX4, mlh1Δ::hphMX4</i>
KTY614	Same as S288c, but <i>SK1-mlh3-32::kanMX4, mlh1Δ::hphMX4</i>
KTY621	Same as S288c, but <i>SK1-mlh3-D523N::kanMX4, mlh1Δ::hphMX4</i>
KTY626	Same as S288c, but <i>mlh3Δ::kanMX4, mlh1Δ::hphMX4</i>

EAY3252, EAY3255 and derivatives, and EAY3486 are SK1 strains that contain spore autonomous fluorescent markers described in Thacker et al. [56]. EAY1112 and EAY2413 contain chromosome XV markers described in Argueso et al. [7]. KTY strains were used for whole genome recombination mapping as described in Krishnaprasad et al. [64] and Oke et al. [34].

Table 2.2. Diploid strains used to measure % tetratype, spore viability, meiotic progression, genetic map distances and for whole genome recombination mapping.

Diploid strain	Relevant genotype
Spore autonomous fluorescent protein assay [56], meiotic progression assay [93], and Sgs1 overexpression.	
EAY3252/EAY3486	<i>MLH3/mlh3Δ</i>
EAY3255/EAY3486	<i>mlh3Δ/mlh3Δ</i>
EAY3490 to EAY3708/EAY3486	<i>mlh3-1 to mlh3-60/mlh3Δ</i>
EAY3339/EAY3552-53	<i>MLH3/mlh3-32</i>
EAY3819-20/EAY3486	<i>mlh3-D523N/mlh3Δ</i>
Tetrad analysis [7]	
EAY1112/EAY2413	<i>MLH3/mlh3Δ</i>
EAY1848/EAY2413	<i>mlh3Δ/mlh3Δ</i>
EAY3713-14/EAY2413	<i>mlh3-6/mlh3Δ</i>
EAY3715-16/EAY2413	<i>mlh3-23/mlh3Δ</i>
EAY3717-18/EAY2413	<i>mlh3-32/mlh3Δ</i>
EAY3719-20/EAY2413	<i>mlh3-42/mlh3Δ</i>
EAY3721-22/EAY2413	<i>mlh3-45/mlh3Δ</i>
EAY3723-24/EAY2413	<i>mlh3-54/mlh3Δ</i>
Whole genome recombination mapping [34, 64]	
KTY618/KTY610	<i>SK1-MLH1/mlh1Δ, SK1-MLH3/mlh3Δ</i>
KTY618/KTY616	<i>SK1-MLH1/mlh1Δ, SK1-mlh3-23/mlh3Δ</i>
KTY618/KTY614	<i>SK1-MLH1/mlh1Δ, SK1-mlh3-32/mlh3Δ</i>
KTY618/KTY621	<i>SK1-MLH1/mlh1Δ, SK1-mlh3-D523N/mlh3Δ</i>
KTY618/KTY626	<i>SK1-MLH1/mlh1Δ, mlh3Δ/mlh3Δ</i>

The indicated haploid strains were mated to form the diploids with the relevant genotype shown.

Table 2.3. Plasmids used in this study.A. *MLH3* mutagenesis plasmids
allele

pEAI254	<i>MLH3</i>
pEAI325	<i>mlh3-1</i>
pEAI326	<i>mlh3-2</i>
pEAI327	<i>mlh3-3</i>
pEAI328	<i>mlh3-4</i>
pEAI329	<i>mlh3-5</i>
pEAI330	<i>mlh3-6</i>
pEAI331	<i>mlh3-7</i>
pEAI332	<i>mlh3-8</i>
pEAI333	<i>mlh3-9</i>
pEAI334	<i>mlh3-10</i>
pEAI335	<i>mlh3-11</i>
pEAI336	<i>mlh3-12</i>
pEAI337	<i>mlh3-13</i>
pEAI338	<i>mlh3-14</i>
pEAI339	<i>mlh3-15</i>
pEAI340	<i>mlh3-16</i>
pEAI341	<i>mlh3-17</i>
pEAI342	<i>mlh3-18</i>
pEAI343	<i>mlh3-19</i>
pEAI344	<i>mlh3-20</i>
pEAI345	<i>mlh3-21</i>
pEAI346	<i>mlh3-22</i>
pEAI347	<i>mlh3-23</i>
pEAI348	<i>mlh3-24</i>
pEAI349	<i>mlh3-25</i>
pEAI350	<i>mlh3-26</i>
pEAI351	<i>mlh3-27</i>
pEAI352	<i>mlh3-28</i>
pEAI353	<i>mlh3-29</i>
pEAI354	<i>mlh3-30</i>
pEAI355	<i>mlh3-31</i>
pEAI356	<i>mlh3-32</i>
pEAI357	<i>mlh3-33</i>
pEAI358	<i>mlh3-34</i>
pEAI359	<i>mlh3-35</i>

Table 2.3. (continued)

pEAI360	<i>mlh3-36</i>
pEAI361	<i>mlh3-37</i>
pEAI362	<i>mlh3-38</i>
pEAI363	<i>mlh3-39</i>
pEAI364	<i>mlh3-40</i>
pEAI365	<i>mlh3-41</i>
pEAI366	<i>mlh3-42</i>
pEAI367	<i>mlh3-43</i>
pEAI368	<i>mlh3-44</i>
pEAI369	<i>mlh3-45</i>
pEAI370	<i>mlh3-46</i>
pEAI371	<i>mlh3-47</i>
pEAI372	<i>mlh3-48</i>
pEAI373	<i>mlh3-49</i>
pEAI374	<i>mlh3-50</i>
pEAI375	<i>mlh3-51</i>
pEAI376	<i>mlh3-52</i>
pEAI377	<i>mlh3-53</i>
pEAI378	<i>mlh3-54</i>
pEAI379	<i>mlh3-55</i>
pEAI380	<i>mlh3-56</i>
pEAI394	<i>mlh3-57</i>
pEAI395	<i>mlh3-58</i>
pEAI396	<i>mlh3-59</i>
pEAI397	<i>mlh3-60</i>
pEAI252	<i>mlh3-D523N</i>

B. Two-hybrid plasmids

	relevant genotype
pBTM116	<i>pADH1-lexA</i> , dummy vector 2μ , <i>TRP1</i>
pEAM105	<i>pADH1-lexA-MLH1_{SK1}</i> 2μ , <i>TRP1</i>
pGAD424	<i>pADH1-GAL4AD</i> 2μ , <i>LEU2</i>
pEAM234	<i>pADH1-GAL4-AD-MLH3_{SK1}</i> 2μ , <i>LEU2</i>
pEAM235	<i>pADH1-GAL4-AD-mlh3-39</i> 2μ , <i>LEU2</i>
pEAM236	<i>pADH1-GAL4-AD-mlh3-40</i> 2μ , <i>LEU2</i>
pEAM237	<i>pADH1-GAL4-AD-mlh3-41</i> 2μ , <i>LEU2</i>
pEAM238	<i>pADH1-GAL4-AD-mlh3-42</i> 2μ , <i>LEU2</i>
pEAM241	<i>pADH1-GAL4-AD-mlh3-45</i> 2μ , <i>LEU2</i>

Table 2.3 (continued)

pEAM242	<i>pADH1-GAL4-AD-mlh3-48</i>	2 μ , LEU2
pEAM244	<i>pADH1-GAL4-AD-mlh3-54</i>	2 μ , LEU2
pEAM245	<i>pADH1-GAL4-AD-mlh3-60</i>	2 μ , LEU2

C. 2 μ URA3 plasmids

	Relevant genotype
pRS426 (pEAO34)	Dummy vector
pEAM266	<i>SGS1</i>
pEAM270	<i>sgs1-K706A (sgs1-hd)</i>

A. All *MLH3* mutagenesis plasmids are derived from pEAI254, a 7.8 kb *MLH3_{SK1}::KANMX* integrating vector. pEAI254 was mutagenized by QuickChange to create the alleles listed. The DNA sequence of the entire ORF, and 70 bp upstream and 150 bp downstream, were confirmed by DNA sequencing using primers EAO318, EAO319, EAO1778 and EAO321. B. For the two-hybrid analysis, pEAM105 contains the entire *MLH1* gene derived from the SK1 strain inserted immediately after the *lexA* binding domain in pBTM116. All *GAL4* activating domain-*mlh3* plasmids are derived from pEAM234, which contains DNA sequence encoding SK1 *MLH3* amino acids 481 to 715 inserted immediately after the *GAL4* activating domain in pGAD424. C. *SGS1* was cloned into pRS426 to make pEAM266 as described in the Methods. Site directed mutagenesis of pEAM266 was performed to make pEAM270 carrying the *sgs1* helicase defective allele as described in the Methods.

Structure-function analysis of Mlh3.

We analyzed the effect of *mlh3* mutations on MMR in vegetatively grown cells and on meiotic COs in diploids induced to undergo sporulation. For MMR we employed the *lys2-A₁₄* reversion assay to assess the mutation rate in *mlh3* haploid strains (Table 2.1; [55]). In this assay, the median reversion rate of *mlh3Δ* is six-fold higher than wild-type (Figure 2.3; Figure 2.4B; Table 2.4; [5, 6]). To measure meiotic crossing over we crossed mutant *mlh3* strains to *mlh3Δ* strains (*MLH3* is haplosufficient [6]) to form diploids that were then sporulated (Table 2.2). The resulting tetrads were directly visualized for chromosome VIII CO events using a spore autonomous fluorescence assay ([56]; Figure 2.4A). In *mlh3Δ* strains we observed a more than two-fold decrease in crossing over, as measured by percent tetratype, compared to wild-type (Figure 2.3; Figure 2.4B; Table 2.4). Similar effects of the *mlh3Δ* mutation on crossing over were seen at other genetic intervals [5-8]. It is important to note that nonparental ditype (NPD) events were not scored because they cannot be distinguished from Meiosis I nondisjunction events [56].

Similar to work performed on a smaller number of *mlh3* alleles and a structure-function analysis of *MLH1*, we found that *MLH3* MMR functions were more sensitive to mutagenesis than CO functions (Figure 2.3; [6, 54]). Phenotypes exhibited by *mlh3* strains containing mutations in the ATP-binding motif suggested that this region plays a more critical role in MMR compared to crossing over. However, a region just beyond the ATP-binding domain appeared insensitive to mutagenesis. A null phenotype for both functions was observed in strains bearing mutations in endonuclease motifs, further confirming that endonuclease activity is essential for MMR and crossing over (Figure 2.3; Table 2.4; [5, 8, 16]).

Comparison of the MMR and CO assay results for each individual allele led to the

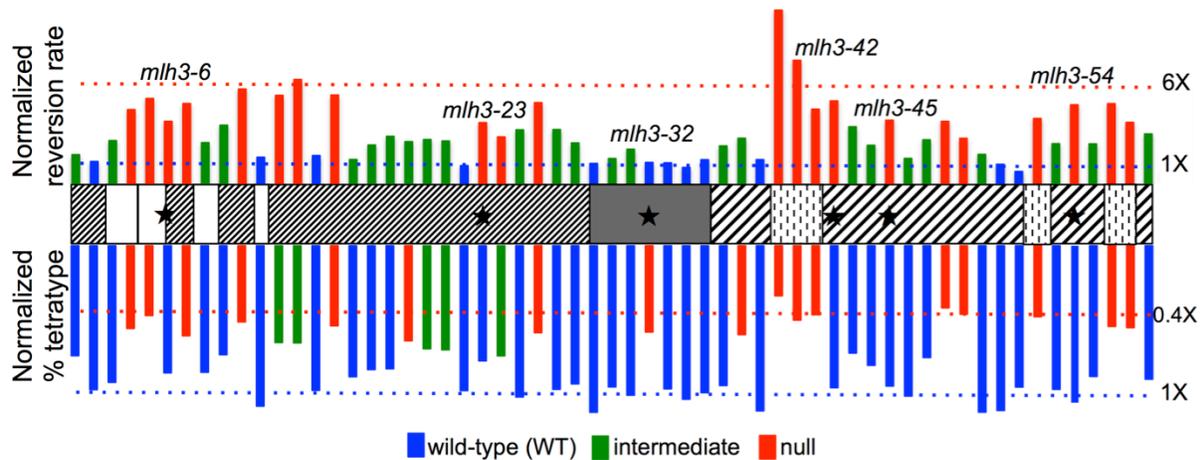


Figure 2.3. Structure-function map of *Saccharomyces cerevisiae* Mlh3.

Mismatch repair (top) and crossing over (bottom) phenotypes of *mlh3* alleles analyzed in this study (except *mlh3-60*). Mismatch repair was measured using the *lys2-A₁₄* reversion assay [55] and crossing over was measured using the assay depicted in Figure 2.4A [56]. Bars represent the median reversion rates and percent tetraplate normalized to *MLH3* (1X). The vertical bars indicate the approximate position of the *mlh3* mutations analyzed in this study with the height of each bar corresponding to the phenotype relative to *MLH3* (1X). Red indicates a null phenotype, blue indicates wild-type (WT), and green indicates intermediate. 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 1X=1.43x10⁻⁶ (n=140). For crossing over (bottom), bars represent percent tetraplate of at least 250 tetrads from two independently constructed strains presented here normalized to *MLH3* percent tetraplate 1X=36.7% (n= 226). Blue and red dotted lines represent *MLH3* and *mlh3Δ* respectively. Black stars indicate separation of function mutants (Table 2.4).

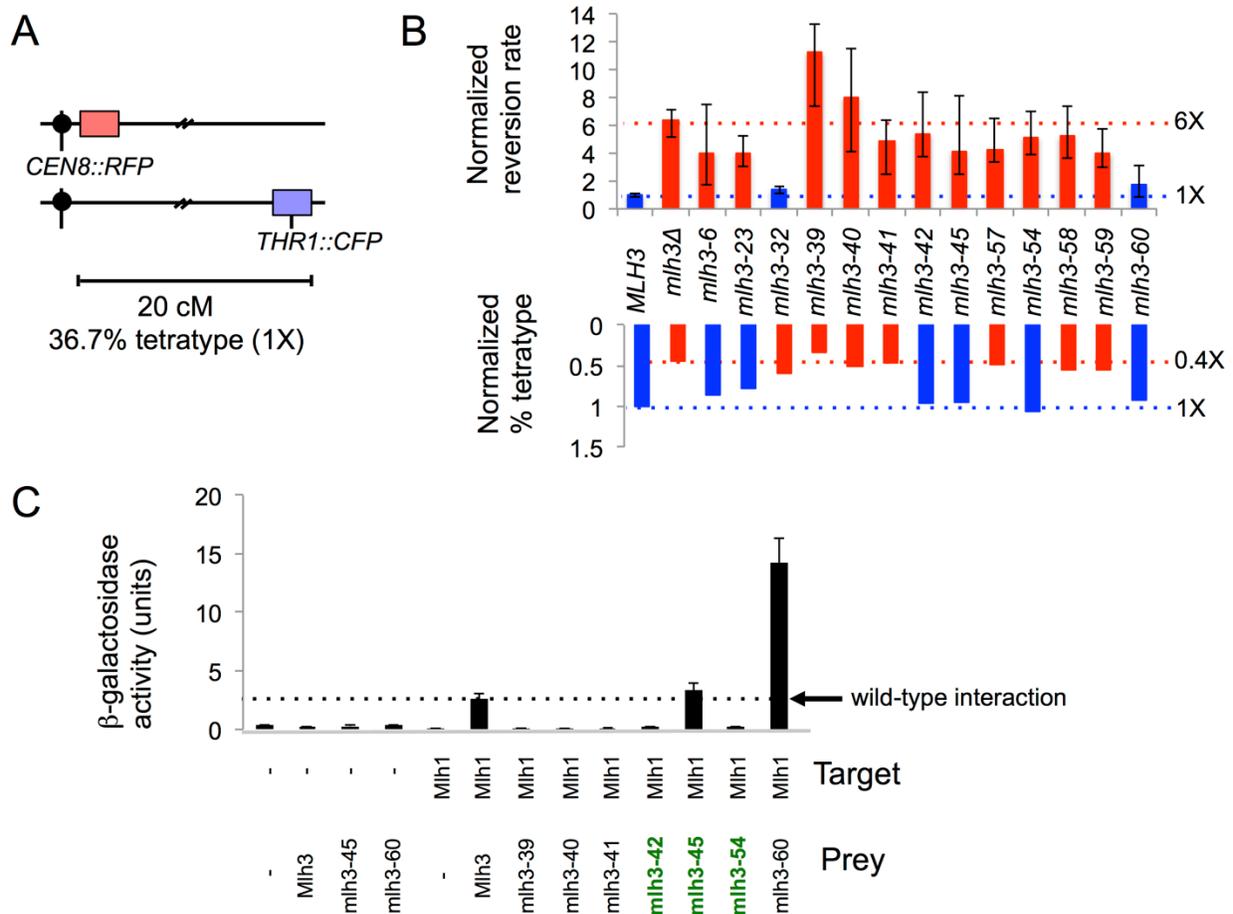


Figure 2.4. Identification and characterization of *mlh3* separation of function alleles.

A. Spore-autonomous fluorescent protein expression was used to quantify crossing over [56]. Shown is the starting parental configuration on chromosome VIII with a map distance of 20 cM separating the red fluorescent protein (RFP) marker and the blue fluorescent protein (CFP) marker. Percent tetratype at this interval in wild-type meiosis is 36.7%. B. MMR (top) and CO (bottom) phenotypes for *MLH3* and *mlh3Δ*, separation of function, endonuclease, and C-terminal tail (*mlh3-60*) mutants. Mismatch repair was measured using the *lys2-A₁₄* reversion assay [55] 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=1.43x10⁻⁶ (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; Table 2.4). Red indicates a null phenotype, blue indicates wild-type. Dotted lines represent *MLH3* and *mlh3Δ*. C. *mlh3-42*, *-54* weaken Mlh1 interaction yet maintain crossover function. Yeast two-hybrid interactions between lexA-Mlh1 (target) and Gal4-Mlh3 (amino acids 481-715; prey) or Gal4-*mlh3-39*, *-40*, *-41*, *-42*, *-45*, *-54*, *-60* derivative constructs, as measured in the ONPG assay for β -galactosidase activity. Error bars indicate standard error of mean from at least three independent assays. *mlh3* separation of function alleles indicated in green font.

Table 2.4. Mismatch repair (MMR) and crossover (CO) phenotypes of the *mlh3* variants as measured in *lys2-A₁₄* reversion and spore autonomous fluorescent assays.

Allele	Reversion rate (x10 ⁻⁶)	95% CI (x10 ⁻⁶)	Relative to WT	% tetatype	Relative to WT	Phenotype	
						MMR	CO
<i>MLH3</i>	1.43	1.23-1.65	1.00	36.70	1.00	+	+
<i>mlh3Δ</i>	9.07	7.4-10.28	6.34	16.10	0.44	-	-
<i>mlh3-1</i>	2.72	2.14-3.13	1.90	27.27	0.74	±	+
<i>mlh3-2</i>	2.09	1.63-2.94	1.46	35.69	0.97	+	+
<i>mlh3-3</i>	4.03	3.21-5.39	2.82	33.86	0.92	±	+
<i>mlh3-4</i>	6.89	4.16-8.68	4.82	20.51	0.56	-	-
<i>mlh3-5</i>	7.93	6.79-9.65	5.55	17.30	0.47	-	-
<i>mlh3-6</i>	5.80	2.44-10.7	4.06	31.52	0.86	-	+
<i>mlh3-7</i>	7.46	5.59-10.16	5.22	22.33	0.61	-	-
<i>mlh3-8</i>	3.84	2.77-5.10	2.68	31.37	0.85	±	+
<i>mlh3-9</i>	5.46	3.78-6	3.82	27.02	0.74	±	+
<i>mlh3-10</i>	8.80	7.72-11.98	6.16	18.89	0.51	-	-
<i>mlh3-11</i>	2.49	1.57-2.83	1.74	39.77	1.08	+	+
<i>mlh3-12</i>	8.22	6.2-14.44	5.75	24.00	0.65	-	±
<i>mlh3-13</i>	9.69	6.08-25.7	6.77	24.10	0.66	-	±
<i>mlh3-14</i>	2.64	1.57-3.99	1.85	35.85	0.98	+	+
<i>mlh3-15</i>	8.25	5.84-9.6	5.77	19.85	0.54	-	-
<i>mlh3-16</i>	2.27	1.7-4.78	1.59	32.49	0.89	±	+
<i>mlh3-17</i>	3.62	2.7-6.39	2.53	30.71	0.84	±	+
<i>mlh3-18</i>	4.42	2.74-6.40	3.09	30.51	0.83	±	+
<i>mlh3-19</i>	3.93	3.65-5.23	2.75	23.57	0.64	±	-
<i>mlh3-20</i>	4.12	3.19-6.05	2.88	25.56	0.70	±	±
<i>mlh3-21</i>	3.99	3.41-5.32	2.79	25.81	0.70	±	±
<i>mlh3-22</i>	1.71	1.17-2.96	1.19	35.92	0.98	+	+
<i>mlh3-23</i>	5.69	4.37-7.56	3.98	28.54	0.78	-	+
<i>mlh3-24</i>	4.37	2.57-10.26	3.05	27.31	0.74	-	±
<i>mlh3-25</i>	5.03	4.57-6.22	3.52	37.50	1.02	±	+
<i>mlh3-26</i>	7.54	3.97-12.11	5.28	21.56	0.59	-	-
<i>mlh3-27</i>	5.05	2.75-6.78	3.53	35.60	0.97	±	+
<i>mlh3-28</i>	3.81	2.49-4.51	2.66	34.25	0.93	±	+
<i>mlh3-29</i>	1.91	0.61-4.69	1.33	41.29	1.13	+	+
<i>mlh3-30</i>	2.36	1.7-3.6	1.65	35.00	0.95	±	+
<i>mlh3-31</i>	3.23	2.01-6.91	2.26	37.06	1.01	±	+
<i>mlh3-32</i>	2.00	1.54-2.22	1.40	21.40	0.58	+	-
<i>mlh3-33</i>	1.97	1.65-2.58	1.38	35.49	0.97	+	+
<i>mlh3-34</i>	1.54	1.25-2.82	1.07	38.06	1.04	+	+

Table 2.4 (continued)

<i>mlh3-35</i>	2.26	0.65-3.21	1.58	36.43	0.99	+	+
<i>mlh3-36</i>	3.54	2.37-5.45	2.47	34.66	0.94	±	+
<i>mlh3-37</i>	4.25	2.54-5.15	2.97	22.04	0.60	±	-
<i>mlh3-38</i>	2.26	1.14-3.94	1.58	40.94	1.12	+	+
<i>mlh3-39</i>	16.11	10.54-19	11.26	12.42	0.34	-	-
<i>mlh3-40</i>	11.46	5.81-16.57	8.02	18.45	0.50	-	-
<i>mlh3-41</i>	6.94	3.56-9.1	4.85	17.11	0.47	-	-
<i>mlh3-42</i>	7.70	5.38-12	5.39	35.20	0.96	-	+
<i>mlh3-43</i>	5.31	4.09-7.23	3.71	26.61	0.73	±	+
<i>mlh3-44</i>	3.59	2.87-4.25	2.51	29.65	0.81	±	+
<i>mlh3-45</i>	5.92	3.49-11.7	4.14	34.79	0.95	-	+
<i>mlh3-46</i>	2.38	1.75-3.01	1.66	37.27	1.02	±	+
<i>mlh3-47</i>	4.10	2.98-5.29	2.87	27.72	0.76	±	+
<i>mlh3-48</i>	5.80	4.01-8.93	4.06	15.41	0.42	-	-
<i>mlh3-49</i>	4.24	3.6-9.72	2.97	16.98	0.46	-	-
<i>mlh3-50</i>	2.75	2.24-3.33	1.92	41.26	1.12	±	+
<i>mlh3-51</i>	1.83	1.04-3.23	1.28	40.84	1.11	+	+
<i>mlh3-52</i>	1.17	0.79-2.7	0.82	35.04	0.95	+	+
<i>mlh3-57</i>	6.07	4.74-9.4	4.25	17.60	0.48	-	-
<i>mlh3-53</i>	3.74	2.52-6.9	2.62	35.61	0.97	±	+
<i>mlh3-54</i>	7.34	5.59-9.97	5.13	38.93	1.06	-	+
<i>mlh3-55</i>	3.72	2.22-5.45	2.60	32.41	0.88	±	+
<i>mlh3-58</i>	7.45	5.23-10.63	5.21	20.00	0.54	-	-
<i>mlh3-59</i>	5.71	4.28-8.19	3.99	20.31	0.55	-	-
<i>mlh3-56</i>	4.65	3.53-6.11	3.25	33.09	0.90	±	+
<i>mlh3-60</i>	2.50	1.18-4.45	1.75	33.83	0.92	+	+

Two independently constructed strains with *mlh3* variants were analyzed in the EAY3255 background which contains the *lys2::insE-A₁₄* for MMR testing and the red fluorescent protein for meiotic testing. Haploid strains were examined for reversion to Lys⁺. At least n=10 reversion assays were performed per allele. Median reversion rates are presented with 95% confidence intervals (CI), and relative reversion rates compared with the wild-type strain are shown. The haploid strains were mated to EAY3486, which contains the blue fluorescent protein to make diploids suitable for meiotic testing. Diploid strains were induced for meiosis and % tetatype was measured. At least 250 tetrads were counted for each allele. WT, wild-type. +, indistinguishable from WT as measured by 95% CI (for reversion rates) or χ^2 (p<0.01, for % tetatype). -, indistinguishable from null as measured by 95% CI or χ^2 (p<0.01). +/-, distinguishable from both wild-type and null as measured by 95% CI or χ^2 (p<0.01).

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 2.4B; Figure 2.3, 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⁺). As indicated below, *mlh3-23* displayed different CO phenotypes in different strain backgrounds, and so its designation as a separation of function allele is less clear.

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 [57, 58].

The endonuclease active sites in Mlh3 and Pms1 appear to be similar.

S. cerevisiae Mlh1-Pms1 and Mlh1-Mlh3 and human MLH1-PMS2 display latent endonuclease activities dependent on the integrity of a highly conserved metal binding motif DQHA(X)₂E(X)₄E [14-17]. This motif is critical for Mlh3's MMR and meiotic functions [5]. Two additional motifs were implicated in MLH family endonuclease function based on sequence alignment: ACR and C(P/N)HGRP [59]. In the Mlh1-Pms1 C-terminal domains crystal structure,

five Pms1 residues, located in the three conserved motifs (H703, E707, C817, C848, H850), form a metal binding site through folding of the Pms1 C-terminal domain (Figure 2.2C; [51]). This organization was also seen in the crystal structure of the C-terminal domain of *B. subtilis* endonuclease MutL (all but H703 are conserved; [60]).

We performed a sequence alignment of Mlh3, Pms1, and *B. subtilis* MutL and found all three possess conserved metal binding motifs, with the following five residues predicted to form the endonuclease active site in Mlh3: H525, E529, C670, C701, and H703. In addition, we constructed a homology model of *S. cerevisiae* Mlh3, and found the C-terminal domain can potentially fold in a similar manner to Pms1 such that these five conserved residues form a single putative metal binding site (Figure 2.2C). These residues were targeted for site directed mutagenesis of *MLH3* (Figure 2.2B, shown in red). As shown in Figure 2.3 (represented in the C-terminal domain by dotted white squares) and Figure 2.4B and Table 2.4, mutations in the putative conserved metal binding motifs of Mlh3 (*mlh3-39(D523A, H525A)*, *-40(D528A, E529A, R530A)*, *-57(C670A)*, *-58(C701A)*, and *-59(H703A)*) conferred null phenotypes for MMR and crossing over, indicating they are essential for Mlh3 function. Thus, these genetic data, combined with the high sequence homology, suggest that H525, E529, C670, C701, and H703 in the C-terminal domain of Mlh3 form the catalytic active site (Figure 2.2C).

The endonuclease motifs in Mlh3 overlap with the C-terminal Mlh1 interaction domain. To determine if mutations in the Mlh3 endonuclease motifs disrupted interaction with Mlh1, three alleles spanning the DQHA(X)₂E(X)₄E endonuclease motif (*mlh3-39*, *-40*, and *-41*) were analyzed by yeast two-hybrid for interaction with Mlh1. We also tested these alleles because a previously characterized mutation in the DQHA(X)₂E(X)₄E endonuclease motif (*mlh3-E529K*) disrupts Mlh1-Mlh3 interactions [5]. As shown in Figure 2.4C, these mutations disrupted Mlh1-

Mlh3 interactions, possibly by altering the endonuclease active site structure. This idea is supported by the Mlh1-Pms1 crystal structure. In this model, heterodimer stability is maintained through interactions between the C-terminal domain of Mlh1 and the endonuclease active site of Pms1 [51]. We cannot rule out the possibility that the null phenotypes observed for MMR and crossing over in *mlh3-39*, *-40*, and *-41* were caused by specifically mutating residues that comprise the Mlh1-Mlh3 dimerization interface without causing a gross disruption in protein folding.

The Mlh1-Pms1 C-terminal domain structure reveals three patches constituting the heterodimerization interface of Mlh1-Pms1 [51]. Patch I is a pseudosymmetric hydrophobic core, Patch II is composed of the last 12 residues of Pms1 and contributes two salt bridges, and Patch III involves the C-terminus of Mlh1 and contributes to the Pms1 metal binding site [51]. Patches I and III are likely maintained in the Mlh1-Mlh3 heterodimerization interface, but Mlh3 lacks the last 11 residues that comprise the bulk of Patch II. This finding gives a likely explanation for partial disruption of the Mlh1-Mlh3 complex when we attempted to analyze it further by gel-filtration [16]. We hypothesized that restoring Patch II to the Mlh1-Mlh3 interaction interface will strengthen this interaction. We engineered a fusion construct of Mlh3 carrying the last 11 residues of Pms1 (*mlh3-60*, Figure 2.2B in red). As shown in Figure 2.4C, when we inserted the last 11 residues of Pms1 after the C-terminal residue of Mlh3, we observed a striking increase in the strength of the interaction between Mlh1 and Mlh3 as measured in the yeast two-hybrid assay (2.6 ± 0.5 Miller units of β -galactosidase activity for wild-type Mlh1-Mlh3 compared to 14.2 ± 2.1 for Mlh1-*mlh3-60*). We initially hypothesized that such an enhanced interaction would be detrimental to MMR because it would sequester Mlh1 from the major MMR endonuclease Pms1. Surprisingly we did not observe a significant effect of the

mlh3-60 mutation on MMR or on crossing over (Figure 2.4B; Table 2.4), suggesting that strengthening the interaction between Mlh1-Mlh3 does not affect formation of the Mlh1-Pms1 heterodimer.

***mlh3-42* and *mlh3-54* weaken interaction with Mlh1 yet maintain wild-type levels of crossing over.**

Three of the five MMR⁻, CO⁺ alleles (*mlh3-42*, *-45*, and *-54*) contained mutations that mapped to the Mlh1 interaction interface. We performed a two-hybrid assay to test whether these mutations affected Mlh1-Mlh3 dimerization. The *mlh3-45* mutation did not alter Mlh1-Mlh3 interactions; however, both the *mlh3-42* and *mlh3-54* mutations disrupted this interaction (Figure 2.4C). While such a result could explain the null MMR phenotype conferred by *mlh3-42* and *mlh3-54*, it does not explain why these strains are functional for meiotic crossing over. One explanation is that additional pro-CO factors act as structural scaffolds to stabilize the weakened Mlh1-mlh3 heterodimer, thus allowing it to perform its function at dHJs. Several observations support this idea: 1. The pro-CO factors Msh4-Msh5, Sgs1-Top3-Rmi1, Zip3, and Exo1 interact with one another and/or with Mlh1-Mlh3 [29, 40, 41, 42, 61, 62]. 2. Studies in mice showed that MLH1 and MLH3 do not form a complex until mid to late pachytene; at early to mid pachytene, only MLH3 foci are seen [37, 44]. 3. Exo1's role in crossing over is independent of its enzymatic activity; it is suggested to play a structural role, acting as a platform for pro-CO factors [29]. Together these observations support the presence of a resolvase complex at CO sites that regulates the endonuclease activity of Mlh1-Mlh3 (see Discussion). Alternatively, a weak Mlh1-Mlh3 interaction defect is sufficient to inhibit a yeast-two hybrid interaction, but not affect meiotic recombination if the strength of the Mlh1-Mlh3 interaction is not a limiting factor

for CO resolution.

Cumulative genetic distance and spore viability measurements confirm fluorescent assay results for *mlh3* separation of function alleles.

The meiotic CO phenotype for the separation of function mutants was determined at four consecutive intervals in chromosome (XV) using traditional genetic map distance analyses (Figure 2.5A; [5-7]). The overall effect of *mlh3* mutations in crossing over on chromosome XV was similar to that determined on chromosome VIII (Figure 2.4B; Figure 2.5B; Table 2.4; Table 2.5). *mlh3-6*, *mlh3-42*, *mlh3-45*, and *mlh3-54* appear similar to wild-type; *mlh3-23* displays an intermediate phenotype, and *mlh3-32* appears similar to the null (Fig 2.5B; Table 2.5). Thus, these mutants, phenotypes of which are summarized in Table 2.6, were confirmed as separation of function alleles, and are candidates for in-depth characterization and high-resolution recombination mapping.

It is important to note that the spore viability and genetic distance measurements of the six separation of function alleles indicated that CO levels, represented by genetic distance measurements in Chr. XV, can be reduced from 115 cM to ~70 cM without compromising spore viability (Figure 2.5C; Table 2.6). A similar observation was made with *msh4/5* hypomorph alleles where crossing over in the four same intervals in Chr. XV could be reduced to ~50 cM without affecting spore viability [63-64]. Together, these observations and the high resolution mapping below indicate that the baker's yeast meiotic cell does not require the full amount of COs maintained by CO homeostasis (~90; see [65]) for accurate chromosome segregation and to form viable spores.

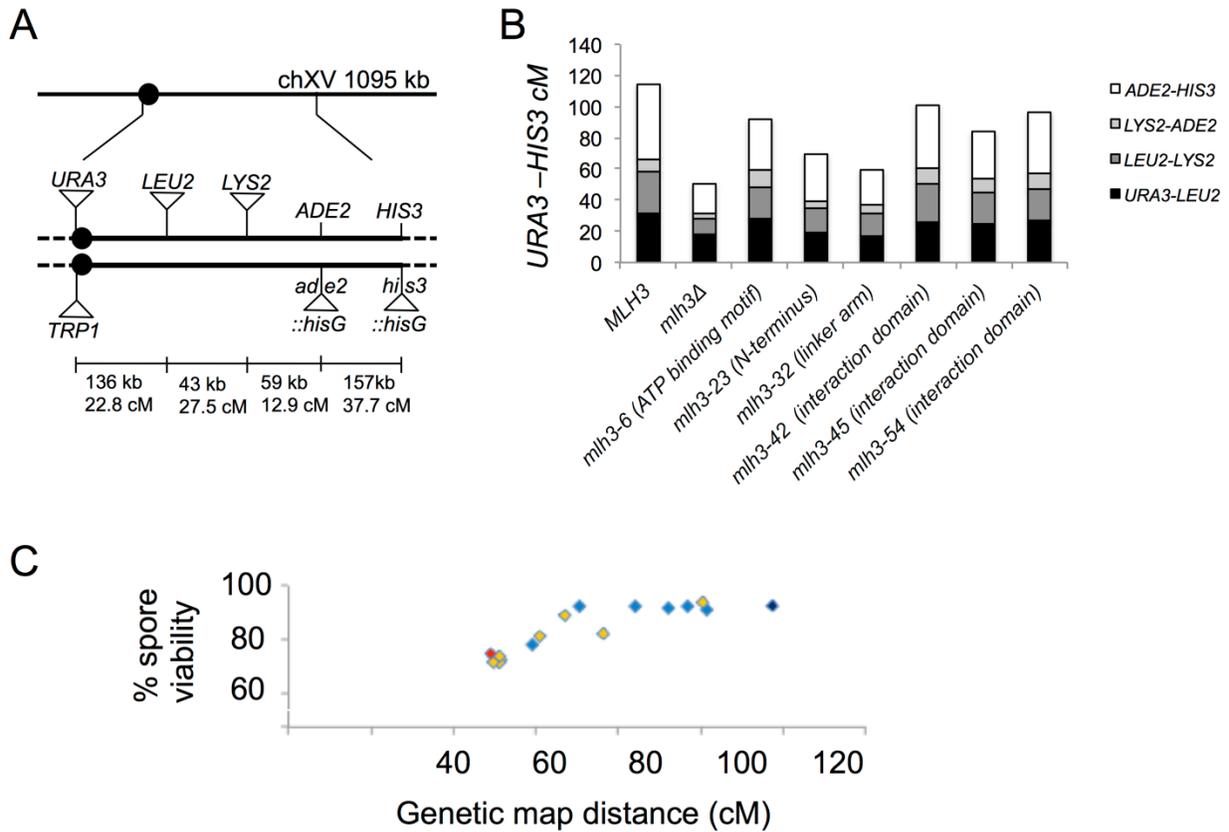


Figure 2.5. Cumulative genetic distance and spore viability of *mlh3* separation of function mutants.

A. Distribution of genetic markers on chromosome XV used to determine genetic distances in the EAY1112/EAY2413 background (Table 2.1). The solid circle indicates the centromere. The distances between markers are not drawn to scale. The actual physical and genetic distances in the wild-type diploid are given numerically for each interval and for the entire region between *CENXV* and *HIS3* [7]. B. Cumulative genetic distances between *URA3* and *HIS3* markers from tetrads of *MLH3* and indicated *mlh3* variants. Each bar is further divided into sectors that correspond to the four genetic intervals that span *URA3-HIS3* (Table 2.5). C. Spore viabilities are plotted vs. genetic map distances from panel B for *MLH3* (dark blue), *mlh3Δ* (red), and the separation of function mutants (light blue). Yellow diamonds represent data from Sonntag Brown et al. [6].

Table 2.5. Genetic map distances for *mlh3* separation of function mutants on chromosome XV from single spores and tetrads.

All *mlh3* mutants are isogenic derivatives of EAY1112/EAY2413 (Table 2.2; Methods). For single spores, recombination frequencies (recombinant spores/total spores) were multiplied by 100 to yield genetic map distances (cM). For tetrads, genetic distance in centimorgans (cM) was calculated using the RANA software without considering aberrant segregants (Argueso et al. [7]). The Stahl Laboratory Online Tools website (<http://molbio.uoregon.edu/~fstahl/>) was used to calculate standard error (SE) around the genetic distance for tetrads. n; number of single spores, N; four spore viable tetrads analyzed; Par, parental single spores; Rec, recombinant single spores.

Table 2.5 (continued)

Genotype	Single Spores				Tetrads					
	n	Par.	Rec.	cM	N	PD	TT	NPD	cM	SE
<i>URA3-LEU2</i>										
<i>MLH3</i>	896	641	255	28.5	205	93	106	4	32.0	3.2
<i>mlh3Δ</i>	1028	876	152	14.8	210	146	61	2	17.5	2.5
<i>mlh3-6</i>	907	691	216	23.8	205	110	84	5	28.6	3.5
<i>mlh3-23</i>	480	391	89	18.5	105	69	32	1	18.6	3.6
<i>mlh3-32</i>	995	836	159	16.0	205	140	61	1	16.6	2.1
<i>mlh3-42</i>	731	551	180	24.6	154	79	73	1	25.8	2.7
<i>mlh3-45</i>	925	709	216	23.4	204	103	99	0	24.5	1.8
<i>mlh3-54</i>	694	533	161	23.2	155	79	70	2	27.2	3.3
<i>LEU2-LYS2</i>										
<i>MLH3</i>	896	680	216	24.1	205	103	99	1	25.9	2.2
<i>mlh3Δ</i>	1028	923	105	10.2	210	166	43	0	10.3	1.4
<i>mlh3-6</i>	907	738	169	18.6	205	129	68	2	20.1	2.6
<i>mlh3-23</i>	480	397	83	17.3	105	69	33	0	16.2	2.3
<i>mlh3-32</i>	995	854	141	14.2	205	144	58	0	14.4	1.6
<i>mlh3-42</i>	731	564	167	22.9	154	81	71	1	25.2	2.7
<i>mlh3-45</i>	925	739	186	20.1	204	123	78	1	20.8	2.2
<i>mlh3-54</i>	694	550	144	20.8	155	90	61	0	20.2	2.0
<i>LYS2-ADE2</i>										
<i>MLH3</i>	896	815	81	9.0	205	168	35	0	8.6	1.3
<i>mlh3Δ</i>	1028	988	40	3.9	210	193	16	0	3.8	0.9
<i>mlh3-6</i>	907	812	95	10.5	205	158	41	0	10.3	1.4
<i>mlh3-23</i>	480	454	26	5.4	105	92	10	0	4.9	1.5
<i>mlh3-32</i>	995	930	65	6.5	205	176	26	0	6.4	1.2
<i>mlh3-42</i>	731	663	68	9.3	154	124	29	0	9.5	1.6
<i>mlh3-45</i>	925	856	69	7.5	204	172	29	1	8.7	1.9
<i>mlh3-54</i>	694	627	67	9.7	155	121	30	0	9.9	1.6
<i>ADE2-HIS3</i>										
<i>MLH3</i>	896	560	336	37.5	205	63	129	11	48.0	4.5
<i>mlh3Δ</i>	1028	829	199	19.4	210	130	79	0	18.9	1.7
<i>mlh3-6</i>	907	665	242	26.7	205	98	95	6	32.9	3.8
<i>mlh3-23</i>	480	352	128	26.7	105	46	55	1	29.9	3.6
<i>mlh3-32</i>	995	811	184	18.5	205	129	70	3	21.8	2.9
<i>mlh3-42</i>	731	491	241	33.0	154	58	89	6	40.8	4.7
<i>mlh3-45</i>	925	650	275	29.7	204	85	116	1	30.2	2.2
<i>mlh3-54</i>	694	469	225	32.4	155	58	88	5	39.1	4.4

Table 2.6. Summary of *mlh3* separation of function phenotypes.

Genotype	%SV	Genetic map distance (cM)	% tetratype	Reversion rate X10 ⁻⁶ (95%CI)	β-gal units (±SEM)
<i>MLH3</i>	93.3	114.5	36.7	1.4 (1.2-1.7)	2.6 (±0.5)
<i>mlh3Δ</i>	74.9	50.5	16.1	9.1 (7.4-10.3)	ND
<i>mlh3-6</i>	91.4	91.9	31.5	5.8 (2.4-10.7)	ND
<i>mlh3-23</i>	90.9	69.6	28.5	5.7 (4.4-7.6)	ND
<i>mlh3-32</i>	78.5	59.2	21.4	2.0 (1.5-2.2)	ND
<i>mlh3-42</i>	91.8	101.3	35.2	7.7 (5.4-12.0)	0.20 (±0.04)
<i>mlh3-45</i>	92.9	84.2	34.8	5.9 (3.5-11.7)	3.3 (±0.6)
<i>mlh3-54</i>	92.3	96.4	38.9	7.3 (5.6-10.0)	0.16 (±0.03)

Data were obtained from Table 2.4 (% tetratype, reversion rate), Table 2.5 (% spore viability (SV), genetic distance in cM), Figure 2.4C (β-gal units in the two-hybrid assay). ND, not determined.

High-resolution recombination maps illustrate unexpected effects of *mlh3* mutants on resolving meiotic recombination intermediates.

We characterized three *mlh3* alleles (*mlh3-23*-MMR⁻, CO⁺, *mlh3-32*-MMR⁺, CO⁻, *mlh3-D523N*) in a genome-wide meiotic recombination assay [66]. *mlh3-23* and *mlh3-32* were chosen as they confer stronger defects in one function relative to the other. The *mlh3-D523N* endonuclease mutation contains an aspartic acid to asparagine substitution in the DQHA(X)₂E(X)₄E metal binding motif of Mlh3. This mutation does not disrupt formation of the Mlh1-Mlh3 complex; however, it conferred a null phenotype for *MLH3* functions in MMR and meiotic CO assays, and the Mlh1-*mlh3-D523N* complex is defective for endonuclease activity [5, 16].

The SK1 *MLH3*, *mlh3Δ*, *mlh3-23*, *mlh3-32*, and *mlh3-D523N* alleles were analyzed in the S288c/YJM789 hybrid background (Figure 2.6A; Table 2.1; Table 2.2; Methods; [64, 66]). To avoid genetic incompatibilities between Mlh1 and Mlh3, we analyzed SK1 *MLH3*, SK1 *mlh3* mutations, and *mlh3Δ* in the presence of SK1 *MLH1* (Methods). The spore viabilities of *mlh3-23* (84%), *mlh3-32* (82%), and *mlh3-D523N* (82%) were similar to *mlh3Δ* (80%) and the wild-type hybrid containing *SK1-MLH1/3* (82%; Table 2.7). Why do wild-type and *mlh3* strains show similar viability in the S288c/YJM789 hybrid? *mlh3Δ* mutants display a range of spore viabilities (70 to 92%) that appear to depend on strain background [5, 6, 67]. This is likely to be a partial explanation; however, another study suggested that sequence divergence present in the hybrid strains can affect spore viability through mismatch repair or rejection mechanisms that act on heteroduplex DNA formed during genetic recombination [68].

Figure 2.6. Genome-wide increase in simple noncrossover events (E1) compared to wild-type in *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ* mutants.

A. Generation of S288c/YJM789 isogenic strains with SK1 *MLH1*, *MLH3* and the *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ* mutant alleles (Methods). B. Cartoon description of simple NCO (E1; 3:1 tract on one chromatid, not within 5 kb of another CO or NCO), simple CO (E2; CO with or without an associated gene conversion (GC) tract, and not within 5 kb of another CO or NCO), simple CO with discontinuous gene conversion tracts (E3; same definition as for E2, except with one or more gene conversions within 5 kb and on one of the same chromatids as the CO chromatid), and discontinuous NCOs (E4; two or more NCOs consecutively on one chromatid, with 2:2 marker segregation separating them) as presented in Oke *et al.* [34]. C. Crossover (CO, E2+E3) and noncrossover (NCO, E1) counts per meiosis for wild-type, *mlh3-23*, *mlh3-32*, *mlh3-D523N*, and *mlh3Δ*. The minimum, first quantile, median, third quantile and maximum count are indicated in the box plot. The ratio of CO to NCO events is presented above the box plots. The proficiency of the *mlh3* alleles in mismatch repair is shown as +, MMR proficient, or –, MMR deficient. D. Average number of simple NCO (E1) and NCO with discontinuous tract (E4) events per tetrad (+/- standard error). * $p \leq 0.05$ compared to wild-type (Table 2.7).

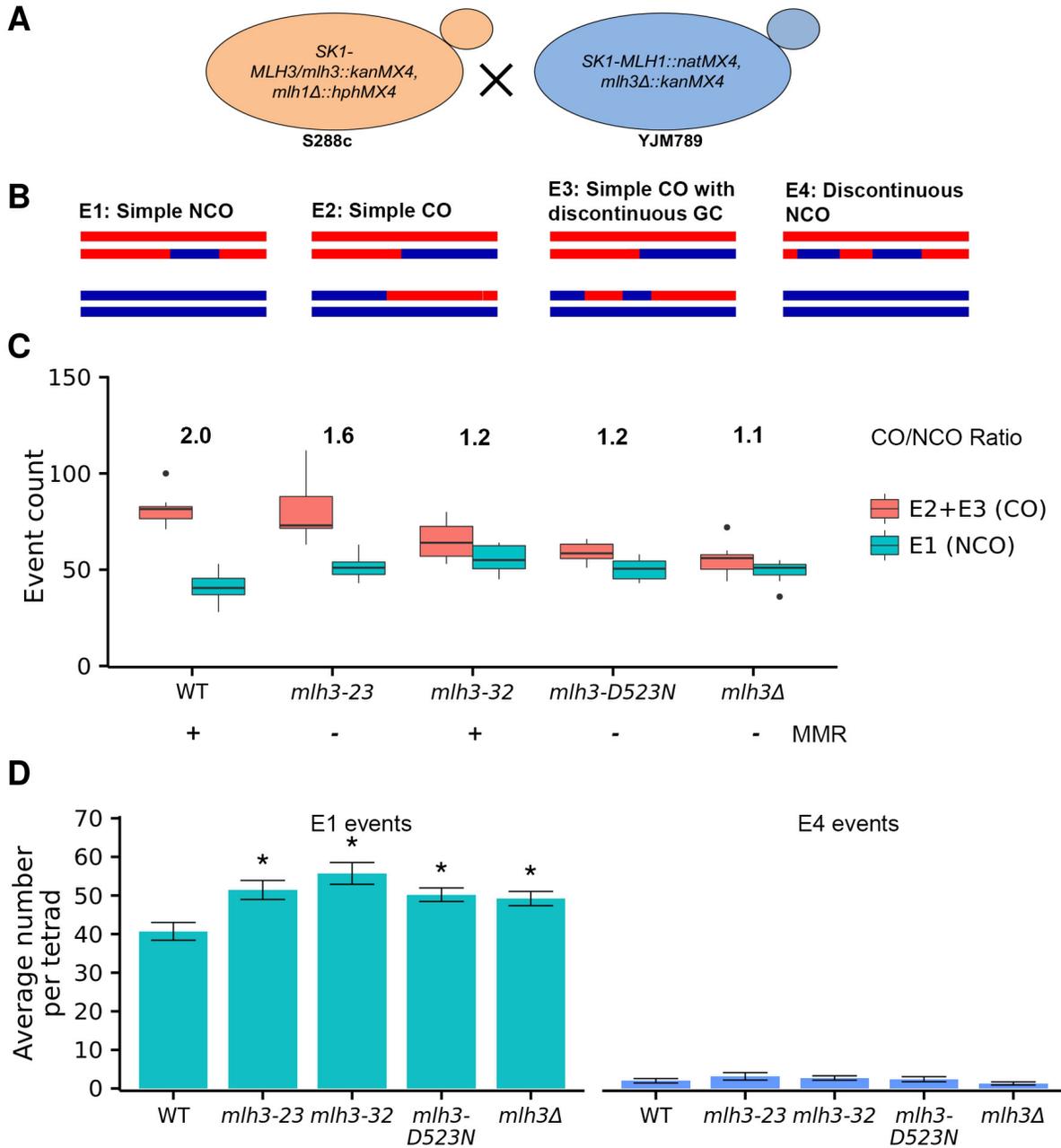


Table 2.7. Spore viability (SV), crossover (CO) and noncrossover (NCO) values for *mlh3-23*, *mlh3-32*, *mlh3-D523N*, and *mlh3Δ* mutants in the S288c/YJM789 hybrid.

Genotype: S288c x YJM789 with SK1- <i>MLH1</i>	N	% SV	Tetrads genotyped	Avg. CO (E2 + E3) ± SEM (Median)	Avg. NCO (E1) ± SEM (Median)	Avg. NCO (E4) ± SEM (Median)
SK1- <i>MLH3</i>	168	82	10	81±3 (82)	41±2 (41)	2.0±0.6 (1.0)
SK1- <i>mlh3-23</i>	80	84	7	81±6 (73)	51±3 (51)	3.1±0.9 (3.0)
SK1- <i>mlh3-32</i>	80	82	7	65±4 (64)	56±3 (55)	2.7±0.6 (3.0)
SK1- <i>mlh3-D523N</i>	88	82	10	59±2 (59)	50±2 (51)	2.4±0.7 (2.0)
<i>mlh3Δ</i>	80	80	10	55±3 (56)	49±2 (51)	1.3±0.4 (1.5)

N = number of tetrads dissected for measuring spore viability (SV). The E classes were all analyzed in groupEvents. E2 + E3 = total crossovers; E1 = simple noncrossover and E4 = discontinuous noncrossover. SEM, standard error of the mean.

Seven four-viable spore tetrads each of *mlh3-23* and *mlh3-32*, and ten four-viable spore tetrads each of *mlh3-D523N*, *mlh3Δ*, and *MLH3* were used for sequencing (Table 2.8). The sequence data for all the above *mlh3* point and null mutants along with the control strain are available from the National Centre for Biotechnology Information Sequence Read Archive under the accession numbers SRP096621 and SRP110341. The segregation of the SNPs in all 44 tetrads is shown in Appendix A. The CO and NCO counts for all tetrads and the average CO and NCO counts per chromosome are shown in Table 2.9 and Table 2.10, respectively.

As described below, high-resolution recombination mapping analysis using groupEvents [34] revealed overall patterns of average CO and NCO events that grouped *mlh3-23* close to wild-type, and *mlh3-D523N* and *mlh3-32* close to *mlh3Δ* (Figure 2.6). Importantly, this analysis showed that *mlh3* point and null mutants displayed, compared to wild-type, genome-wide increases in NCO events. We did not obtain any evidence that the number of DSBs increased in *mlh3* point mutants based on measuring the total number of interhomolog (IH) events (Figure 2.7A); such an increase would have provided a simple explanation for why an increase in NCO events was observed, though it is important to note that small changes in total events are challenging to interpret, partly because NCO events can escape detection due to a lack of a polymorphic marker, and we do not have a good estimate of how many such invisible events exist, or how they could be altered by differences in conversion tract length. Thus, we recognize that analyzing total events does not provide a particularly good proxy for DSB (e.g. [69]) levels. Together, these data provide evidence for *mlh3* mutants altering the resolution of meiotic recombination intermediates (see Discussion).

Table 2.8. Sequencing statistics for spores derived from wild type, *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3* Δ in S288c/YJM789 hybrid bearing SK1-*MLH1/MLH3* alleles.

Spore ID	SNPs genotyped	Coverage Depth (X)
<i>wt_2a</i>	73593	66.40
<i>wt_2b</i>	74058	79.95
<i>wt_2c</i>	73682	70.88
<i>wt_2d</i>	73565	73.79
<i>wt_5a</i>	71120	31.09
<i>wt_5b</i>	74269	71.53
<i>wt_5c</i>	71985	40.23
<i>wt_5d</i>	71278	39.96
<i>wt_6a</i>	73967	31.05
<i>wt_6b</i>	73838	39.35
<i>wt_6c</i>	73779	40.65
<i>wt_6d</i>	72914	30.12
<i>wt_7a</i>	72106	17.09
<i>wt_7b</i>	70109	10.83
<i>wt_7c</i>	71589	12.98
<i>wt_7d</i>	73637	48.78
<i>wt_8a</i>	73934	26.02
<i>wt_8b</i>	72929	27.20
<i>wt_8c</i>	72853	23.29
<i>wt_8d</i>	72057	20.39
<i>wt_9a</i>	72729	25.17
<i>wt_9b</i>	73773	34.92
<i>wt_9c</i>	73366	29.59
<i>wt_9d</i>	72637	27.90
<i>wt_10a</i>	73326	88.63
<i>wt_10b</i>	73176	25.16
<i>wt_10c</i>	73198	30.37
<i>wt_10d</i>	73612	22.72
<i>wt_11a</i>	73020	40.30
<i>wt_11b</i>	73309	43.98
<i>wt_11c</i>	73553	31.89
<i>wt_11d</i>	73034	37.50
<i>wt_12a</i>	72904	30.86
<i>wt_12b</i>	73006	20.06
<i>wt_12c</i>	74485	29.35
<i>wt_12d</i>	74593	28.81
<i>wt_13a</i>	73772	39.24
<i>wt_13b</i>	73934	46.79
<i>wt_13c</i>	73051	37.72
<i>wt_13d</i>	73012	29.61
<i>mlh3_23_6a</i>	72821	26.00

Table 2.8 (continued)

<i>mlh3_23_6b</i>	73731	30.46
<i>mlh3_23_6c</i>	74242	74.35
<i>mlh3_23_6d</i>	73587	38.45
<i>mlh3_23_7a</i>	73307	31.01
<i>mlh3_23_7b</i>	72910	30.35
<i>mlh3_23_7c</i>	74113	26.86
<i>mlh3_23_7d</i>	73731	33.87
<i>mlh3_23_8a</i>	73652	42.29
<i>mlh3_23_8b</i>	73500	23.72
<i>mlh3_23_8c</i>	73523	35.44
<i>mlh3_23_8d</i>	73019	27.51
<i>mlh3_23_9a</i>	72890	24.92
<i>mlh3_23_9b</i>	72762	39.19
<i>mlh3_23_9c</i>	73009	23.54
<i>mlh3_23_9d</i>	73289	28.28
<i>mlh3_23_10a</i>	74434	34.56
<i>mlh3_23_10b</i>	71249	13.69
<i>mlh3_23_10c</i>	72944	27.91
<i>mlh3_23_10d</i>	73138	36.67
<i>mlh3_23_11a</i>	74791	44.61
<i>mlh3_23_11b</i>	72578	39.03
<i>mlh3_23_11c</i>	73377	30.72
<i>mlh3_23_11d</i>	74015	35.07
<i>mlh3_23_12a</i>	72112	20.22
<i>mlh3_23_12b</i>	72717	19.77
<i>mlh3_23_12c</i>	72680	23.84
<i>mlh3_23_12d</i>	72243	21.08
<i>mlh3_32_6a</i>	73361	36.57
<i>mlh3_32_6b</i>	73183	37.66
<i>mlh3_32_6c</i>	72332	24.60
<i>mlh3_32_6d</i>	71679	22.96
<i>mlh3_32_7a</i>	72832	198.87
<i>mlh3_32_7b</i>	71095	11.60
<i>mlh3_32_7c</i>	73330	31.58
<i>mlh3_32_7d</i>	73238	26.67
<i>mlh3_32_8a</i>	72758	217.13
<i>mlh3_32_8b</i>	70392	29.44
<i>mlh3_32_8c</i>	70315	27.33
<i>mlh3_32_8d</i>	72879	149.71
<i>mlh3_32_9a</i>	74387	98.47
<i>mlh3_32_9b</i>	63895	45.17
<i>mlh3_32_9c</i>	72482	141.67
<i>mlh3_32_9d</i>	73950	204.92

Table 2.8 (continued)

<i>mlh3_32_11a</i>	73202	29.09
<i>mlh3_32_11b</i>	73770	27.74
<i>mlh3_32_11c</i>	72274	34.69
<i>mlh3_32_11d</i>	73102	31.94
<i>mlh3_32_12a</i>	73497	41.42
<i>mlh3_32_12b</i>	72933	31.95
<i>mlh3_32_12c</i>	73648	32.22
<i>mlh3_32_12d</i>	73847	32.89
<i>mlh3_32_13a</i>	73579	34.68
<i>mlh3_32_13b</i>	73554	34.46
<i>mlh3_32_13c</i>	74400	71.42
<i>mlh3_32_13d</i>	72551	20.67
<i>mlh3DN_1a</i>	74093	53.98
<i>mlh3DN_1b</i>	74212	54.27
<i>mlh3DN_1c</i>	74196	52.60
<i>mlh3DN_1d</i>	74018	78.47
<i>mlh3DN_2a</i>	73782	67.29
<i>mlh3DN_2b</i>	73893	98.45
<i>mlh3DN_2c</i>	73947	104.45
<i>mlh3DN_2d</i>	74717	131.11
<i>mlh3DN_3a</i>	74808	91.32
<i>mlh3DN_3b</i>	73523	53.45
<i>mlh3DN_3c</i>	73649	55.91
<i>mlh3DN_3d</i>	73680	51.84
<i>mlh3DN_4a</i>	73994	46.26
<i>mlh3DN_4b</i>	74042	63.36
<i>mlh3DN_4c</i>	74334	85.10
<i>mlh3DN_4d</i>	74398	51.04
<i>mlh3DN_5a</i>	74793	49.82
<i>mlh3DN_5b</i>	74596	60.02
<i>mlh3DN_5c</i>	74230	47.14
<i>mlh3DN_5d</i>	73957	58.53
<i>mlh3DN_6a</i>	74602	57.33
<i>mlh3DN_6b</i>	74329	72.08
<i>mlh3DN_6c</i>	74203	48.69
<i>mlh3DN_6d</i>	74376	77.95
<i>mlh3DN_7a</i>	74205	48.52
<i>mlh3DN_7b</i>	74246	58.80
<i>mlh3DN_7c</i>	74130	60.66
<i>mlh3DN_7d</i>	73308	65.19
<i>mlh3DN_8a</i>	73246	36.71
<i>mlh3DN_8b</i>	74400	55.49
<i>mlh3DN_8c</i>	74116	78.77

Table 2.8 (continued)

<i>mlh3DN_8d</i>	73848	62.79
<i>mlh3DN_9a</i>	74698	64.18
<i>mlh3DN_9b</i>	74851	76.91
<i>mlh3DN_9c</i>	73978	50.63
<i>mlh3DN_9d</i>	73810	66.36
<i>mlh3DN_10a</i>	73932	66.84
<i>mlh3DN_10b</i>	74272	89.49
<i>mlh3DN_10c</i>	74783	91.61
<i>mlh3DN_10d</i>	74421	78.58
<i>mlh3null_1a</i>	72746	23.75
<i>mlh3null_1b</i>	71493	66.08
<i>mlh3null_1c</i>	73461	27.09
<i>mlh3null_1d</i>	73062	36.04
<i>mlh3null_2a</i>	73687	30.53
<i>mlh3null_2b</i>	73350	28.38
<i>mlh3null_2c</i>	70835	70.07
<i>mlh3null_2d</i>	71034	18.25
<i>mlh3null_3a</i>	72707	26.01
<i>mlh3null_3b</i>	73321	29.99
<i>mlh3null_3c</i>	74005	37.19
<i>mlh3null_3d</i>	72088	24.08
<i>mlh3null_4a</i>	74052	32.98
<i>mlh3null_4b</i>	72938	49.10
<i>mlh3null_4c</i>	74692	52.23
<i>mlh3null_4d</i>	73792	32.23
<i>mlh3null_5a</i>	73803	31.86
<i>mlh3null_5b</i>	74048	33.79
<i>mlh3null_5c</i>	72848	33.27
<i>mlh3null_5d</i>	72772	21.49
<i>mlh3null_6a</i>	71196	69.15
<i>mlh3null_6b</i>	73701	38.79
<i>mlh3null_6c</i>	71783	76.77
<i>mlh3null_6d</i>	71928	19.94
<i>mlh3null_7a</i>	73097	36.67
<i>mlh3null_7b</i>	72210	36.42
<i>mlh3null_7c</i>	75003	60.87
<i>mlh3null_7d</i>	73297	35.11
<i>mlh3null_8a</i>	73597	44.31
<i>mlh3null_8b</i>	73050	56.57
<i>mlh3null_8c</i>	73092	40.96
<i>mlh3null_8d</i>	73375	41.63
<i>mlh3null_9a</i>	73689	38.75
<i>mlh3null_9b</i>	72537	41.30

Table 2.8 (continued)

<i>mlh3null_9c</i>	73632	36.13
<i>mlh3null_9d</i>	73622	34.73
<i>mlh3null_10a</i>	74117	34.20
<i>mlh3null_10b</i>	71858	30.54
<i>mlh3null_10c</i>	73608	39.18
<i>mlh3null_10d</i>	73109	36.95

Table 2.9. Crossovers (CO) and noncrossovers (NCO) in tetrads of wild-type, *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ*.

Tetrad ID	CO (E2+E3)	NCO (E1)
<i>wt 2</i>	73	53
<i>wt 5</i>	82	43
<i>wt 6</i>	71	37
<i>wt 7</i>	100	47
<i>wt 8</i>	78	37
<i>wt 9</i>	76	44
<i>wt 10</i>	83	46
<i>wt 11</i>	81	28
<i>wt 12</i>	85	34
<i>wt 13</i>	82	38
<i>mlh3 23 6</i>	92	63
<i>mlh3 23 7</i>	73	53
<i>mlh3 23 8</i>	63	43
<i>mlh3 23 9</i>	84	49
<i>mlh3 23 10</i>	73	55
<i>mlh3 23 11</i>	112	46
<i>mlh3 23 12</i>	70	51
<i>mlh3 32 6</i>	64	51
<i>mlh3 32 7</i>	72	45
<i>mlh3 32 8</i>	58	64
<i>mlh3 32 9</i>	56	64
<i>mlh3 32 11</i>	53	50
<i>mlh3 32 12</i>	73	55
<i>mlh3 32 13</i>	80	61
<i>mlh3DN 1</i>	61	45
<i>mlh3DN 2</i>	51	53
<i>mlh3DN 3</i>	58	49
<i>mlh3DN 4</i>	66	43
<i>mlh3DN 5</i>	58	55
<i>mlh3DN 6</i>	66	58
<i>mlh3DN 7</i>	55	57
<i>mlh3DN 8</i>	64	44
<i>mlh3DN 9</i>	52	46
<i>mlh3DN 10</i>	59	52
<i>mlh3null 1</i>	44	54
<i>mlh3null 2</i>	36	56
<i>mlh3null 3</i>	47	46
<i>mlh3null 4</i>	53	49
<i>mlh3null 5</i>	48	44
<i>mlh3null 6</i>	55	57

Table 2.9 (continued)

<i>mlh3null</i> 7	55	58
<i>mlh3null</i> 8	52	60
<i>mlh3null</i> 9	50	72
<i>mlh3null</i> 10	52	56

Crossovers are E2+E3 events and noncrossovers are E1 events [34].

Table 2.10. Average crossovers (CO) and noncrossovers (NCO) per chromosome for WT (wild type), *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ* mutants.

Chromosome	WT	<i>mlh3-23</i>	<i>mlh3-32</i>	<i>mlh3-D523N</i>	<i>mlh3Δ</i>
I	1.6	1.4	1.4	1.1	1.7
II	4.3	5.9	5.0	4.7	3.4
III	3.1	3.0	1.7	2.4	1.9
IV	9.5	8.6	8.1	6.1	6.3
V	3.4	4.4	2.9	2.2	2.5
VI	2.3	2.1	1.9	2.0	1.6
VII	7.0	8.0	4.0	5.3	4.4
VIII	4.7	4.1	3.1	2.7	3.1
IX	3.2	2.9	3.0	2.5	2.3
X	5.3	4.7	3.1	3.1	3.3
XI	4.7	5.0	4.0	3.8	3.5
XII	7.3	7.0	5.3	5.2	4.8
XIII	6.5	5.0	6.0	4.3	4.3
XIV	4.6	5.3	4.1	3.7	2.8
XV	7.5	7.0	5.7	4.8	5.1
XVI	6.1	6.6	5.7	5.1	4.2
Total_CO	81.0	81.0	65.0	59.0	55.0

Table 2.10 (continued)

Chromosome	WT	<i>mlh3-23</i>	<i>mlh3-32</i>	<i>mlh3-D523N</i>	<i>mlh3Δ</i>
I	1.2	1.3	1.0	1.0	1.4
II	4.0	4.1	4.0	4.1	3.3
III	0.8	0.7	2.1	1.9	0.9
IV	4.5	7.1	7.3	5.9	5.6
V	2.6	3.3	2.3	3.7	2.9
VI	1.3	1.4	1.1	0.8	1.4
VII	4.3	4.0	5.9	4.1	5.9
VIII	1.4	1.4	3.1	2.3	1.8
IX	1.0	1.9	2.0	1.9	1.7
X	1.9	3.6	3.7	3.1	3.5
XI	1.8	2.1	3.3	2.4	2.1
XII	2.5	3.9	5.7	4.1	4.3
XIII	3.3	5.7	3.7	3.2	3.4
XIV	3.2	2.6	3.0	2.9	2.9
XV	3.4	4.9	4.1	6.1	4.8
XVI	3.5	3.4	3.3	2.7	3.3
Total_NCO	41.0	51.0	56.0	50.0	49.0

Crossovers are E2+E3 events and noncrossovers are E1 events [34].

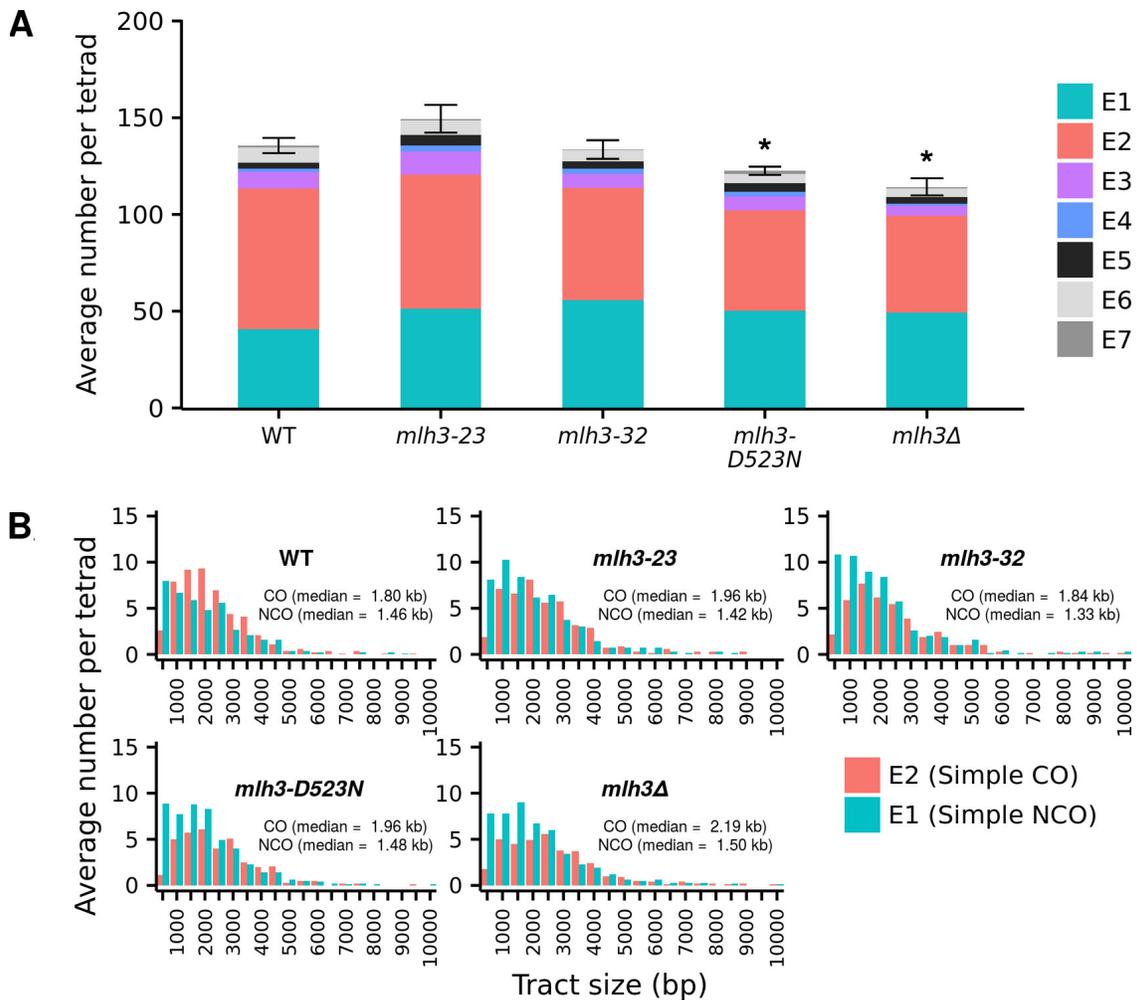


Figure 2.7. Total interhomolog events and distribution of gene conversion tract lengths associated with NCO and CO events in wild-type, *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ* mutants.

A. Total average inter-homolog events (IH; Table 2.11), and each event type (E1-E7) as a fraction of the total, in *MLH3*, *mlh3-23*, *mlh3-32*, *mlh3-D523N*, and *mlh3Δ* mutants. * $p \leq 0.05$ compared to wild-type. B. Average number of tracts ordered by size for simple CO (E2) and simple NCO (E1) events. Median CO and NCO tract sizes are also presented, and events were assigned as described by Oke et al. [34]. The Y-axis represents pooled data from all tetrads, normalized by dividing the number of tetrads.

i. *mlh3-23*, *mlh3-32*, and *mlh3-D523N* display distinct CO phenotypes. We used groupEvents ([34]; 5 kb threshold for merging nearby events) to categorize all IH events, focusing primarily on the major CO (E2, E3) and NCO events (E1; Figure 2.1; Figure 2.6B), but also examining minority NCO (E4) events that can provide information on whether events resulted from defective mismatch repair or multiple invasion intermediates (Figure 2.6; Figure 2.7; Figure 2.8; Table 2.7; Table 2.11; Methods).

Wild-type and *mlh3-23* displayed the same average number of CO events (E2+E3) ($p=0.99$), and *mlh3Δ*, *mlh3-D523N*, and *mlh3-32* displayed CO levels lower than wild-type ($p < 0.01$). *mlh3-D523N* and *mlh3Δ* displayed similar average CO levels ($p = 0.23$), but *mlh3-32* displayed levels higher than *mlh3Δ* ($p=0.05$) (Figure 2.6C and Table 2.7). An analysis of average CO counts per chromosome was consistent with these observations; *mlh3-23* had a CO distribution similar to wild-type, and *mlh3-32* and *mlh3-D523N* had distributions similar to *mlh3Δ* (Figure 2.9). Significant reduction in crossovers were observed primarily on medium (0.5 to 0.9 Mb-II, V, VIII, X, XI, XIV) and large (>0.9 Mb-IV, VII, XII, XIII, XV, XVI) chromosomes in *mlh3Δ* (III, VIII, X, XIV, XIII, XVI, XII, VII, XV, IV), *mlh3-D523N* (VIII, X, XIII, VII, XV, IV) and *mlh3-32* (X, VII; Figure 2.9B; [70]). Together, these observations are consistent with the analysis of map distances at specific loci, though the *mlh3-23* mutant displayed an improved CO phenotype in the S288c/YJM789 background compared to the SK1 background (Figure 2.4; Figure 2.5; Figure 2.6C).

The median gene conversion tract lengths associated with COs (E2 events; Figure 2.7B; Figure 2.8; Table 2.12) were significantly longer ($p \leq 0.03$) for *mlh3Δ* (2.19 kb) compared to *MLH3* (1.79 kb) and *mlh3-32* (1.84 kb), but were not significantly longer ($p \geq 0.18$) for *mlh3-23* (1.96 kb), and *mlh3-D523N* (1.96 kb). We found this interesting because *MLH3* and *mlh3-32*

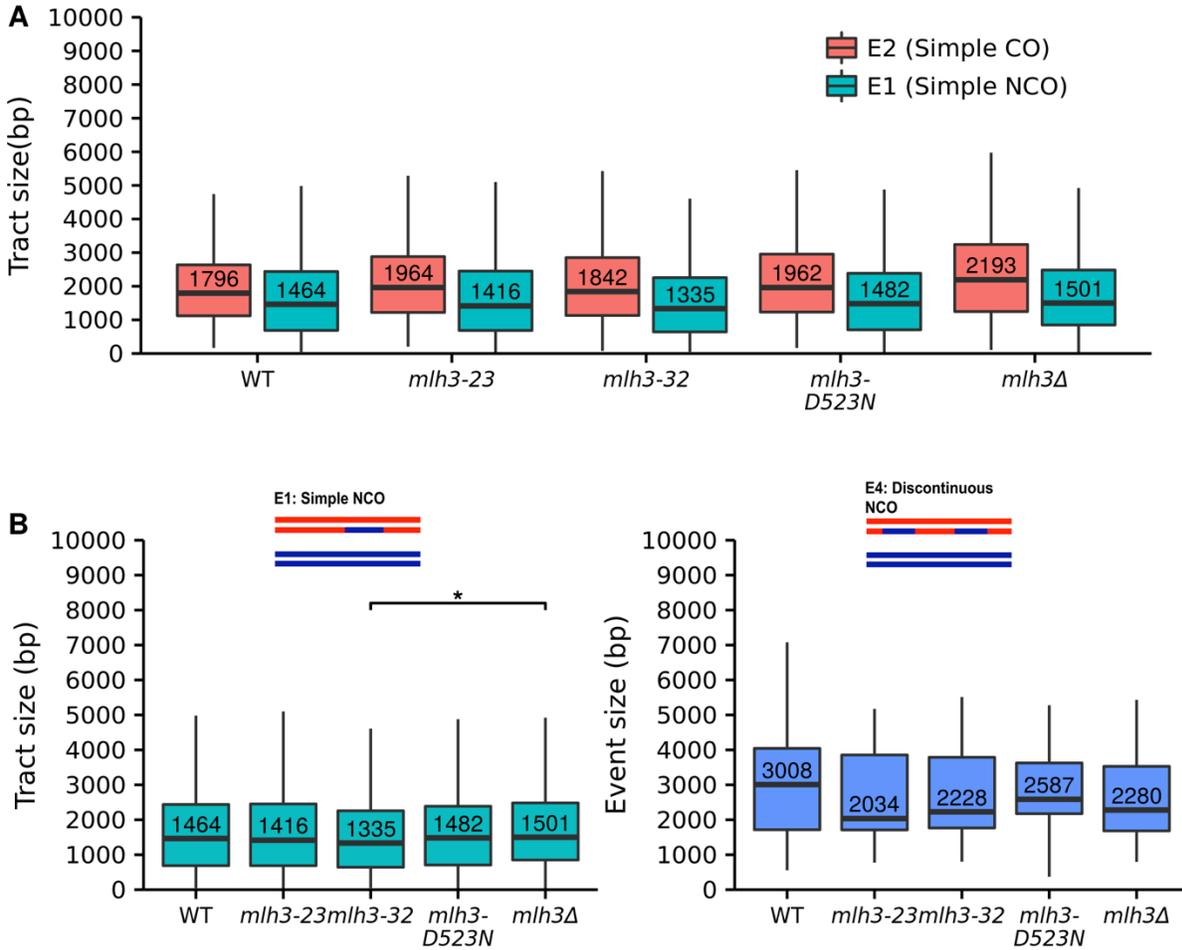


Figure 2.8. Distribution of gene conversion tract lengths associated with crossovers and noncrossovers (NCOs) for wild-type, *mlh3-23*, *mlh3-32*, *mlh3-D523N*, and *mlh3Δ*.

A. Distribution of gene conversion tract lengths associated with simple NCO (E1) and CO (E2) events. B. Distribution of gene conversion tract lengths associated with simple NCO (E1) and discontinuous NCO (E4) events. * $p \leq 0.05$ comparing *mlh3Δ* to *mlh3-32*. In A and B the minimum, first quantile, median, third quantile, and maximum count are indicated in the box plot. Outlier points are not shown.

Table 2.11. Summary of all event categories, as defined in Oke et al. [34], in wild-type (WT), *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ*.

	WT	<i>mlh3-23</i>	<i>mlh3-32</i>	<i>mlh3-D523N</i>	<i>mlh3Δ</i>
Number of tetrads	10.00	7.00	7.00	10.00	10.00
Total E1	407.00	360.00	390.00	502.00	492.00
E1 per tetrad	40.70	51.43	55.71	50.20	49.20
E1 length mean	1821.22	1804.54	1748.45	1811.19	1895.31
E1 length median	1464.50	1415.75	1335.00	1482.50	1501.00
Total E2	728.00	485.00	407.00	519.00	501.00
E2 per tetrad	72.80	69.29	58.14	51.90	50.10
E2 length mean	1418.21	1530.45	1512.03	1570.09	1775.29
E2 length median	1166.00	1146.50	1153.00	1263.00	1433.50
E2 length(>0) median	1795.00	1964.00	1842.00	1962.00	2192.00
Total E3	83.00	82.00	49.00	71.00	51.00
E3 per tetrad	8.30	11.71	7.00	7.10	5.10
E3 length mean	3154.80	3253.14	2565.66	3588.28	3396.47
E3 length median	3069.00	3036.25	2545.50	3163.00	2870.00
E3 track length mean	1254.65	1094.16	800.45	1079.31	1258.72
E3 tract length median	974.00	920.50	709.00	850.25	966.50
E3 GC length mean	669.32	772.35	641.97	741.70	542.78
E3 GC length median	340.50	304.25	189.00	331.00	0.00
E3 gap length mean	1352.59	1520.39	1247.16	1718.63	1474.08
E3 gap length median	870.50	1128.00	972.00	1488.25	1352.00
Total E4	20.00	22.00	19.00	24.00	13.00
E4 per tetrad	2.00	3.14	2.71	2.40	1.30
E4 length mean	3113.55	2924.32	2762.08	2961.08	2514.19
E4 length median	3008.00	2034.50	2227.50	2587.00	2279.50
E4 track length mean	1182.68	1006.43	806.79	810.73	907.46
E4 tract length median	802.00	649.50	661.50	493.50	553.50
E4 gap length mean	572.66	681.32	741.15	968.88	584.50
E4 gap length median	440.75	484.00	520.75	701.50	276.75
Total E5	29.00	39.00	27.00	45.00	34.00
E5 per tetrad	2.90	5.57	3.86	4.50	3.40
E5 length mean	3358.71	3685.97	3389.41	3616.73	3718.44
E5 length median	2438.50	3654.50	3305.00	3682.50	3364.00
Total E5A	17.00	16.00	16.00	23.00	21.00
E5A per tetrad	1.70	2.29	2.29	2.30	2.10
E5A length mean	2550.94	3695.19	3305.63	3260.50	3714.02
E5A length median	2408.00	3350.50	3211.25	3571.50	3360.00
Total E6	80.00	53.00	40.00	50.00	43.00
E6 per tetrad	8.00	7.57	5.71	5.00	4.30
E6 length mean	3724.04	4055.31	3863.69	3992.64	4050.65
E6 length median	3199.75	3663.00	3368.25	3520.50	3829.00
Total E7	10.00	5.00	3.00	15.00	8.00

Table 2.11 (continued)

E7 per tetrad	1.00	0.71	0.43	1.50	0.80
E7 length mean	3499.25	3163.50	4589.33	3299.13	5029.06
E7 length median	3832.00	2661.00	3392.50	3832.00	4572.75
Total JM	930.00	664.00	526.00	700.00	637.00
Total JM per tetrad	93.00	94.86	75.14	70.00	63.70
Total NCO	427.00	382.00	409.00	526.00	505.00
Total NCO per tetrad	42.70	54.57	58.43	52.60	50.50
Total IH	1357.00	1046.00	935.00	1226.00	1142.00
Total IH per tetrad	135.70	149.43	133.57	122.60	114.20

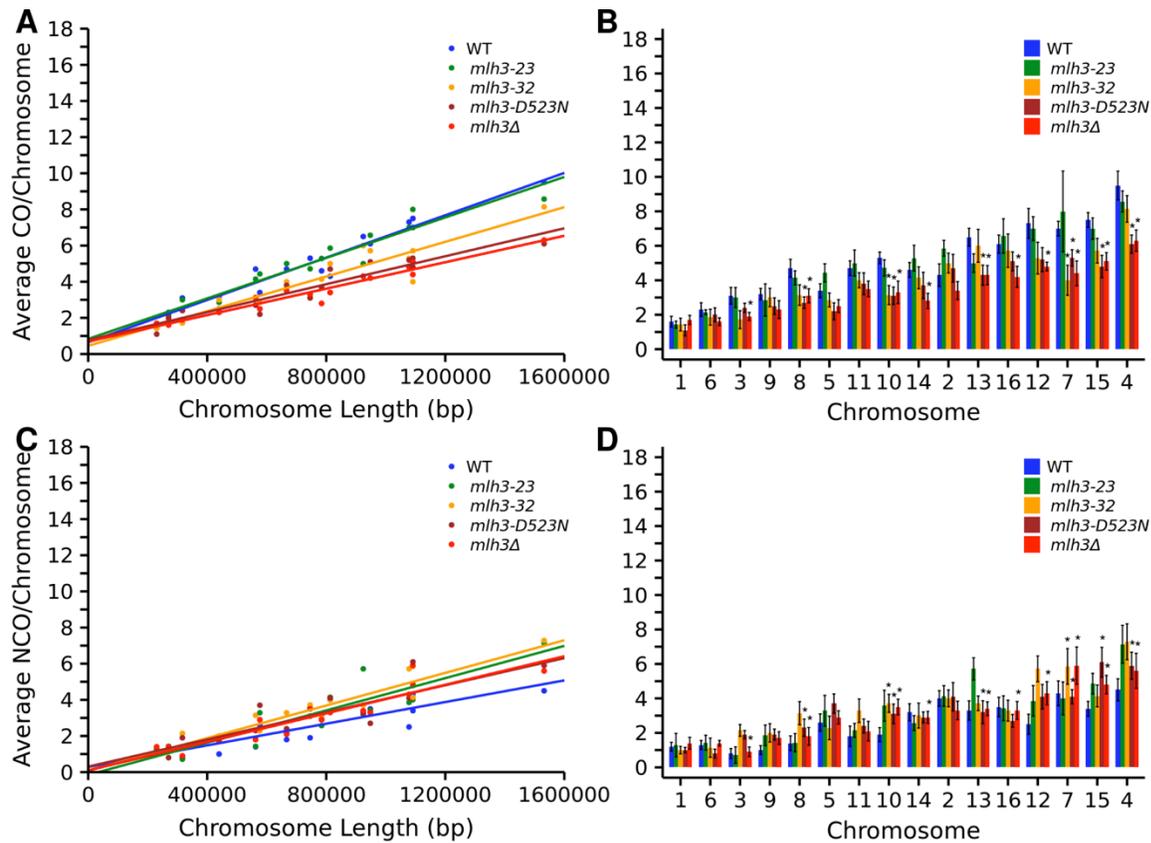


Figure 2.9. Crossover and noncrossover distribution on chromosomes for wild-type, *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ*.

A. and C. Scatter plot of average crossover (CO, E2+E3) and noncrossover (NCO, E1) counts per chromosome against chromosome size [34]. The equations for the regression lines are: wild-type (CO= 5.85×10^{-6} x chr. size + 0.65; NCO= 2.99×10^{-6} x chr. size + 0.28); *mlh3-23* (CO= 5.6×10^{-6} x chr. size + 0.84; NCO= 4.46×10^{-6} x chr. size - 0.15); *mlh3-32* (CO= 4.79×10^{-6} x chr. size + 0.45; NCO= 4.51×10^{-6} x chr. size + 0.08); *mlh3-D523N* (CO= 3.8×10^{-6} x chr. size + 0.77; NCO= 3.75×10^{-6} x chr. size + 0.31); *mlh3Δ* (CO= 3.65×10^{-6} x chr. size + 0.69; NCO= 3.94×10^{-6} x chr. size + 0.1).

B. and D. Bar plot of average crossover and noncrossover counts per chromosome. The asterisk symbol (*) marks chromosomes that have significant difference (two tailed t-test for difference in mean; $P < 0.05$) in crossover / noncrossover counts compared to wild-type. Chromosomes are arranged by size from left to right. Error bars are mean \pm standard error.

Table 2.12. Gene conversion tract lengths in *mlh3* Δ , wild-type, *mlh3-23*, *mlh3-32* and *mlh3-D523N* mutants.

Strain	E2 median (kb)	<i>P</i> value	E1 median (kb)	<i>P</i> value	E4 median (kb)	<i>P</i> value
<i>mlh3</i> Δ	2.19		1.50		2.28	
Wild-type	1.79	0.005	1.46	0.51	3.01	0.37
<i>mlh3-23</i>	1.96	0.18	1.42	0.37	2.03	0.69
<i>mlh3-32</i>	1.84	0.034	1.34	0.045	2.23	0.65
<i>mlh3-D523N</i>	1.96	0.49	1.48	0.43	2.59	0.40

E1, E2, E4 are classified by groupEvents [34]. P-values show statistical significance of difference in median CO and NCO gene conversion tract lengths of wild-type and *mlh3* alleles compared to *mlh3* Δ using the Wilcoxon rank sum test (Methods).

mutants appeared similarly functional for MMR, whereas *mlh3Δ*, *mlh3-23*, and *mlh3-D523N* all displayed MMR defects (Table 2.4). These observations encouraged us to test if Mlh1-Mlh3 MMR functions affected the processing of gene conversion tracts (see below and Discussion).

ii. Genome-wide increase in noncrossovers in *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ*

mutants. The average numbers of NCOs (E1 events) in *mlh3Δ* (49), *mlh3-23* (51), *mlh3-32* (56), and *mlh3-D523N* (50) were significantly higher than the wild-type value (41; t-test, $p = 0.01, 0.0067, 0.0012, 0.0044$, respectively; Table 2.7; Figure 2.6D; Figure 2.9C). This was unexpected because CO levels were similar to wild-type for *mlh3-23* mutants and similar to *mlh3Δ* for *mlh3-32* and *mlh3-D523N* mutants (see Discussion). Compared to *mlh3Δ*, the average number of NCO events in *mlh3-23*, *mlh3-32*, and *mlh3-D523N* was similar ($p = 0.48, 0.080, 0.70$, respectively). Significant increases in NCOs on chromosomes were observed for *mlh3-32* (X and VII), *mlh3-D523N* (VIII, X, XIII, VII, XV, IV), *mlh3Δ* (III, VIII, X, XIV, XIII, XVI, XII, VII, XV, IV). No increase in NCOs on specific chromosomes was observed for *mlh3-23* (Figure 2.9B).

It is possible that some of the increase in NCOs seen in *mlh3* mutants was due to the presence of discontinuous gene tracts that arose from the same initiating DSB. Such discontinuous tracts could have resulted from defective MMR but are not included in the E1 NCO class above. The E4 class (Figure 2.6B) is a minority discontinuous NCO event that is hypothesized to arise from defective MMR or multiple invasion intermediates [34]. As shown in Figure 2.6D, none of the E4 events in the *mlh3* mutants were significantly different from wild-type ($p=0.33$ to 0.80). This suggests that discontinuous tracts are unlikely to significantly contribute to the increased level of NCO events seen in *mlh3* mutants.

The average NCO (E1) tract lengths were similar to wild-type (1.46 kb) in all of the *mlh3* mutants analyzed (1.34 to 1.50 kb; $p = 0.23$ to 0.91 ; Figure 2.8). The NCO tract length distribution suggested an enrichment of short NCO tract lengths (<1 kb) in the *mlh3* mutants compared to wild-type (Figure 2.7B). We found this potentially interesting because Oke et al. [34] identified such an enrichment in *sgs1* mutants; they hypothesized that the short NCO class arose from joint molecules or joint molecule-like intermediates that were resolved by the activity of structure-selective nucleases (SSNs). However, both Kolmogorov–Smirnov and Wilcoxon Rank-Sum tests indicated that an apparent enrichment of short NCO tract lengths in *mlh3* mutants, compared to wild-type, was not significant. A similarly weak pattern (an insignificant increase in *mlh3* mutants compared to wild-type) was seen for the E5A class (Table 2.11, [34]), which is thought to represent double COs and is predicted to arise if joint molecule resolution is not biased (Class II COs resolved by SSNs; Figure 2.1).

One explanation for the increase in NCO events seen in *mlh3Δ*, *mlh3-23*, *mlh3-32* and *mlh3-D523N* mutants is that these mutants experienced meiotic progression delays that resulted in the continued accumulation of NCOs, possibly through increased DSB formation. Increases in NCO events and a meiotic delay were observed in *ndt80* and the ZMM *zip1*, *zip3* and *msh5* mutants as a result of impeding feedback circuits that inhibit DSB formation [27, 30, 49, 71, 72]. To investigate this possibility, we examined total interhomolog (IH) events, which consists of the average of all joint molecule (E2, E3, E5, E6, E7) and NCO (E1, E4) events, and serves as a rough estimate of double-strand break levels [34]. As shown in Figure 2.7A and Table 2.11, the IH value for wild-type was 135.7 events per tetrad. None of the *mlh3* mutants displayed a significantly higher value than wild-type, but *mlh3-D523N* (122.6 events, $p = 0.011$) and *mlh3Δ* (114.2 events, $p = 0.002$) displayed lower values.

We also examined meiotic progression in *MLH3*, *mlh3Δ*, *mlh3-32*, *mlh3-23* and *mlh3-D523N* SK1 strains by measuring the completion of the first meiotic division. This would be difficult to do in S288c/YJM789 strains because they do not show the highly synchronous and efficient meiotic progression profile seen in SK1. As shown in Figure 2.10, *MLH3*, *mlh3Δ*, *mlh3-32*, *mlh3-23*, and *mlh3-D523N* mutants showed similar kinetics for completion of at least the first meiotic division (MI+MII). These data suggest that the increase in NCO events in *mlh3Δ*, *mlh3-32*, *mlh3-23*, and *mlh3-D523N* cannot simply be explained by an increase in DSB levels, MMR defects, or a meiotic progression delay.

In summary, the increase in NCO events in *mlh3* mutants is consistent with a possible role for SSNs and/or the STR complex (see Discussion), and is unlikely to be explainable by an increase in DSBs or MMR defects.

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

We 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 (Figure 2.4C). As shown in Figure 2.11 and Figure 2.12, all three mutant complexes purified as heterodimers and display endonuclease activities similar to wild-type. When this work was initiated we thought that separation of function mutant complexes might show endonuclease defects

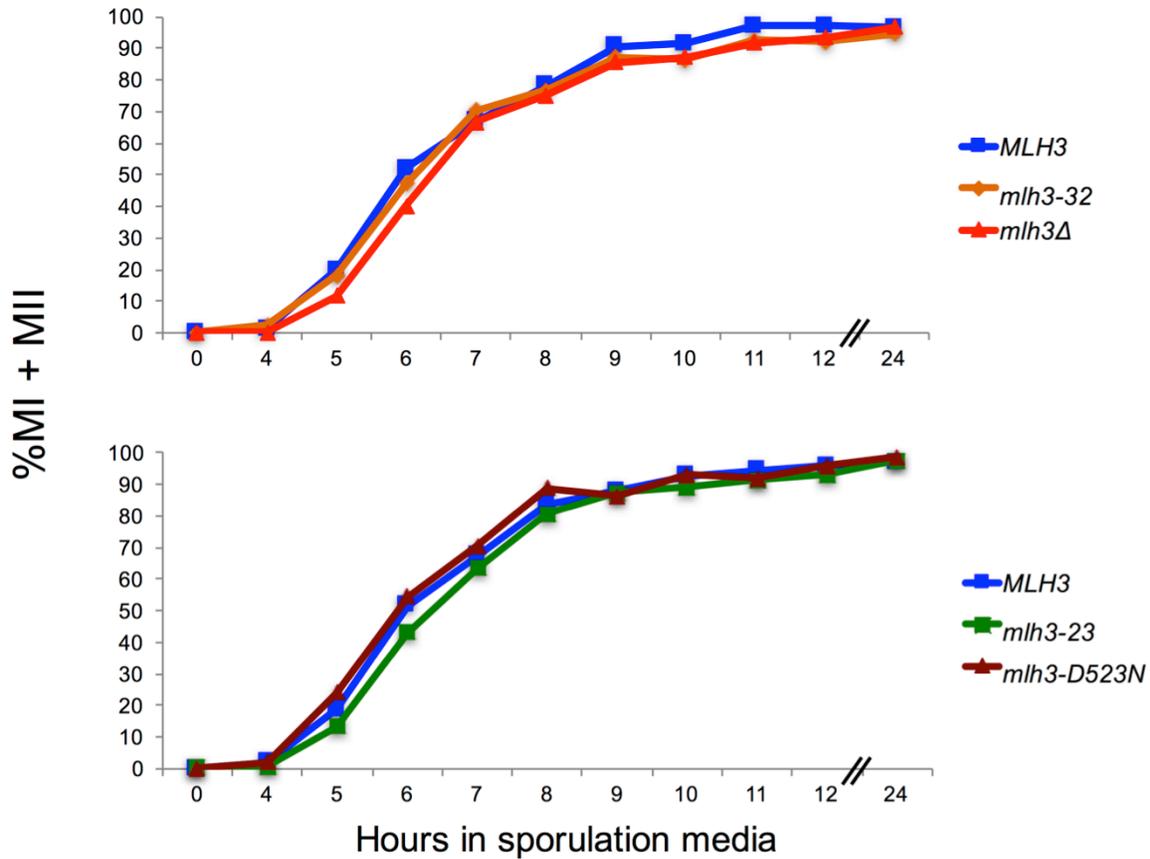


Figure 2.10. *mlh3* mutants display normal meiotic prophase progression as measured by the completion of the first meiotic division.

Representative time courses showing the completion of the MI division (MI+MII) in wild-type, *mlh3-23*, *mlh3-32*, *mlh3-D523N* and *mlh3Δ* strains. Cells with two, three, or four nuclei were counted as having completed MI (MI+MII). All strains for a single time course were grown in the same batch of media under identical conditions. Two independent transformants were measured per allele.

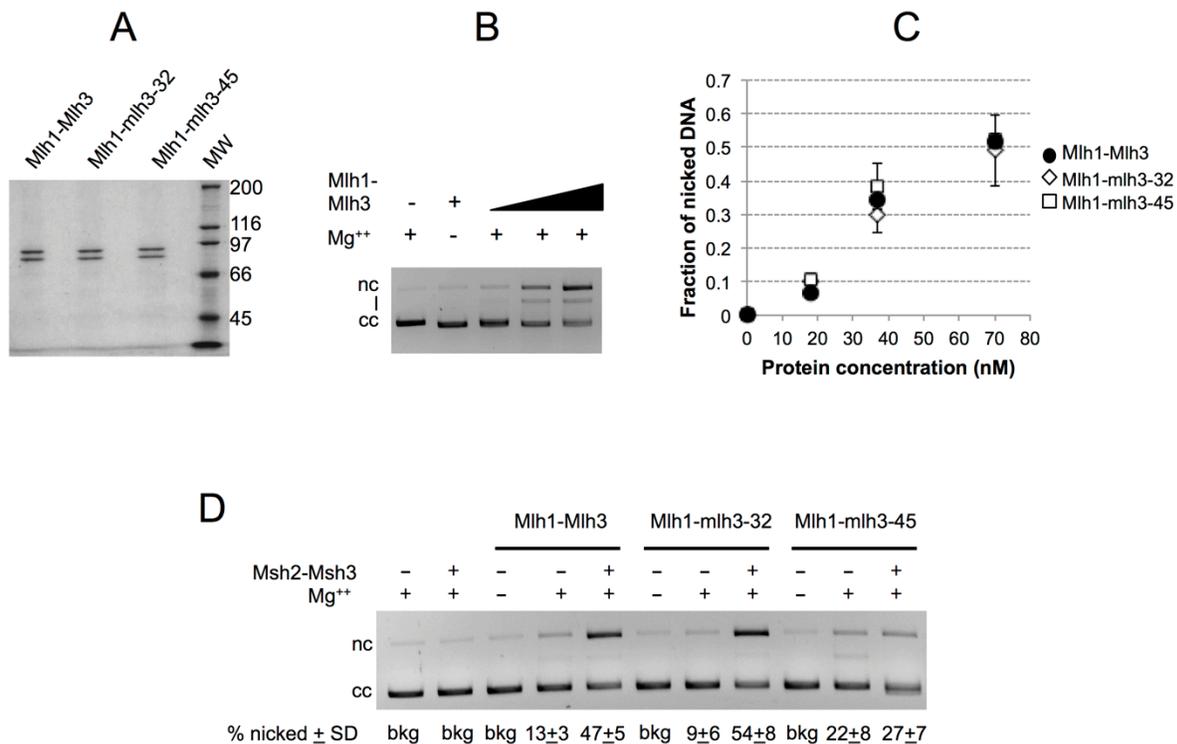


Figure 2.11. Mlh1-mlh3-32 and Mlh1-mlh3-45 display wild-type endonuclease activities that are differentially stimulated by Msh2-Msh3.

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 μ g 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 \pm -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 \pm -S.D. below the gel. (bkg) background, (cc) closed circular DNA, (nc) nicked DNA.

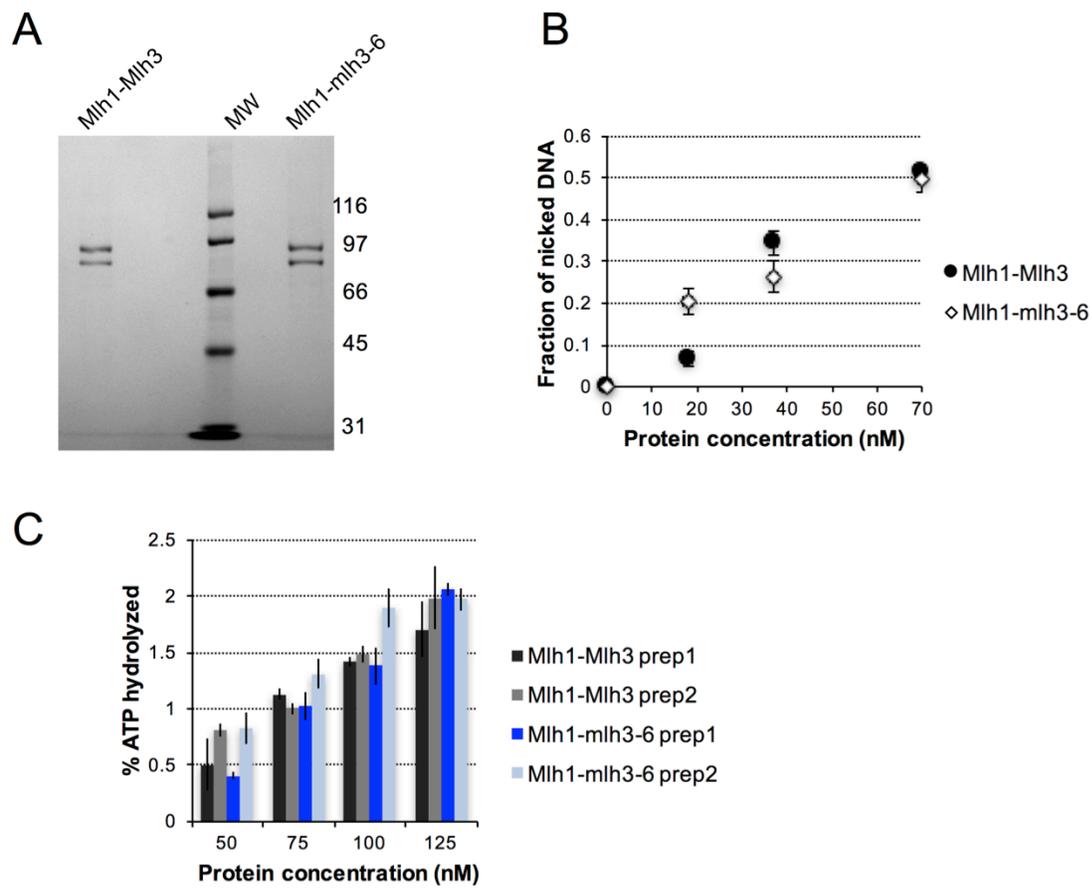


Figure 2.12. Mlh1-mlh3-6 exhibits wild-type endonuclease and ATPase activity.

A. SDS-PAGE analysis of purified Mlh1-Mlh3 and Mlh1-mlh3-6. Coomassie Blue R250-stained 8% Tris-glycine gel. 0.5 μ g 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. [16], but contained the indicated amounts of Mlh1-Mlh3 and Mlh1-mlh3-6 incubated with 100 μ M 32 P-g-ATP. Reactions were performed in duplicate for two separate purifications of each, and the average values, +/-SD, are presented.

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. Our 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 (Figure 2.3). We then tested whether the mutant complex displayed a defect in ATPase activity. As shown in Figure 2.12C, Mlh1-Mlh3 and Mlh1-*mlh3-6* displayed similar ATPase activities.

Because Mlh1-Mlh3's endonuclease activity is enhanced by Msh2-Msh3 [16], we 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 2.11D, 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. In a first step to test if *mlh3-32* mutants are defective in interactions with meiotic CO factors, we determined if the *mlh3-32* mutation is dominant; such a phenotype could provide hints on the nature of the *mlh3-32* meiotic defect. We mated EAY3552 (*mlh3-32*, *CEN8Tomato::LEU2*) to EAY3339 (relevant genotype *MLH3*, *THR1::m-Cerulean-TRP1*; Table 2.1; Table 2.2), and the sporulated progeny displayed a tetratype frequency similar to wild-type (40.9%, n = 252 tetrads), indicating that *mlh3-32* is recessive.

Sgs1 but not sgs1-hd overexpression differentially affects spore viability in *mlh3Δ* vs. *mlh3-32*.

As presented in Figure 2.1, the STR complex can act as both a negative and positive regulator of CO formation in meiotic prophase [8, 30, 31, 46, 49]. In its role as a negative regulator, STR is thought to prevent the formation of aberrant recombination structures by disassembling branched recombination intermediates to form early NCOs via synthesis dependent strand annealing (SDSA), or by re-forming the DSB intermediate. In its role as a pro-CO factor STR promotes stabilization of ZMM complexes on recombination intermediates, leading to the resolution of dHJs by an interference-dependent CO pathway (class I) that requires the Mlh1-Mlh3 endonuclease. In *sgs1Δ* mutants COs have been shown to be ZMM independent [73]. Strand invasion intermediates that escape STR disassembly are thought to be resolved as COs or NCOs using an alternative interference-independent CO pathway (class II) that involves the SSNs Mus81-Mms4, Yen1, and Slx1-Slx4.

The increase in NCO events seen in *mlh3* mutants is suggestive of their formation through SSN or STR activities. To test for genetic interactions between *SGS1* and *MLH3*, we expressed *SGS1* via its native promoter on a 2μ multi-copy vector. Sgs1 overexpression enhanced the *mlh3Δ* spore viability defect (Figure 2.13: 76% in *mlh3Δ+2μ* vs. 69% in *mlh3Δ+SGS1-2μ*; $p < 0.05$, χ^2 test) and conferred a more apparent MI nondisjunction pattern (an excess of 4, 2, 0 viable spore tetrads compared with 3 and 1 viable tetrads ([6]; Figure 2.13). In contrast, Sgs1 overexpression modestly suppressed the *mlh3-32* spore viability defect (Figure 2.13: 76% in *mlh3-32+2μ* vs. 84% in *mlh3-32+SGS1-2μ*; $p < 0.05$, χ^2 test). Suppression or enhancement of phenotypes was not observed upon overexpression of the *sgs1-K706A* helicase defective mutant protein (*sgs1-hd*), indicating that these effects were dependent on Sgs1 helicase

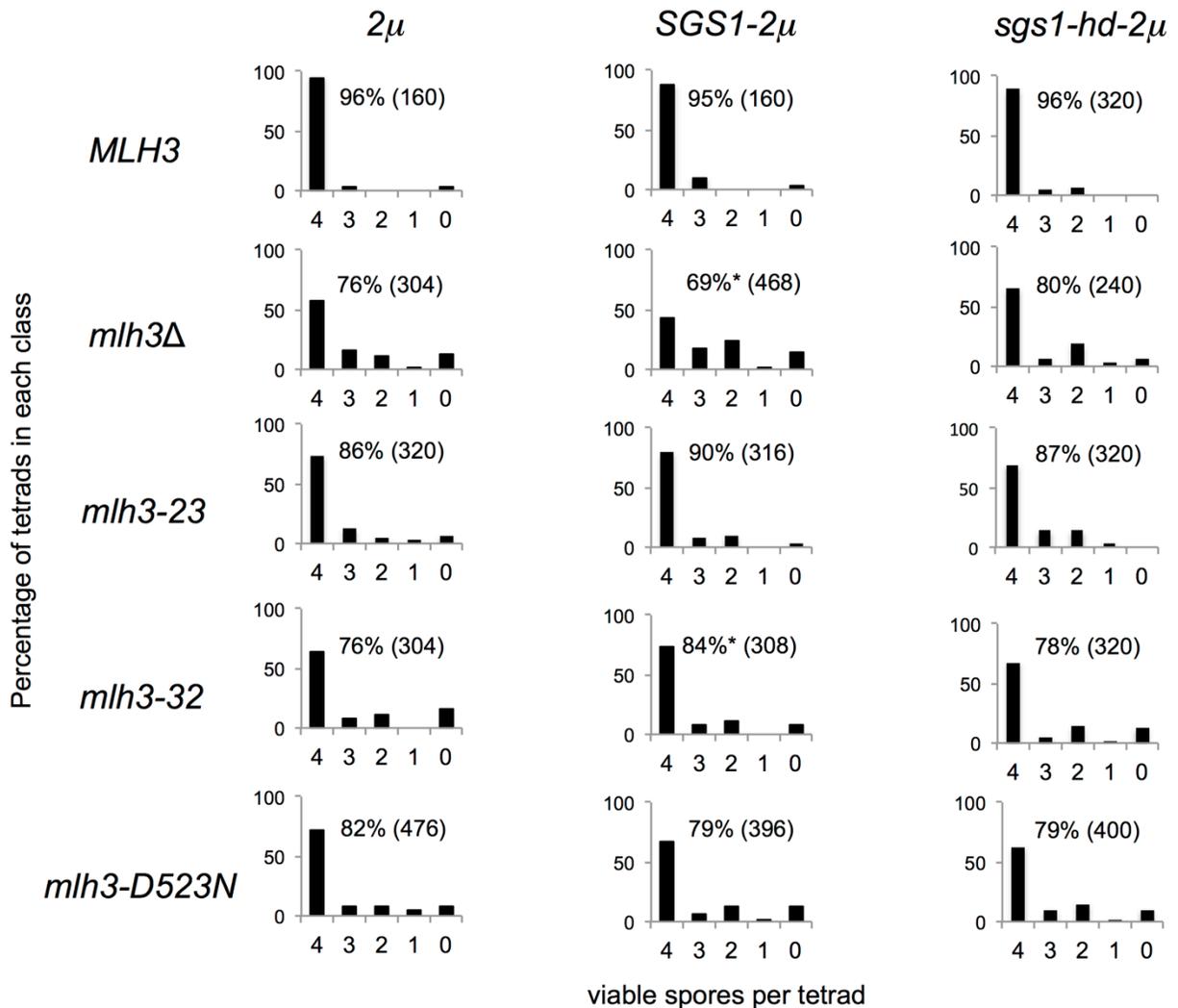


Figure 2.13. Sgs1 but not sgs1-hd overexpression differentially affects spore viability in *mlh3Δ* vs. *mlh3-32*.

Distribution of viable spores in tetrads of *MLH3*, *mlh3Δ*, *mlh3-23*, *mlh3-32*, and *mlh3-D523N* strains containing no insert (2μ), *pSGS1-2μ*, or *psgs1-hd-2μ* (helicase defective mutant). In all plots, the horizontal axis corresponds to the classes of tetrads with 4, 3, 2, 1 and 0 viable spores, and the vertical axis corresponds to the frequency of each class given in percentage. The overall spore viability (SV) and the total number of spores counted (n) are shown. * $p < 0.05$ (χ^2 test), comparing spore viability for the indicated strain transformed with a 2μ *SGS1* plasmid to that of the same strain transformed with a 2μ no insert plasmid.

activity (Figure 2.13: 80% in *mlh3Δ+sgs1-hd-2μ* and 78% in *mlh3-32+sgs1-hd-2μ*). In addition, no effect was seen upon overexpression of Sgs1 or *sgs1-hd* in *MLH3* (Figure 2.13: 96% in *MLH3+2μ* vs. 95% in *MLH3+SGS1-2μ* and 96% in *MLH3+sgs1-hd-2μ*), *mlh3-23* (86% in *mlh3-23+2μ* vs. 90% in *mlh3-23+SGS1-2μ* and 87% in *mlh3-23+sgs1-hd-2μ*) and *mlh3-D523N* (82% in *mlh3-D523N+2μ* vs. 79% in *mlh3-D523N+SGS1-2μ*, and 79% in *mlh3-D523N+sgs1-hd-2μ*) strains. Although *mlh3-32* and *mlh3-D523N* both confer null-like meiotic defects, they did not confer phenotypes similar to *mlh3Δ* in response to Sgs1 overexpression. This observation suggests potential roles for Mlh1-Mlh3 that are independent of its enzymatic activity.

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; Mus81-Mms4, Slx1-Slx4, and Yen1) in terms of homology and intrinsic behavior *in vitro* (reviewed in [50]). 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, 50, 74]. 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 [74]. 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 (Table 2.6; Figure 2.11). Mlh1-Mlh3 has been shown genetically to act downstream of Msh4-Msh5 [40, 41, 44, 47]; this order of events is analogous to steps in DNA MMR where MLH acts following MSH recognition [14, 75]. 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 [50]). 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 act in concert to strengthen a possibly weakened Msh4-Msh5-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.

Mlh1-Mlh3 in DNA mismatch repair

In MMR the asymmetric loading of PCNA by the RFC complex is thought to direct the endonuclease activity of MutLa (Mlh1-Pms1 in *S. cerevisiae*, MLH1-PMS2 in humans) to act in

strand-specific repair [76]. Additional studies suggest that specific protein-protein interactions influence and activate MLH endonuclease activity, and direct nicking to a specific location. For example, *in vitro* studies performed with yeast proteins showed that RFC-loaded PCNA can activate Mlh1-Pms1 but not Mlh1-Mlh3 endonuclease on circular plasmids (Mlh3 lacks a PCNA binding motif present in Pms1; [15-17; 74]. Our finding that endonuclease active site residues are highly conserved between Mlh1-Mlh3 and Mlh1-Pms1, which has no role in meiotic crossing over, suggests that the different functions of the two complexes are a result of the different protein-protein interactions.

Data from the Crouse lab suggest that Mlh1-Mlh3 acts in conjunction with Mlh1-Pms1 in Msh2-Msh3 dependent MMR [4]. This observation helps address how Mlh1-Mlh3 is involved in MMR in the absence of a PCNA interaction. In one scenario consistent with the above, Mlh1-Mlh3 is recruited and activated by Msh2-Msh3, and also forms a complex with Mlh1-Pms1, which can be directed by PCNA to promote efficient repair. If dimerization between Mlh3 and Mlh1 is weakened, the ability to be recruited by Msh2-Msh3 and interact with Mlh1-Pms1 is likely inhibited, creating a defect in MMR that is minor because Mlh1-Pms1 is still active. For meiotic crossing over, a relatively slow process compared to MMR at the replication fork, we suggest that a weakened dimer can be compensated for by interactions with other meiotic factors (e.g. Msh4-Msh5, Exo1 and STR). Thus, our work provides further motivation to examine Mlh1-Mlh3 activity on recombination substrates in the presence of pro-CO factors.

Mlh1-Mlh3 in meiotic recombination

Current models for meiotic DSB repair postulate an enzymatic role for Mlh1-Mlh3 in the class I CO pathway after DSB intermediates have been captured and stabilized by the ZMM

proteins [8, 30, 49]. In these models DSB intermediates that escape capture by ZMM proteins are resolved into class II COs or NCOs by SSNs, though NCOs mostly arise from the action of the STR complex through synthesis dependent strand annealing (SDSA; [30, 49]; Figure 2.1).

We observed a genome-wide increase in NCOs compared to wild-type in *mlh3-23*, *mlh3-32*, *mlh3-D523N*, and *mlh3Δ* mutants. This suggested to us that in mutants defective in Mlh1-Mlh3 functions, dHJs formed in the ZMM pathway become susceptible to the actions of SSNs to form COs and NCOs. They could also be acted upon by STR through unwinding and convergent migration of the two HJs until a single pair of crossing strands in a hemicatenane can be removed by topoisomerase activity, giving rise to NCOs [30,31]. In this model Mlh1-Mlh3, in concert with the ZMM proteins, acts to protect recombination intermediates from STR. In support of this idea, we identified genetic interactions between Mlh1-Mlh3 and Sgs1 that suggest an active interplay (Mlh1 and Mlh3 both interact with Sgs1 [42,77]) between these factors: Sgs1 overexpression decreased the spore viability of *mlh3Δ* strains but not *MLH3*, *mlh3-23*, or *mlh3-D523N* strains, where a structurally intact Mlh1-Mlh3 complex is thought to be present, and modestly increased the spore viability of *mlh3-32* mutants (Figure 2.13).

We also observed an increase in tract length for gene conversion events associated with COs in *mlh3Δ* compared to wild-type (Figure 2.8; Table 2.11; Table 2.12). The latter observation is interesting because recent work from Duroc et al. [78] provided evidence that another MLH complex, Mlh1-Mlh2, which lacks endonuclease activity, acts to limit the extent of meiotic gene conversion. They found that gene conversion tract lengths associated or not associated with COs increased from ~1 kb in wild-type to ~2 kb in *mlh2Δ*. Thus, regulating gene conversion tracts may be a common property of the MLH complexes that act in meiotic recombination, and is consistent with the fact that Mlh1 has been shown to physically interact with helicases that can

modulate heteroduplex extension/migration.

An alternative explanation for the increase in NCO events in *mlh3* mutants is that delays in meiotic progression result in the accumulation of NCO events as the result of increased DSB formation [72]. However, we did not observe such delays in any of the *mlh3* mutant backgrounds or any dramatic change in DSB formation based on an analysis of total interhomolog events in *mlh3* mutants, though a direct measure of DSB levels is critical to resolve this issue (Figure 2.7A; Figure 2.10; Table 2.11; [69]).

Mlh3's linker arm is critical for its meiotic function.

MLH proteins act as dimers and contain long unstructured linkers that connect the N- and C-terminal domains of each subunit. These linkers vary in length and are resistant to amino acid substitutions [54]. Previous work showed that the Mlh1-Pms1 heterodimer undergoes large global conformational changes in an ATP binding and hydrolysis cycle [79]. In this cycle the linkers act as arms that can switch between extended and condensed states. These conformational changes are hypothesized to be important to expose different domains of the heterodimer for new protein-protein or protein-DNA interactions in addition to mediating the timing of these interactions [79], and have also been implicated in *B. subtilis* MutL for “licensing” its latent endonuclease activity [60]. In addition, a series of truncation mutants in Mlh1-Pms1 indicate that the Pms1 linker arm appears more important than the Mlh1 linker arm for DNA binding [80]. Extending these ideas to Mlh1-Mlh3, it is interesting to note that the MMR^+, CO^- *mlh3-32* allele maps to the unstructured linker, suggesting that this domain is particularly important in crossing over (Figure 2.3; Table 2.4), possibly facilitating interactions with CO promoting factors that in turn direct and position Mlh3's endonuclease activity on recombination substrates. It is

important to note that Claeys Bouuaert and Keeney [81] identified mutations in the Mlh3 linker domain based on a biochemical analysis of Mlh1-Mlh3 that overlap with residues mutated in the *mlh3-32* allele. Interestingly, the mutations that they identified also conferred a greater defect in crossing over than in MMR, consistent with our analysis of *mlh3-32*. In addition, they found that mutations within and near the *mlh3-32* allele compromised DNA binding activity of Mlh1-Mlh3, suggesting that DNA binding within the linker region may be important for meiotic functions, though we did not detect any apparent defect in the endonuclease activity of Mlh1-*mlh3-32*.

Alanine-scan mutagenesis of Mlh1 [54] and Mlh3 have provided us with additional information regarding the unstructured linkers in Mlh proteins. Previously we used protein structure prediction and molecular analyses to map the Mlh1 unstructured linker to amino acids 336 to 480 [80]; a similar analysis mapped the Mlh3 unstructured linker to amino acids 373 to 490 [16]. As in the analysis of the Mlh3 random coil, few mutations were identified in the Mlh1 unstructured linker that conferred defects in MMR and crossing over. For example, in Mlh1, no mutations were identified between amino acids 427 and 490 that conferred mutator phenotypes. However, similar to results seen for Mlh3 (Figure 2.3; Table 2.4), mutations were identified just before the unstructured linker in Mlh1 (253-312) that conferred strong mutator phenotypes [54]. Curiously, the corresponding region in MutL contains residues that have been linked through crystallographic analysis to DNA binding [82], suggesting that the organization of the DNA binding and unstructured linker domains in the MLH proteins is conserved. Finally, in both Mlh1 and Mlh3, a localized set of mutations within the center of the unstructured linker (390-403 in Mlh1, 414-416 in Mlh3) affect function, suggesting that this specific region is likely to play an important function beyond serving as a random coil.

Closing thoughts.

Mlh1-Mlh3 appears to be acting in CO resolution through a novel mechanism distinct from that of known structure-selective endonucleases [75]. Mlh1-Mlh3 does not share conservation with the known endonuclease superfamilies (XPF, URI-YIG, Rad2/XPG), and does not appear capable of resolving model HJ substrates [50, 74]. As mentioned previously, dHJ resolution by Mlh1-Mlh3 results in only CO products whereas the interference-independent CO pathway, which is dependent on Mus81-Mms4, resolves dHJs into a mixture of CO and NCO products [8]. Thus, Mlh1-Mlh3's distinct activity suggests that its nicking is positioned by pro-CO factors such as Msh4-Msh5, Zip3, the STR complex, and Exo1. Such factors are likely to orient Mlh1-Mlh3 to promote asymmetric cleavage of dHJs in a highly regulated and coordinated manner. Thus our work provides further motivation to examine Mlh1-Mlh3 activity on recombination substrates in the presence of pro-CO factors.

Polymorphisms in human *MLH3* genes have been associated with male and female infertility [83-85], and errors in meiotic chromosome segregation are considered a leading cause of spontaneous miscarriages and birth defects [13]. It is interesting to note that the *mlh3-23* mutation, which only weakly affected crossing over, conferred an alteration in meiotic recombination outcomes that was similar to that seen in *mlh3* mutants that conferred more severe defects (Figure 2.6). This observation suggests that some polymorphisms in meiotic recombination genes could cause more severe defects on human fertility than expected.

Methods

Media.

S. cerevisiae SK1, S288c, and YJM789 strains were grown on either yeast extract-peptone-

dextrose (YPD) or minimal complete media at 30°C [86]. 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 [87, 88]. 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 (Table 2.1). The single-step integration vector (pEAI254), containing the SK1 *MLH3* gene with a *KANMX4* selectable marker inserted 40 bp downstream of the stop codon [5], was used as a template to create plasmids (Table 2.3) 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.

Mlh3 homology model.

The amino acid sequence of *S. cerevisiae* Mlh3 (YJM789) was used to construct a homology model from HHpred (<http://toolkit.tuebingen.mpg.de/hhpred>) and Modeller software. PyMOL was used for imaging.

Construction of strains to measure meiotic crossing over and MMR.

The SK1 strain EAY3255 (Table 2.1) was constructed to allow for the simultaneous analysis of *mlh3* MMR and meiotic crossing over phenotypes. It carries a spore autonomous fluorescent protein marker (RFP) on chromosome VIII to monitor chromosome behavior (crossing over and nondisjunction; [56]) as well as the *lys2::InsE-A₁₄* cassette to measure reversion to Lys⁺ [55]. pEAI254 and mutant derivatives described above and in Table 2.3 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. [89]. 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 (Table 2.1). These haploid strains were used to measure the effect of *mlh3* mutations on reversion rate and were mated to EAY3486, an *mlh3Δ* strain containing the CFP marker, resulting in diploid strains suitable for analysis of crossing over (Table 2.2). 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. [55]. At least 10 independent cultures were analyzed for each mutant allele alongside wild-type or *mlh3Δ* controls. Analyses were performed for two independent transformants per allele. Reversion rates were measured as described [90, 91], 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 [92].

Spore autonomous fluorescent protein expression to measure percent tetraptype.

Diploids in the EAY3255/EAY3486 background described above (Table 2.2) 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 [56]. At least 250 tetrads for each *mlh3* allele were counted to determine the % tetraptype. Two independent transformants were measured per allele. A statistically significant difference ($p < 0.01$) from wild-type and *mlh3Δ* controls based on χ^2 analysis was used to classify each allele as exhibiting a wild-type, intermediate, or null phenotype.

Meiotic time courses.

Meiotic time course were performed as described in Sonntag Brown et al. [93] for the diploid strains EAY3252/EAY3486 (*MLH3*), EAY3255/EAY3486 (*mlh3Δ*), EAY3534-35/EAY3486 (*mlh3-23*), EAY3552-53/EAY3486 (*mlh3-32*), and EAY3819-20/EAY3486 (*mlh3-D523N*; Table 2.2). Strains in single time courses were grown in the same batch of media under identical conditions. Aliquots of cells at specific time points were stained with DAPI to determine the percentage of cell that completed the first meiotic division (cells in which 2, 3, or 4 nuclei were observed by DAPI staining, presented as MI+MII). Cells were visualized using a Zeiss Axio Imager M2 microscope equipped with a DAPI filter. At least 150 cells were counted for each time point. Two independent transformants were analyzed per allele.

Yeast two-hybrid analysis.

The L40 strain [94] was co-transformed with bait and target vectors. Residues 481-715 of the Mlh3 C-terminus were PCR amplified from pEAI254 (SK1 *MLH3* described above) and mutant derivatives, and then sub-cloned into the target vector pEAM98 (S288C *MLH3*) to make pEAM234 (SK1 *MLH3*) and mutant derivatives (Table 2.3). pEAM98 contains a fusion between the GAL4 activation domain in pGAD10 and residues 481-715 of the Mlh3 C-terminus [5, 36]. The resulting SK1 derived target vectors were confirmed by sequencing (Sanger method). The bait vector used was pEAM105 (pBTM-MLH1-SK1). Expression of the *LACZ* reporter gene was determined by the ortho-nitrophenyl- β -D-galactopyranoside (ONPG) assay as described in [95].

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

Mlh1-Mlh3 and Mlh1-mlh3 mutant derivatives were purified from Sf9 cells infected with Bacto-Bac baculovirus expression system using pFastBacDual constructs [16]. 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 (Qiagen) and heparin sepharose (GE Healthcare) column purifications. Mlh1-Mlh3 and mutant derivative yields were ~150 μ g per 5×10^8 cells; aliquots from the final heparin purification were frozen in liquid N₂ and stored at -80°C. Protein concentrations were determined by Bradford assay [96] using BSA standard. The *mlh3-6*, *mlh3-32* and *mlh3-45* mutations were introduced into pEAE358 (pPH-His₁₀-MLH3-HA pFastBacDual construct; Rogacheva et al. [16]) by Quick Change (Stratagene). *His₁₀-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 *His₁₀-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 [97].

Endonuclease assay on supercoiled plasmid DNA and ATPase assay.

Mlh1-Mlh3 nicking activity was assayed on supercoiled pBR322 or pUC18 (Thermo Scientific). DNA (2.2 nM) was incubated in 20 μ l reactions containing indicated amounts of Mlh1-Mlh3 and Msh2-Msh3 [97] in 20 mM HEPES-KOH pH 7.5, 20 mM KCl, 0.2 mg/ml BSA, 1% glycerol, and 1 mM MgCl₂ for 1 h at 37°C. Reactions were quenched by incubation for 20 min at 37°C 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 μ g/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. *bk* indicates that amount of nicked product was not above background levels established by negative controls. ATPase assays were performed as described [16].

Genetic map distance analysis.

Diploids from the SK1 congenic strain background EAY1112/1108 [5-7] were used for genetic map distance analyses. EAY1112/2413 (*MLH3/mlh3 Δ ::NATMX*) and EAY1848/2413 (*mlh3 Δ ::KANMX/mlh3 Δ ::NATMX*) were used as wild-type and null controls respectively (Table 2.2). *mlh3* alleles of interest were integrated into EAY3712 (same as EAY1112 but *mlh3 Δ ::URA3*) using standard techniques [89]. The resulting haploid strains (EAY3713-

EAY3724) were mated to EAY2413 (*mlh3Δ::NATMX*) giving rise to diploids carrying markers suitable for genetic distance measurements (Table 2.2). Two independent transformants were analyzed per allele. Diploids were selected on media lacking the appropriate nutrients and maintained as stable strains. Diploids were sporulated as described [7]. Tetrads were dissected on synthetic complete media and germinated at 30°C after an incubation of 2-3 days. Spore clones were then replica-plated on various selective media to be scored after 1 day of incubation at 30°C. Chromosome behavior was analyzed using the recombination analysis software RANA to measure genetic map distances and spore viabilities [7]. Genetic map distances ± SE were calculated using the formula of Perkins [98] through the Stahl Laboratory Online Tools portal (<http://molbio.uoregon.edu/~fstahl/>).

Construction of strains for whole genome sequencing.

Whole genome sequencing was performed using S288c/YJM789 hybrid diploids. The Mlh1 protein sequence has four amino acid differences between SK1 and YJM789 strains and five amino acid differences between SK1 and S288c strains. The SK1 Mlh3 protein has 11 amino acid differences with respect to S288c Mlh3 and seven with respect to YJM789 Mlh3. To avoid genetic incompatibilities between Mlh1 and Mlh3, we analyzed the SK1 *mlh3* mutations, including the wild-type and *mlh3Δ* controls in the presence of SK1 *MLH1* in the S288c/YJM789 hybrid. Strain genotypes are shown in Table 2.1 and Table 2.2. SK1-*MLH1* and *MLH3* alleles were introduced into wild-type YJM789 and S288c *mlh3Δ::natMX4*, respectively, using plasmids pEAA214 and pEAI254. The SK1 *MLH1* specific SNPs were confirmed by Sanger sequencing. The *mlh3-23::kanMX4*, *mlh3-32::kanMX4*, and *mlh3-D523N::kanMX4* mutations were introduced using plasmids pEAI347, pEAI356, and pEAI252, respectively into the S288c

mlh3Δ::natMX4 background. The S288c *mlh1Δ::hphMX4* and YJM789 *mlh3Δ::kanMX4* strains were made using an *mlh1Δ::hphMX4* construct amplified by PCR and the *mlh3Δ::kanMX4* deletion construct pEAI168, respectively.

Genome wide mapping of meiotic recombination events in the S288c/YJM789 hybrid.

Genomic DNA was extracted from spore colonies of four viable spore tetrads of the *mlh3* point mutants, isogenic wild type, and isogenic *mlh3Δ* mutants as described previously [64]. Whole genome sequencing on the Illumina Hi-Seq 2500 platform was performed at Fasteris, Switzerland. Raw sequence reads were processed and SNPs genotyped as described in Krishnaprasad et al. [64] and Chakraborty et al. [70], with 4:0 marker segregations shared between tetrads removed from the analysis. Analysis of genome wide recombination events was performed as described in Oke et al. [34] (Tables 2.8 to 2.12). Custom R scripts were used to generate the segregation file (input file for the CrossOver program (v6.3) of the ReCombine [99] program suite (v2.1)). The output of the CrossOver program was used as input for groupEvents (v15). Median crossover associated tract lengths were calculated from E2 events with non-zero tracts [34]. For all count data, the t-test was used to examine the difference in the mean between samples. For tract lengths, the Wilcoxon Rank-Sum test was used to test the difference in the median between samples.

The segregation files, CrossOver program output files, groupEvents output files and the custom R scripts are available online at the Dryad digital repository (<http://dx.doi.org/10.5061/dryad.bb702>). Sequence data are available from the National Centre for Biotechnology Information Sequence Read Archive (Accession numbers SRP096621 and SRP110341).

Overexpression of Sgs1 and sgs1-hd.

SGS1 (native promoter, ORF, and termination sequence) was PCR amplified from SK1 genomic DNA obtained from NKY730 (*MATa/alpha, ura3Δ::hisG/ura3Δ::hisG, leu2::hisG/leu2::hisG, lys2/lys2*) and cloned into the high copy vector pRS426 (2μ , *amp^R*, *URA3*) to make pEAM266 (Table 2.3). The correct DNA sequence was confirmed by Sanger sequencing the entire insert. *sgs1-K706A* the helicase defective variant of *SGS1* (*sgs1-hd*) was made by site-directed mutagenesis of pEAM266 (New England BioLabs Q5 kit). The correct *sgs1-hd*, 2μ , and *URA3* sequences were confirmed by Sanger sequencing. Primer sequences used to make these constructs are available upon request. The high copy vector (with or without the *SGS1* or *sgs1-hd* insert) was transformed into stable diploids of EAY3534-35/EAY3468 (*mlh3-23/mlh3Δ*), EAY3552-53/EAY3486 (*mlh3-32/mlh3Δ*), EAY3819-20/EAY3486 (*mlh3-D523N/mlh3Δ*), EAY3252/EAY3486 (*MLH3/mlh3Δ*), and EAY3255/EAY3486 (*mlh3Δ/mlh3Δ*) (Table 2.2). Meiosis was induced as described in Argueso et al. [7] and vector selection was maintained by growing the diploid strains on minimal media lacking uracil prior to sporulation. In addition, sporulation media lacked uracil. For spore viability measurements, tetrads were dissected on synthetic complete media and germinated at 30°C after an incubation of 2-3 days. At least two independent transformants were analyzed per high copy vector. Differences in spore viability were assessed for significance using the χ^2 test.

Acknowledgements

We are grateful to Michael Lichten, Nancy Kleckner, Alba Guarne, Duyen Bui, Ujani Chakraborty, and Christopher Furman for fruitful discussions, Scott Keeney for reagents, Nathan Kruer-Zerhusen for help constructing the *S. cerevisiae* Mlh3 homology model, and Erika Mudrak

from the Cornell Statistical Consulting Unit for help with some of the statistical analyses. We also thank Scott Keeney and Corentin Claeys Bouuaert for insightful comments and for sharing unpublished data, anonymous reviewers for extremely helpful comments, and Carol Anderson and Jennifer Fung for providing us with advice on using groupEvents.

References

- [1] Kunkel TA, Erie DA (2015) Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu Rev Genet* 49: 291-313.
- [2] Flores-Rozas H, Kolodner RD (1998) The *Saccharomyces cerevisiae* *MLH3* gene functions in *MSH3*-dependent suppression of frameshift mutations. *Proc Natl Acad Sci USA* 95: 12404-12409.
- [3] Harfe BD, Minesinger BK, 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: 145-148.
- [4] Romanova NV, Crouse GF (2013) Different roles of eukaryotic MutS and MutL complexes in repair of small insertion and deletion loops in yeast. *PLoS Genet* 9: e1003920.
- [5] Nishant KT, Plys AJ, 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: 747-755.
- [6] Sonntag Brown M, Lim E, Chen C, Nishant KT, Alani E (2013) Genetic analysis of *mlh3* mutations reveals interactions between crossover promoting factors during meiosis in baker's yeast. *G3* 3: 9-22.
- [7] Argueso JL, Wanat J, Gemici Z, Alani E (2004) Competing crossover pathways act during

meiosis in *Saccharomyces cerevisiae*. *Genetics* 168: 1805-1816.

[8] Zakharyevich K, Tang S, Ma Y, Hunter N (2012) Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* 149: 334-347.

[9] Hunter N (2006) Meiotic Recombination, in: *Mol Genet Recomb Top Curr Genet*, 17th ed., pp. 381–442.

[10] Zickler D (2006) From early homologue recognition to synaptonemal complex formation. *Chromosoma* 115: 158-174.

[11] Champion MD, Hawley RS (2002) Playing for half the deck: The molecular biology of meiosis. *Nat Cell Biol* 4 Suppl: s50-56.

[12] Maguire MP (1974) Letter: The need for a chiasma binder. *J Theor Biol* 48: 485-487.

[13] Hassold T, Hunt P (2001) To err (meiotically) is human: the genesis of human aneuploidy. *Nat Rev Genet* 2: 280-291.

[14] Kadyrov FA, Dzantiev L, Constantin N, Modrich P (2006) Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* 126: 297-308.

[15] Kadyrov FA, Holmes SF, Arana ME, Lukianova OA, O'Donnell M, Kunkel TA, et al. (2007) *Saccharomyces cerevisiae* MutLalpha is a mismatch repair endonuclease. *J Biol Chem* 282: 37181–37190.

[16] Rogacheva MV, Manhart CM, Chen C, Guarne A, Surtees J, Alani E (2014) Mlh1-Mlh3, A meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated endonuclease. *J Biol Chem* 289: 5664–5673.

[17] Ranjha L, Anand R, Cejka P (2014) The *Saccharomyces cerevisiae* Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday Junctions. *J Biol Chem* 289: 5674–5686.

- [18] Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* 33: 25–35.
- [19] Robine N, Uematsu N, Amiot F, Gidrol X, Barillot E, et al. (2007) Genome-wide redistribution of meiotic double-strand breaks in *Saccharomyces cerevisiae*. *Mol Cell Biol* 27: 1868-1880.
- [20] Cao L, Alani E, Kleckner N (1990) A pathway for generation and processing of double-strand breaks during meiotic recombination in *S. cerevisiae*. *Cell* 61: 1089–1101.
- [21] Gilbertson LA, Stahl FW (1996) A test of the double-strand break repair model for meiotic recombination in *Saccharomyces cerevisiae*. *Genetics* 144: 27-41.
- [22] Keeney S, Giroux CN, Kleckner N (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* 88: 357-384.
- [23] Chen SY, Tsubouchi T, Rockmill B, Sandler JS, Richards DR, et al. (2008) Global analysis of the meiotic crossover landscape. *Dev Cell* 15: 401-415.
- [24] Schwacha A, Kleckner N (1995) Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* 83: 783-791.
- [25] Hunter N, Kleckner N (2001) The single-end invasion: an asymmetric intermediate at the double-strand break to double-Holliday Junction transition of meiotic recombination. *Cell* 106: 59-70.
- [26] Allers T, Lichten M (2001a) Intermediates of yeast meiotic recombination contain heteroduplex DNA. *Mol Cell* 8: 225-231.
- [27] Allers T, Lichten M (2001b) Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* 106: 47-57.

- [28] Börner GV, Kleckner N, Hunter N (2004) Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* 117: 29-45.
- [29] Zakharyevich K, Ma Y, Tang S, Hwang PY, 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: 1001-1015.
- [30] Kaur H, De Muyt A, Lichten M (2015) Top3-Rmi1 DNA single-strand decatenase is integral to the formation and resolution of meiotic recombination intermediates. *Mol Cell* 57: 583–594.
- [31] Tang, S Wu MKY, Zhang R, Hunter N (2015) Pervasive and essential roles of the Top3-Rmi1 decatenase orchestrate recombination and facilitate chromosome segregation in meiosis. *Mol Cell* 57: 607–621.
- [32] de Los Santos T, Hunter N, Lee, C., Larkin B, Loidl J, et al. (2003) The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote distinct subset of crossovers during meiosis in budding yeast. *Genetics* 164: 81-94.
- [33] Matos J, Blanco MG, Maslen S, Skehel JM, West SC (2011) Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis. *Cell* 147: 158-172.
- [34] Oke A, Anderson CM, Yam P, Fung JC (2014) Controlling meiotic recombinational repair - specifying the roles of ZMMs, Sgs1 and Mus81/Mms4 in crossover formation., *PLoS Genet* 10: e1004690.
- [35] Hunter N, Borts RH (1997) Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. *Genes Dev* 11: 1573-1582.

- [36] Wang TF, Kleckner N, 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 USA* 96: 13914-13919.
- [37] Lipkin SM, Moens PB, Wang V, Lenzi M, Shanmugarajah D, Gilgeous A, et al. (2002) Meiotic arrest and aneuploidy in MLH3-deficient mice. *Nat Genet* 31: 385–390.
- [38] Woods LM, Hodges CA, Baart E, Baker SM, Liskay M, Hunt PA (1999). Chromosomal influence on meiotic spindle assembly: abnormal meiosis I in female Mlh1 mutant mice. *J Cell Biol* 145: 1395-1406.
- [39] Svetlanov A, Baudat F, Cohen PE, de Massy B (2008) Distinct functions of MLH3 at recombination hot spots in the mouse. *Genetics* 178: 1937-1945.
- [40] 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: 1539-1547.
- [41] 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: 1697-1706.
- [42] Wang T-F, Kung W-M Kung (2002) Supercomplex formation between Mlh1-MLh3 and Sgs1-Top3 heterocomplexes in meiotic yeast cells. *Biochem Biophys Res Commun* 296: 949–953.
- [43] Snowden T, Acharya S, Butz C, Berardini M, Fishel R (2004) hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Mol Cell* 15: 437-451.

- [44] Kneitz B, Cohen PE, Avdievich E, Zhu L, Kane MF, 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: 1085-1097.
- [45] Hoffmann ER, Borts RH (2004) Meiotic recombination intermediates and mismatch repair proteins. *Cytogenet Genome Res* 107: 232-248.
- [46] Fasching CL, Cejka P, Kowalczykowski SC, Heyer W-D (2015) Top3-Rmi1 Dissolve Rad51-Mediated D Loops by a Topoisomerase-Based Mechanism, *Mol Cell* 57: 595–606.
- [47] Kolas NK, Cohen PE (2004) Novel and diverse functions of the DNA mismatch repair family in mammalian meiosis and recombination. *Cytogenet Genome Res* 107: 216-231.
- [48] Kolas NK, Svetlanov A, Lenzi ML, Macaluso FP, Lipkin SM, Liskay RM, et al. (2005) Localization of MMR proteins on meiotic chromosomes in mice indicates distinct functions during prophase I. *J Cell Biol* 171: 447–458.
- [49] de Muyt A, Jessop L, Kolar E, Sourirajan A, Chen J et al. (2012) BLM helicase ortholog Sgs1 is a central regulator of meiotic recombination intermediate metabolism. *Mol. Cell* 46: 42-53.
- [50] Manhart CM, 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.
- [51] Gueneau E, Dherin C, Legrand P, Tellier-Lebegue C, Gilquin B, Bonnesoeur P, et al. (2013) Structure of the MutL α C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site., *Nat Struct Mol Biol* 20: 461–468.
- [52] Bass SH, Mulkerrin MG, Wells JA (1991) A systematic mutational analysis of hormone-binding determinants in the human growth hormone receptor. *Proc Natl Acad Sci USA* 88: 4498-4502.

- [53] Bennett WF, Paoni NF, Keyt BA, Botstein D, Jones AJ, et al. (1991) High resolution analysis of functional determinants on human tissue-type plasminogen activator. *J Biol Chem* 266: 5191-5201.
- [54] Argueso JL, Kijas AW, Sarin A, 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: 873-886.
- [55] Tran HT, Keen JD, Kricker M, Resnick MA, Gordenin DA (1997) Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Mol Cell Biol* 17: 2859-2865.
- [56] Thacker D, Lam I, Knop M, Keeney S. (2011) Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in *Saccharomyces cerevisiae*. *Genetics* 189: 423-439.
- [57] Groothuizen FS, Winkler I, Cristóvão M, Fish A, Winterwerp HH, Reumer A, Marx AD, Hermans N, Nicholls RA, Murshudov GN, Lebbink JH, Friedhoff P, Sixma TK (2015) MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *Elife* 4: e06744.
- [58] Hargreaves VV, Putnam CD, Kolodner RD (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: 41232-41244. doi: 10.1074/jbc.M112.402495.
- [59] Kosinski J, Plotz G, Guarné A, Bujnicki JM, Friedhoff P (2008) The PMS2 subunit of human MutL α contains a metal ion binding domain of the iron-dependent repressor protein family. *J Mol Biol* 382: 610-627.

- [60] Pillon MC, Lorenowicz JJ, Uckelmann M, Klocko AD, Mitchell RR, Chung YS, et al. (2010) Structure of the endonuclease domain of MutL: unlicensed to cut. *Mol Cell* 39: 145–151.
- [61] Argueso JL, Smith D, Yi J, Waase M, Sarin S, Alani E (2002) Analysis of conditional mutations in the *Saccharomyces cerevisiae MLH1* gene in mismatch repair and in meiotic crossing over., *Genetics* 160: 909–921.
- [62] Agarwal S, Roeder GS (2000) Zip3 provides a link between recombination enzymes and synaptonemal complex proteins., *Cell* 102: 245–255.
- [63] Nishant KT, Chen C, Shinohara M, Shinohara A, Alani E (2010) Genetic analysis of baker's yeast Msh-Msh5 reveals a threshold crossover level for meiotic viability. *PLoS Genet* 6: e1001083.
- [64] Krishnaprasad GN, Anand MT, Lin G, Tekkedil MM, Steinmetz LM, Nishant KT (2015). Variation in crossover frequencies perturb crossover assurance without affecting meiotic chromosome segregation in *Saccharomyces cerevisiae*. *Genetics* 199: 399-412.
- [65] Martini E, Diaz RL, Hunter N, Keeney S (2006) Crossover homeostasis in yeast meiosis. *Cell* 126: 285-295.
- [66] Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM (2008) High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature* 454: 479-485.
- [67] Cotton VE, Hoffmann ER, Borts RH (2010) Distinct Regulation of Mlh1p Heterodimers in Meiosis and Mitosis in *Saccharomyces cerevisiae*. *Genetics* 185: 459-467.
- [68] Greig D, Travisano M, Louis EJ, Borts RH (2003) A role for the mismatch repair system during incipient speciation in *Saccharomyces*. *J Evol Biol* 16: 429–437.
- [69] Pan J, Sasaki M, Kniewel R, Murakami H, Blitzblau HG, Tischfield SE, Zhu X, Neale MJ, Jasin M, Socci ND, Hochwagen A, Keeney S (2011) A hierarchical combination of factors

shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* 144: 719-731. doi: 10.1016/j.cell.2011.02.009.

[70] Chakraborty P., Ajith V.P., Lin G, Dutta A., Krishnaprasad G.N., Tekkedil M.M., Shinohara, A., Steinmetz, L.M. and Nishant K.T (2017) Modulating crossover frequency and interference for obligate crossovers in *Saccharomyces cerevisiae* meiosis. *G3* 7: 1511-1524.

[71] Thacker D, Mohibullah N, Zhu X, Keeney S (2014) Homologue engagement controls meiotic DNA break number and distribution. *Nature* 510: 241-246. doi: 10.1038/nature13120.

[72] Gray S, Allison RM, Garcia V, Goldman AS, Neale MJ (2013) Positive regulation of meiotic DNA double-strand break formation by activation of the DNA damage checkpoint kinase Mec1(ATR). *Open Biol* 3: 130019. doi: 10.1098/rsob.130019.

[73] Jessop L1, Rockmill B, Roeder GS, Lichten M (2006) Meiotic chromosome synapsis-promoting proteins antagonize the anti-crossover activity of Sgs1. *PLoS Genet* 2: e155.

[74] Manhart CM, Ni X, White MA, Ortega J, Surtees JA, Alani E (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.

[75] Hombauer H, Campbell CS, Smith CE, Desai A, Kolodner RD (2011) Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 147: 1040-1053. doi: 10.1016/j.cell.2011.10.025.

[76] Pluciennik A, Dzantiev L, Iyer RR, Constantin N, Kadyrov FA, Modrich P (2010) PCNA function in the activation and direction of MutLa endonuclease in mismatch repair. *Proc Natl Acad Sci USA* 107: 16066-16071.

- [77] Pedrazzi G, Perrera C, Blaser H, Kuster P, Marra G, Davies SL, Ryu GH, Freire R, Hickson ID, Jiricny J, Stagljar I (2001) Direct association of Bloom's syndrome gene product with the human mismatch repair protein MLH1. *Nucleic Acids Res* 29: 4378-4386.
- [78] Duroc Y, Kumar R, Ranjha L, Adam C, Guérois R, Md Muntaz K, et al. (2017) Concerted action of the MutL β heterodimer and Mer3 helicase regulates the global extent of meiotic gene conversion. *Elife* 6: e21900. doi: 10.7554/eLife.21900.
- [79] Sacho EJ, Kadyrov FA, Modrich P, Kunkel TA, Erie DA (2008) Direct visualization of asymmetric adenine-nucleotide-induced conformational changes in MutL alpha. *Mol Cell* 29: 112-121.
- [80] Plys AJ, Rogacheva MV, Greene EC, Alani E (2012) The unstructured linker arms of Mlh1-Pms1 are important for interactions with DNA during mismatch repair. *J Mol Biol* 422: 192–203.
- [81] Claeys Bouuaert C, Keeney S (2017) Distinct DNA-binding surfaces in the ATPase and linker domains of MutL γ determine its substrate specificities and exert separable functions in meiotic recombination and mismatch repair. *PLoS Genet* 13: e1006722. doi: 10.1371/journal.pgen.1006722.
- [82] Ban C, Junop M, Yang W (1999) Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell* 97: 85-97.
- [83] Pashaiefar H, Sheikhha MH, Kalantar SM, Jahaninejad T, Zaimy MA, Ghasemi N (2013) Analysis of *MLH3* C2531T polymorphism in Iranian women with unexplained infertility, *Iranian Journal of Reproductive Medicine*, 11: 19–24.

- [84] Markandona O, Dafopoulos K, Anifandis G, Messini CI, Dimitraki M, Tsezou A, Georgoulas P, Messinis IE (2015) Single-nucleotide polymorphism rs 175080 in the *MLH3* gene and its relation to male infertility. *J Assist Reprod Genet* 32: 1795-1799.
- [85] Zhang X, Ding M, Ding X, Li T, Chen H (2015) Six polymorphisms in genes involved in DNA double-strand break repair and chromosome synapsis: association with male infertility. *Syst Biol Reprod Med* 61: 187-193.
- [86] Rose MD, Winston F, Hieter P (1990) *Methods in yeast genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [87] Wach A, Brachat A, Pohlmann R, Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* 10: 1793-1808.
- [88] Goldstein AL, McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* 15: 1541-1553.
- [89] Gietz RD, Schiestl RH, Willems AR, Woods RA (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11: 355-360.
- [90] Heck JA, Argueso JL, Gemici Z, Reeves RG, 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 USA* 103: 3256-3261.
- [91] Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA* 88: 7160-7164.
- [92] Gibbons JD, Chakraborti S (2014) *Nonparametric Statistical Inference*, Fourth Edition. Marcel Dekker, Inc., New York.

- [93] Sonntag Brown M, Zanders S, Alani E (2011) Sustained and rapid chromosome movements are critical for chromosome pairing and meiotic progression in budding yeast. *Genetics* 188: 21–32.
- [94] Pang Q, Prolla TA, Liskay, RM (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: 4465-4473.
- [95] Gietz RD, Woods RA (2002) Screening for protein-protein interactions in the yeast two-hybrid system. *Methods Mol Biol* 185: 471-486.
- [96] Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248-254.
- [97] Surtees JA, 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: 523-536.
- [98] Perkins DD (1949) Biochemical mutants in the smut fungus *Ustilago maydis*. *Genetics* 34: 607-626.
- [99] Anderson CM, Chen SY, Dimon MT, Oke A, DeRisi JL, Fung JC (2011) ReCombine: a suite of programs for detection and analysis of meiotic recombination in whole-genome datasets. *PLoS One* 6: e25509.

CHAPTER 3

Future Directions

Working model of Mlh1-Mlh3 function

Mlh1-Mlh3 is a minor mismatch repair endonuclease with (MMR) and major role in meiotic recombination. Yet the mechanism by which one endonuclease contributes minimally to MMR and centrally to crossing over is not well understood. Mlh1-Mlh3 shares little in common with the well-characterized structure-selective nucleases (SSNs; Mus81-Mms4, Slx1-Slx4, and Yen1) in terms of homology or intrinsic behavior *in vitro* (reviewed in [1]). For example, the SSNs have been shown to bind and cleave branched DNA structures such as a single Holliday junction (HJ) *in vitro*. Whereas, Mlh1-Mlh3 can bind such branched substrates, but has not been shown to cleave them [2], [3]. Furthermore, resolution of double Holliday Junctions (dHJs) by SSNs results in a mixture of crossover (CO) and noncrossover (NCO) products, while Mlh1-Mlh3 gives rise to CO products exclusively. Thus, Mlh1-Mlh3 appears to be acting in a novel mechanism that does not conform to models set by previously characterized dHJ resolvases [3].

Interestingly, recent work in our lab suggests that multiple Mlh1-Mlh3 heterodimers are required to activate the endonuclease activity and that the complex is at least partially inhibited by incorporation of a DNA secondary structure that is not part of a continuous homoduplex DNA substrate [3]. Taken together, these data point towards a model in which Mlh1-Mlh3 is loaded onto DNA to form an activated polymer that, positioned and directed by other recombination proteins, cleaves recombination intermediates to form COs. This behavior is consistent with activities seen for the MMR endonuclease Mlh1-Pms1. In addition, the endonuclease activity of human MLH1-PMS2 (Mlh1-Pms1 in yeast) is strand specifically activated by PCNA [4], which

suggests that protein factors could dictate Mlh1-Mlh3 specificity in a similar manner. It is tempting to speculate that such MLH conserved properties are utilized by Mlh1-Mlh3 in meiosis to resolve dHJs into COs. The multifaceted assembly of the dHJ involves distinct meiotic factors that create and stabilize sequential joint molecule intermediates resulting in formation of the final dHJ substrate. This is in support of the working model that Mlh1-Mlh3 likely relies on other protein factors, such as Msh4-Msh5 (and other ZMM proteins), for recruitment, positioning, and coordination of its endonuclease activity.

In Chapter 1, I summarized data that outlines the two different functions of Mlh1-Mlh3. In addition, I discussed models, which are not mutually exclusive, that can explain how Mlh1-Mlh3 is regulated to preform its dual role. In Chapter 2, I presented studies, performed by myself in collaboration with others, that support the model described above in which protein-protein interactions with MMR or CO specific protein factors are key to Mlh1-Mlh3 activity. This work was based on a structure-function map of Mlh3 that I built from the analysis of 60 new *mlh3* alleles in MMR and crossing over (Chapter 2, Figures 2.2 and 2.3; Table 2.4). From this map, I identified six separations of function mutants, five of these alleles specifically disrupted the MMR function of Mlh3 (MMR⁻, CO⁺; *mlh3-6*, *mlh3-23*, *mlh3-42*, *mlh3-45*, and *mlh3-54*), while one (*mlh3-32*) specifically disrupted crossing over (MMR⁺, CO⁻) (Chapter 2, Figure 2.4; Table 2.6). Biochemical, genetic, and yeast two-hybrid analyses suggest that the defects observed in these mutants can be explained by disrupted protein-protein interactions that are tolerated in meiosis but not MMR. Thus, I hypothesize that for meiotic CO resolution, Mlh1-Mlh3 acts as part of a resolvase complex that facilitates access to a unique DNA substrate that can be recognized and nicked by Mlh1-Mlh3. Based on these and other findings, my work supports the working model that the Mlh1-Mlh3 endonuclease is recruited and coordinated by distinct MMR

or CO promoting factors which recognize and bind the DNA substrate to spatially and temporally coordinate Mlh1-Mlh3 nicking on a given structure.

In this Chapter, in section one, I will discuss immediate future plans (and ongoing work) based on the mutants presented in Chapter 2. In the second section, I will present two future projects that remain an active area of study in the meiotic recombination field. I will propose experiments, based on work presented in this dissertation, that will contribute significantly to these projects. The first project is the complete biochemical reconstitution of the Mlh1-Mlh3 reaction *in vitro*. The second is exploring the disease relationship of the meiotic recombination defects presented in Chapter 2.

1- Remaining questions and immediate future plans

In Chapter 2, I identified an MMR⁺, CO⁻ allele, *mlh3-32*, that displayed an *in vitro* endonuclease activity similar to wild-type, and can similarly be stimulated by the MMR factor Msh2-Msh3 (Chapter 2, Figure 2.11). This indicates that the meiotic defect observed in this mutant is likely the result of a disrupted interaction with a specific pro-crossover factor. To test this idea, I created overexpression constructs of a large number of meiotic genes, including *SGS1*, *MSH4-MSH5*, *EXO1*, *ZIP3*, and *MER3*. These constructs were transformed into *MLH3*, *mlh3Δ*, and *mlh3-32* backgrounds to test for suppression of the meiotic defect. Interestingly, as shown in Chapter 2 (Figure 2.13), I observed that Sgs1 overexpression enhanced the spore viability defect in *mlh3Δ* but modestly suppressed the spore viability defect in *mlh3-32*. These effects were dependent on the helicase activity of Sgs1 and were not observed upon overexpression of a helicase defective variant of Sgs1. Mer3 is a helicase that acts as part of the ZMM family of proteins (*Zip1-4*, *Mer3*, and *Msh4-Msh5*) important in the Class I CO pathway

[5]–[8]. It would be interesting to test how the effects of overexpressing wild-type and helicase-defective *MER3* in *MLH3*, *mlh3Δ*, and *mlh3-32* backgrounds would differ from *SGS1* overexpression. Sgs1 helicase acts as a master regulator of meiotic double strand break (DSB) repair at steps prior to the involvement of Mer3 and the formation of the ZMM-protein protected joint molecule intermediate [9], [10]. Preliminary analysis shows that *MER3* overexpression also suppresses the spore viability defect of *mlh3-32*, but has no effect on the spore viability of *mlh3Δ*.

In Chapter 2, I show that introducing the last 11 residues of Pms1 onto the C-terminal domain of Mlh3 (*mlh3-60*) strengthens the Mlh1-Mlh3 interaction (Chapter 2, Figure 2.4C). This effect is likely the result of two additional salt bridges contributed by the last 11 residues of Pms1, that are not conserved in Mlh3, as shown in the crystal structure of Mlh1-Pms1 [11]. I hypothesized that strengthening the interaction between Mlh1-Mlh3, via the *mlh3-60* mutant, would be detrimental to MMR as it may hinder the formation of the Mlh1-Pms1 heterodimer that is more critical in this process. Curiously, I did not observe an effect on MMR or crossing over in an *mlh3-60* background (Chapter 2, Figure 2.4B). This suggests that strengthening the interaction between Mlh1-Mlh3 does not sequester Mlh1 from Pms1. Consequently it would be interesting to test an *mlh3-60 pms1Δ* double mutant. In this strain background, I hypothesize that strengthening the Mlh1-Mlh3 interaction will partially alleviate the MMR defects observed as a result of *pms1* deletion.

2- Future projects

a. complete biochemical reconstitution of the Mlh1-Mlh3 reaction in vitro

The segregation of chromosomes at the first meiotic division is dependent in most

organisms on at least one CO event between homologous chromosomes. Homologs that do not receive a CO frequently undergo nondisjunction at the first meiotic division, yielding aneuploid gametes. Such events have been linked to human disease and infertility [15]. As described above, the Mlh1-Mlh3 complex is an endonuclease involved in resolving recombination intermediates into COs and appears to be acting through a novel mechanism distinct from that of known SSNs [3]. SSNs have been shown to recognize and symmetrically cleave branched DNA structures such as single HJs *in vitro* (reviewed in [1]). Although, Mlh1-Mlh3 can bind branched DNA substrates, it has not been shown to cleave them. *In vitro*, Mlh1-Mlh3 displays a non-specific nicking activity on circular plasmid DNA ([2] and [12]) and does not conform to paradigms set by previously characterized dHJ resolvases [3]. Rather, data presented in Chapter 2 highlights the importance of protein-protein interactions in presenting a unique substrate that Mlh1-Mlh3 can recognize and act on. Thus, the full biochemical reconstitution of the Mlh1-Mlh3 reaction *in vitro* remains an essential goal in the meiotic recombination field.

In Chapter 2, I presented two MMR⁻, CO⁺ alleles (*mlh3-42* and *mlh3-54*) that weaken the interaction with Mlh1 (Figure 2.4; Table 2.6). The disrupted Mlh1-*mlh3-42* and *-54* interaction provides a likely explanation for the null-like phenotype observed for these alleles with regard to their MMR function, but it is surprising that crossing over in these mutants is maintained at wild-type levels. I hypothesize that during meiotic recombination additional pro-crossover factors present at the dHJ act as a scaffold to structurally support a weakened Mlh1-*mlh3* interaction. In support of this, the pro-crossover factors Sgs1-Top3-Rmi1 (STR), Zip3, Msh4-Msh5, and Exo1 have all been shown to interact with one another and/or with Mlh1-Mlh3 [1]. To test this hypothesis, I propose deleting one copy of different meiotic genes (listed below), involved in various stages of meiotic recombination, in a diploid mutant strain background. If my hypothesis

is correct, and the dose of meiotic proteins, that are serving the role of structural support to Mlh1-mlh3, is lowered, then the Mlh1-mlh3-42 or -54 complex will fall apart and, consequently, CO levels will decrease. I have assembled a comprehensive list of such meiotic genes to test including those presented in Chapter 1, Table 1.1 such as the ZMM family of proteins, Exo1, and the STR complex. In addition to the cohesin protein, Rec8, and the meiotic checkpoint protein Cdc5. Furthermore, it will be interesting to test proteins involved in strand invasion (Rad51, Rad52, Rad54, Dmc1) and synaptonemal complex formation (Red1, Ecm11, Gmc2). Such a comprehensive analysis will lend support to the model that Mlh1-Mlh3 is directed in CO resolution as part of a resolvase complex and ultimately contribute to the biochemical reconstitution of an active complex *in vitro*. In addition, this analysis can potentially reveal novel genetic interactions between proteins involved in various stages of meiotic Prophase I.

In Chapter 2, I identified one MMR^+ , CO^- mutation (*mlh3-32*) in the linker arm of Mlh3 (Figure 2.3). Further characterization of *mlh3-32* using biochemical and genetic analyses did not reveal why this mutation specifically disrupts crossing over. Interestingly, *mlh3-32* lies within the linker arm of Mlh3 and the linker arms of Mlh1-Mlh3 have been implicated in mediating conformational changes via an ATPase cycle. Such conformational changes expose different regions of the heterodimer to modulate diverse protein-protein or protein-DNA interactions in addition to mediating the timing of these interactions. I hypothesize that *mlh3-32* obstructs the formation of a configuration that may be more important for facilitating interactions with recombination specific protein factors. Overexpression of *MSH4-MSH5*, *EXO1*, and *ZIP3* in an *mlh3-32* background did not suppress the CO defect, and only mild suppression was observed upon *SGS1* overexpression. Such a selective testing approach, in which a limited number of meiotic genes were chosen, was unlikely to reveal the causative genetic interactions.

Furthermore, the complete protein interaction network of wild-type Mlh1-Mlh3 during meiotic recombination remains largely unknown. Thus, for the purpose of identifying novel Mlh1-Mlh3 meiotic interacting partners in addition to those likely defective in Mlh1-mlh3-32, I propose an immunoprecipitation approach. For this analysis, purified wild-type Mlh1-Mlh3 protein and purified Mlh1-mlh3-32 can be used to pull down their respective interacting partners from meiotic extracts, followed by mass spectrometry. Such an approach can be extended to extracts from vegetative cells to reveal the full complement of Mlh1-Mlh3 interactions in MMR. If my hypothesis is correct, a mutant that disrupts CO specific functions, such as *mlh3-32*, will be defective in pulling down a CO specific interacting partner (observed in wild-type pull-downs) and likewise for a mutant that specifically disrupts MMR.

In mouse spermatocytes, MSH4 co-localizes with MLH1 at early to mid pachytene [13]. Yeast Msh4 has been observed to localize on chromosomes predominantly during pachytene [14]. It would be interesting to use an immunofluorescence assay to investigate the temporal and dynamic relationship of Msh4-Msh5 and Mlh1-Mlh3 localization in yeast meiosis. In particular, it would be useful to compare MMR⁺, CO⁻ mutants to wild-type localization patterns. This will provide an *in vivo* perspective and validation of the immunoprecipitation experiments outlined above. I hypothesize that wild-type Mlh1-Mlh3 will co-localize with Msh4-Msh5 but not the MMR⁺, CO⁻ separation of function mutant likely due to defective interactions with recruiting factors.

Taken together the experiments proposed above will help build the Mlh1-Mlh3 meiotic interactome and will prove essential in bringing the Mlh1-Mlh3 community one step closer to the ultimate goal of reconstituting the dHJ resolution reaction *in vitro*.

b. clinical impact and association to human infertility

Errors in meiotic chromosome segregation are considered a leading cause of spontaneous miscarriages and birth defects [15]. Furthermore, polymorphisms in human *MLH3* genes have been associated with male and female infertility [16]–[18]. The following section will describe experiments aimed at investigating the disease relationship of the alleles presented in Chapter 2.

I propose, in collaboration with the Schimenti lab, utilizing Baker’s yeast as a model for a quick functional interrogation of potentially deleterious nonsynonymous SNPs identified in human populations [19]. I hypothesize that a CO defective phenotype in yeast, resulting from a mutation based on human SNPs in a residue that is highly conserved from yeast to human, will likely result in infertility if mimicked in mice. This method can be used to quickly screen in yeast disease causing genetic variants from human populations to be modeled and validated in mice. In collaboration with the Schimenti lab, *MLH3* SNPs found in humans were identified in amino acid residues that are highly conserved from yeast to human and lie within the functionally important ATPase and endonuclease domains of Mlh3. I hypothesize yeast bearing these “humanized” alleles will likely result in a CO defect. More importantly, if my hypothesis is correct, such variants can be chosen for modeling in mice and will likely give rise to an infertility phenotype thus validating disease causing variants in humans. Of particular interest are SNPs that align with *mlh3-40* and *mlh3-41* that gave a null phenotype in both MMR and crossing over as shown in Chapter 2 (Figure 2.3; Table 2.4).

In Chapter 2 I presented work, in collaboration with K.T. Nishant’s lab, that implicates *mlh3* mutants in altering the fate of meiotic recombination intermediates (Figure 2.6; Table 2.7). High-resolution recombination maps revealed that *mlh3* mutants increase NCOs genome wide. Of particular interest is *mlh3-23* (MMR⁻, CO⁺), because although it did not significantly affect

CO levels, as compared to wild-type, it displayed increased NCO levels similar to those observed in *mlh3* mutants with more severe CO defects (Figure 2.6 Table 2.7). I propose DNA physical assays, briefly described below [20], to determine which specific steps of meiotic recombination are altered in this mutant. DNA physical assays can be used to monitor recombination at a given hotspot by analyzing DNA fragments diagnostic for DSBs, joint molecules, COs, and NCOs [20]. In addition, three types of joint molecules can be detected: early single-end invasion intermediates, dHJs, and multichromatid joint molecules. Such insight would be valuable in an *mlh3-23* strain to determine the possible formation of defective joint molecule intermediates in this background and the relative timing of their formation. This will provide information on the particular stage during the formation of joint molecule intermediates that is likely disrupted by *mlh3-23*.

Finally, mouse *Mlh3*^{-/-} mutants are viable but infertile due to meiotic defects [21]. Mutations such as *mlh3-23* suggest that some mutants in meiotic recombination genes could have more severe defects on fertility than others. The separation of function mutants identified in Chapter 2, such as *mlh3-23* and *mlh3-32*, are present in regions conserved between yeast *MLH3* and mouse *Mlh3*. It would be of interest to investigate the effects of equivalent mutations in mouse *Mlh3* on mammalian meiosis and infertility.

References

- [1] C. M. Manhart and E. Alani, "Roles for mismatch repair family proteins in promoting meiotic crossing over," *DNA Repair (Amst)*., vol. 38, pp. 84–93, 2015.
- [2] M. V Rogacheva, C. M. Manhart, C. Chen, A. Guarne, J. Surtees, and E. Alani, "Mlh1-Mlh3, a meiotic crossover and DNA mismatch repair factor, is a Msh2-Msh3-stimulated

- endonuclease.,” *J. Biol. Chem.*, vol. 289, no. 9, pp. 5664–73, Feb. 2014.
- [3] C. M. Manhart, X. Ni, M. A. White, J. Ortega, J. A. Surtees, and E. Alani, “The mismatch repair and meiotic recombination endonuclease Mlh1-Mlh3 is activated by polymer formation and can cleave DNA substrates in trans,” *PLoS Biol.*, p. In press, 2017.
- [4] A. Pluciennik, L. Dzantiev, R. R. Iyer, N. Constantin, F. A. Kadyrov, and P. Modrich, “PCNA function in the activation and strand direction of MutL α endonuclease in mismatch repair.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 37, pp. 16066–71, 2010.
- [5] O. M. Mazina, A. V. Mazin, T. Nakagawa, R. D. Kolodner, and S. C. Kowalczykowski, “*Saccharomyces cerevisiae* Mer3 helicase stimulates 3'-5' heteroduplex extension by Rad51: Implications for crossover control in meiotic recombination,” *Cell*, vol. 117, no. 1, pp. 47–56, 2004.
- [6] T. Nakagawa and R. D. Kolodner, “The MER3 DNA Helicase Catalyzes the Unwinding of Holliday Junctions,” *J. Biol. Chem.*, vol. 277, no. 31, pp. 28019–28024, 2002.
- [7] T. Nakagawa and R. D. Kolodner, “*Saccharomyces cerevisiae* Mer3 Is a DNA Helicase Involved in Meiotic Crossing Over,” *Mol. Cell. Biol.*, vol. 22, no. 10, pp. 3281–3291, 2002.
- [8] Y. Duroc, R. Kumar, L. Ranjha, C. Adam, R. Guérois, K. Md Muntaz, M.-C. Marsolier-Kergoat, F. Dingli, R. Laureau, D. Loew, B. Llorente, J.-B. Charbonnier, P. Cejka, and V. Borde, “Concerted action of the MutL β heterodimer and Mer3 helicase regulates the global extent of meiotic gene conversion,” *Elife*, vol. 6, pp. 1–24, 2017.
- [9] H. Kaur, A. De Muyt, and M. Lichten, “Top3-Rmi1 DNA Single-Strand Decatenase Is Integral to the Formation and Resolution of Meiotic Recombination Intermediates,” *Mol. Cell*, vol. 57, no. 4, pp. 583–594, 2015.

- [10] A. De Muyt, L. Jessop, E. Kolar, A. Sourirajan, J. Chen, Y. Dayani, and M. Lichten, “BLM helicase ortholog Sgs1 is a central regulator of meiotic recombination intermediate metabolism.,” *Mol. Cell*, vol. 46, no. 1, pp. 43–53, Apr. 2012.
- [11] E. Gueneau, C. Dherin, P. Legrand, C. Tellier-Lebegue, B. Gilquin, P. Bonnesoeur, F. Londino, C. Quemener, M.-H. Le Du, J. a Márquez, M. Moutiez, M. Gondry, S. Boiteux, and J.-B. Charbonnier, “Structure of the MutL α C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site.,” *Nat. Struct. Mol. Biol.*, vol. 20, no. 4, pp. 461–8, 2013.
- [12] L. Ranjha, R. Anand, and P. Cejka, “The *Saccharomyces cerevisiae* Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions.,” *J. Biol. Chem.*, vol. 289, no. 9, pp. 5674–86, Mar. 2014.
- [13] S. Santucci-Darmanin, D. Walpita, F. Lespinasse, C. Desnuelle, T. Ashley, and V. Paquis-Flucklinger, “MSH4 acts in conjunction with MLH1 during mammalian meiosis.,” *FASEB J.*, vol. 14, no. 11, pp. 1539–47, Aug. 2000.
- [14] G. S. Roeder, “Mutation of a Meiosis-Specific MutS decreases Crossing Over Not Mismatch Correction,” vol. 79, pp. 1069–1080, 1994.
- [15] T. Hassold and P. Hunt, “To err (meiotically) is human: the genesis of human aneuploidy.,” *Nat. Rev. Genet.*, vol. 2, no. 4, pp. 280–91, Apr. 2001.
- [16] H. Pashaiefar, M. H. Sheikha, S. M. Kalantar, T. Jahaninejad, M. A. Zaimy, and N. Ghasemi, “Analysis of MLH3 C2531T polymorphism in Iranian women with unexplained infertility,” *Iran. J. Reprod. Med.*, vol. 11, no. 1, pp. 19–24, 2013.
- [17] O. Markandona, K. Dafopoulos, G. Anifandis, C. I. Messini, M. Dimitraki, A. Tsezou, P. Georgoulas, and I. E. Messinis, “Single-nucleotide polymorphism rs 175080 in the

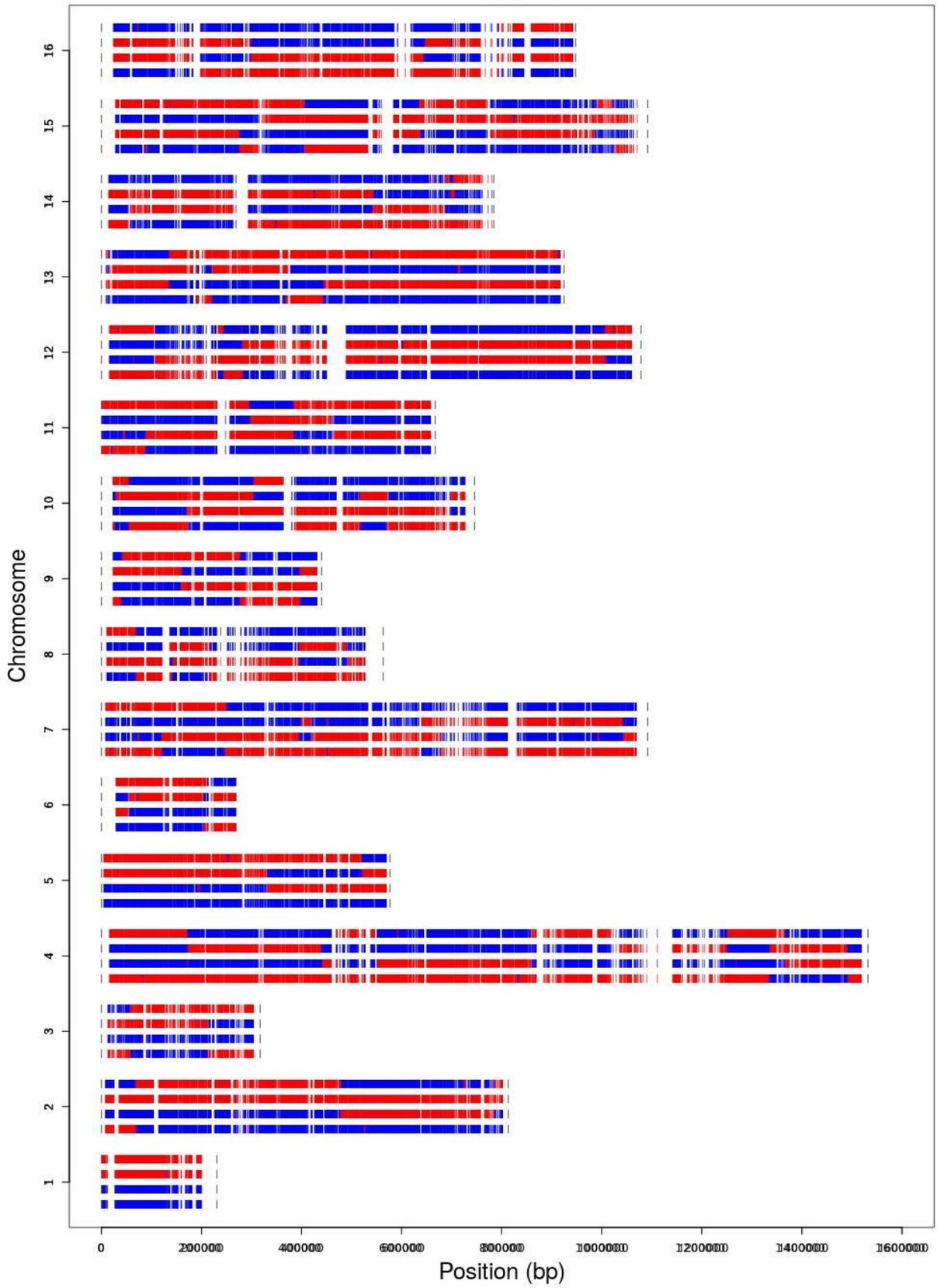
- MLH3 gene and its relation to male infertility,” *J. Assist. Reprod. Genet.*, vol. 32, no. 12, pp. 1795–1799, 2015.
- [18] X. Zhang, M. Ding, X. Ding, T. Li, and H. Chen, “Six polymorphisms in genes involved in DNA double-strand break repair and chromosome synapsis: association with male infertility,” *Syst. Biol. Reprod. Med.*, vol. 61, no. 4, pp. 187–93, 2015.
- [19] P. Singh and J. C. Schimenti, “The genetics of human infertility by functional interrogation of SNPs in mice,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 33, pp. 10431–6, 2015.
- [20] K. Zakharyevich, S. Tang, Y. Ma, and N. Hunter, “Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase,” *Cell*, vol. 149, no. 2, pp. 334–47, Apr. 2012.
- [21] S. M. Lipkin, P. B. Moens, V. Wang, M. Lenzi, D. Shanmugarajah, A. Gilgeous, J. Thomas, J. Cheng, J. W. Touchman, E. D. Green, P. Schwartzberg, F. S. Collins, and P. E. Cohen, “Meiotic arrest and aneuploidy in MLH3-deficient mice,” *Nat. Genet.*, vol. 31, no. 4, pp. 385–90, Aug. 2002.

APPENDIX A

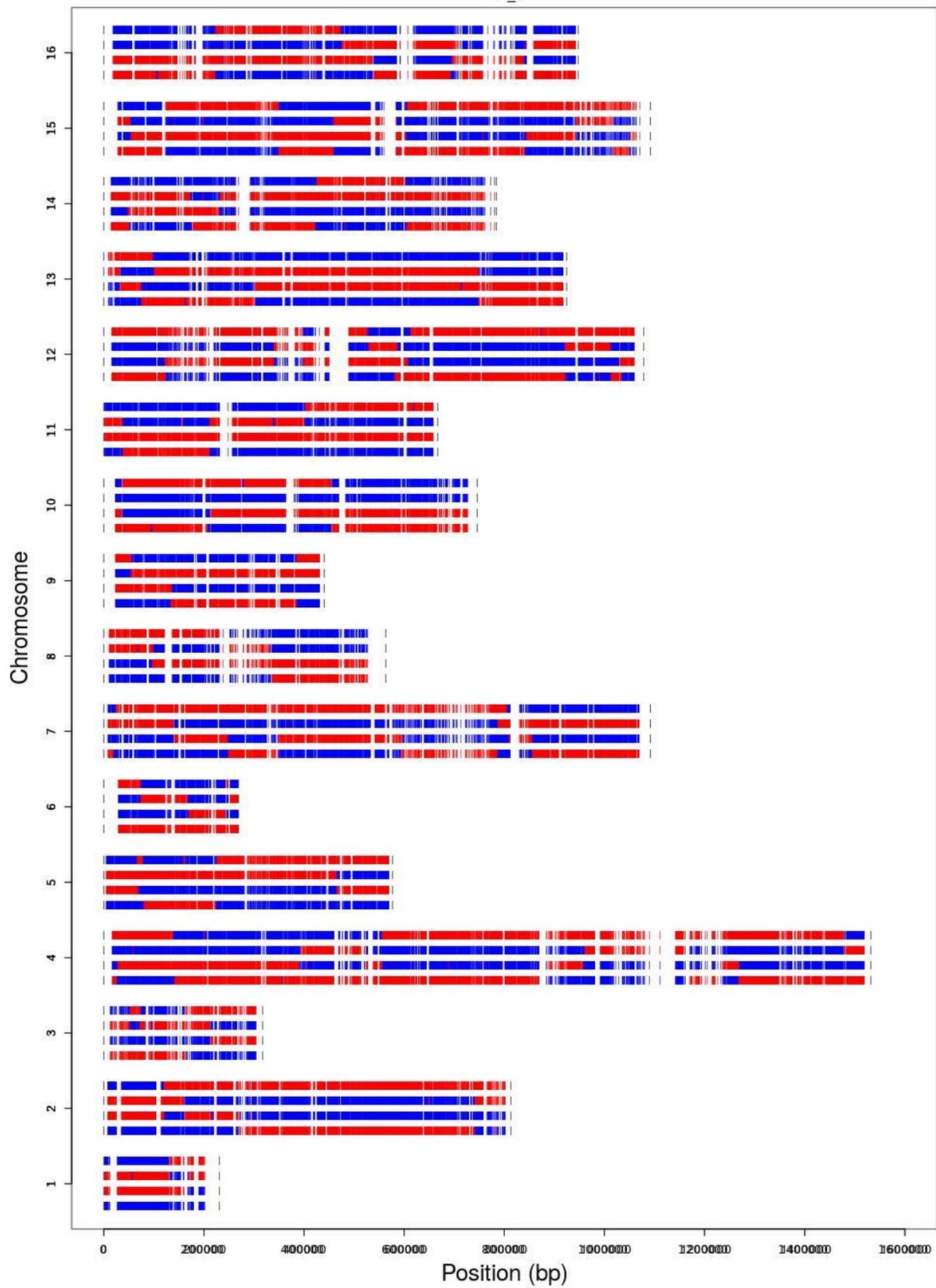
Segregation of SNPs in all 44 tetrads as labeled in Chapter 2, Table 2.9.

S288c and YJM789 SNPs are shown in red and blue, respectively.

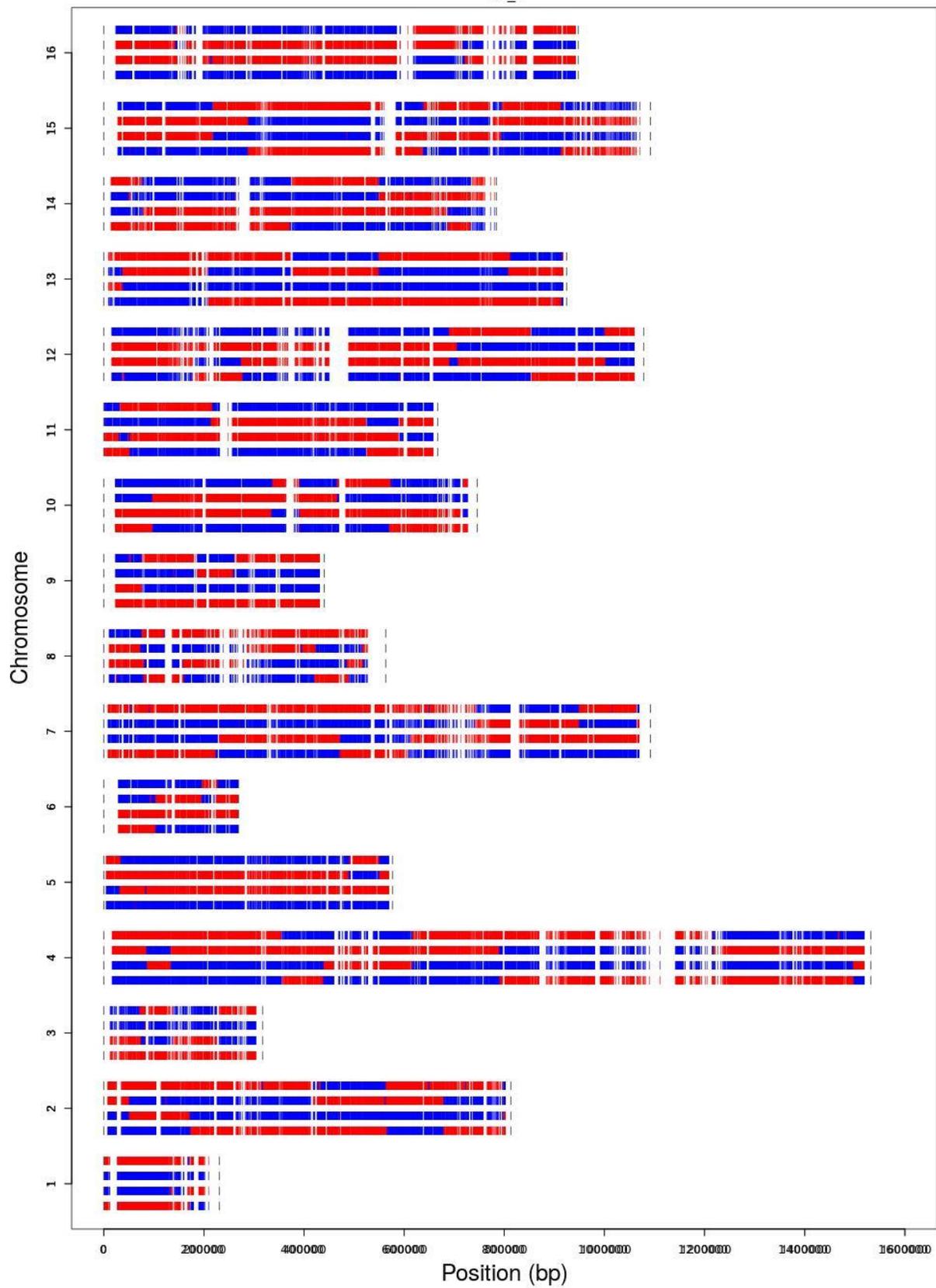
WT_2



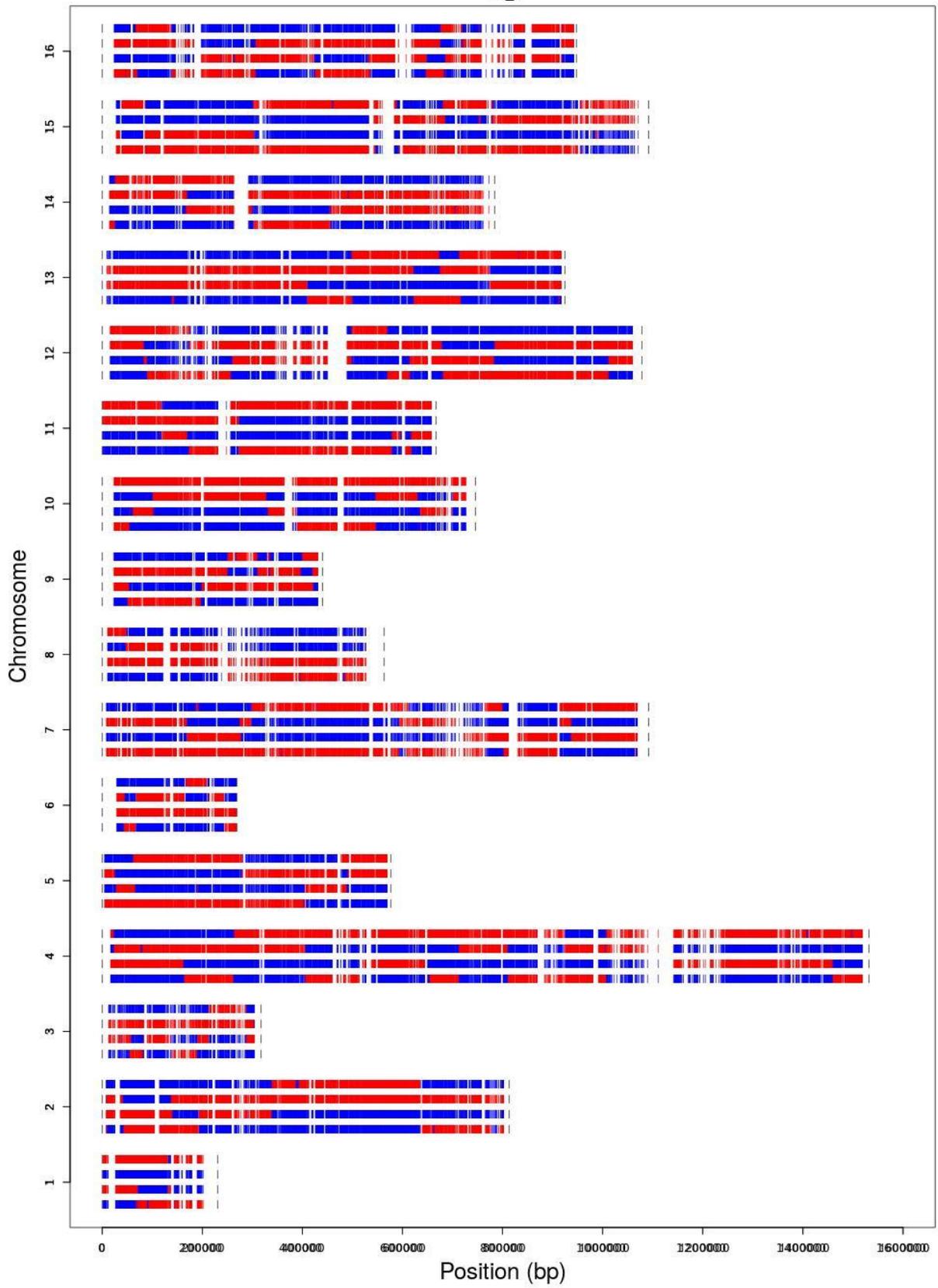
WT_5



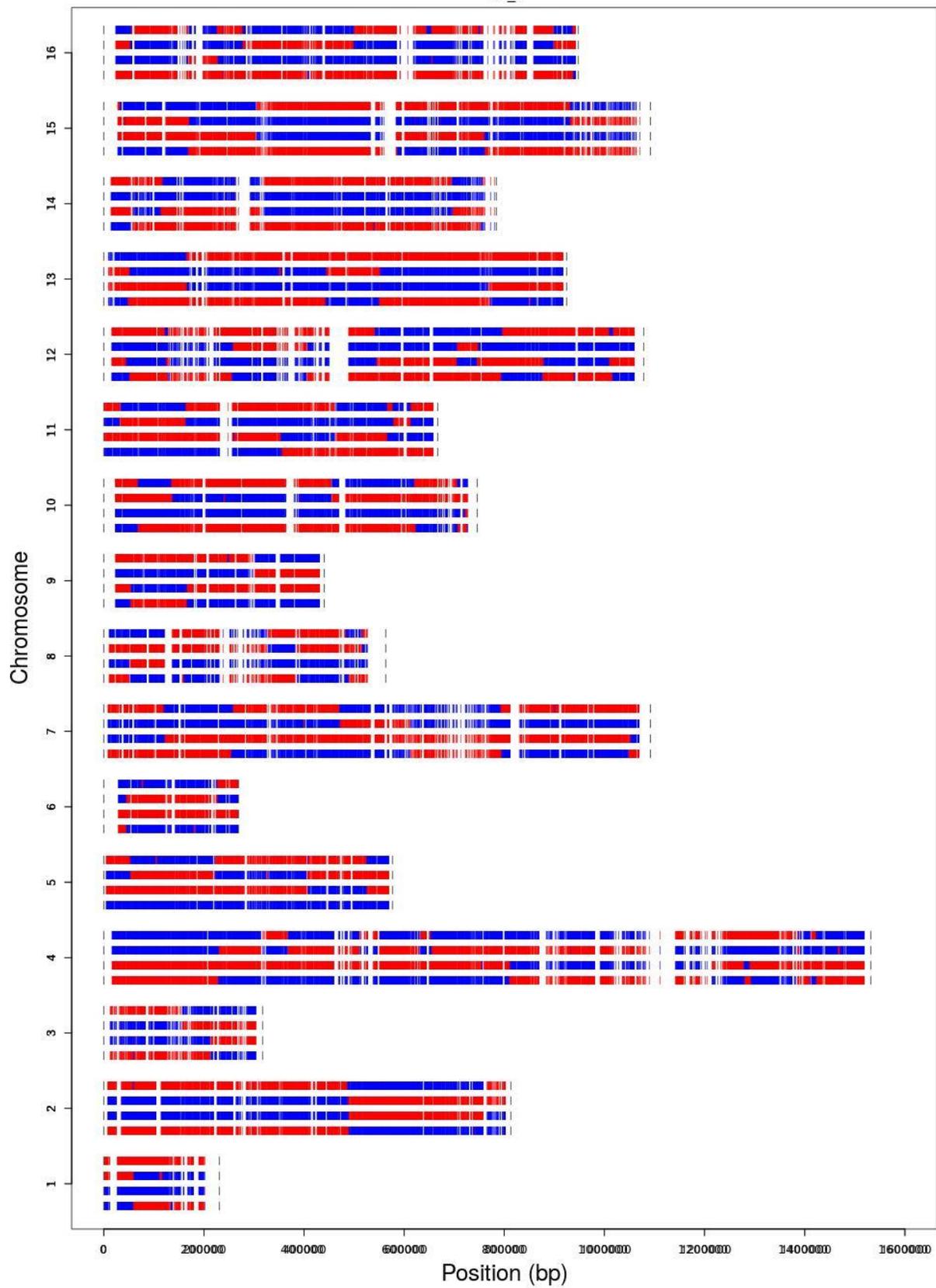
WT_6



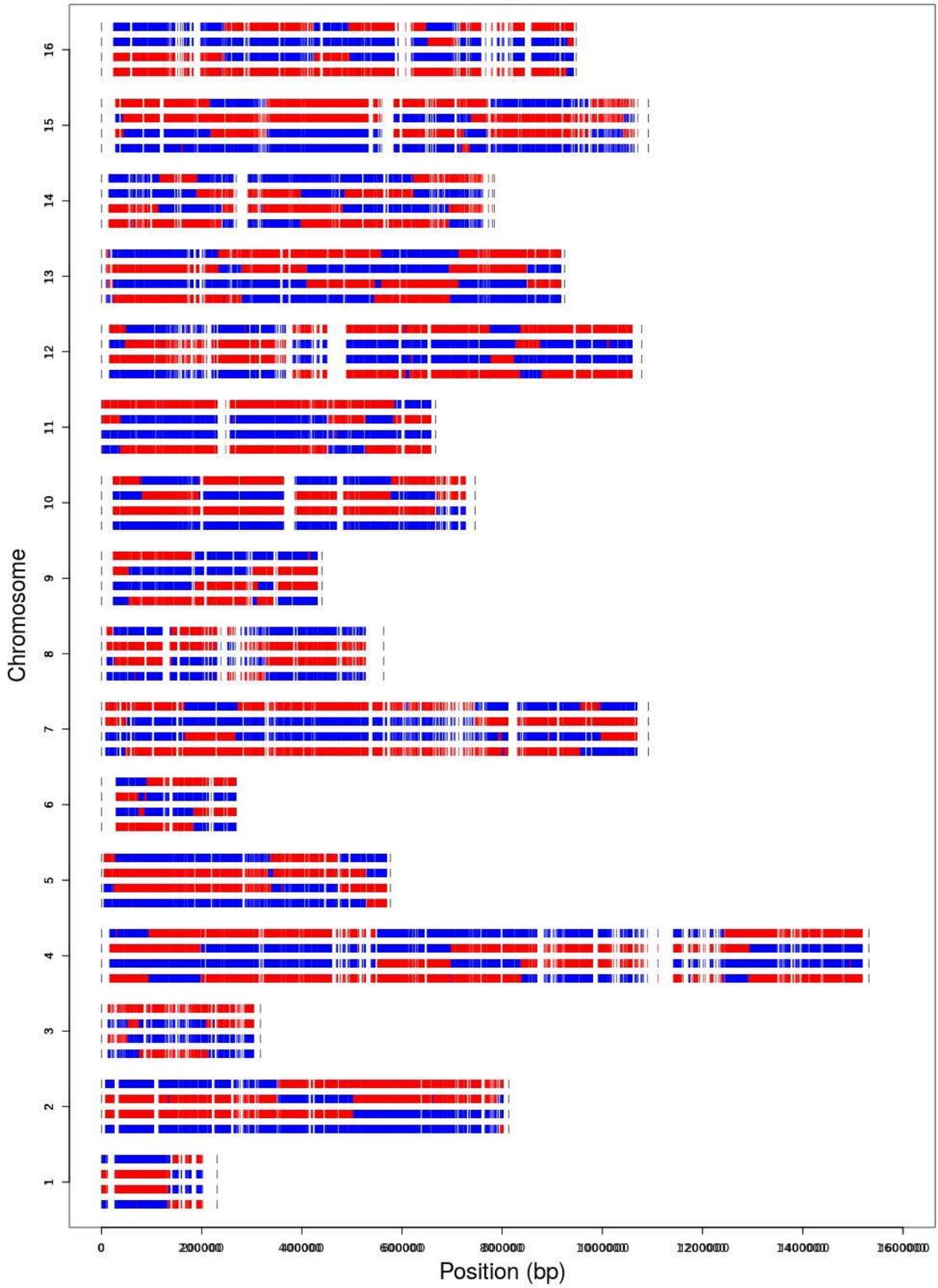
WT_7



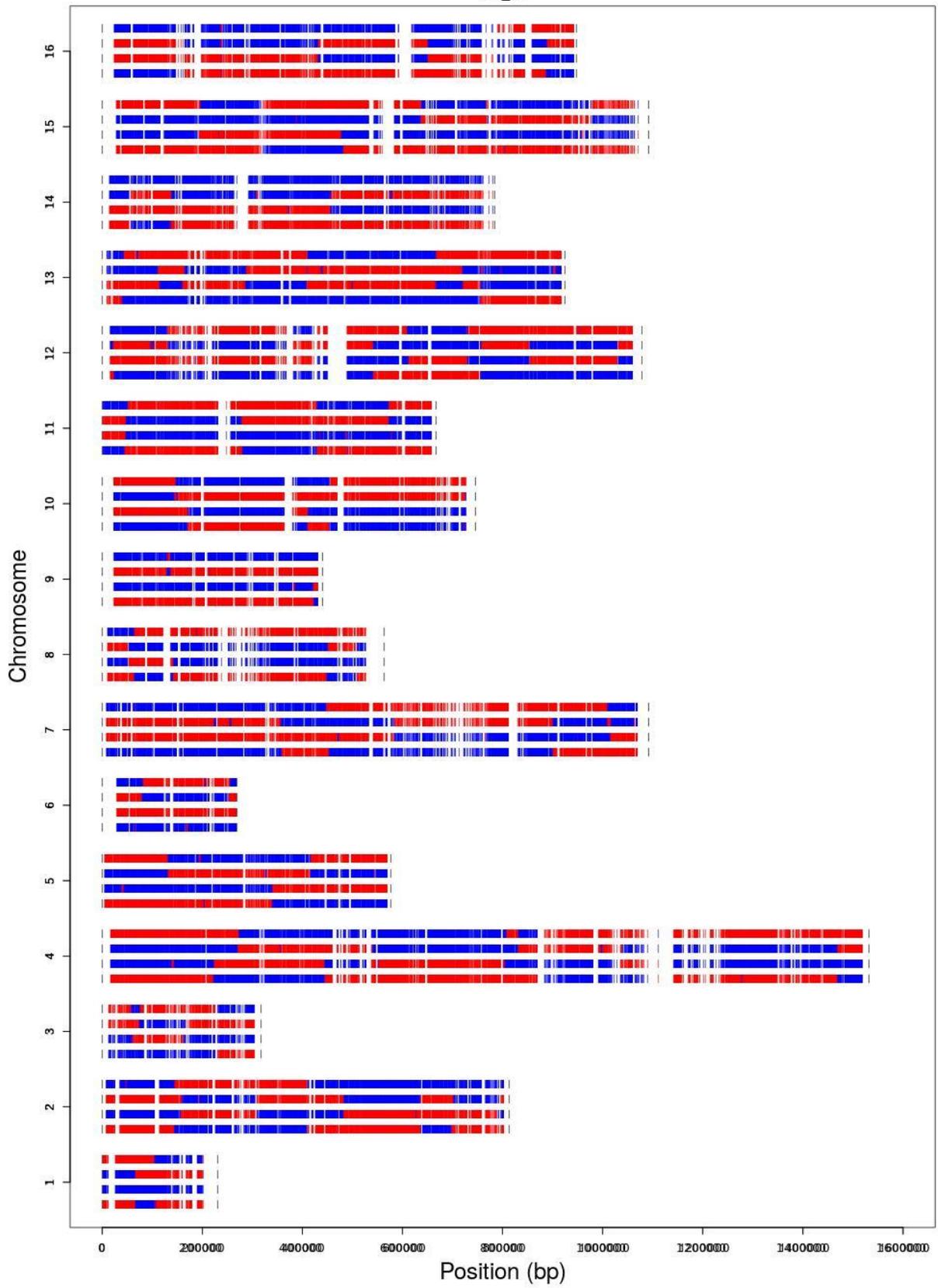
WT_8



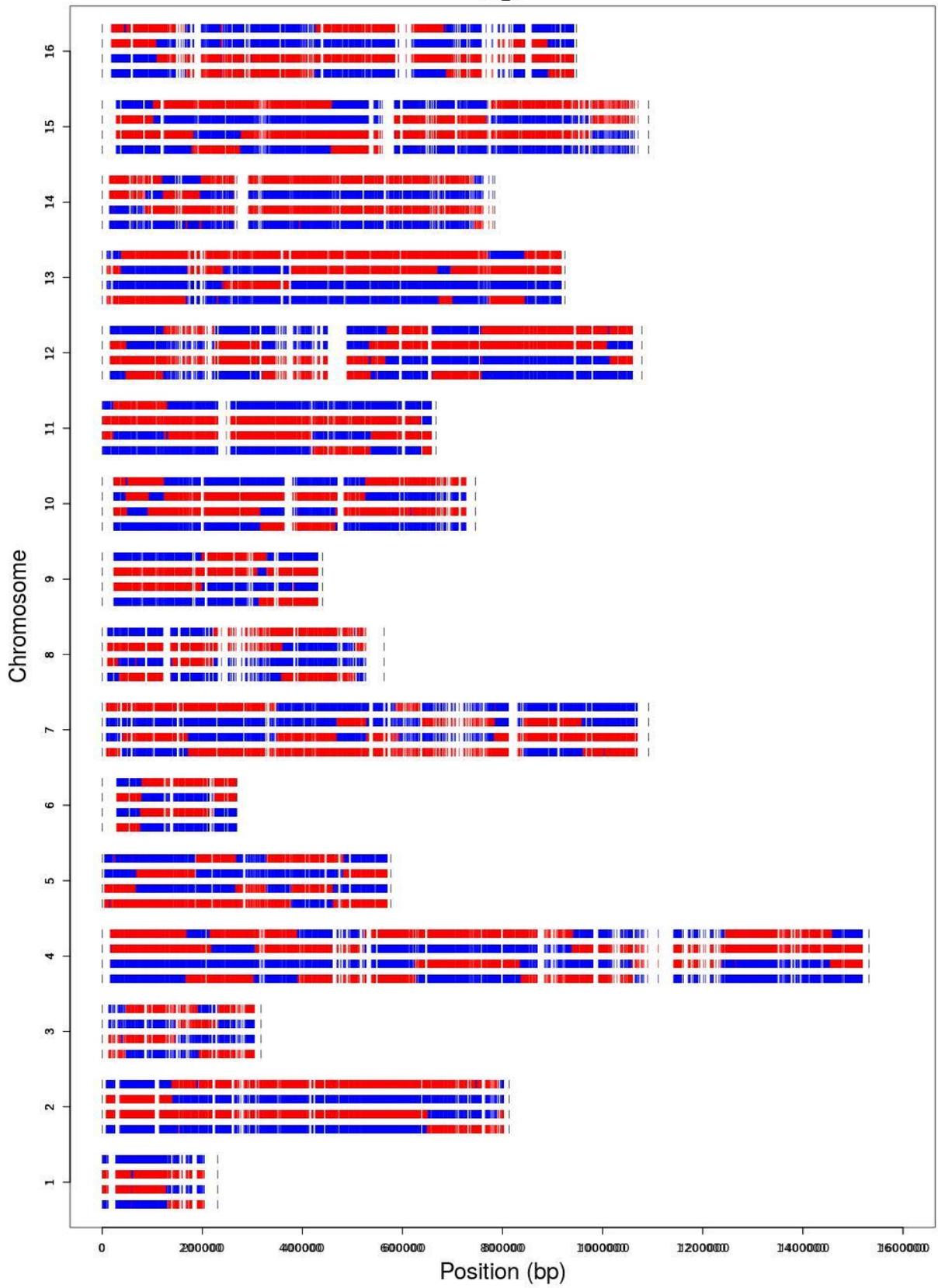
WT_9



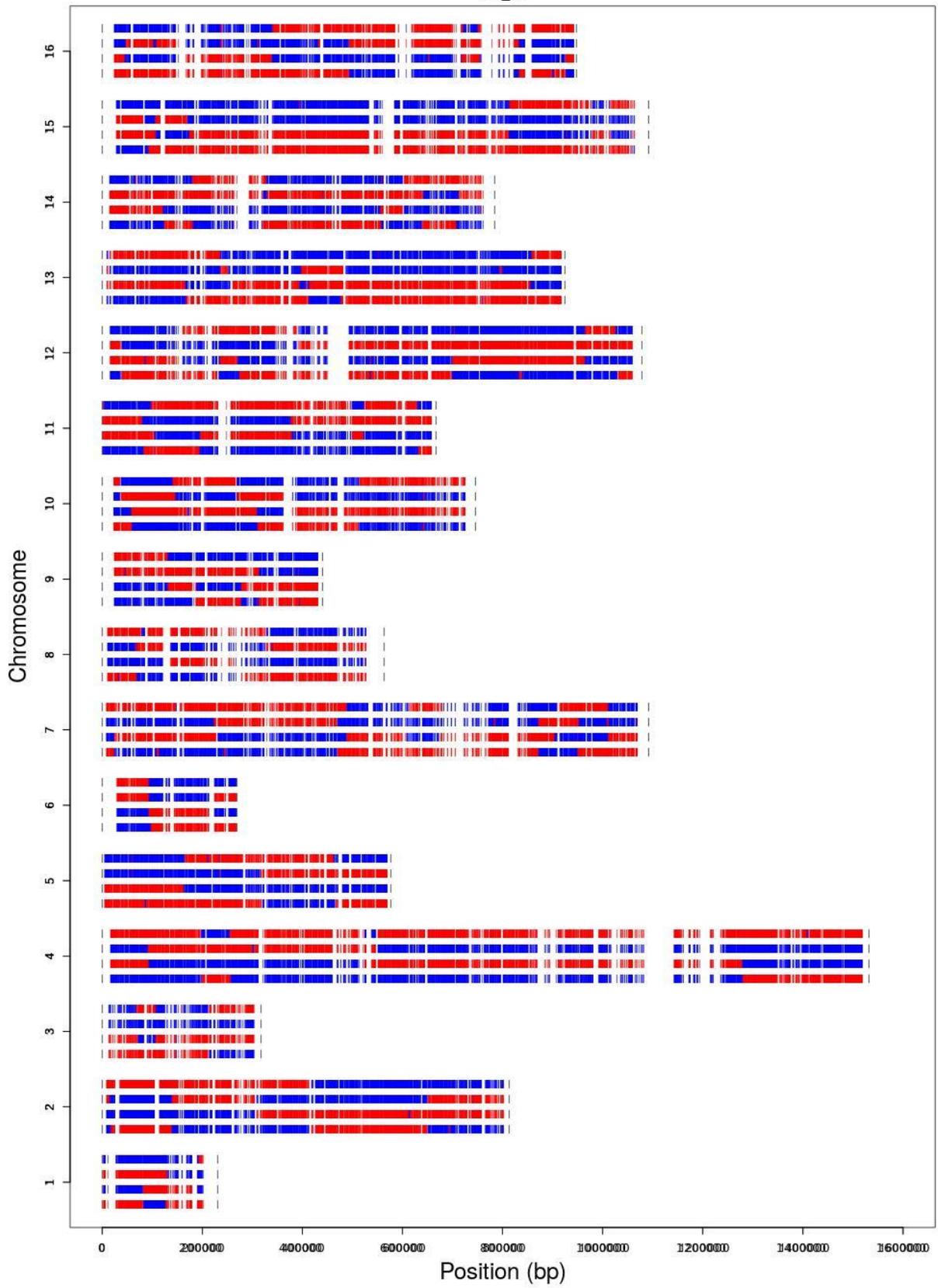
WT_10



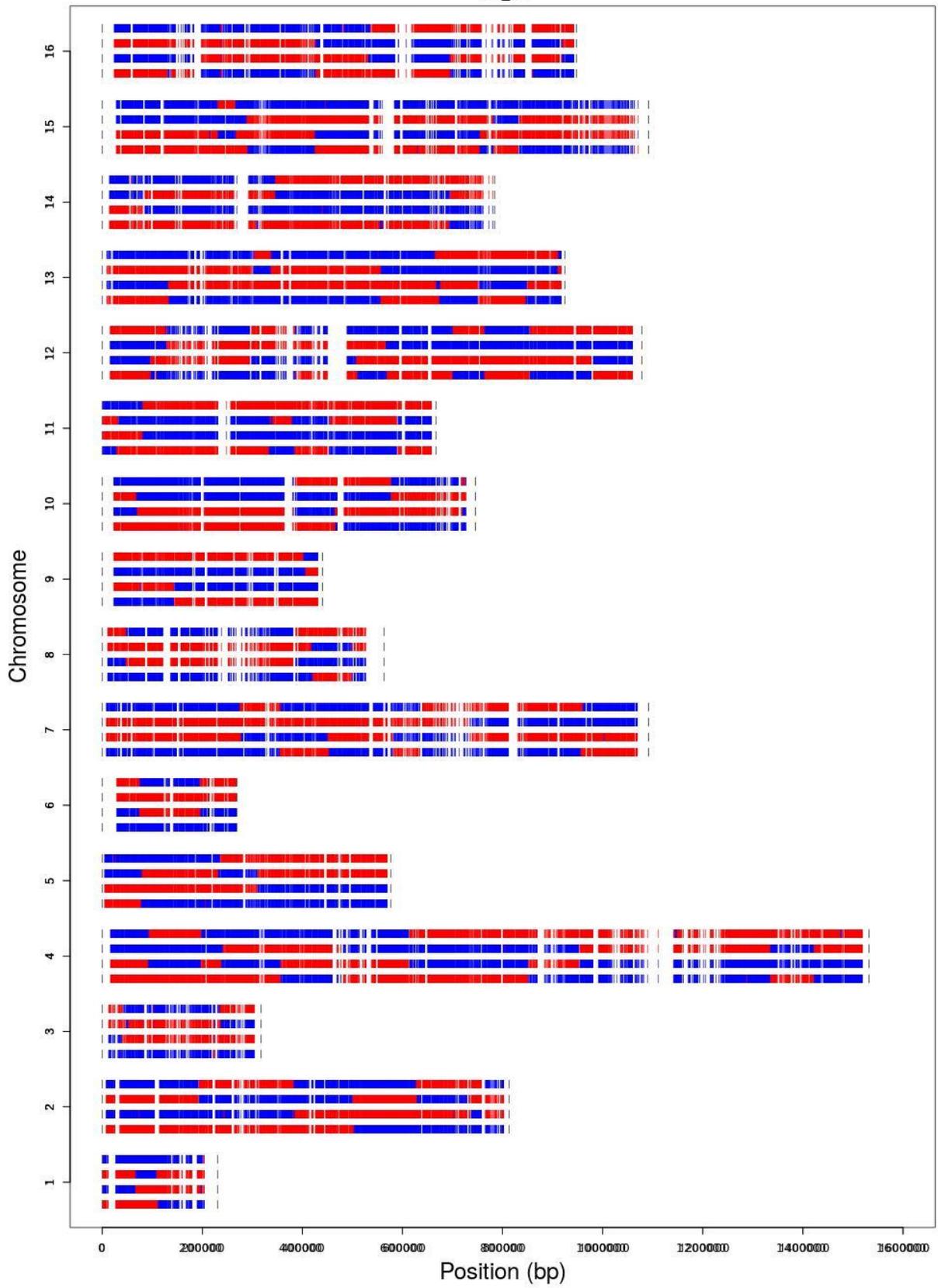
WT_11

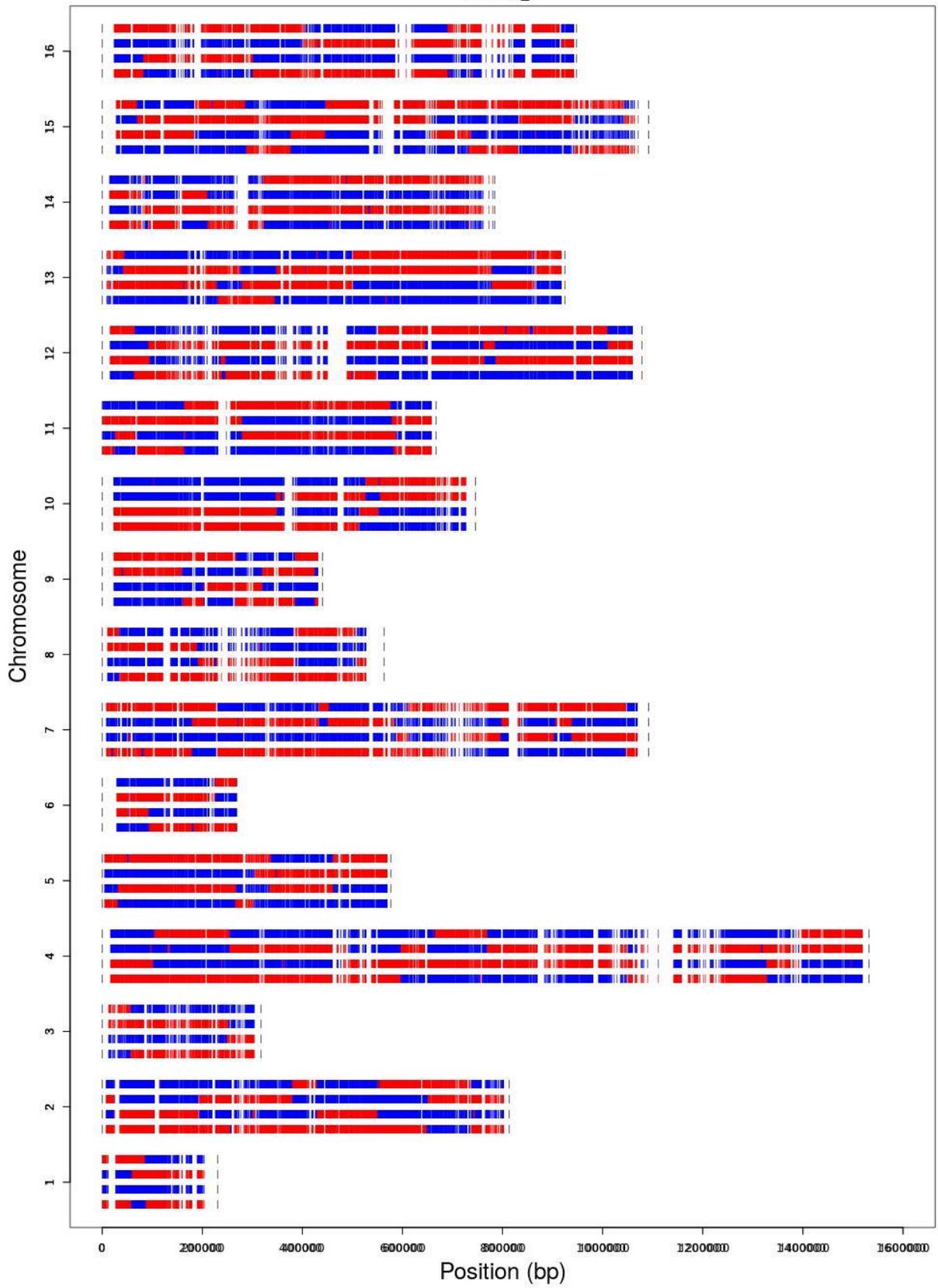


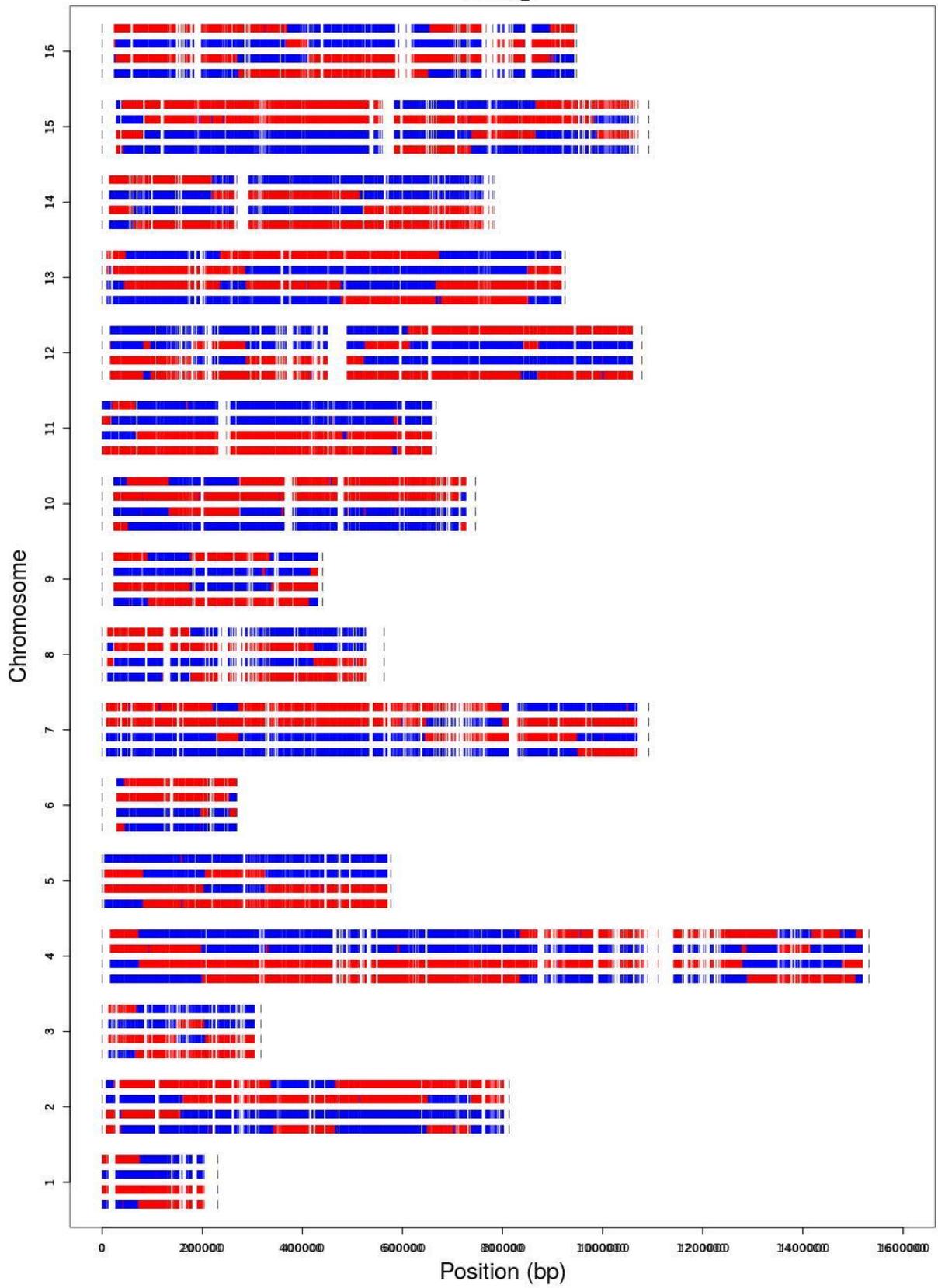
WT_12

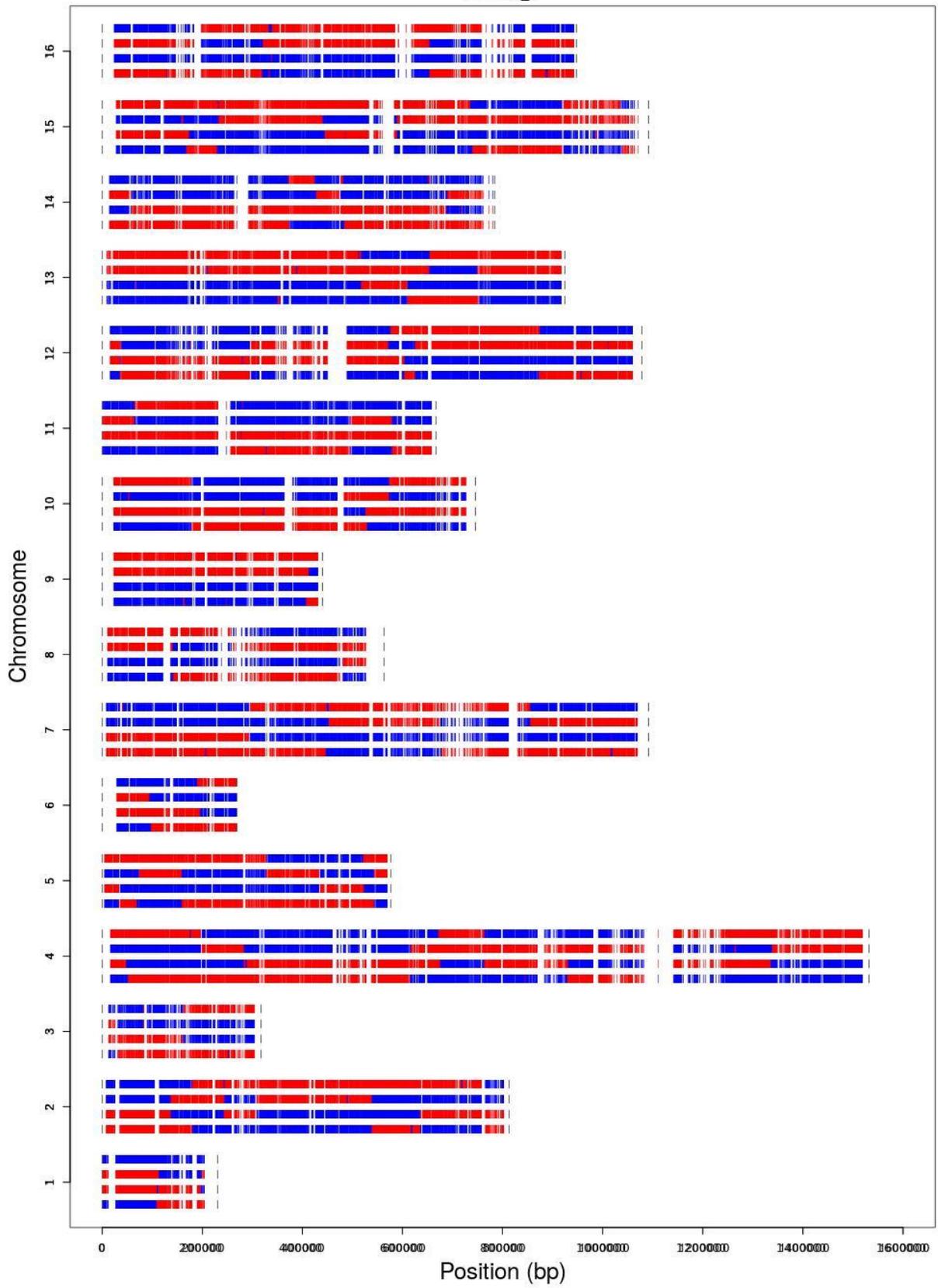


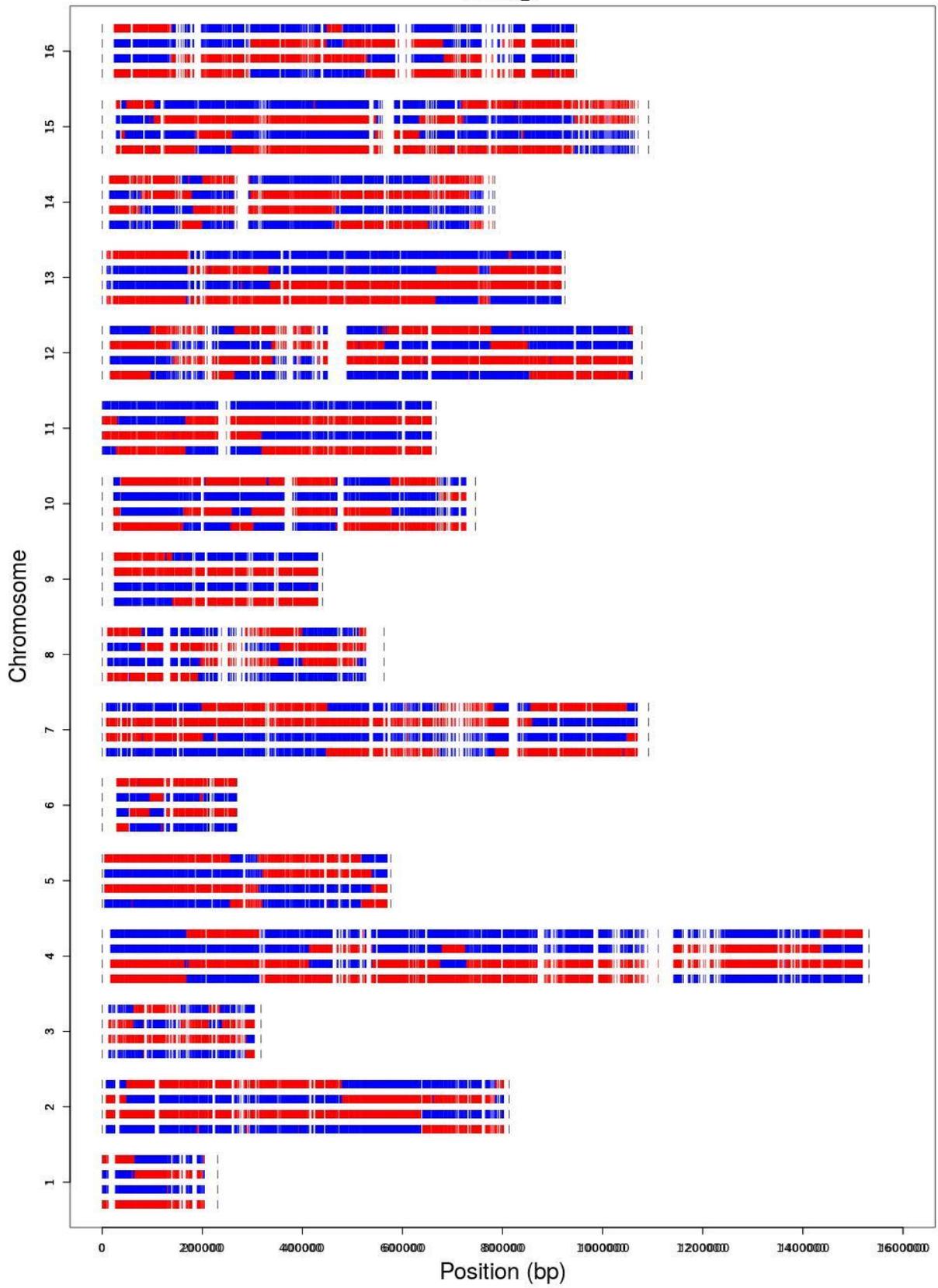
WT_13

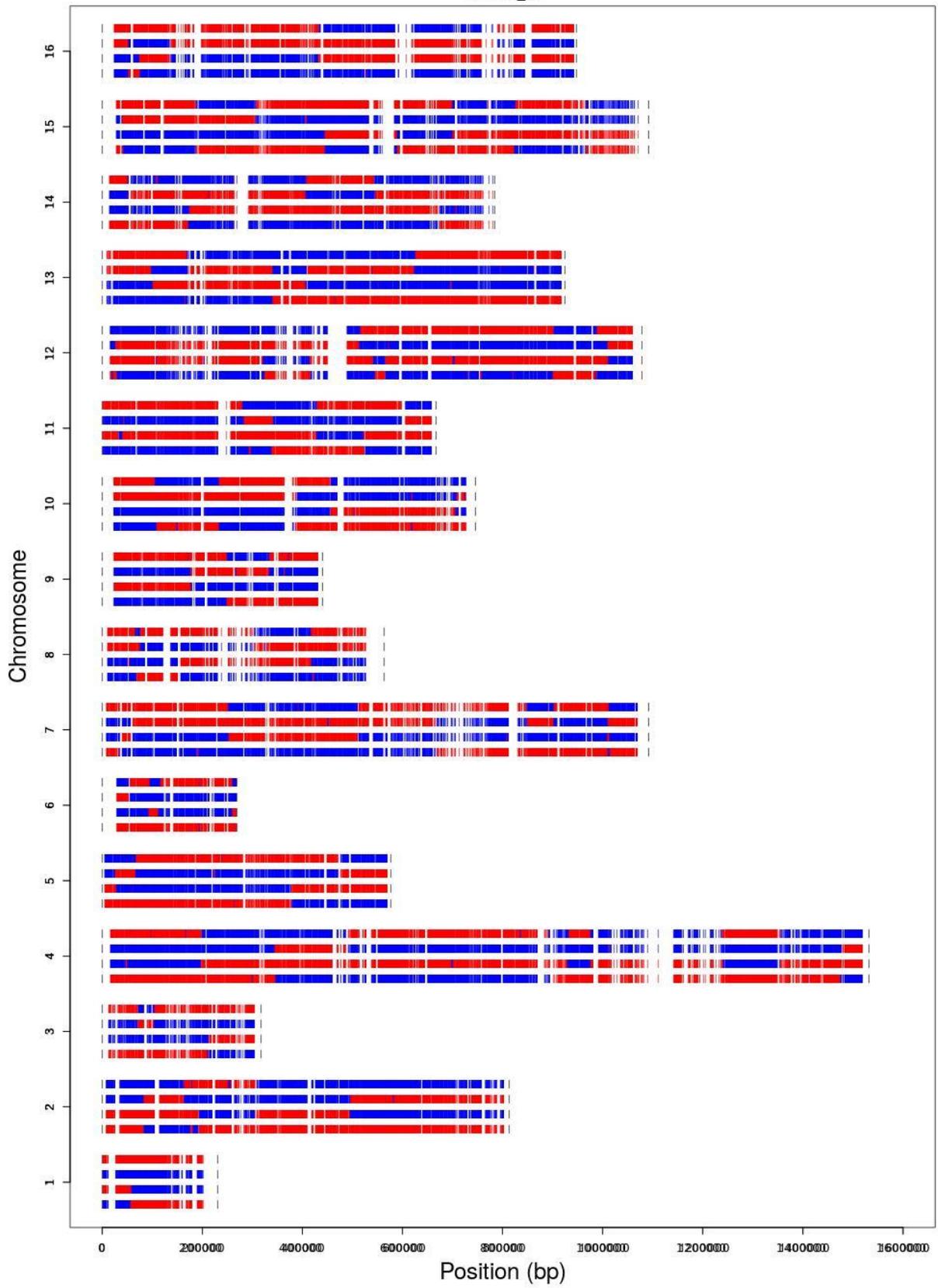


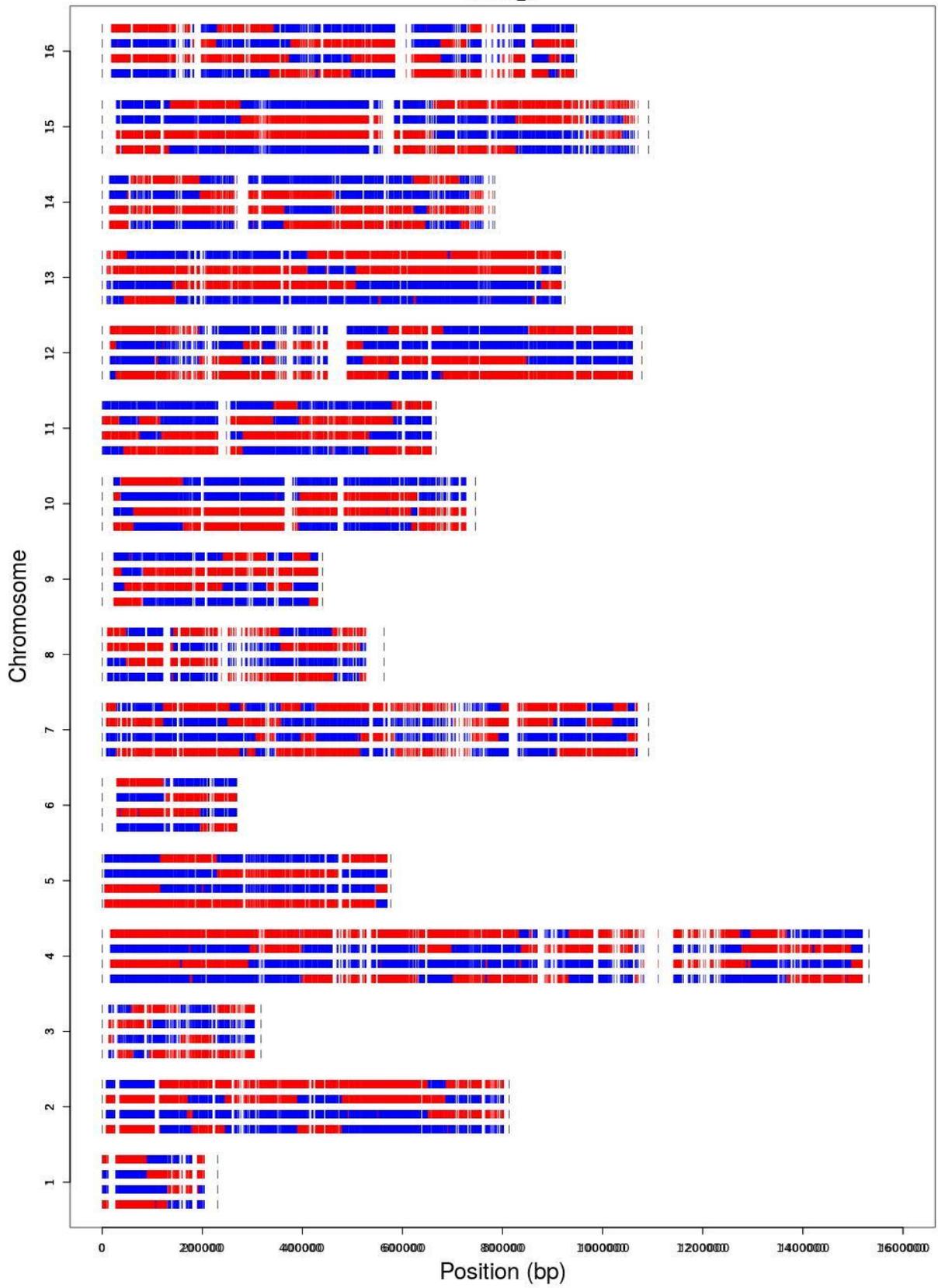


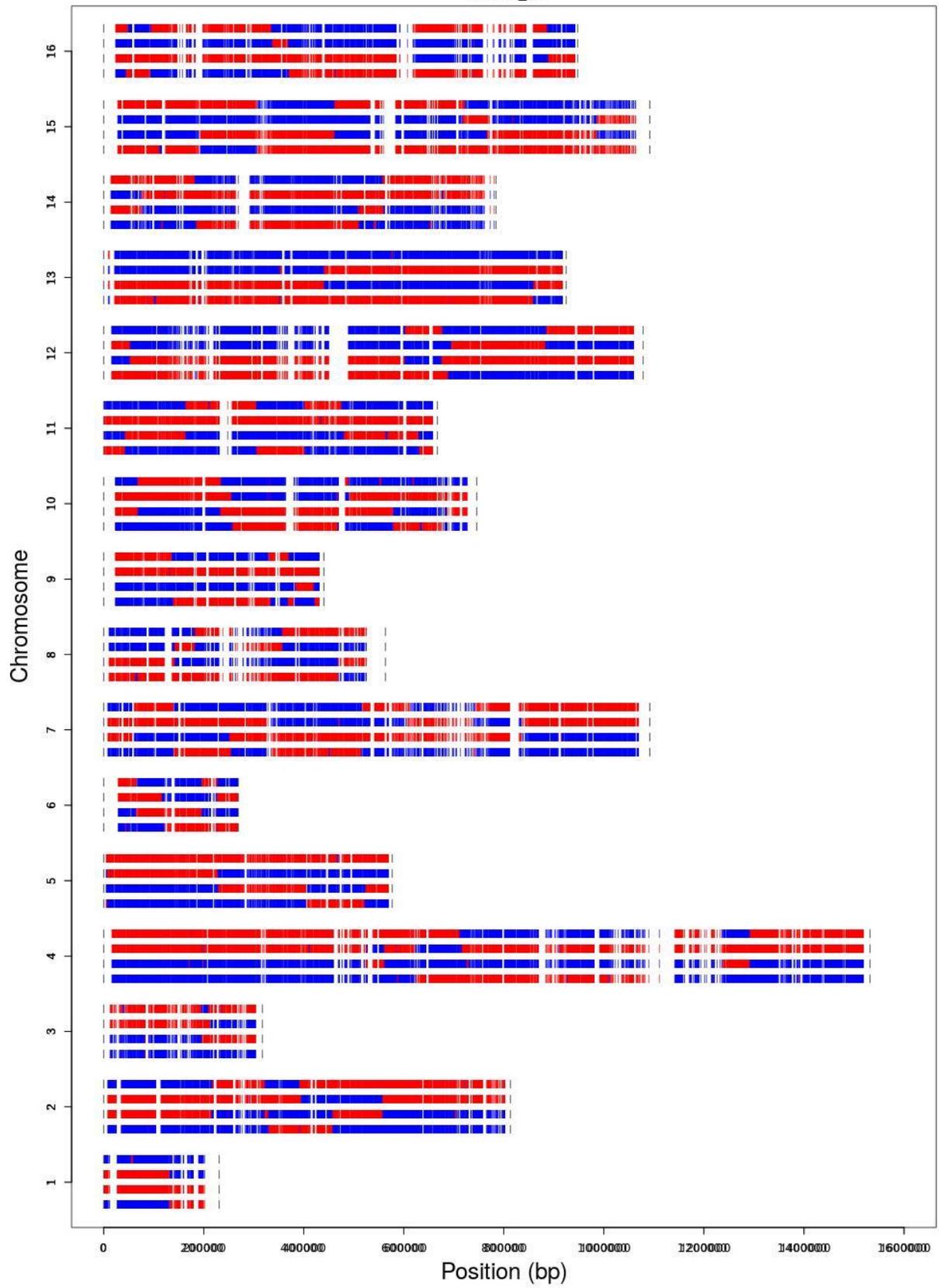


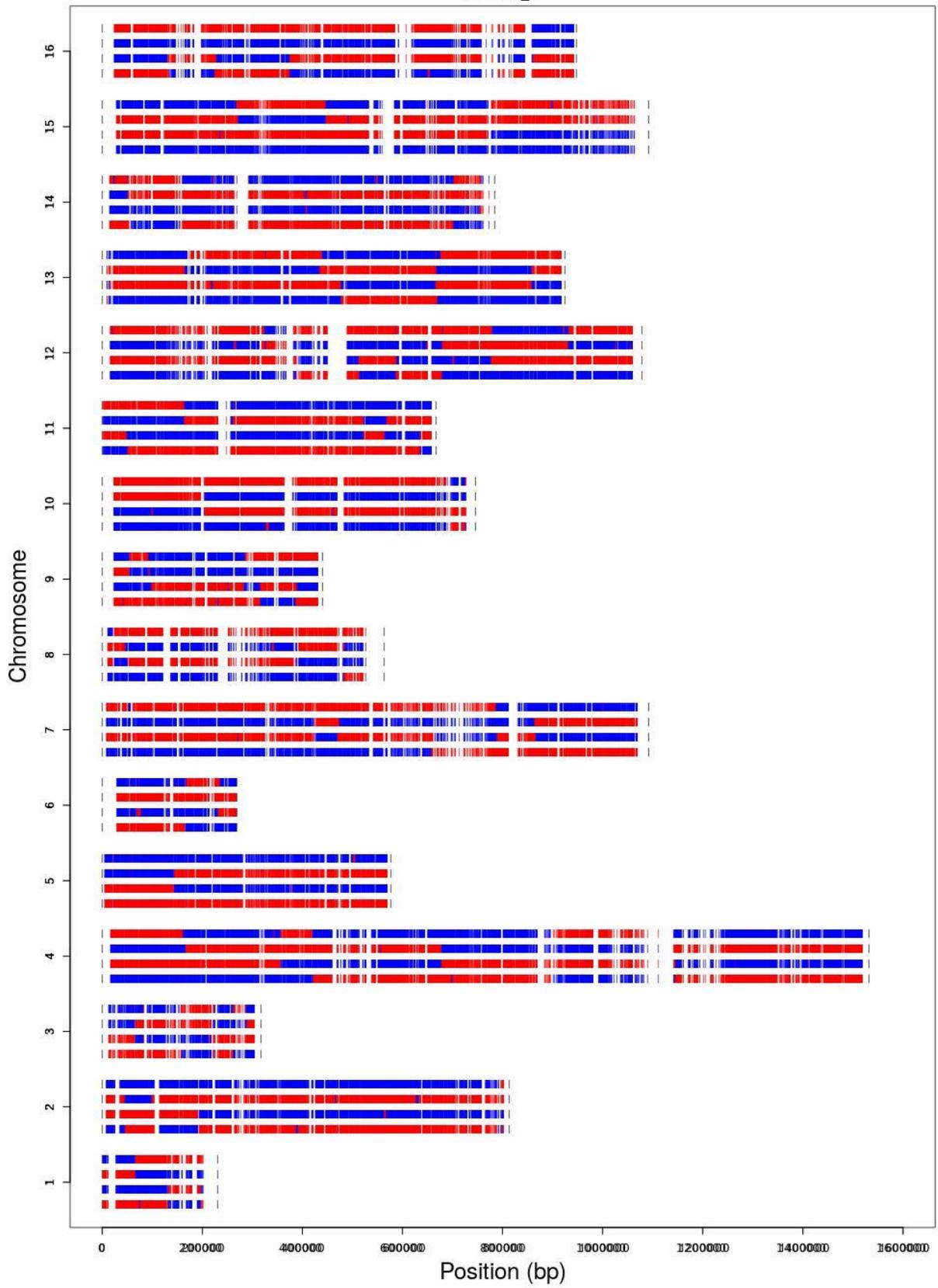




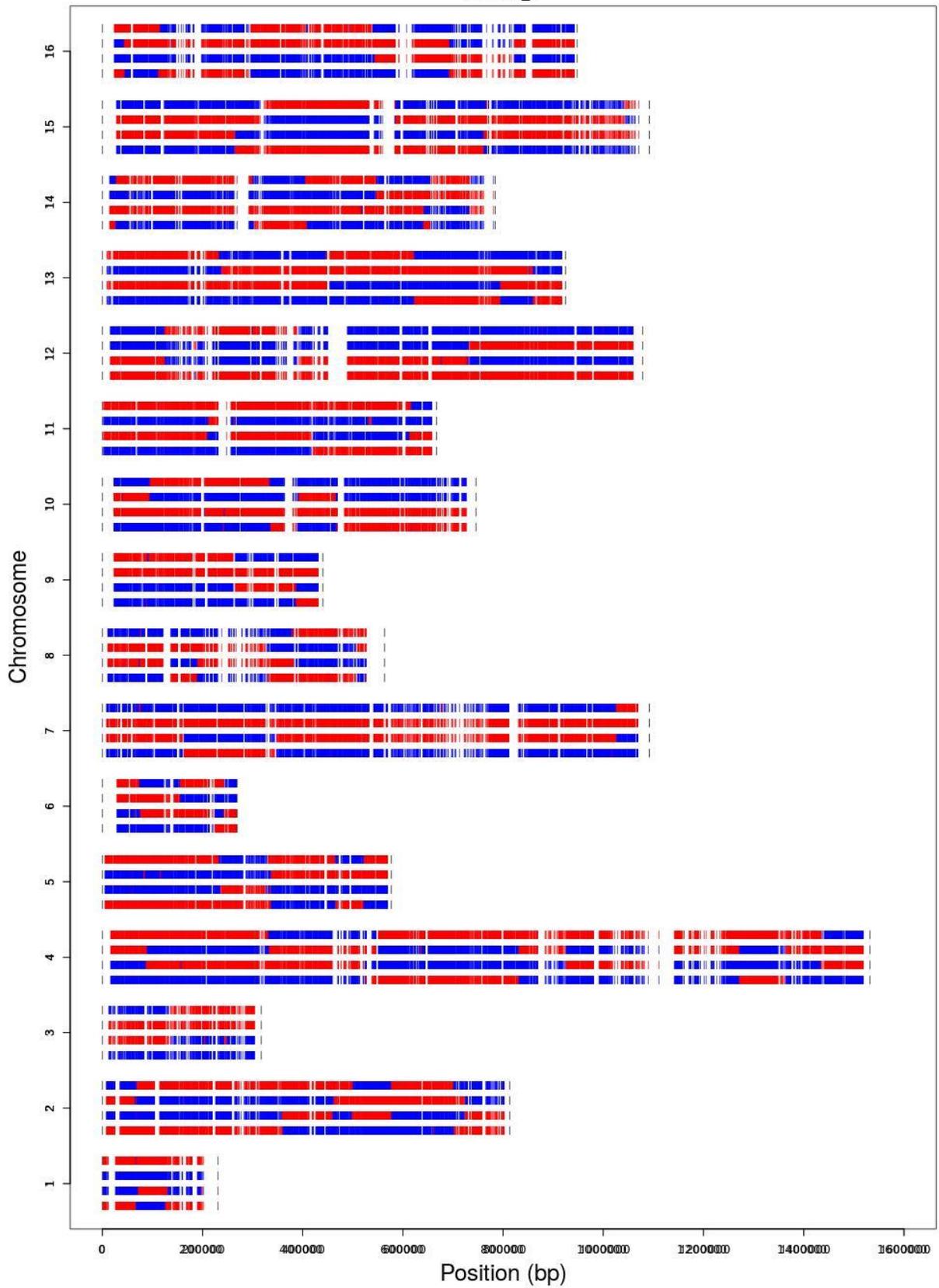


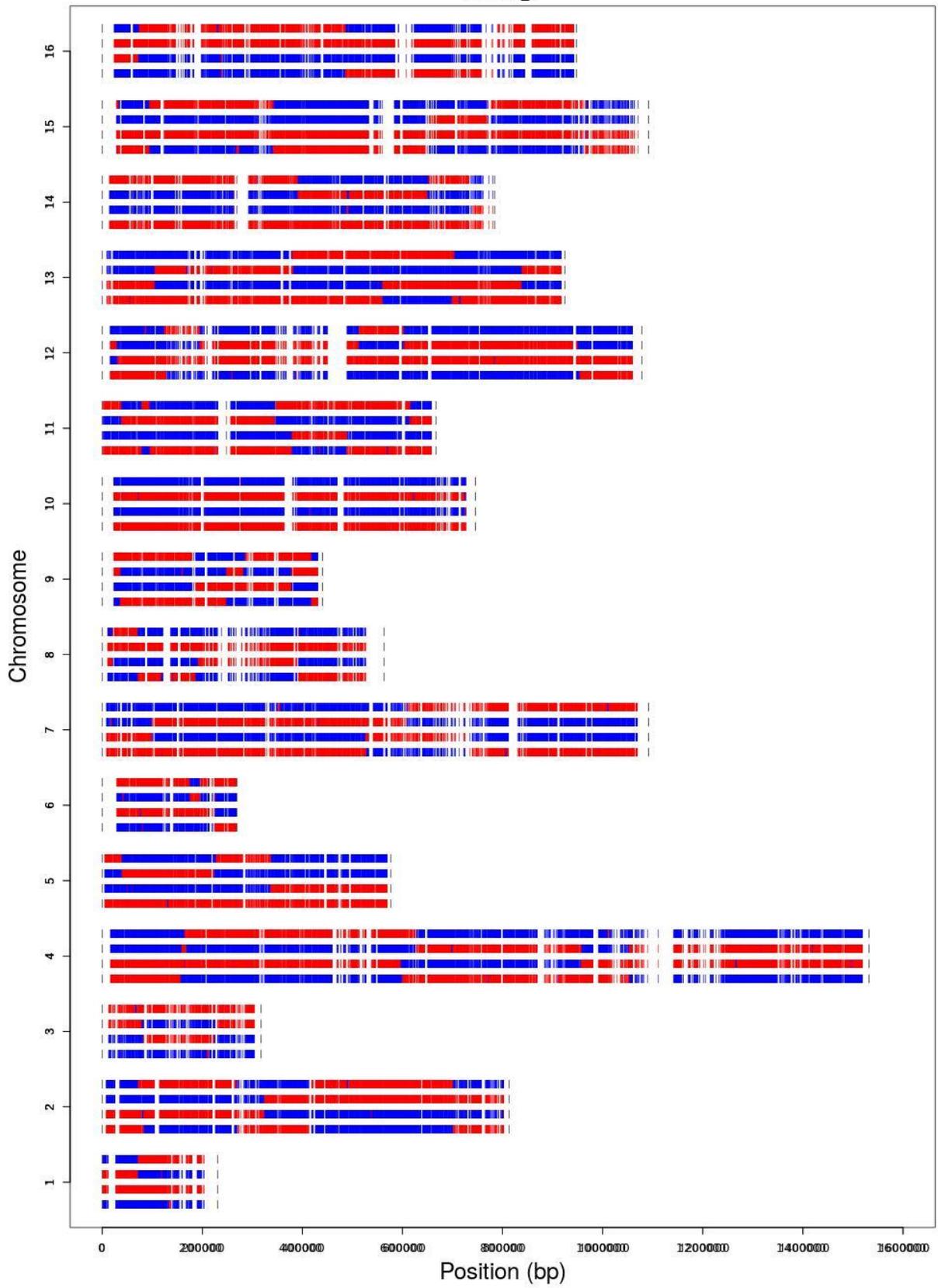


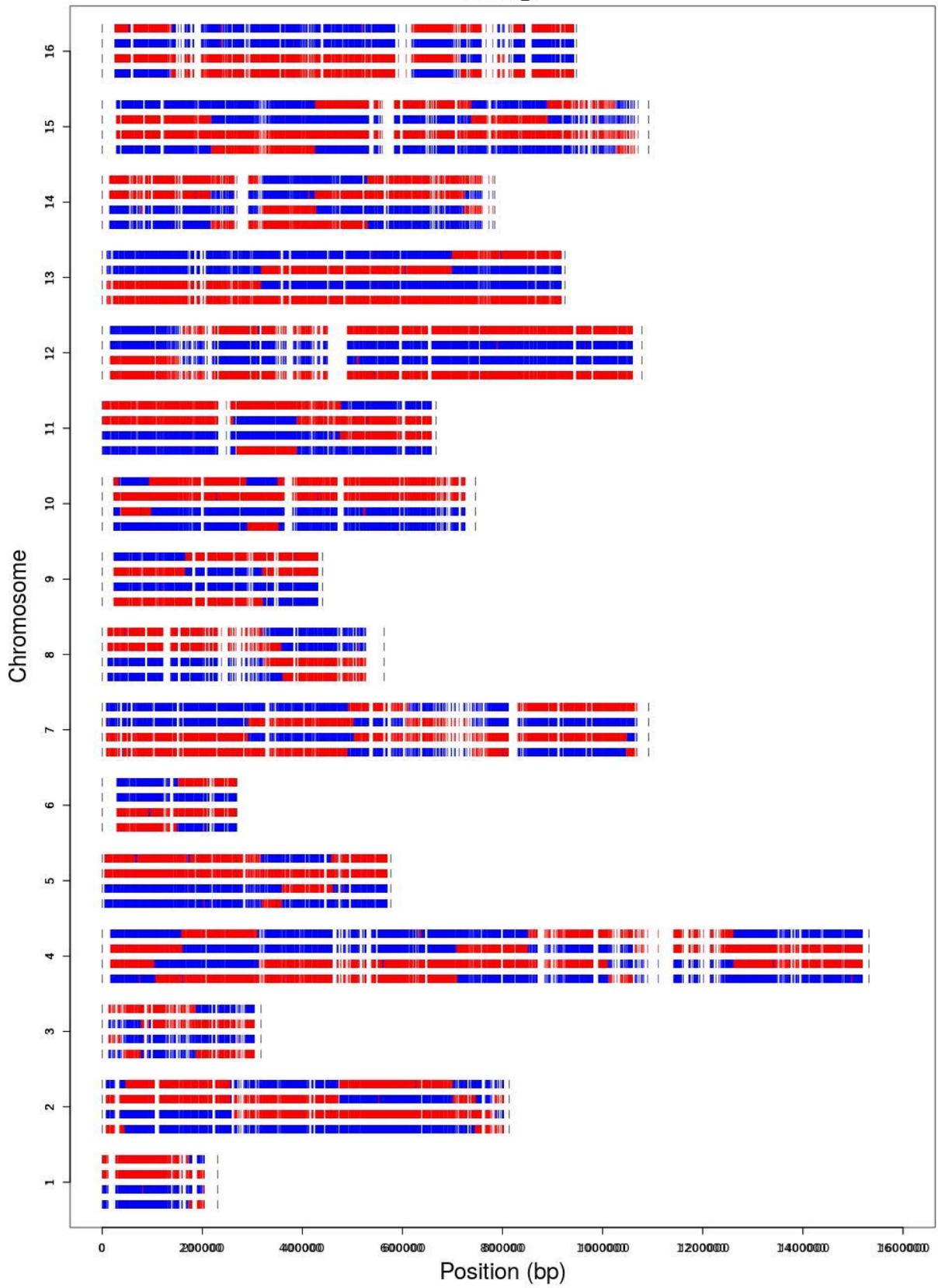


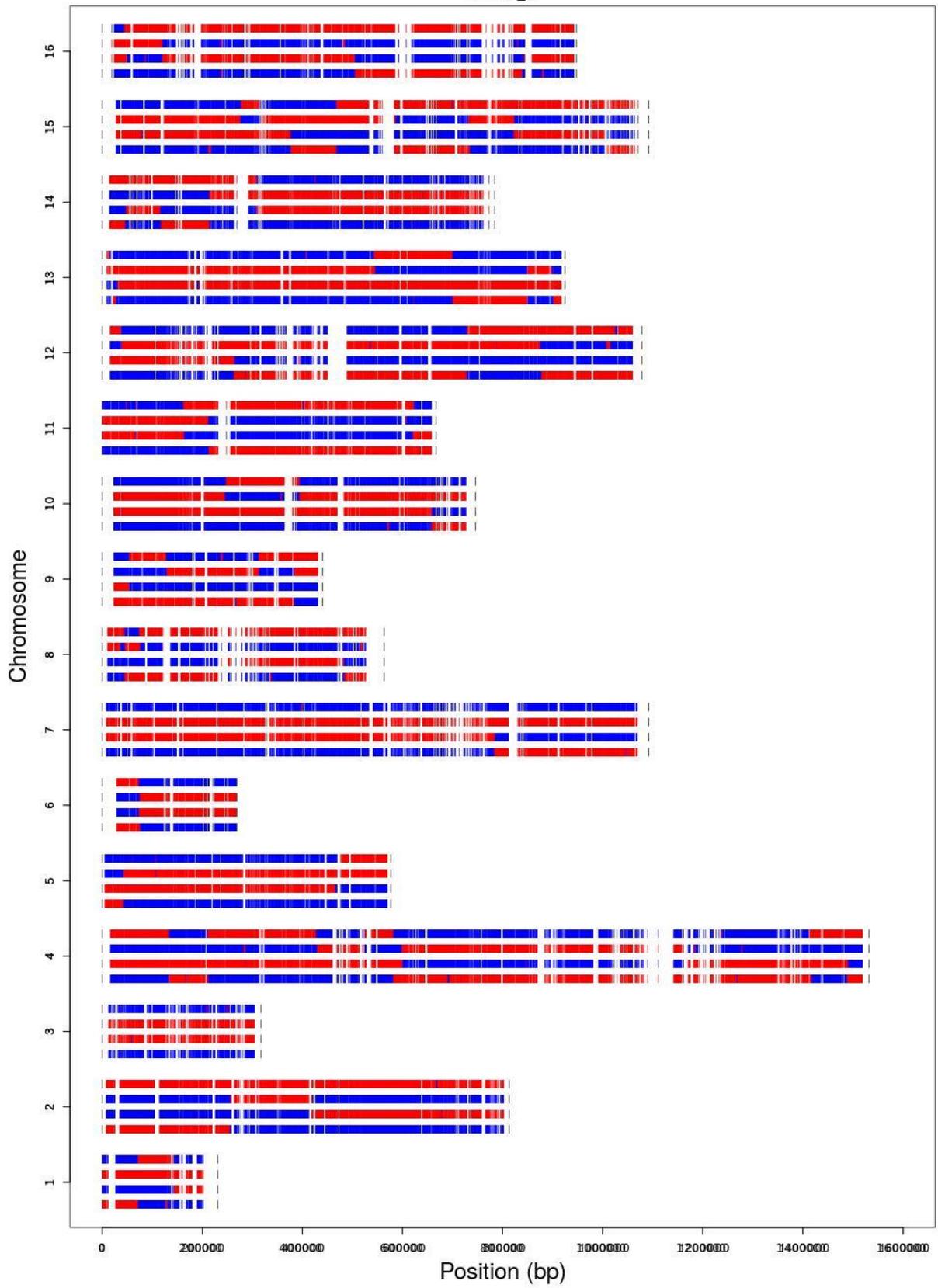


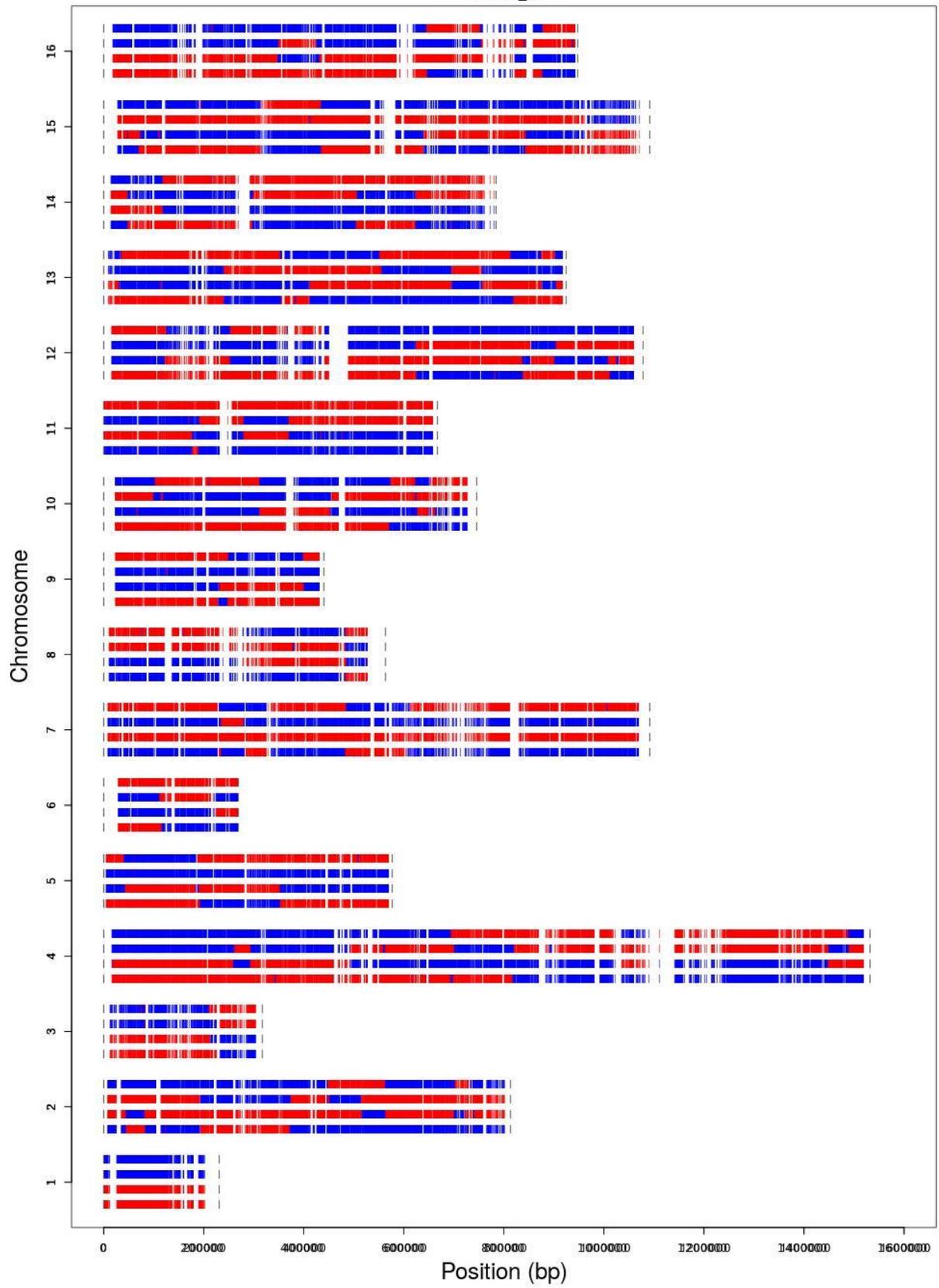
mlh3-32_7

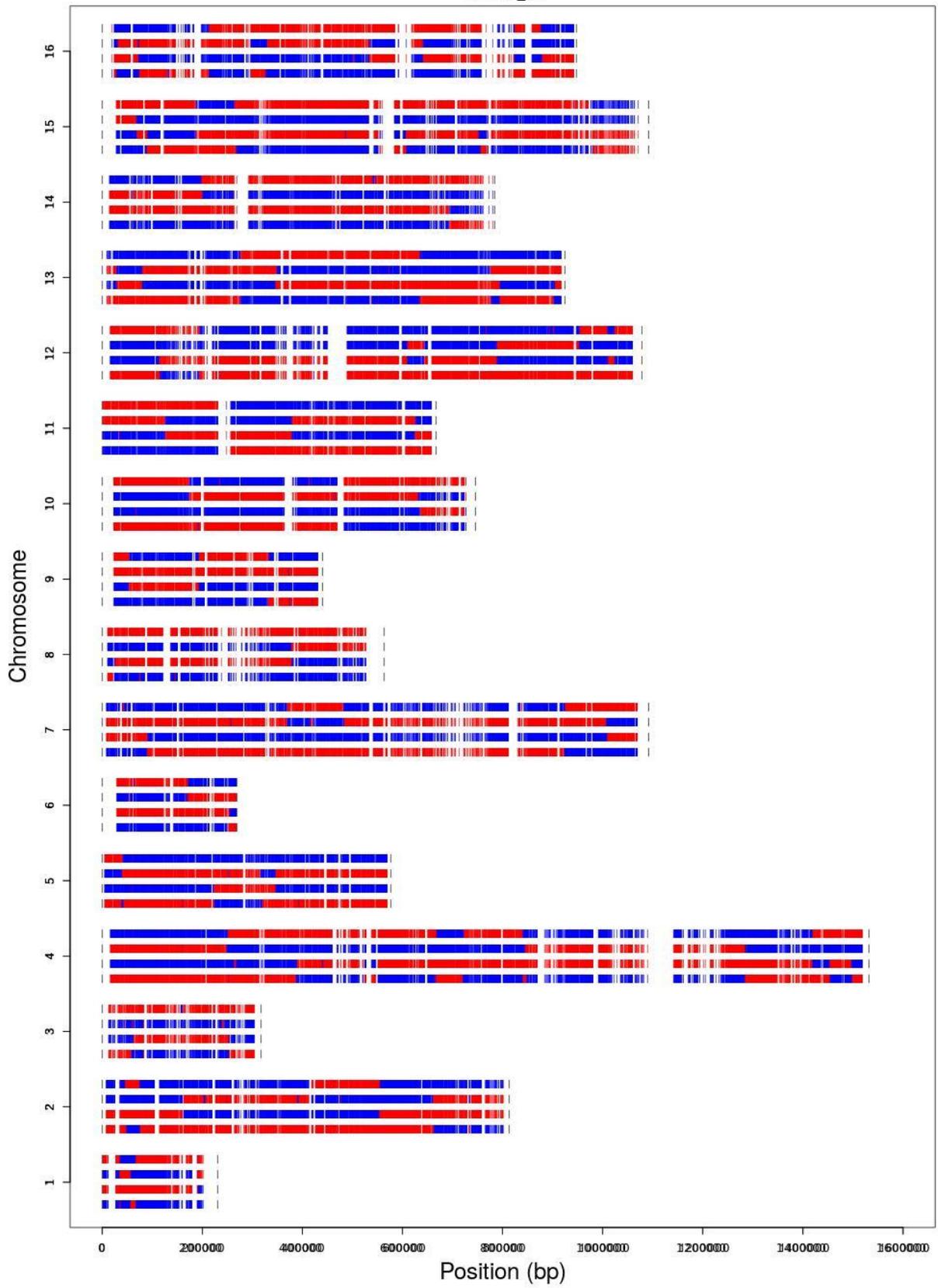




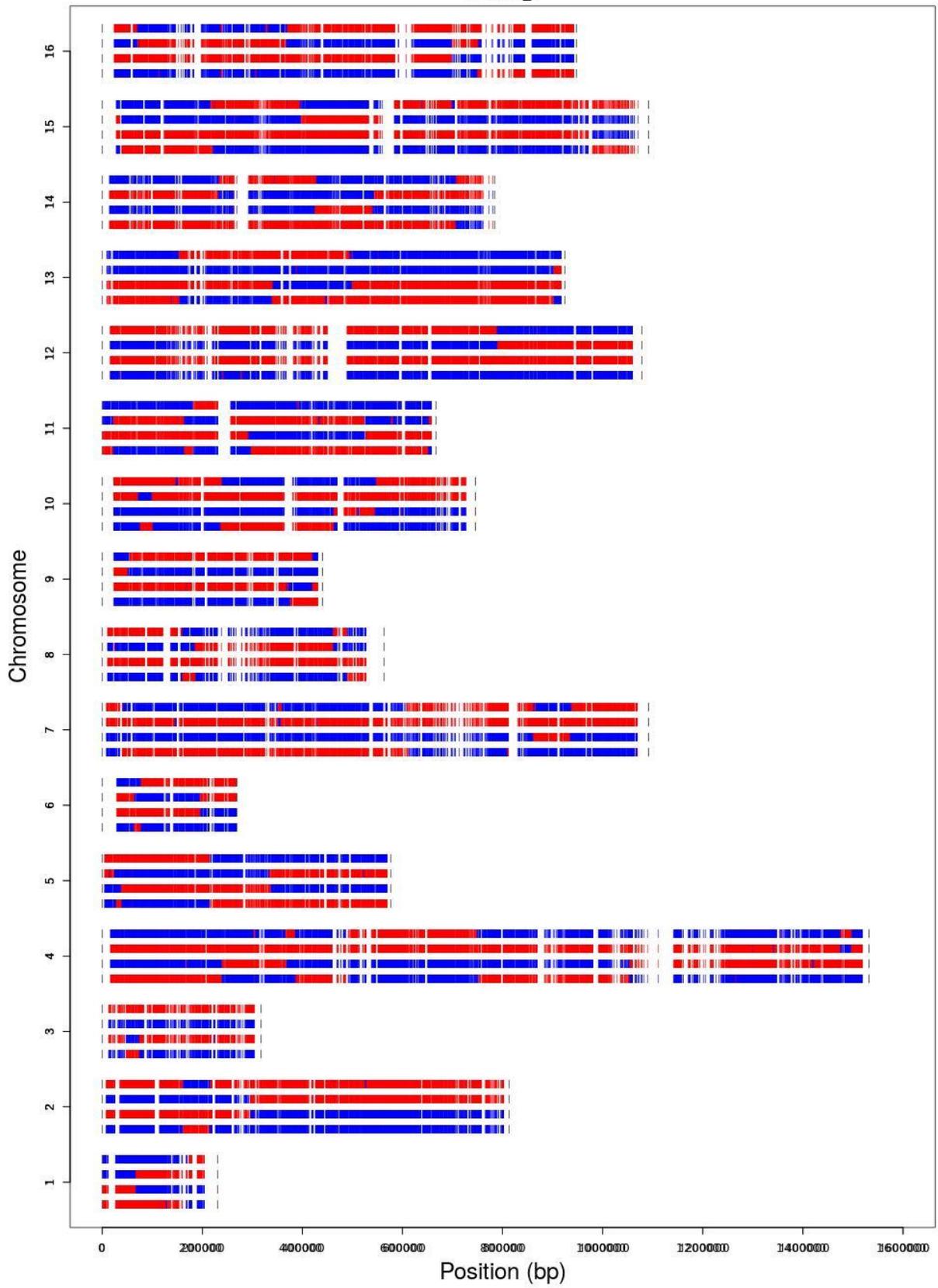




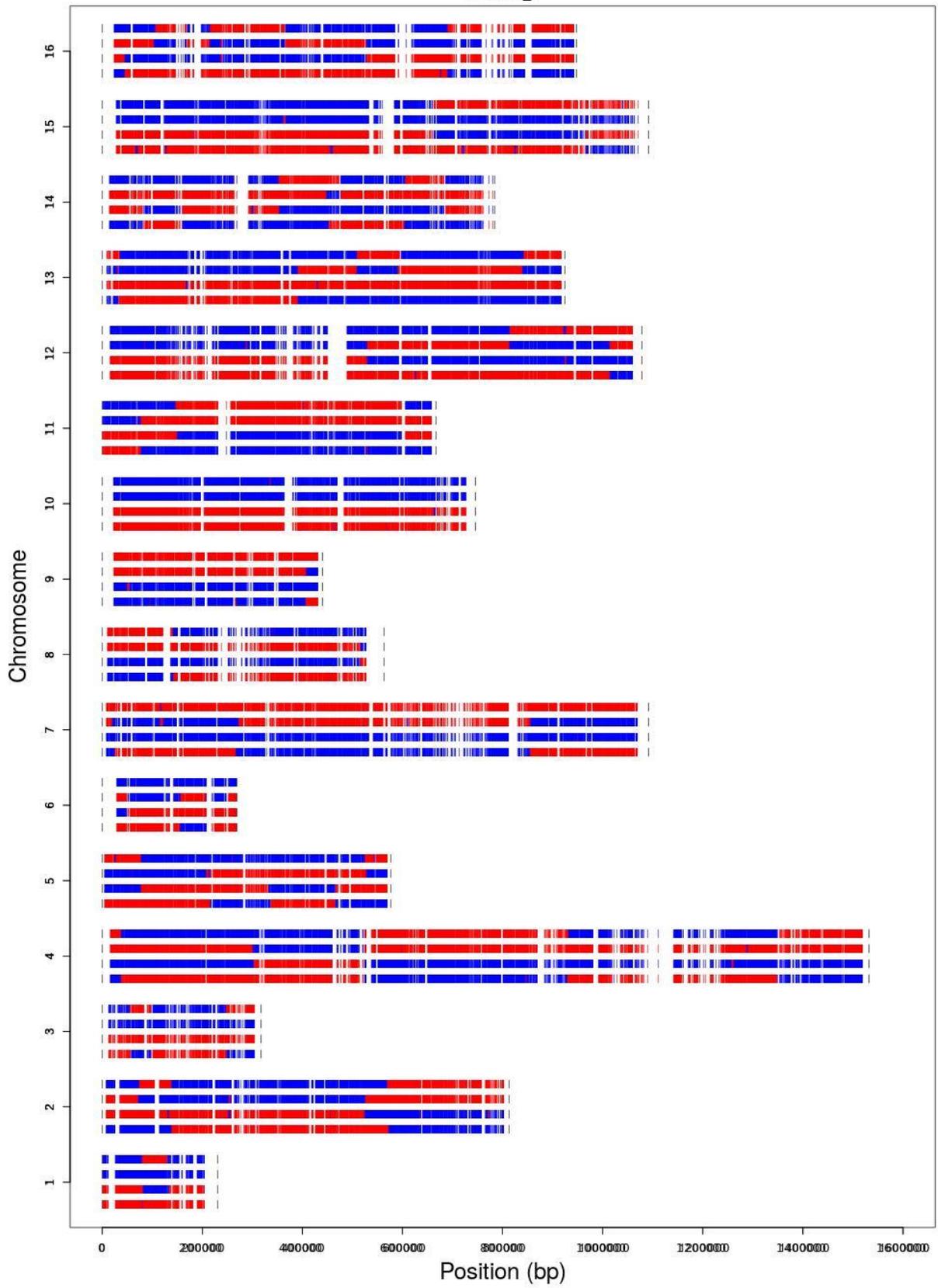




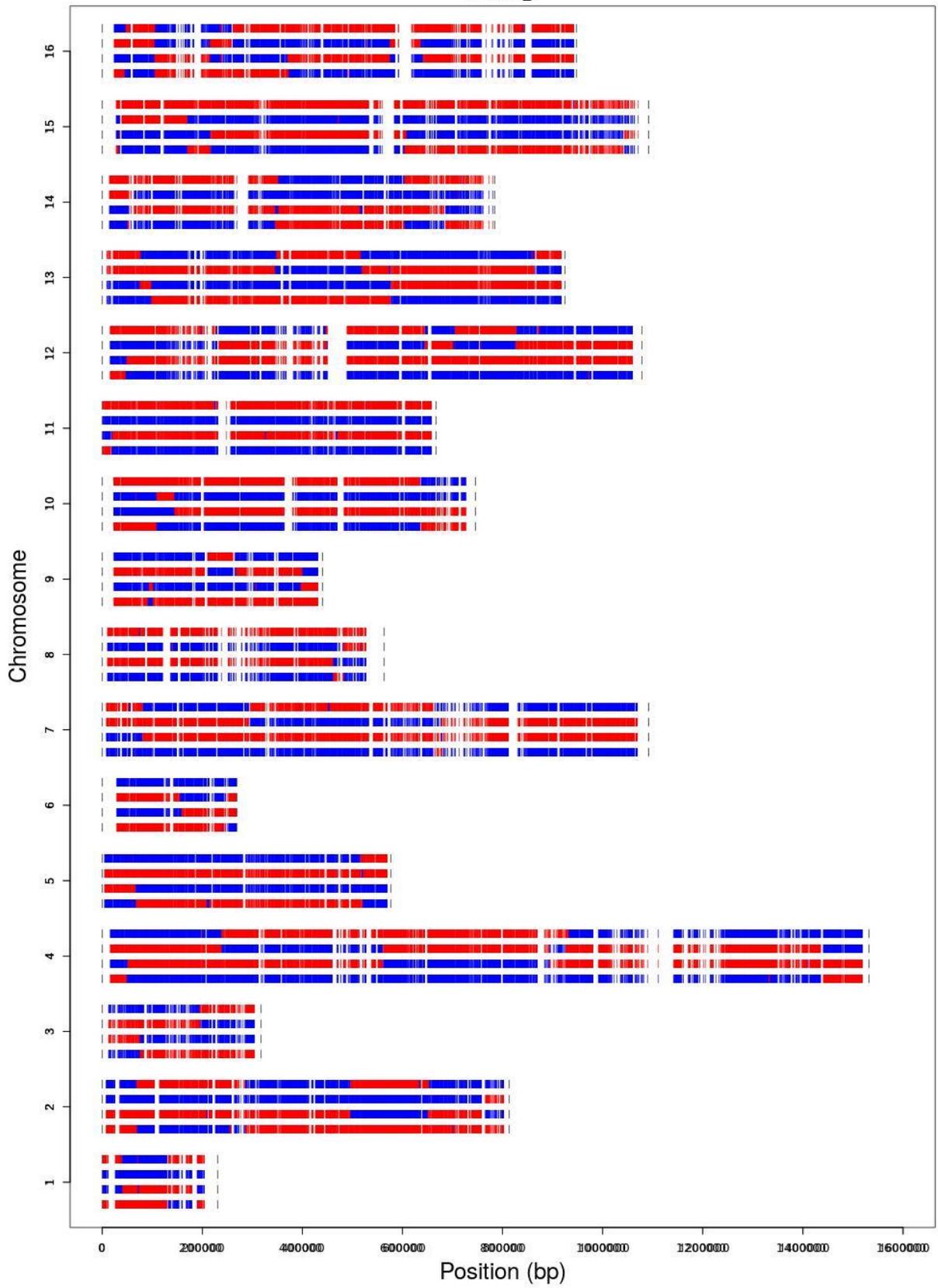
mlh3-DN_1



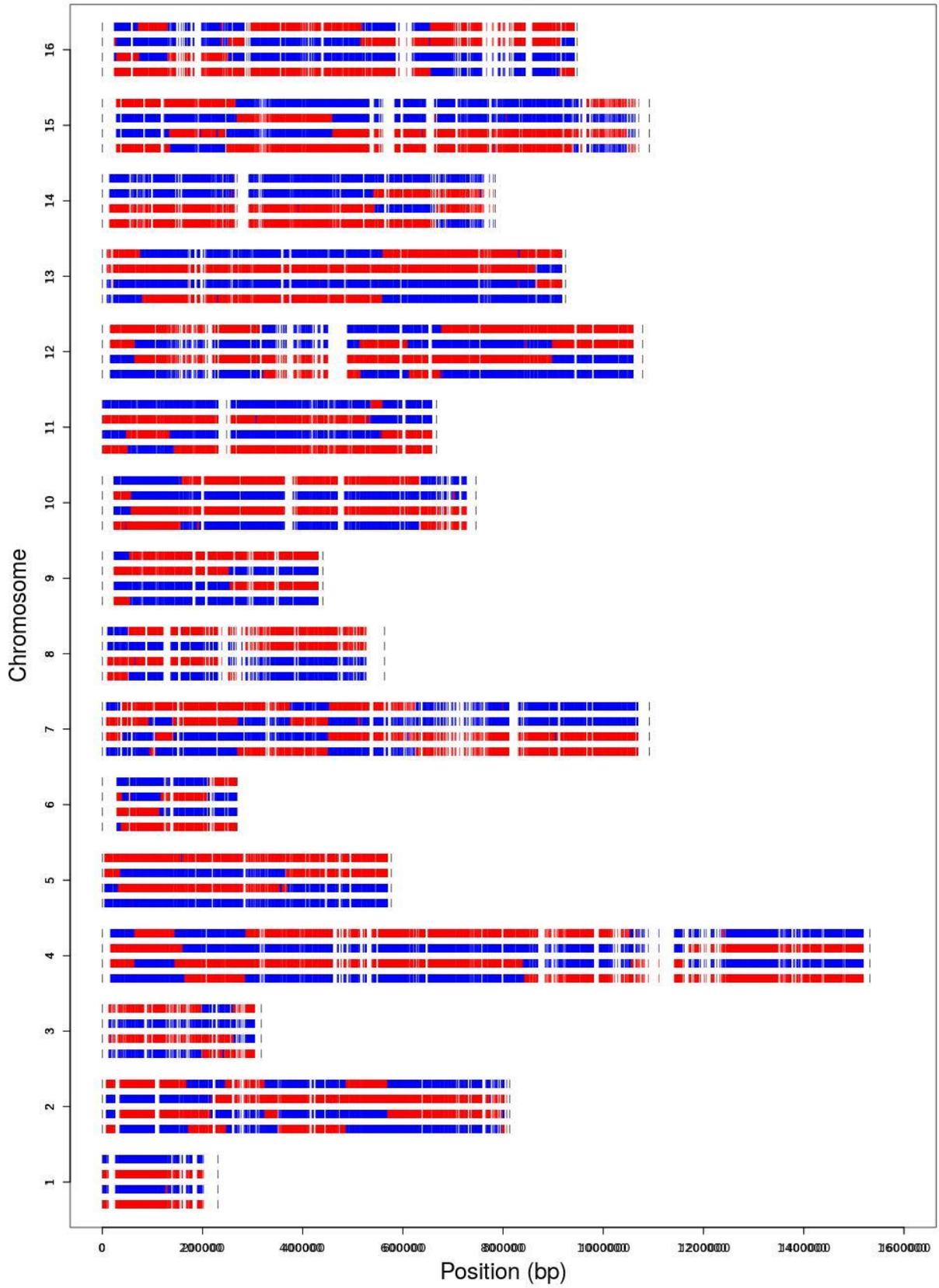
mlh3-DN_2



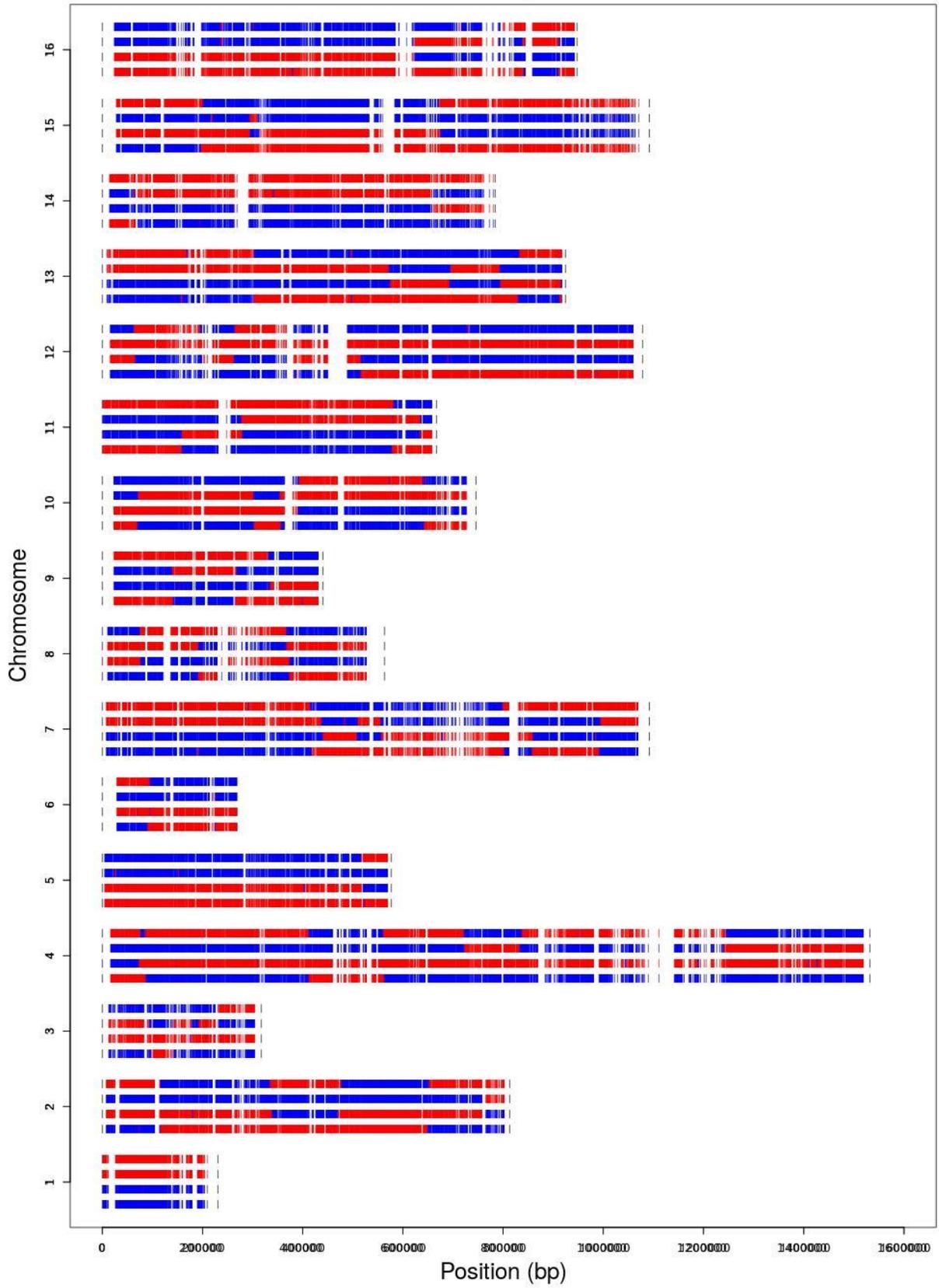
mlh3-DN_3



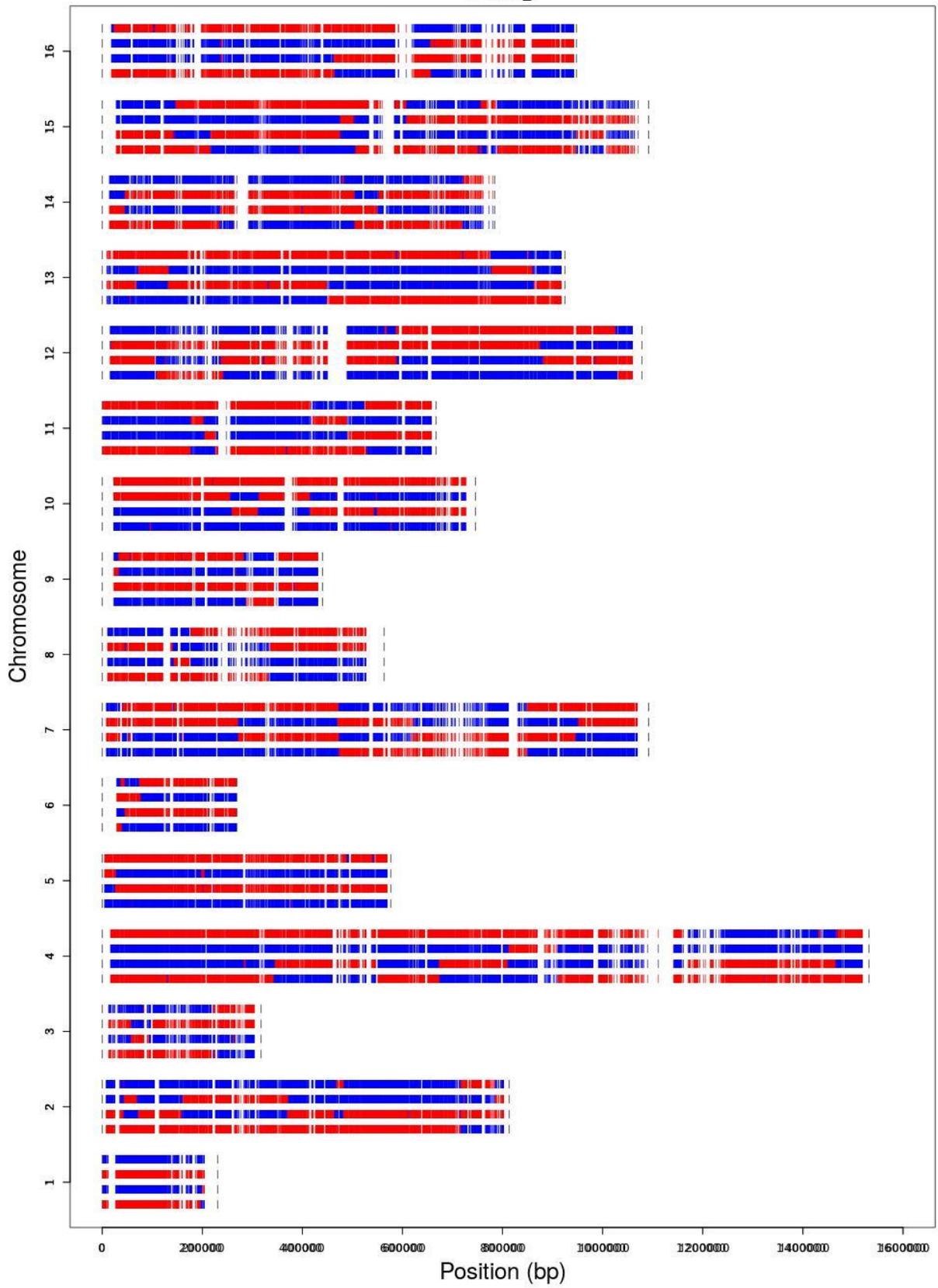
mlh3-DN_4



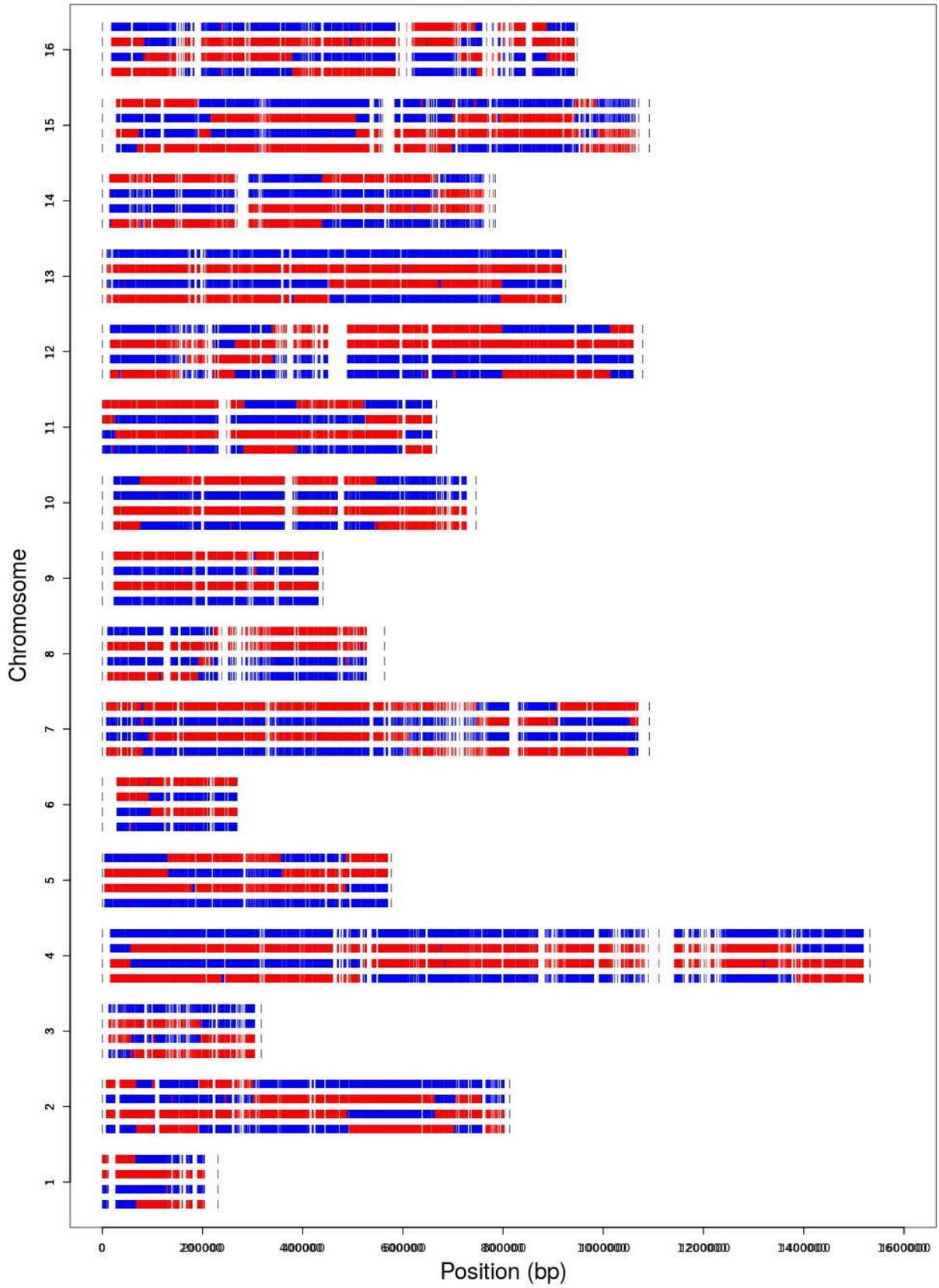
mlh3-DN_5



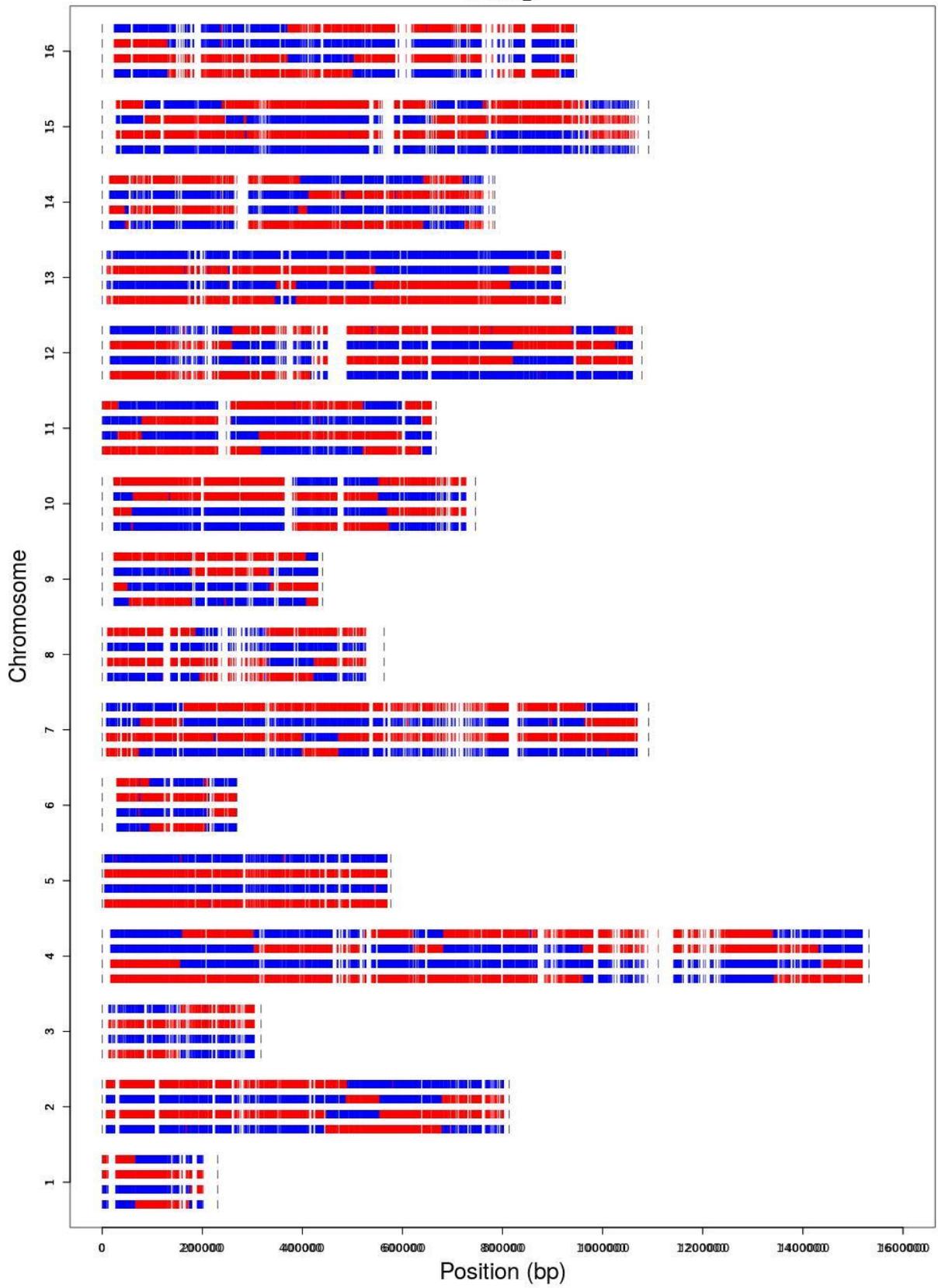
mlh3-DN_6



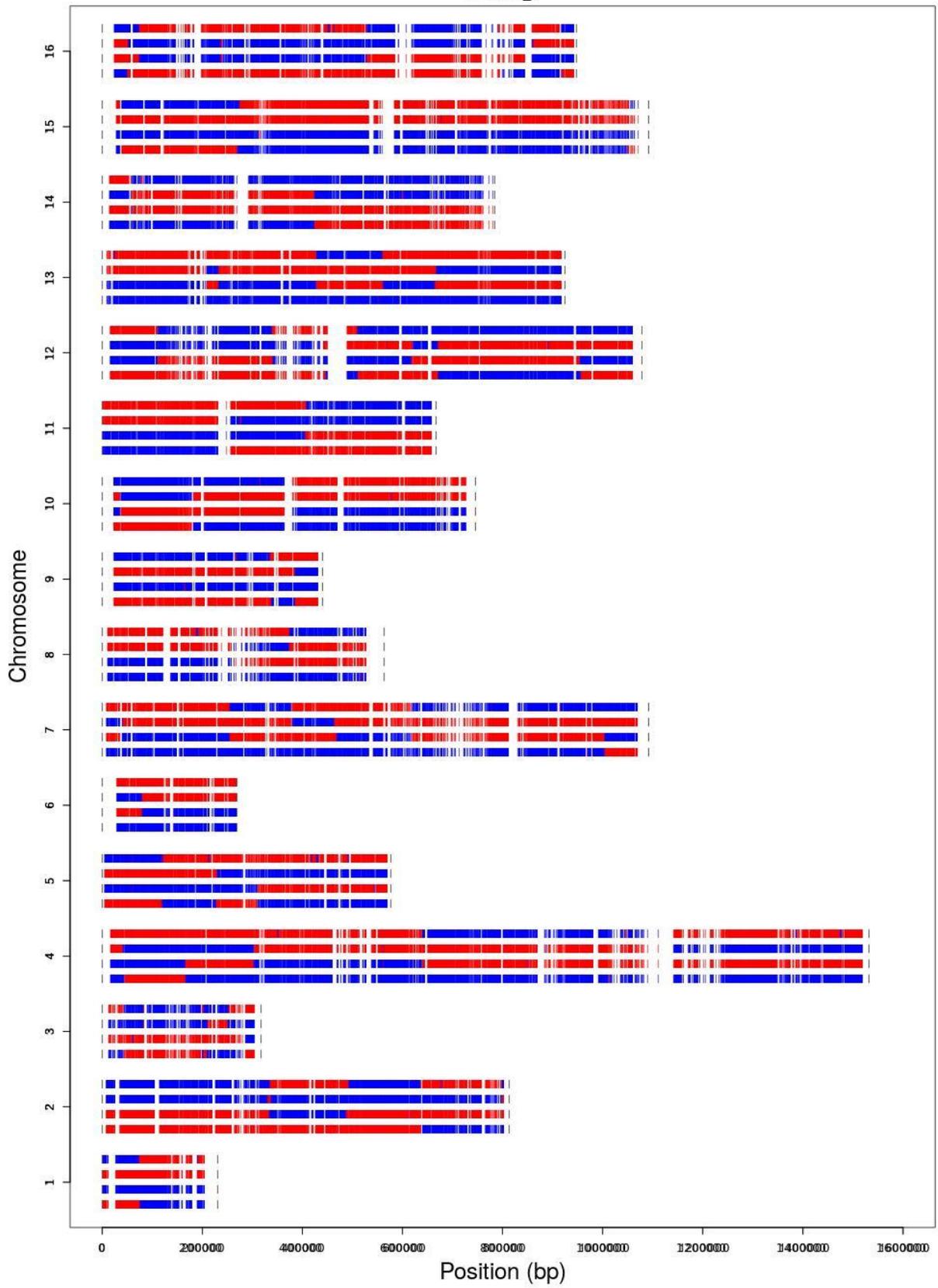
mlh3-DN_7



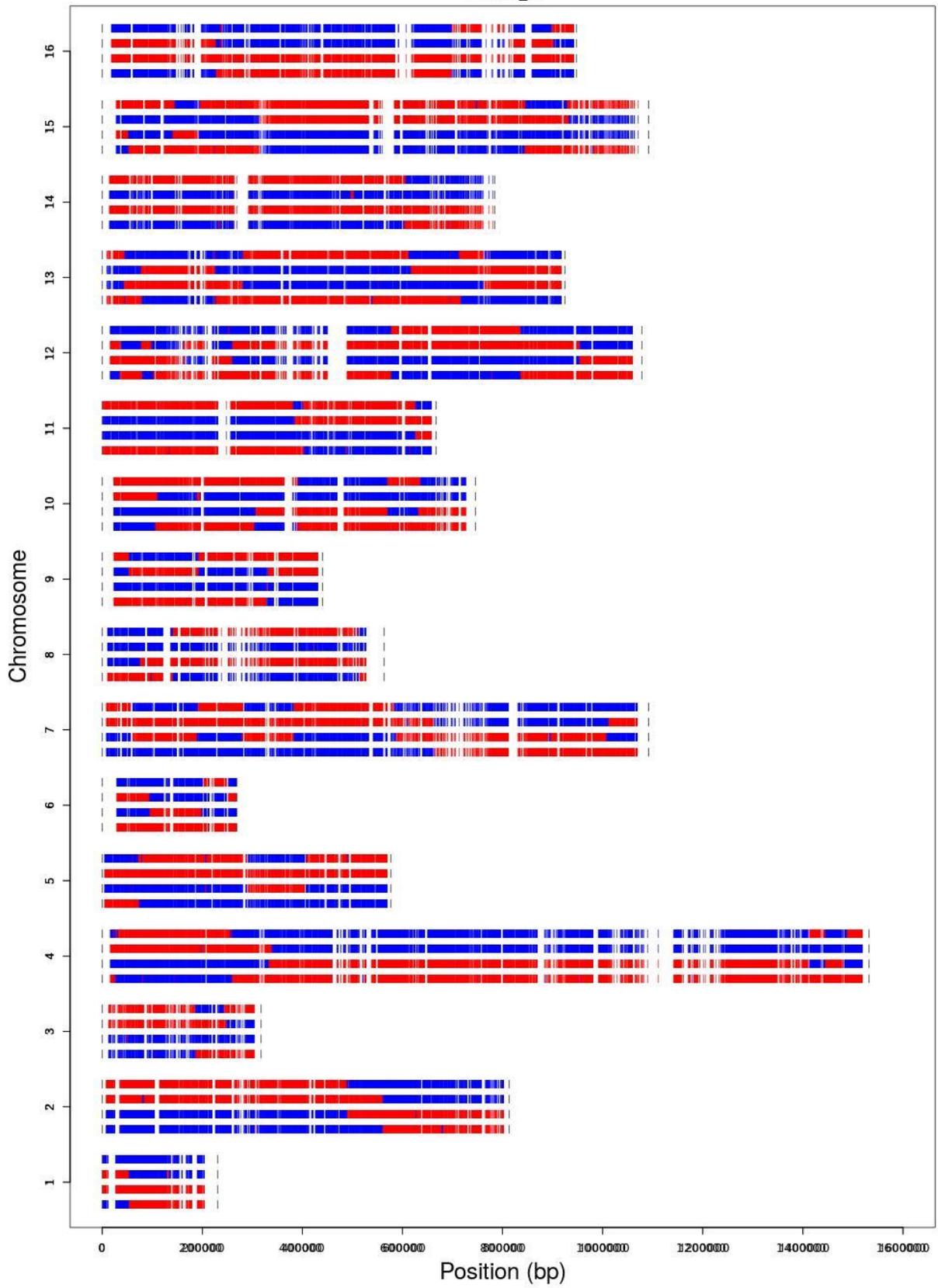
mlh3-DN_8

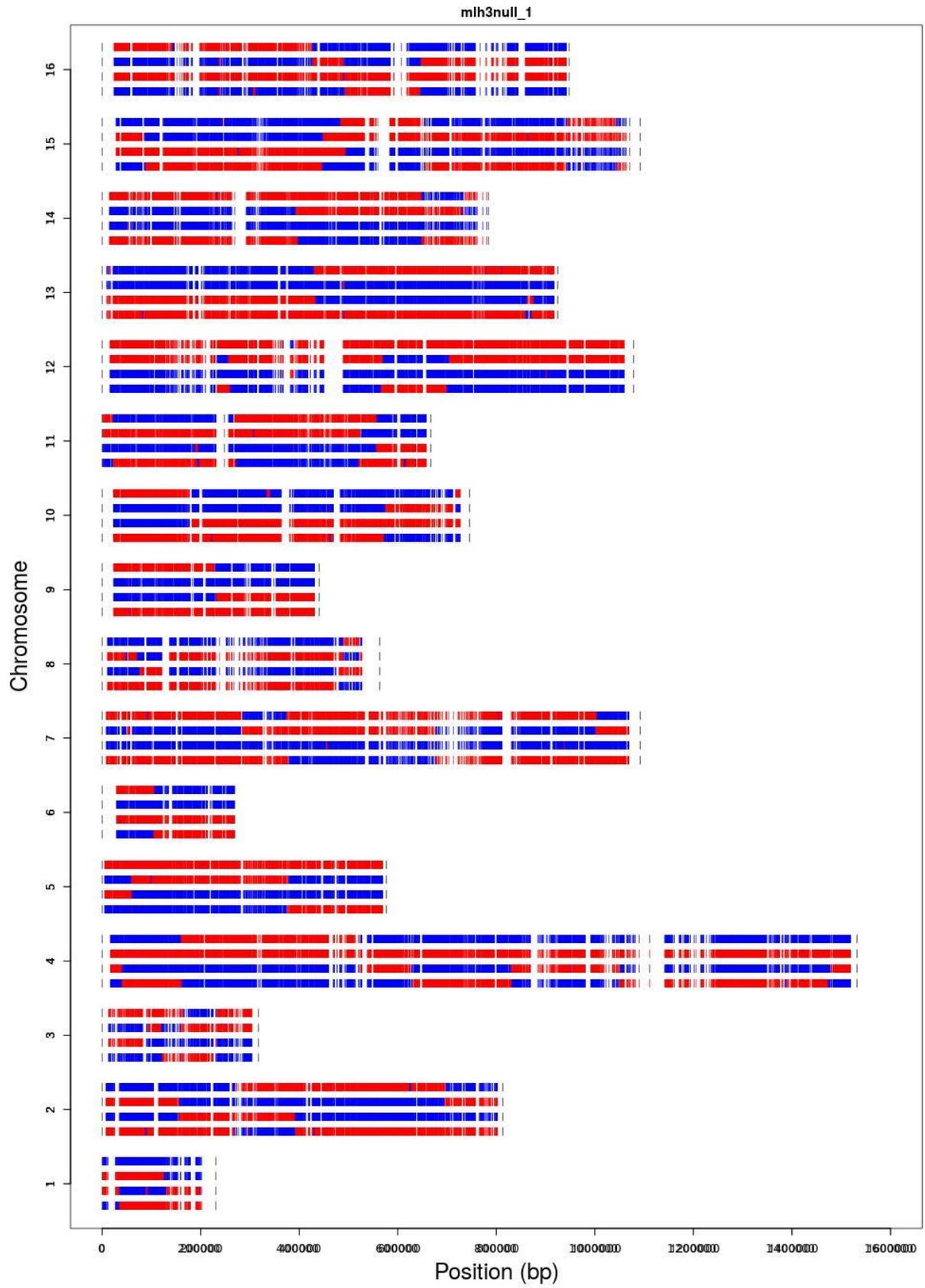


mlh3-DN_9

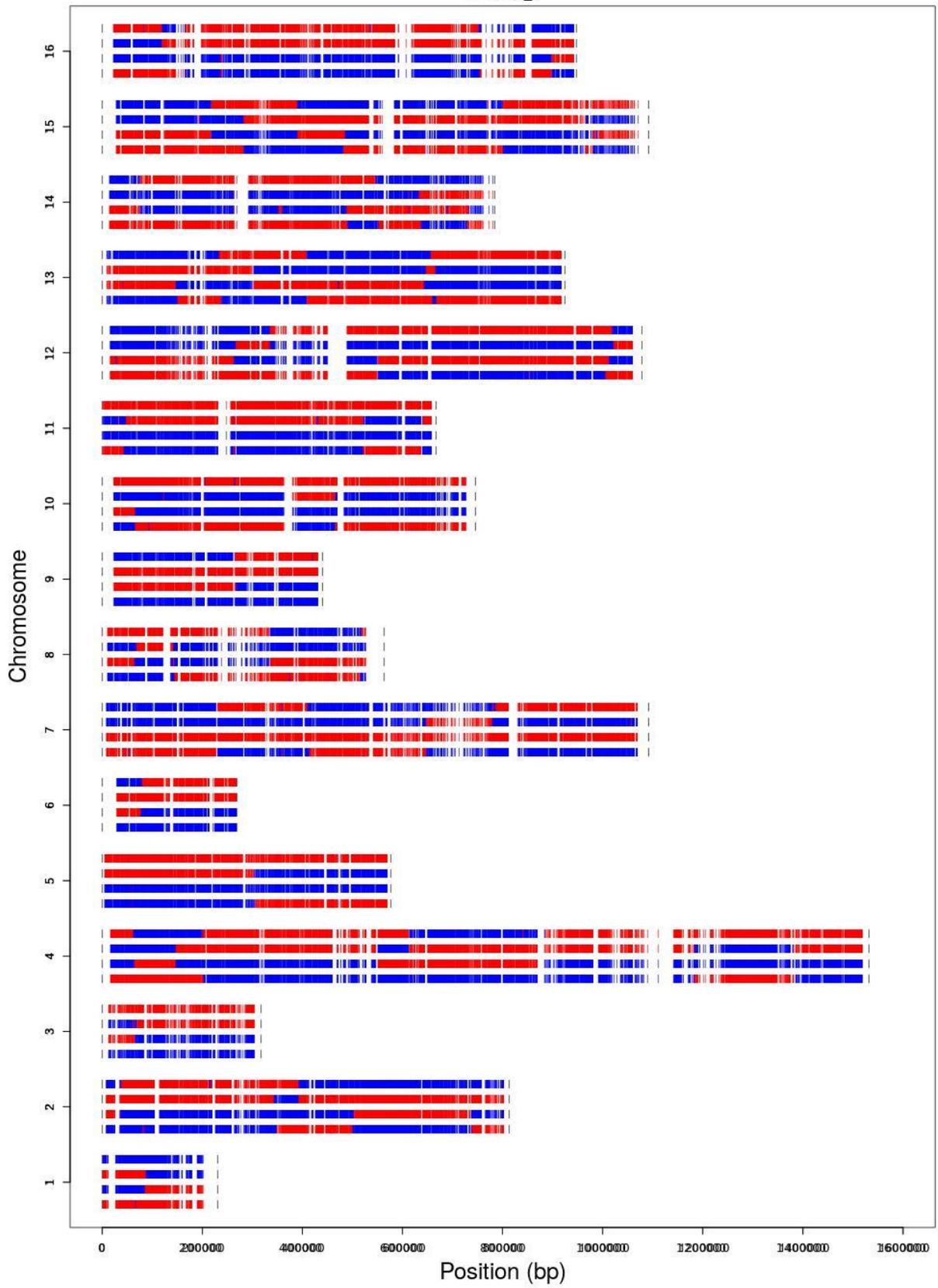


mlh3-DN_10

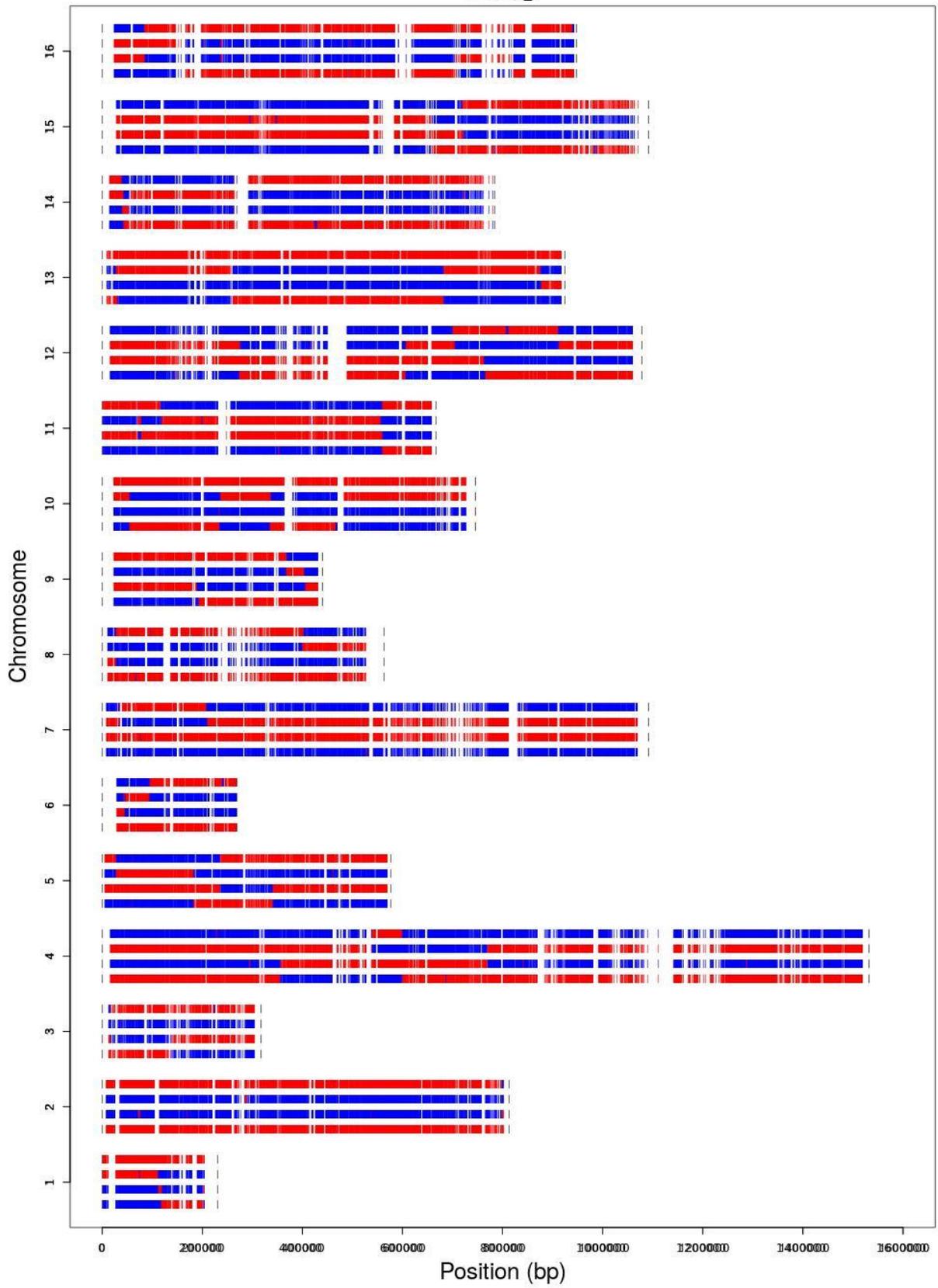




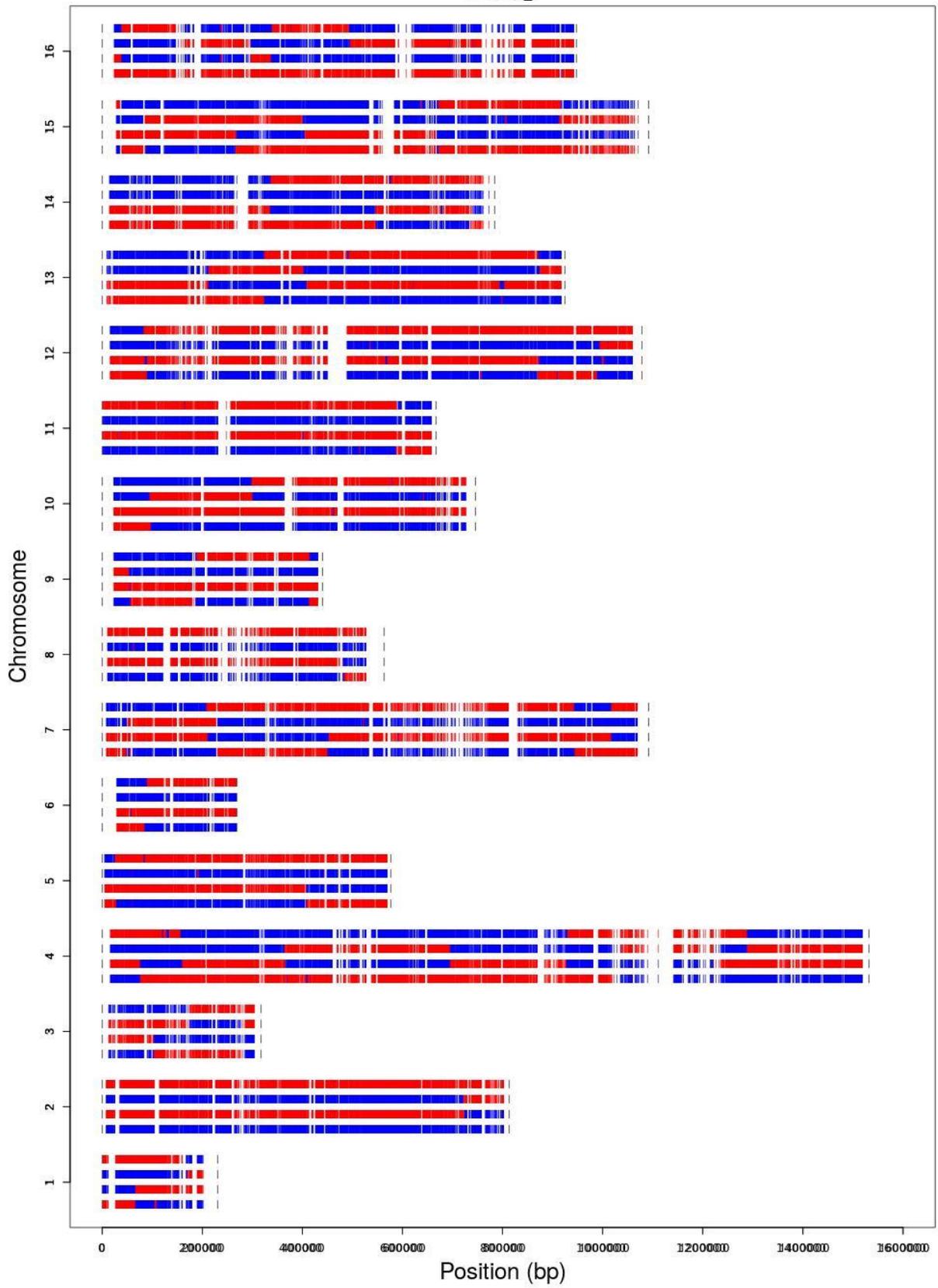
mlh3null_2



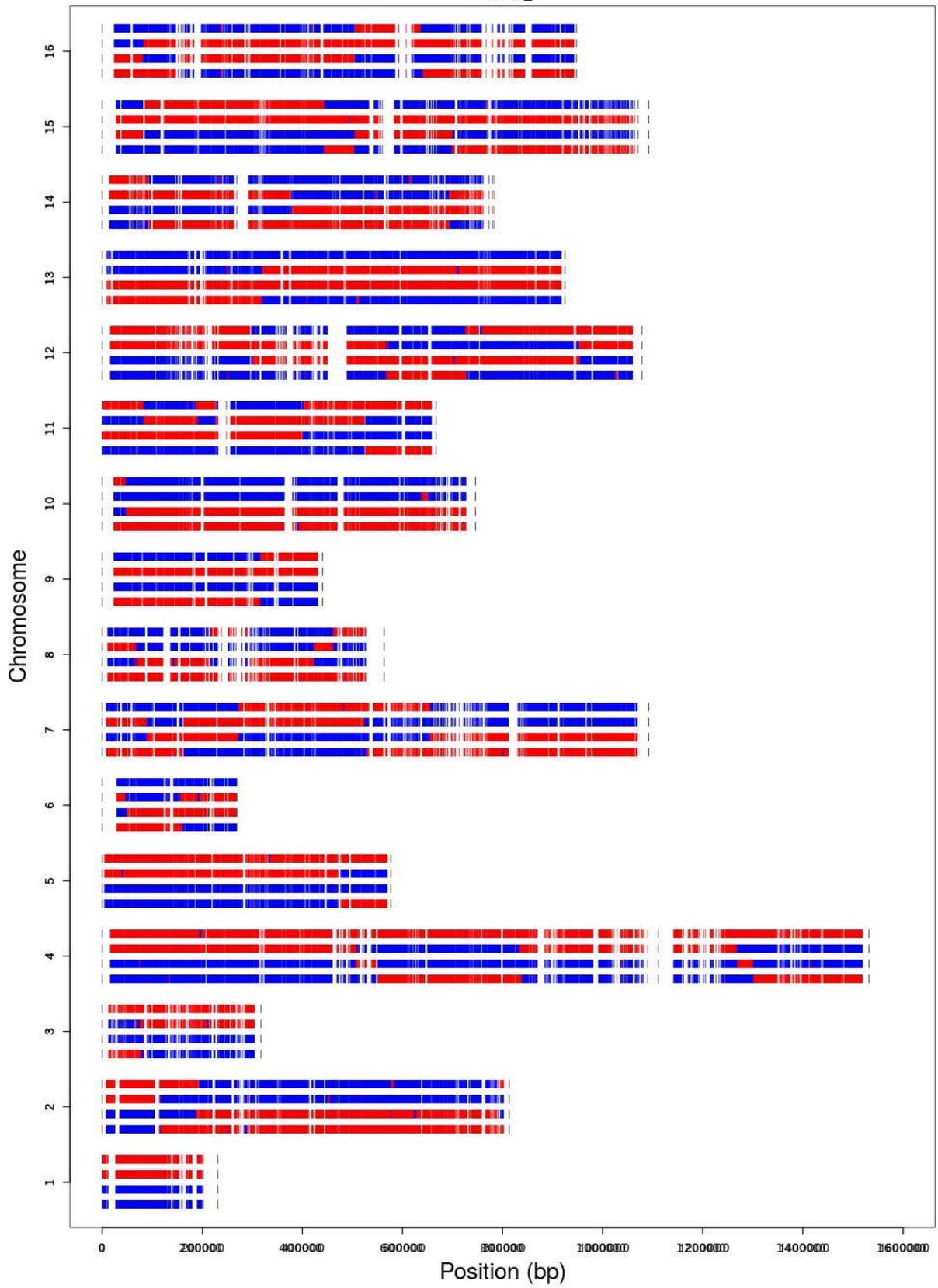
mlh3null_3



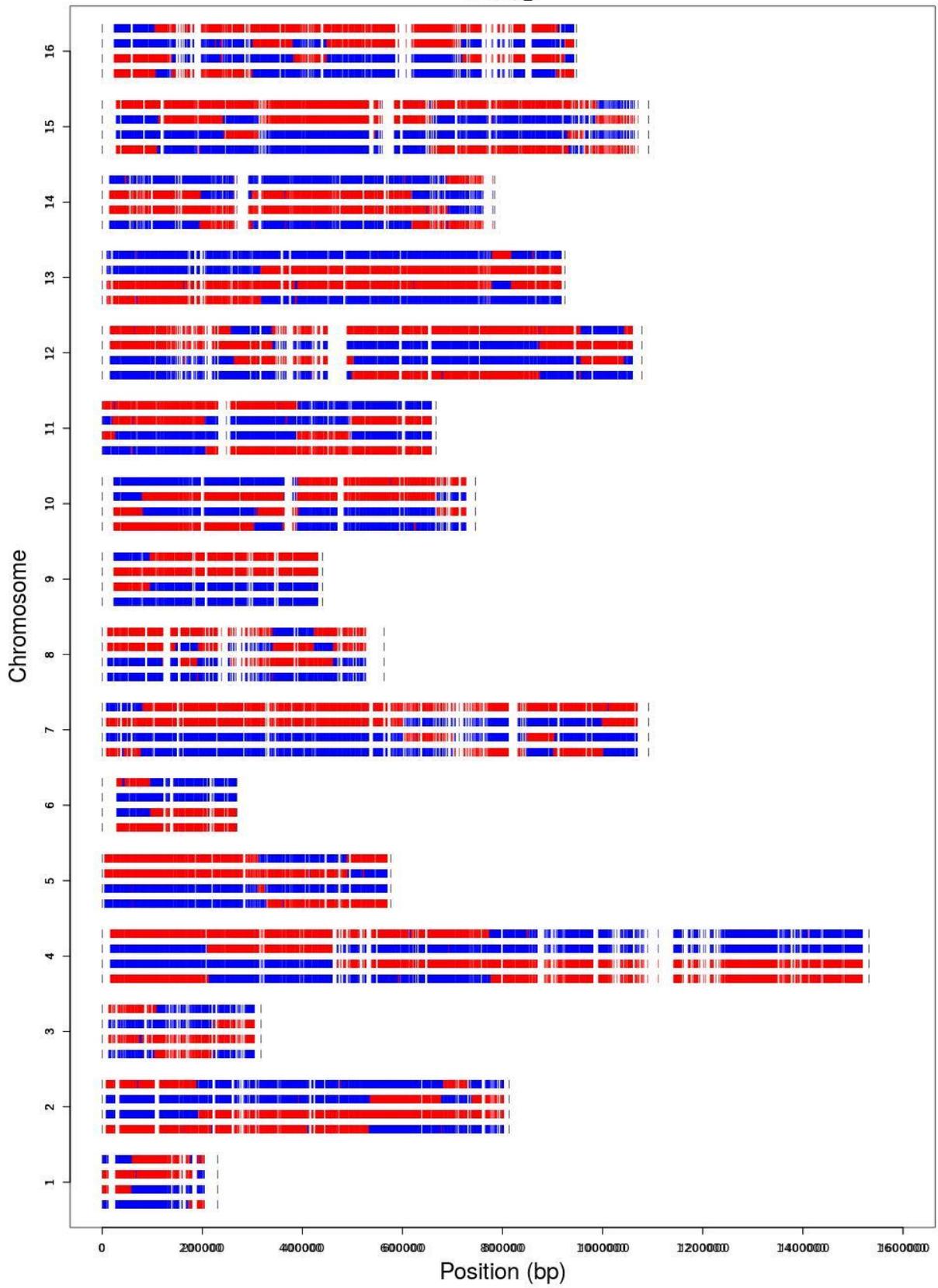
mlh3null_4



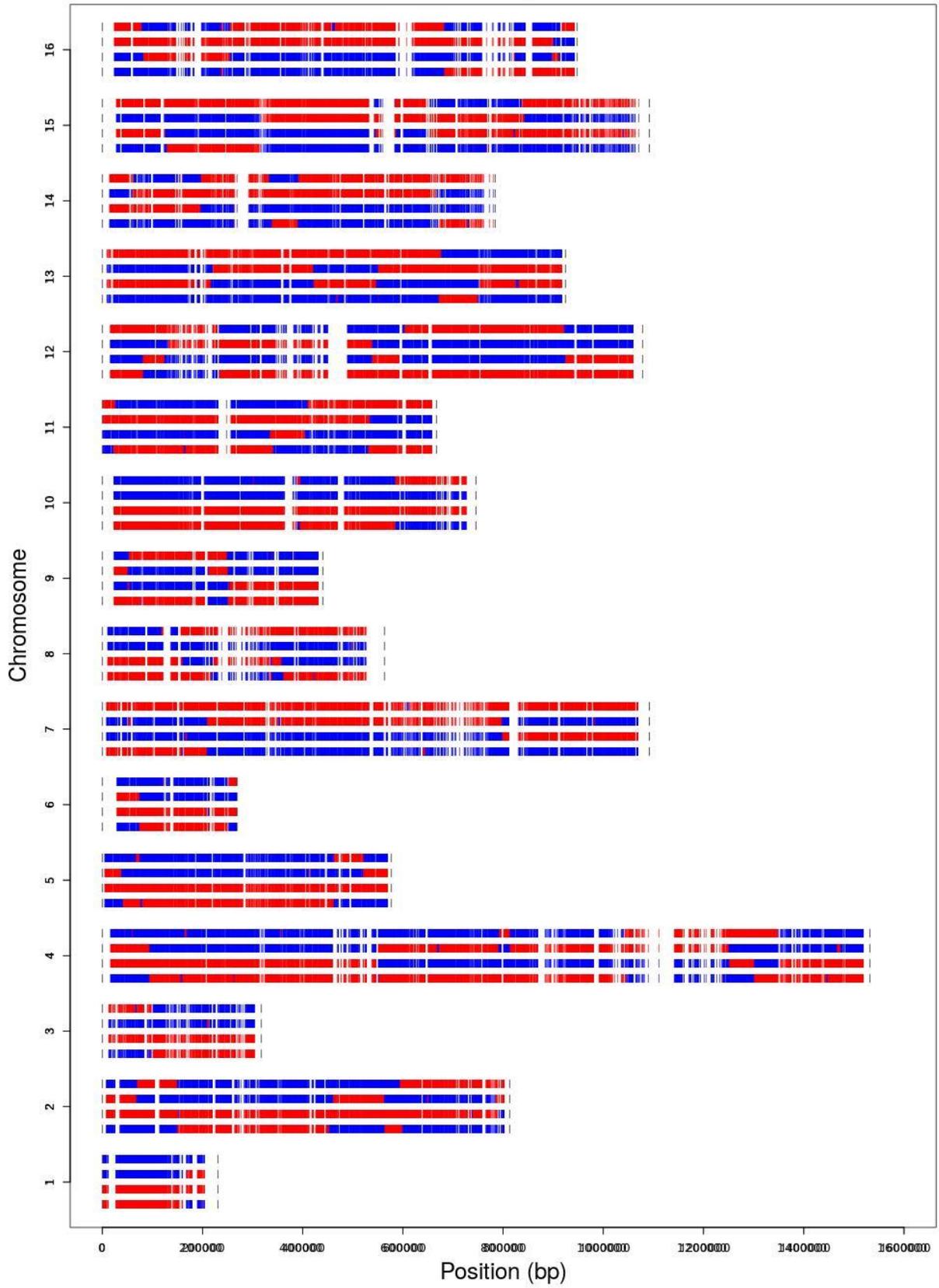
mlh3null_5



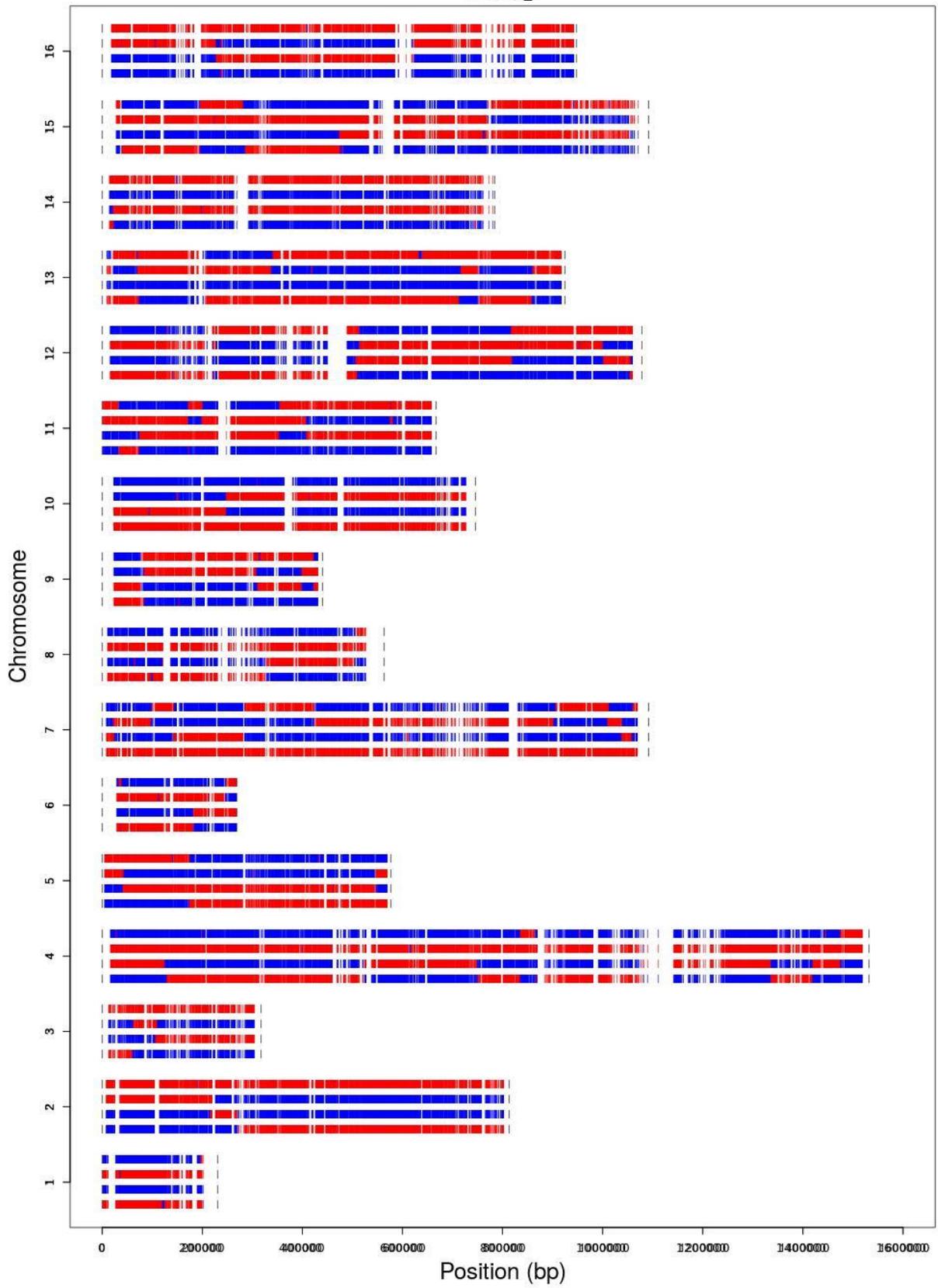
mlh3null_6



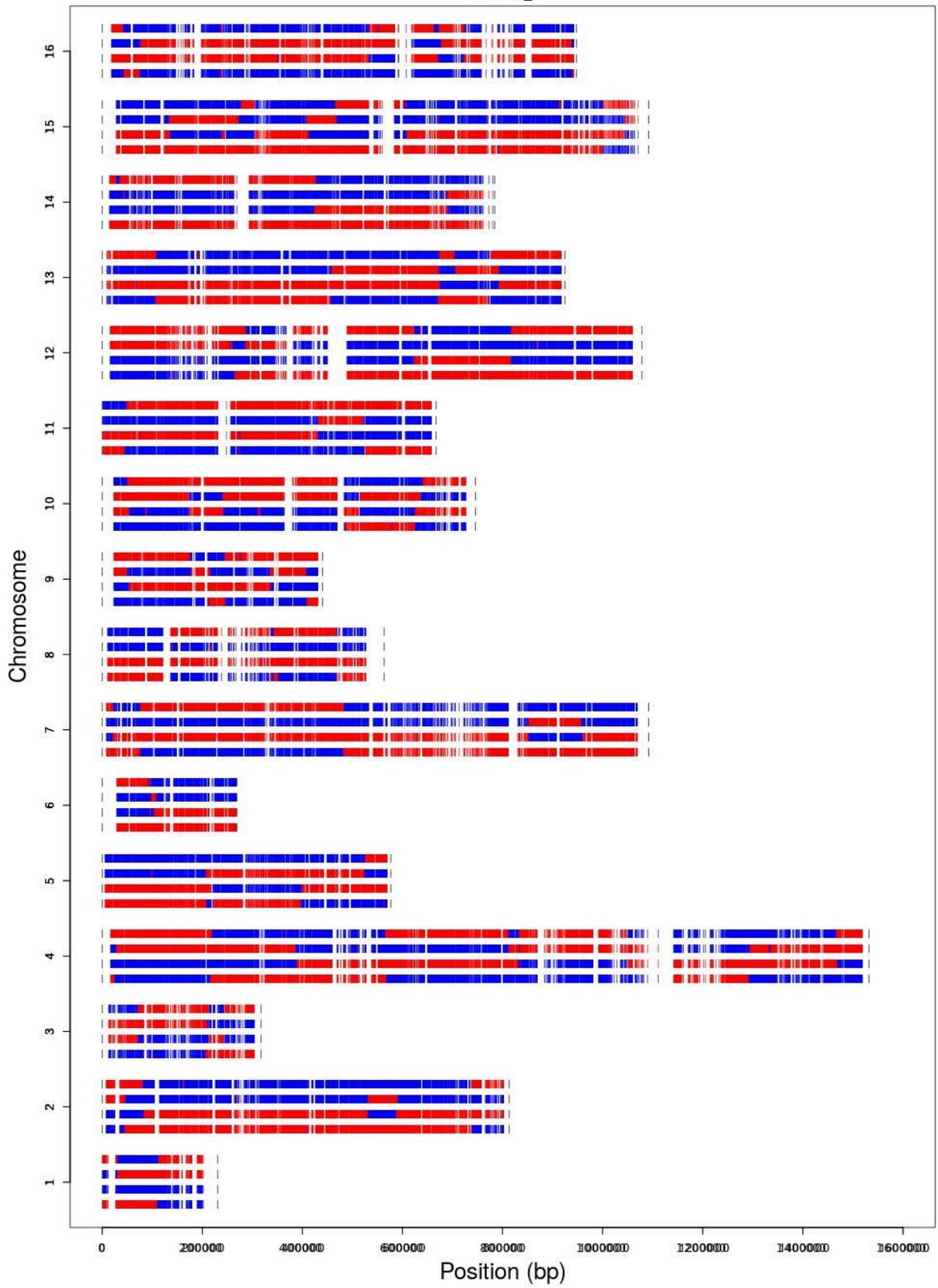
mlh3null_7

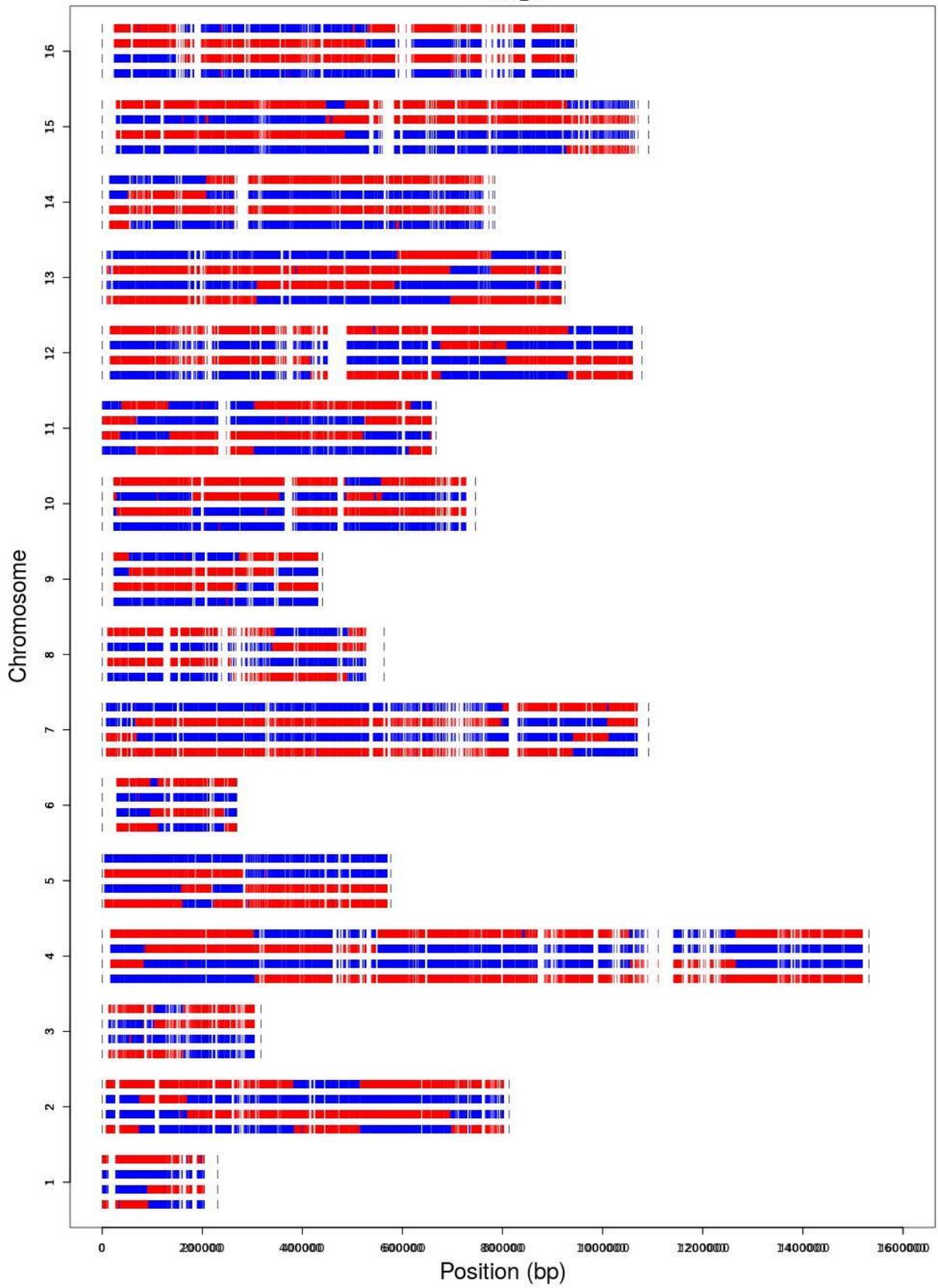


mlh3null_8



mlh3null_9





APPENDIX B

Mismatch repair incompatibilities in diverse yeast populations

Duyen T. Bui*[&], Anne Friedrich,⁺ Najla Al-Sweel*, Gianni Liti[#], Joseph Schacherer,⁺ Charles F. Aquadro,* and Eric Alani*

*Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York, United States of America

⁺Université de Strasbourg, CNRS, GMGM UMR 7156, F-67000, Strasbourg, France

[#]Institute for Research on Cancer and Ageing of Nice (IRCAN), 06107, Nice, France

[&]Present address: Department of Microbiology and Immunology, University of California, San Francisco, CA 94143-2200

This work was published in *Genetics* on 2/13/17:

D. T. Bui, A. Friedrich, N. Al-Sweel, G. Liti, J. Schacherer, C. F. Aquadro, and E. Alani,

“Mismatch Repair Incompatibilities in Diverse Yeast Populations,” *Genetics*, 205(4):1459-1471.

Feb. 2017.

Contributions: N.A.S. performed the work, shown in Figure 6 and Table S7, to determine that the suppression of incompatibility in YJM555 appears to involve more than one locus. All other work was contributed by D.T.B., A.F., G.L., J.S., C.F.A., and E.A.

Abstract

An elevated mutation rate can provide cells with a source of mutations to adapt to changing environments. We identified a negative-epistatic interaction involving naturally occurring variants in the *MLH1* and *PMS1* mismatch repair (MMR) genes of *Saccharomyces cerevisiae*. We hypothesized that this MMR incompatibility, created through mating between divergent *S. cerevisiae* strains, yields mutator progeny that can rapidly but transiently adapt to an environmental stress. Here we analyzed the *MLH1* and *PMS1* genes across 1,010 *S. cerevisiae* natural isolates spanning a wide range of ecological sources (tree exudates, *Drosophila*, fruits, various fermentation and clinical isolates) and geographical sources (Europe, America, Africa, and Asia). We identified one homozygous clinical isolate and eighteen heterozygous isolates containing the incompatible MMR genotype. The *MLH1-PMS1* gene combination isolated from the homozygous clinical isolate conferred a mutator phenotype when expressed in the S288c laboratory background. Using a novel reporter to measure mutation rates, we showed that the overall mutation rate in the homozygous incompatible background was similar to that seen in compatible strains, indicating the presence of suppressor mutations in the clinical isolate that lowered its mutation rate. This observation and the identification of eighteen heterozygous isolates, which can lead to MMR incompatible genotypes in the offspring, are consistent with an elevated mutation rate rapidly but transiently facilitating adaptation. To avoid long-term fitness costs the incompatibility is apparently buffered by mating or by acquiring suppressors. These observations highlight effective strategies in eukaryotes to avoid long-term fitness costs associated with elevated mutation rates.

Introduction

Most spontaneous mutations that occur in natural populations are neutral or deleterious (Kimura 1967; Eyre-Walker and Keightley 2007). However, in changing environments bacteria and yeast can display mutator phenotypes, often through the loss of DNA mismatch repair (MMR) function, that provide a competitive advantage by increasing the chance of obtaining the first adaptive mutation(s) (LeClerc *et al.* 1996; Boe *et al.* 2000; Denamur *et al.* 2000; Taddei *et al.* 1997; Tanaka *et al.* 2003; Townsend *et al.* 2003; Giraud *et al.* 2001; Chao and Cox 1983; Bui *et al.* 2015). Ultimately, such benefits are lost due to long-term fitness costs caused by the accumulation of deleterious mutations. However, in the case of MMR defective bacteria, once adapted to an environment, cells can recover MMR functions through horizontal gene transfer (Denamur *et al.* 2000). In support of this idea, Taddei *et al.* (1997) suggested that mutators are likely to be common in natural asexual populations, but then disappear to avoid long term fitness costs once favorable mutations reach fixation (Taddei *et al.* 1997; Giraud *et al.* 2001; Wielgoss *et al.* 2013).

While it might be common for a eukaryotic organism such as yeast to have lost MMR functions (see Thompson *et al.* 2006; Raynes *et al.* 2012), there is little evidence that these functions can be recovered through horizontal gene transfer (Liti and Louis 2005). One approach that yeast could employ to avoid long-term fitness effects associated with a mutator phenotype is mating to a non-mutator strain, followed by sporulation and segregation of genotypes so that the mutator locus and a beneficial mutation are no longer linked. Such a strategy is considered effective because outcrossing has the potential to generate new genotypes at a much higher frequency than spontaneous mutation (Ruderfer *et al.* 2006). This strategy predicts that a hybrid diploid in which the mutator phenotype is recessive would be observed in

nature.

We developed a model to explain how a genetic incompatibility involving the DNA mismatch repair genes *MLH1* and *PMS1* could arise in yeast (Figure 1; Bui *et al.* 2015; Heck *et al.* 2006; Demogines *et al.* 2008). We define incompatibility as a negative epistatic interaction between *MLH1* and *PMS1* genes that results in an elevated mutation rate. In this model, based on ideas first proposed by Dobzhansky and Muller as a mechanism for speciation (Muller and Pontecorvo 1940; Muller 1939; Dobzhansky 1936; Orr 1995), a common ancestral state gives rise to derived strains (S288c, SK1 in our case) that acquire neutral or beneficial mutations in the *MLH1* and *PMS1* MMR genes. Mating between these divergent populations would create a hybrid genotype. In laboratory experiments MMR incompatible spore progeny of the hybrid combination were shown to confer a higher mutation rate, and provide a fitness advantage to adapt to high salt (Bui *et al.* 2015; Heck *et al.* 2006). These studies suggested that such an advantage would be transient because the mutator genotype would result in a long-term fitness cost (Heck *et al.* 2006). Consistent with a fitness cost, none of 65 wild, lab, and clinical strains/isolates examined displayed the incompatible genotype (Demogines *et al.* 2008). However, the number of strains/isolates examined was too small to determine if their absence in the incompatibility group was lower than expected by chance if there were no epistatic interactions involving *MLH1* and *PMS1* genes. An attractive explanation for this observation is that incompatible isolates exist transiently to provide an adaptive advantage, but mechanisms such as outcrossing restore the compatible genotype. Consistent with this idea is a sequence analysis of a 32 kb genomic region

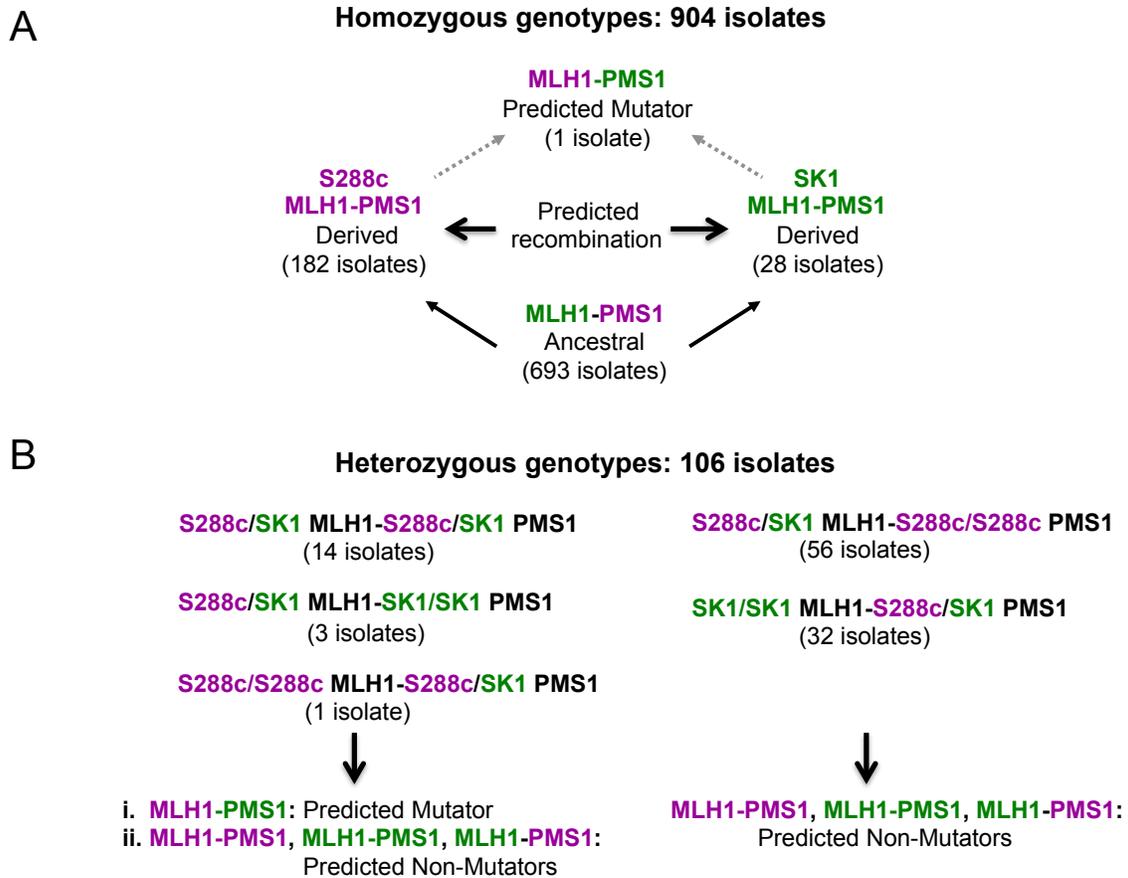


Figure 1. A model proposing that an incompatibility involving the MMR genes *MLH1* and *PMS1* drives adaptive evolution. In this model (Panel A), an ancestral isolate bearing *MLH1* Gly 761 and *PMS1* Arg 818/822 alleles acquire neutral or beneficial mutations that lead to the derived S288c (purple, Asp 761, Arg 818/822) and SK1 (green, Gly 761, Lys 818/822) group isolates. Mating between the derived isolates can yield an allele combination (*MLH1* Asp 761, *PMS1* Lys 818/822) that displays negative epistasis and thus a mutator phenotype. Previous work suggested that recombination has occurred between the two derived classes, leading to exchange of genetic information and a hybrid mutator genotype that can also recombine to reconstruct derived or ancestral genotypes (Heck *et al.* 2006). As described in the text, sequences of *MLH1* and *PMS1* genes from 1,010 *S. cerevisiae* from worldwide collections genomes were grouped according to their amino acid residues 761 (G or D) in *MLH1* and 818 (R or K) in *PMS1* (File S1; File S2). One isolate maps to the predicted hybrid mutator category. Of the 1,010 isolates, 904 are homozygous for the *MLH1* 761 and *PMS1* 818/822 genotypes. The distribution in each genotype is shown here and a list of isolates in the S288c and SK1 derived classes is shown in Table S2. Panel B. 106 isolates are heterozygous for the *MLH1* 761 and/or *PMS1* 818/822 genotypes, with the distribution shown. On the left are the genotypes which can lead to both mutator and non-mutator progeny (see Table 1 and text). On the right are genotypes that can lead only to non-mutator progeny.

Table 1. Nineteen natural isolates that contain incompatible *MLH1-PMS1* combinations.

Isolate name	Genotype		Ploidy (n)	Origin (ecological, geographical)
	<i>MLH1</i>	<i>PMS1</i>		
CLIB324_2	SK1/S288c	SK1/S288c	2	bakery, Vietnam
CBS4455	S288c	SK1/S288c	3	beer, South Africa
CECT1462	SK1/S288c	SK1/S288c	4	beer, United Kingdom
NCYC_2780	SK1/S288c	SK1/S288c	4	human clinical, Belgium
2680	SK1/S288c	SK1/S288c	4	human clinical, Spain
YS8(E)	SK1/S288c	SK1/S288c	4	bakery, Unknown
SD-15	SK1/S288c	SK1/S288c	4	bakery, Italy
WLP001	SK1/S288c	SK1/S288c	4	beer, USA
WLP013	SK1/S288c	SK1	4	beer, United Kingdom
WLP006	SK1/S288c	SK1	4	beer, Unknown
UCD_06-645	SK1/S288c	SK1/S288c	4	fruit, Davis, California
Win-8B	SK1/S288c	SK1	2	beer, United Kingdom
Ponton 11	SK1/S288c	SK1/S288c	2	human clinical, Spain
CLI_16	SK1/S288c	SK1/S288c	4	human clinical, France
CLI_19	SK1/S288c	SK1/S288c	2	human clinical, France
CLI_23	SK1/S288c	SK1/S288c	4	human clinical, France
CLI_26	SK1/S288c	SK1/S288c	4	human clinical, France
YJM521	SK1/S288c	SK1/S288c	nt	human clinical, California
YJM523	S288c	SK1	nt	human clinical, California

1,010 isolates were genotyped at amino acids 761 for *MLH1* and 818 in *PMS1*. Isolates were genotyped as S288c if they contained the S288c amino acids (D761 in *MLH1*, R818 in *PMS1*), and SK1 if they contained the SK1 amino acids (G761 in *MLH1*, K822 in *PMS1*). SK1/S288c indicates that both SK1 and S288c alleles were detected in the strain at the indicated locus. We analyzed the yeast genomes of the 1,010 isolates by placing them into clades. We then computed the mean, median, standard deviation and standard error for the number of singletons within these clades and compared the number of singletons of the 19 isolates to their associated clades. We did not observe an excess of singletons for any of the isolates (1002 Yeast Genomes Project - <http://1002genomes.u-strasbg.fr/>). Ploidy was determined by FACS, and all but CLIB324_2, WLP001, and Win-8B are capable of sporulating (Materials and Methods). nt = not tested.

of yeast (Chr. XIV, 445 to 477 kb) which provided evidence for recombination between the S288c and SK1 groups Heck *et al.* 2006; Demogines *et al.* 2008; Steinmetz *et al.* 2002).

A challenge to the above model is that several studies have estimated that the number of clonal generations that two yeast strains would have experienced prior to outcrossing is very high, one outcrossing event per 12,500 to 62,500 generations (Ruderfer *et al.* 2006; Magwene *et al.* 2011). While such observations suggest that yeast has primarily an asexual lifestyle, random mating between natural strains can be achieved at high rates in the lab and can be elevated in different environments, such as in the gut of the fruit fly (Murphy and Zeyl 2010; Reuter *et al.* 2007). It is also worth noting that population structure analyses of domestic and wild yeasts showed evidence for many different natural recombinant/mosaic genomes (Liti *et al.* 2009; Strobe *et al.* 2015), and outcrossing rate is likely to be different depending on the ecological origin of the isolates and their stress conditions; e.g. a high outcrossing rate was observed for domestic isolates (reviewed in Marsit and Dequin 2015). Thus isolates heterozygous for MMR genotypes could in stress conditions undergo sporulation and yield incompatible haploid progeny.

Previously, we showed that incompatible strains display a transient fitness advantage Bui *et al.* 2015). In the present study, we looked to see if evolved incompatible strains maintain fitness in non-stress and stress conditions. Our work showed that such strains rapidly displayed fitness defects in non-stress conditions. This observation encouraged us to survey 1,010 *S. cerevisiae* wild, clinical, industrial and lab strains from worldwide collections for their *MLH1* and *PMS1* genotypes. We identified and carefully characterized one clinical isolate (YJM523) that displayed the homozygous incompatible *MLH1-PMS1* genotype. Long-term fitness costs appeared to have been avoided in this strain through extragenic mutations that suppressed its

mutation rate. Based on the above we hypothesize that in natural populations a mutator state can be created to transiently promote adaptive evolution, but can then be suppressed once beneficial mutations go to fixation.

Materials and Methods

Strains and media

Saccharomyces cerevisiae spore clones from the S288c strain background were analyzed in competition assays (Table S1; Figure 2; Winston *et al.* 1995). These strains were grown in YPD (yeast extract, peptone, dextrose) and YPD + 1.2 M NaCl (Rose *et al.* 1990). Natural and S288c derived strains transformed with pEAA613 (*ARS-CEN NatMX*, *kanMX::insE-A14*; Table S1; Table S2) were maintained in YPD media containing 50 µg/ml nourseothricin. Yeast transformations were performed as described (Gietz and Schiestl 2007).

EAY1365 (*MATa*, *ura3-52*, *leu2Δ1*, *trp1Δ63*, *his3Δ200*, *lys2::insE-A14*, *mlh1Δ::KanMX4*, *pms1Δ::KanMX4*), an S288c derived strain, was used to measure reversion to Lys⁺. This strain was transformed with *ARS-CEN*, *LEU2*, *cMLH1* (pEAA213 and derivatives) and *ARS-CEN*, *HIS3*, *cPMS1* (pEAA238 and derivatives) plasmids, and were grown in minimal selective media lacking histidine and leucine (Table S1; Rose *et al.* 1990). In this paper 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*). YJM555 is a homozygous diploid strain derived from the natural isolate YJM523 (see details in (Strope *et al.* 2015). YJM523 and YJM555 were obtained from the FGSC collection (<http://www.fgsc.net/>).

Plasmids

pEAA213 (c*MLH1*, *ARSH4 CEN6*, *LEU2*) and pEAA214 (k*MLH1*, *ARSH4 CEN6*, *LEU2*) were described previously (Argueso *et al.* 2003; Heck *et al.* 2006; Table S3). *MLH1* expression is driven in both plasmids by the S288c *MLH1* promoter. *MLH1* from YJM555 was cloned into pEAA213 by amplifying *MLH1* from genomic DNA (Rose *et al.* 1990; Hoffman and Winston 1987) using Roche high fidelity polymerase (Roche) and primers AO324 and AO821 (Table S4). The PCR amplified product containing the entire *MLH1* open reading frame was digested with *Bam*H1-and *Nhe*I and inserted into the corresponding sites of pEAA213. The entire PCR fragment was DNA sequenced. All of the resulting constructs expressed *MLH1* via the S288c *MLH1* promoter.

pEAA238 (c*PMS1*, *ARSH4*, *CEN6*, *HIS3*) and pEAA239 (k*PMS1*, *ARSH4*, *CEN6*, *HIS3*) were described previously (Table S3; Argueso *et al.* 2003; Heck *et al.* 2006). In both plasmids *PMS1* expression is driven by the S288c *PMS1* promoter. *PMS1* from YJM555 was cloned into pEAA238 by amplifying *PMS1* from genomic DNA using Roche high fidelity polymerase and primers AO548 and AO481 (Table S4). The PCR amplified product containing the entire *PMS1* open reading frame was digested with *Aat*II-and *Mlu*I and inserted into the corresponding sites of pEAA238. The entire PCR fragment was DNA sequenced. All of the resulting constructs expressed *PMS1* via the S288c *PMS1* promoter.

pEAA613 contains the *URA3 promoter-kanMX::insE-A₁₄* reversion reporter constructed using overlap PCR (Ho *et al.* 1989). Briefly, 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*; Tran *et al.* 1997) was inserted immediately after the *URA3* ATG, followed by codons 18 to 269 of the *KanMX* open reading frame derived from pFA6-

KanMX, and 159 bp of *KanMX* downstream sequence that contains a transcription termination sequence (also derived from pFA6-KanMX; Goldstein and McCusker 1999). This reporter construct was inserted into pLZ259 (*ARS-CEN*, *NatMX*, kindly provided by Dr. Lu Zhu). Derivatives of pEAA613 that contain homopolymeric tracts of 10 A (in frame, pEAA611) and 11 A residues (+1 out of frame, pEAA612) residues were also constructed. Finally, a set of reporter constructs (pEAA614-616) were built in which the *kanMX::insE-A₁₄* reporter was expressed via the *LEU2* promoter (-308 to the ATG start site of the *LEU2* promoter fused to the *KanMX* ORF).

Spore clone competitions

Clones derived from incompatible strains (*cMLH1::KanMX*, *kPMS1::HIS3*, *pmr1*) evolved in high salt for 70 generations (EAY3688, EAY3689, EAY3690 containing *pmr1* mutations *T459A*, *T412C*, *T2G*, respectively; Bui *et al.* 2015) were mated with the unevolved and compatible strain EAY3241 (*kPMS1::HIS3*, *kMLH1::NatMX*, *PMR1*; Table S1). Diploids were selected on YPD plates containing clonNAT (100 µg/ml) and genecitin (200 µg/ml) and then sporulated on 2% agar media containing 1% potassium acetate. The resulting spore clones were genotyped for *PMR1* by isolating chromosomal DNA (Holm *et al.* 1986) from individual clones and sequencing the PCR amplified *PMR1* locus. These clones were genotyped for *MLH1* by testing for resistance to genecitin and clonNAT, and for the *MAT* locus by mating to tester strains.

Spore clones of the same mating type and of the four possible genotypes (*kMLH1-kPMS1*(compatible)-*PMR1*, *kMLH1-kPMS1*(compatible)-*pmr1*, *cMLH1-kPMS1*(incompatible)-*PMR1*, and *cMLH1-kPMS1*(incompatible)-*pmr1*) were grown overnight in YPD and then mixed in equal proportions (Transfer 0). 4, 3, and 3 spore clones of each genotype were competed from

the matings involving *pmr1-T459A*, *pmr1-T412C*, *pmr1-T2G*, respectively. Ten competitions in total were performed, and different spore clones obtained from different tetrads were pooled to minimize effects of background mutations that might arise in evolved populations. 2×10^7 cells of the initial mixed culture (Transfer 0) were transferred into 6 ml of YPD or YPD + 1.2 M NaCl and then grown for 24 hrs at 30°C (~7 generations of growth). The same amount of cells (to achieve an initial OD₆₀₀ of 0.1, Shimadzu UV-1201 spectrophotometer) was used in subsequent transfers, with cells grown under the same conditions. At transfer numbers indicated in Figure 3, cells in the culture were genotyped for compatibility (*kPMS1::HIS3*, *kMLH1::NatMX*) and incompatibility (*kPMS1::HIS3*, *cMLH1::KanMX*) on YPD plates containing geneticin (200 µg/ml) or clonNAT (100 µg/ml).

DNA sequencing analysis

1,010 *S. cerevisiae* isolates were investigated in the context of the 1002 yeast genome project (File S1; File S2). Illumina reads were mapped against the *S. cerevisiae* 288c reference genome R64-1-1 with bwa 0.7.4-r385 and the sequence of the *MLH1* and *PMS1* genes were inferred for all isolates with GATK (FastaAlternateReferenceMaker).

The 2.3 kb *MLH1* and 2.6 kb *PMS1* open reading frames from YJM521 and YJM523 (Table 1, Tables S1, S2) were determined by sequencing PCR amplified DNA from chromosomal DNA (Hoffman *et al.* 1987) using Expand High Fidelity Polymerase (Roche Life Sciences). Primers AO324 and AO821 were used to amplify the *MLH1* open reading frame and primers AO481 and AO548 were used to amplify the *PMS1* open reading frame (Table S4). PCR products were sequenced by the Sanger method in the Cornell BioResource Center. Duplicate clones from the strain stocks were sequenced for confirmation purposes. To further

show that YJM523 was not heterozygous for sequences at the Mlh1 (amino acid position 761) and Pms1 (amino acid position 818) incompatibility loci, 20 different PCR products, derived from unique PCR primer combinations, were sequenced using primer AO328 for *MLH1* and primer AO485 for *PMS1* (Table S4). In no case was heterozygosity detected at the *MLH1* or *PMS1* incompatibility loci.

Flow cytometry

Cells from isolates shown in Table 1 were fixed overnight in cold ethanol (70% final), washed, and re-suspended at 1.2×10^7 cells/ml in 50 mM sodium citrate (pH 7). They were then treated for 1 hr with RNase A at a final concentration of 1 mg/ml, after which they were stained with propidium iodide at a final concentration of 40 μ g/ml. Flow cytometry was performed on a Cyflow Space, Partec.

Sporulation test

Isolates shown in Table 1 were sporulated for 2–3 days on sporulation medium (10 g/L potassium acetate, 20 g/L agar) at 30°C.

***lys2-A₁₄* reversion assays**

Independent colonies of EAY1365 (relevant genotype *lys2-A₁₄*) containing the *ARS-CEN*, *MLH1* and *ARS-CEN*, *PMS1* plasmids presented in Table S3 were inoculated YPD overnight 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 (Bui *et al.* 2015; Tran *et al.* 1997). The 95% confidence intervals were determined as described by Dixon and Massey (1969). The

Mann-Whitney U test (Mann and Whitney 1947) was performed to determine the significance of the differences in median reversion rates.

***kanMX::insE-A₁₁₋₁₄* reversion assays**

EAY1369 (*cMLH1-cPMS1*, compatible), EAY1370 (*cMLH1-kPMS1*, incompatible), EAY1372 (*msh2Δ*), and YJM555 were transformed with pEAA613 or pEAA616 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), genecitin (300 μg/ml). These strains were analyzed for reversion to resistance to genecitin using methods described previously (Bui *et al.* 2015; Tran *et al.* 1997). The 95% confidence intervals were determined as described by Dixon and Massey (1969). Pair-wise Kruskal–Wallis tests were performed to determine the significance of the differences in median reversion rates. Mann-Whitney U tests were performed to determine the significance of the differences in median reversion rates (Mann and Whitney 1947; Wilcoxon 1945).

As controls, we measured resistance to genecitin (G418) for EAY1369 and YJM555 lacking the *kanMX::insE-A₁₁₋₁₄* reporter plasmid. Ten independent overnight cultures were analyzed for each strain. We did not observe resistant colonies in any of the cultures and thus estimate median frequencies of resistance to be $< 9.7 \times 10^{-10}$ for EAY1369, and $< 6.9 \times 10^{-10}$ for YJM555; these values correspond to mutation rates that are $< 2 \times 10^{-10}$, indicating that spontaneous reversion to G418^r would not interfere with the detection of G418^r using the *kanMX::insE-A₁₁₋₁₄* reporter plasmids.

YJM x EAY3235 cross

To determine if a single locus is present in YJM555 that suppressed *MLH1-PMS1* incompatibility, we mated meiotic spores of YJM555 to the incompatible strain EAY3235 (*MATa*, *lys2-A₁₄*, *ura3*, *trp1*, *leu2*, *MLH1_{S288c}::KanMX*, *PMS1_{S288c}-R818K::HIS3*). Four independent Ura⁺ G418^r colonies were sporulated and tetrad dissected. The spore clones from four-viable spore tetrads all displayed 2:2 segregation for *lys2-A₁₄/LYS2*, *trp1/TRP1*, and *ura3/URA3*. The sporulation viability of the diploid was 82%, with 62% of tetrads displaying four viable spores (79 tetrads dissected).

42 *lys2-A₁₄* spore clones (one Lys⁻ spore clone chosen per tetrad) were tested using a patch assay for reversion to Lys⁺. The 42 spore clones all displayed ~1:1 ratios for *MLH1_{S288c}::KanMX/MLH1_{YJM555}* segregation (the open reading frames for *MLH1_{YJM555}* and *MLH1_{S288c}* are identical; Figure 4), *PMS1_{YJM555}/PMS1_{S288c}R818K::HIS3* (genotyped using PCR primers that distinguished between the presence or absence of the *HIS3* insertion), *trp1/TRP1*, *leu2/LEU2*, and *ura3/URA3* (Table S4). To confirm phenotype testing, the *PMS1* gene was PCR amplified from three *PMS1_{YJM555}* and two *PMS1_{S288c}R818K::HIS3* spore clones. 700 bp of the *PMS1* gene surrounding the incompatibility site at amino acid 818 were sequenced; all five spore clones displayed DNA sequences that corresponded to the assigned phenotype. To determine mutator phenotypes, each of the 42 *lys2-A₁₄* spore clones was struck to single colonies on YPD media. Five ~2 mm in diameter colonies from each spore clone were then patched onto 1.5 cm x 1.5 cm squares on lysine dropout plates. It is important to note that colony morphology and size varied somewhat between spore clones, and for four spore clones, the colonies grew to ~ 1 mm in diameter. Plates were scored for Lys⁺ reversion colonies after 3 days of growth at 30 C, with the median number of colonies recorded per patch.

Strains and plasmids are available upon request, and the DNA sequences of the *MLH1* and *PMS1* genes from 1010 isolates is present in Files S1 and S2. Supporting information contains all detailed descriptions of all supplemental files.

Results

Incompatibility confers a fitness cost in evolved strains in non-stress conditions

In Bui *et al.* (2015), we showed that haploid yeast containing the incompatible S288c *MLH1-SKI* *PMS1* genotype (abbreviated as *cMLH1-kPMS1*) displayed a transient growth advantage in high salt that resulted from mutations in the *PMR1* gene. This observation encouraged us to determine the time frame in which incompatible strains that display an initial transient adaptive advantage can maintain fitness in non-stress and stress conditions. To answer this, we mated three different high-salt evolved clones, constructed to contain the incompatible *cMLH1-kPMS1* combination, to the unevolved compatible strain EAY3241. These incompatible strains, EAY3688 (*pmr1-T459A*), EAY3689 (*pmr1-T412C*), and EAY3690 (*pmr1-T2G*), were evolved for 70 generations in high salt (Bui *et al.* 2015; Table S1). This window was chosen because incompatible strains displayed an adaptive advantage over compatible strains when evolved for 70 generations that was lost after 120 generations (Bui *et al.* 2015). The diploids were then sporulated and the resulting spores genotyped for incompatibility, mating type, and *PMR1* (Figure 2). We then created for each *pmr1* allele mixed cultures that contained equal proportions of the same mating type spore clones that were *PMR1* compatible, *pmr1* compatible, *PMR1* incompatible, or *pmr1* incompatible, and grew them overnight, followed by subsequent transfers, in YPD or YPD + 1.2 M NaCl (roughly seven generations of growth per transfer; Materials and

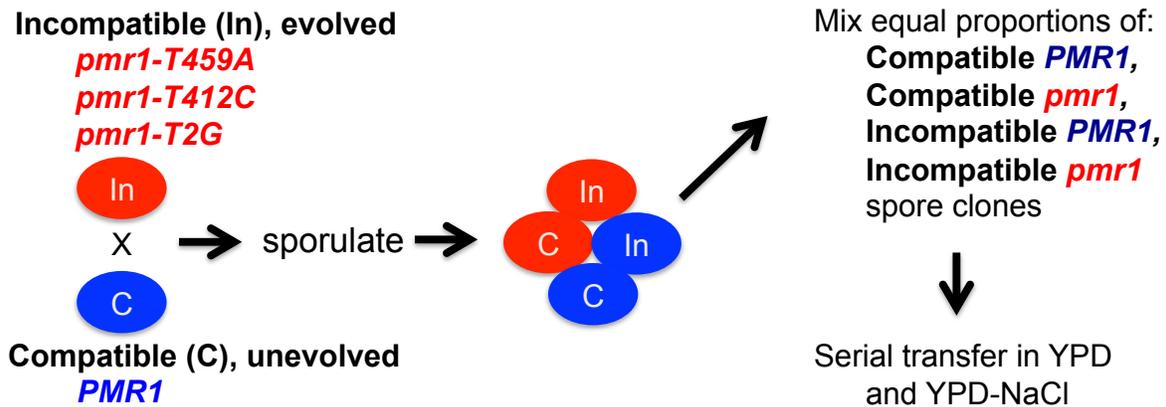


Figure 2. Competition experiments of spore clones from post adaptation strains. Clones derived from incompatible strains evolved in high salt for 70 generations (EAY3688-*pmr1-T459A*, EAY3689-*pmr1-T412C*, EAY3690-*pmr1-T2G*; Bui *et al.* 2015) were mated with the unevolved and compatible strain EAY3241 (Table S1). The resulting diploids were sporulated and genotyped for *PMR1*, *MLH1*, *PMS1*, and mating type. Spore clones of the same mating type and of the four possible genotypes (*kMLH1-kPMS1*(compatible)-*PMR1*, *kMLH1-kPMS1*(compatible)-*pmr1*, *cMLH1-kPMS1*(incompatible)-*PMR1*, and *cMLH1-kPMS1*(incompatible)-*pmr1*) were grown overnight in YPD and then mixed in equal proportions (Transfer 0). 4, 3, and 3 spore clones of each of the genotypes (12 to 16 in total per competition) were competed from the matings involving *pmr1-T459A*, *pmr1-T412C*, *pmr1-T2G*, respectively. 2×10^7 cells of the initial mixed culture (Transfer 0) were transferred into 6 ml of YPD or YPD + 1.2 M NaCl and then grown for 24 hrs (~7 generations of growth). The same amount of cells was used in subsequent transfers. At the indicated transfers cells were genotyped for compatibility (*kPMS1::HIS3*, *kMLH1::KanMX*) and incompatibility (*kPMS1::HIS3*, *cMLH1::KanMX*) on YPD plates containing geneticin or nourseothricin.

Methods). Our goal in these studies was to determine if the incompatible strains would ultimately display a fitness defect in rich media due to the accumulation of recessive mutations, as predicted by Heck *et al.* (2006). We also tested whether the rapidly evolved incompatible strains would continue to show an adaptive advantage in high salt media.

We measured in rich media fitness differences between incompatible and compatible strains. As shown in Table 2, Table S5, and Figure 3, we observed a gradual decrease in fitness of incompatible cells. This observation supports the idea that being incompatible is not favorable shortly after adaptation (Bui *et al.* 2015). At the end of the transfer experiment most competitions resulted in a higher proportion of compatible cells and all 28 independent clones (half compatible, half incompatible) isolated from Transfer 24 were *PMRI*, indicating that all three *pmr1* mutations were deleterious in rich media.

For competitions in YPD +1.2 M NaCl, we observed a general decline in the fitness of incompatible strains compared to compatible strains in the first 50 generations of competition. In the next 100 generations some incompatible strains displayed increased fitness (Figure 3; Table 2; Table S5). Sequencing of 24 clones (12 compatible, 12 incompatible clones) at Transfer 24 showed that all retained their initial adaptive *pmr1* mutations. These results showed that further adaptive advantages beyond the initial transient one were likely countered by the accumulation of deleterious mutations.

Natural isolates were identified that display the incompatible genotype

Previously Heck *et al.* (2006) identified one site each in MLH1 (aa 761, Asp (D) in S288c, Gly (G) in SK1) and PMS1 (aa 818/822, Arg (R) in S288c, Lys (K) in SK1) that accounted for the MMR incompatibility seen in progeny derived from crossing S288c and SK1 strains, with only

Table 2. Fitness of evolved incompatible and compatible strains as a function of transfer in YPD and YPD + NaCl media.

Transfer	Fitness, $w \pm$ SEM		n
	YPD	YPD+NaCl	
3	0.98± 0.01	0.98± 0.01	10
4	0.97± 0.02	0.95± 0.02	6
8	0.95± 0.01	0.93± 0.03	10
12	0.93± 0.03	1.06± 0.09	6
14	0.92± 0.02	1.06± 0.05	10
16	0.92± 0.02	1.05± 0.06	10
20	0.91± 0.02	1.06± 0.06	10
24	0.90± 0.03	1.09± 0.05	10

The proportion of compatible and incompatible genotypes was determined after the indicated number of transfers for spore clone competitions performed in YPD and YPD+ 1.2 M NaCl (Figure 3). Fitness (w) values were calculated as $w = ((p_t/q_t)/(p_0/q_0))^{1/t}$, where t equals the number of generations after T transfers (7 generations per transfer; Hartl and Clark 2007), p_0 and q_0 are the number of incompatible and compatible cells, respectively at Transfer 0, and p_t and q_t are the number of incompatible and compatible cells, respectively, at the indicated Transfer. n is the number of unique competitions performed for each data set. In this table the average fitness of the three *pmr1* genotypes is presented. In Table S5, the fitness values for each *pmr1* genotype are shown.

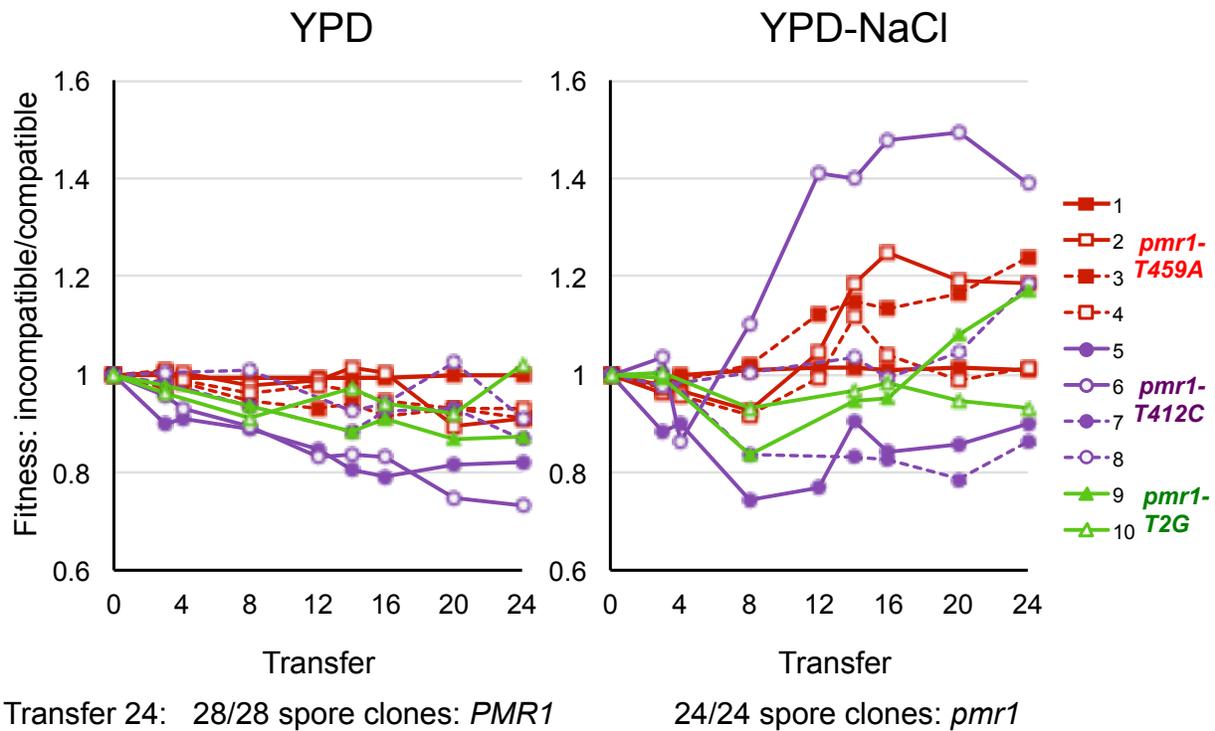


Figure 3. Fitness of post-adaptation incompatible and compatible strains in YPD and YPD-1.2 M NaCl. Fitness (w) was determined in cultures genotyped for *MLH1-PMS1* compatible and incompatible clones in YPD and YPD + 1.2 M NaCl as described in the Materials and Methods and Table 2 and Figure 2. Ten same mating-type competitions (competitions 1-4 for *pmr1-T459A* evolved, 5-8 for *pmr1-T412C* evolved, and 9-10 for *pmr1-T2G* evolved) are shown. For the YPD transfer experiments, 14 incompatible and 14 compatible clones at Transfer 24 (6, 4, 4, each from *pmr1-T459A*, *pmr1-T412C* and *pmr1-T2G*, respectively) were genotyped for *PMR1*. For the YPD + 1.2 M NaCl transfer experiments, 12 incompatible and 12 compatible clones at Transfer 24 (four each from the *pmr1-T459A*, *pmr1-T412C* and *pmr1-T2G*) were genotyped for *PMR1*.

MLH1-D761 / PMS1-K818 conferring a mutator phenotype (Figure 1; Heck *et al.* 2006; Demogines *et al.* 2008). We took advantage of a recent whole-genome sequencing analysis of a world-wide collection of 1,010 *Saccharomyces cerevisiae* isolates to determine if the incompatible combination exists in nature (File S1; File S2; full genomes and their analyses will be presented in the 1002 Yeast Genomes Project - <http://1002genomes.u-strasbg.fr/>). All 1,010 isolates displayed D and/ or G at amino acid position 761 in MLH1 and R and/or K at position 818/822 in PMS1.

In total, 904 of the 1,010 isolates are homozygous for polymorphisms in MLH1 and PMS1. As shown in Figure 1 and Table S2, 693 of these isolates (MLH1-G761 / PMS1-R818) mapped to the ancestral group, 182 to the S288c group (MLH1-D761 / PMS1-R818), 28 to the SK1 group (MLH1-G761 / PMS1-K818), and one (the YJM523 clinical isolate) to the incompatible group (MLH1-D761 / PMS1-K818). The isolates in both the S288c and SK1 groups are not closely related, suggesting that the incompatible genotype could arise independently, or are different combinations of mosaic genomes derived from a common event (Strope *et al.* 2015; Schacherer *et al.* 2009). A Fisher's Exact test of the observed number of isolates in each category showed that the number of isolates in the incompatible group was lower than expected by chance if there were no negative epistatic interactions involving this combination of MLH1-PMS1 ($p = 0.018$). These data suggest that there may be active selection against the incompatible genotype. We hypothesize that because of the long-term fitness cost associated with the incompatibility, there would be few if any natural isolates maintained as incompatible.

Among the 1,010 isolates, 106 are heterozygous for the *MLH1* 761 and/or *PMS1* 818/822 genotypes. Previously we showed in experiments where both incompatible and compatible

genotypes were present that the incompatible genotype is recessive (Argueso *et al.* 2003). In total, 18 of the 106 isolates have the potential to generate incompatible progeny from meiotic spores, assuming that they are diploids competent to undergo meiosis and form haploid spores (Table 1). Of the eighteen, twelve are tetraploid, four are diploid and one is triploid, with most capable of sporulating. Eight were obtained from human clinical samples, six from breweries, and three from bakeries, and one from fruit.

14 of the 18 isolates display heterozygous S288c/SK1 *MLHI* (amino acid 761) and S288c/SK1 *PMSI* (amino acid 818/822) genotypes, suggesting that they could have been created through matings between homozygous strains containing S288c and SK1 *MLHI-PMSI* genotypes. Sporulation of isolates with these genotypes would yield meiotic progeny in which one quarter would have the incompatible genotype (see YJM521 in Figure 4A). One isolate (CBS4455) displays an S288c *MLHI*-S288c/SK1 *PMSI* genotype and three an S288c/SK1 *MLHI*-SK1 *PMSI* genotype, possibly generated through crosses between compatible and incompatible strains. If sporulated, half of the meiotic progeny from these four isolates would display the incompatible genotype. These results suggest that the potential to generate incompatible progeny is present in compatible natural isolates and this potential may facilitate adaptive evolution by providing a brief increase in the mutation rate.

The above results encouraged us to look more closely at YJM523, which is the only one of 1,010 isolates homozygous for the incompatible genotype (Figure 1; Figure 4; File S1; File S2). YJM523 is a clinical isolate sampled from Stanford University Hospital (Stanford, California; Strobe *et al.* 2015; deposited in the Phaff Yeast Culture Collection, University of California, Davis). A previous genome-wide phylogenetic analysis of 93 *S. cerevisiae* strains showed that a spore clone derivative of YJM523, YJM555, is closely related to the compatible

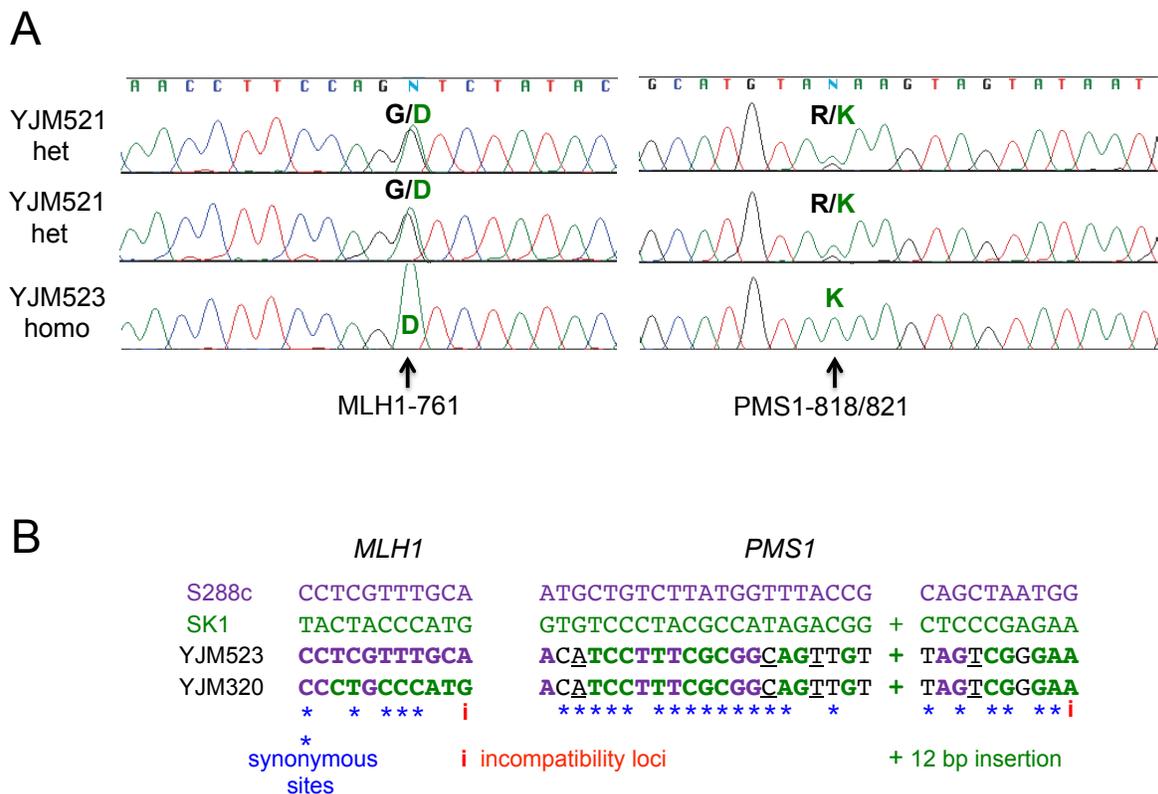


Figure 4. Polymorphisms observed in the closely related YJM555 and YJM320b *S. cerevisiae* strains. **A.** DNA sequence, as shown by chromatogram traces, of the *MLH1* and *PMS1* incompatibility sites in YJM521 and YJM523. DNA sequencing reactions are shown for two *MLH1* and *PMS1* sequences that were independently PCR amplified (using primers AO324, A0821 for *MLH1*, A0548, AO481 for *PMS1*; Table S4) from YJM521. Sequencing reactions were performed on ten *MLH1* and *PMS1* DNA fragments independently PCR amplified from YJM523; one such reaction is shown for *MLH1* and *PMS1* (see Materials and Methods and Table S4 for details). Amino acids in the *MLH1* and *PMS1* ORFs are shown that correspond to the presented DNA sequences. These are Gly (G) or Asp (D) at amino acid 761 in *MLH1*, Arg (R) or Lys (K) at amino acid 818/822 in *PMS1*. Homo = homozygous for genetic information at an incompatibility site; het = heterozygous. **B.** Haplotype structure of *MLH1* (11 polymorphic sites) and *PMS1* (32 polymorphic sites) based on the genome sequence of S288c, SK1, YJM320 and YJM523. YJM555 is a strain derived from the isolate YJM523 (Strope *et al.* 2015), and YJM523 and YJM555 contain identical and homozygous *MLH1* and *PMS1* sequences. Sequences identical to S288c (purple font) and SK1 (green font) are indicated. Synonymous sites are shown as blue asterisks. Sites in *MLH1* and *PMS1* that contribute to MMR incompatibility are shown by the red “i”. Sequences in black font were not found in S288c or SK1; those in black and underlined were not seen in 1007 of the other 1008 yeast isolates that were sequenced (they were seen in the closely related strain YJM554 (Strope *et al.* 2015)). The 12 bp insertion seen in the *SKI* *PMS1* open reading frame (Table 3) is indicated by a +. For *MLH1* nucleotide position of each polymorphic site in the open reading, from left to right, are: 486, 720, 812, 834, 997, 1044, 1237, 1875, 2032, 2108, 2282. For *PMS1* the positions are: 122, 177, 210, 213, 258, 333, 335, 465, 552, 558, 708, 711, 810, 855, 858, 939, 1150, 1175, 1191, 1199, 1201, 1249 (insertion), 1329, 1538, 1575, 1691, 1782, 1821, 2303, 2322, 2364, 2453.

strain YJM320 (Strope *et al.* 2015; Table 3; Figure 4B). We speculate that YJM523 and YJM320 are recently derived from the same origin, perhaps created by a single mating event between S288c and SK1 group isolates (see below and Discussion). It is important to note that at the genome-wide level, it does not appear that YJM523 is the result of a “direct” cross between S288c and SK1. If this was the case, then the YJM523 genome would show a mosaic of large regions identical to S288c interspersed with regions of high sequence divergence corresponding to SK1, but this was not observed (A. Friedrich and J. Schacherer, 1002 Yeast Genomes Project - <http://1002genomes.u-strasbg.fr/>).

Previously we showed that the distribution of sequence polymorphisms among ten *S. cerevisiae* strains for a 32 kb region indicated evolutionarily recent recombination among S288c and SK1 class strains, with YJM320 showing extensive evidence of recombination with the S288c group (Heck *et al.* 2006). At that time we speculated that the incompatible genotype was generated through inter-strain crosses but was not observed in our earlier sampling due to fitness costs. The phylogenetic and single nucleotide polymorphism (SNP) data presented in Strope *et al.* (2015) and Table 3 and Figure 4 strengthen this model, suggesting that YJM523 resulted from a cross between S288c and SK1 group isolates that yielded the incompatible genotype. Consistent with this hypothesis is the finding that YJM523 displayed a *PMS1* SNP pattern consistent with an SK1 class origin (sixteen nucleotide positions identical to the SK1 sequence vs. six identical to S288c), and an *MLH1* SNP pattern consistent with an S288c class origin (all eleven sites identical to S288c sequence). YJM320 is consistent with an SK1 class origin for both *PMS1* (sixteen nucleotide positions identical to the SK1 sequence vs. six identical to S288c) and *MLH1* (eight sites identical to the SK1 sequence vs. three identical to the S288c sequence).

Table 3. Non-synonymous substitutions in MLH1 and PMS1.

Strain	Amino acid position in MLH1					
	240	271*	333	678	703	761
S288c	S	L	E	D	P	D
SK1	R	P	K	N	L	G
YJM523	S	L	E	D	P	D
YJM320	S	P	E	N	L	G

Strain	Amino acid position in PMS1										
	41	112	384	392	400	401	416**	513	564	768	818
S288c	N	I	F	E	T	A	T	Y	A	K	R
SK1	S	T	V	E	S	A	TCEGT	F	A	K	K
YJM523	N	I	V	V	S	S	TCEGT	Y	V	R	K
YJM320	N	I	V	V	S	S	TCEGT	Y	V	R	K

All substitutions that result in amino acid changes are shown relative to the S288c strain. S288c, SK1, YJM320 are compatible for the *MLH1-PMS1* genotype, YJM523 are incompatible.

*Proline residue at MLH1-271 partially suppresses the *cMLH1-kPMS1* incompatibility (Demogines *et al.* 2008). **Compared to S288c, SK1, YJM523 and YJM320b strains/isolates contain a four amino acid insertion (CEGT) after T416 in *PMS1*.

That the pattern for synonymous sites matches the non-synonymous sites gives us confidence that it reflects evolutionary history and not selective convergence of amino acid residues.

The MLH1-PMS1 incompatible combination derived from a natural isolate confers a mutator phenotype in the S288c strain background

We used the highly sensitive *lys2-A₁₄* reversion assay to assess the mutator phenotype of the *MLH1* and *PMS1* gene combinations from YJM523. The *lys2-A₁₄* allele contains a homopolymeric run of 14 A residues inserted out of frame within the *LYS2* gene. In this assay wild-type and MMR null strains display a roughly 10,000-fold difference in the rate of reversion to Lys⁺ (Tran *et al.* 1997). The *MLH1* and *PMS1* genes from YJM523 were cloned into *ARS-CEN* vectors and transformed into an *mlh1Δ pms1Δ* S288c strain that was used previously to characterize *MLH1-PMS1* combinations from 65 natural isolates (Demogines *et al.* 2008). The YJM523 *MLH1*-YJM523 *PMS1* combination confers a 196-fold higher mutation rate compared to the *cMLH1-cPMS1* combination (Table 4). This mutation rate is higher than that seen for the *cMLH1-kPMS1* incompatibility (75-fold), indicating that there are likely to be other polymorphisms present in the two genes that enhance the incompatibility (Demogines *et al.* 2008). Together these data indicate that the *MLH1-PMS1* combination from YJM523 displays a strong incompatibility in the S288c strain background.

Strain containing the MLH1-PMS1 incompatible combination displays a low mutation rate

Table 4 shows that the YJM523/555 *MLH1-PMS1* gene combination confers a mutator phenotype in the S288c strain background. However, neighbor-joining analysis of the genomes of this and other strains presented by Strobe *et al.* (2015) shows that the branch on the tree

Table 4. Mutation rates in an S288c strain containing *MLH1* and *PMS1* genes derived from S288c, SK1 and YJM523.

<i>MLH1-PMS1</i> genotype	Lys ⁺ reversion rate (10 ⁻⁷), (95% CI (10 ⁻⁷))	Relative rate	(n)
S288c-S288c, compatible	4.1 (1.7-13.8)	1	13
S288c-SK1, incompatible	311 (111-919)*	75	16
YJM523/555-YJM523/555	808 (566-4,450)*	196	39
<i>mlh1Δ, pms1Δ</i>	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 strain or isolate (Tables S1, S2). 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. *Significantly different from S288c-S288c (p < 0.001, Mann Whitney test).

leading to YJM555, a spore clone derivative of YJM523, is not unusually long, and is similar to that seen for compatible strains, suggesting that YJM523 may not have been a mutator for a long evolutionary period. One way to reconcile this observation is that the *MLH1-PMS1* incompatibility formed in YJM523 facilitated a rapid increase in mutation rate that was then returned to normal by the occurrence of suppressor mutations that restored normal MMR functions. To further test this idea, we constructed a set of vectors (Figure 5; Table S3; Materials and Methods) that can be transformed into natural yeast strains to serve as a proxy to estimate genome mutation rate. We introduced homopolymeric A sequences into the *KanMX* open reading frame immediately after its methionine 17 codon, with the goal of creating a sensitive mutation detection assay. Previous work showed that homopolymeric runs undergo DNA slippage in a variety of DNA repair mutant backgrounds (Tran *et al.* 1997), and such events occur at an especially high frequency in MMR defective strains. For example, *msh2Δ* strains display a 10,000-fold increase in mutation rate compared to wild-type in strains bearing the *lys2A₁₄* reporter (Tran *et al.* 1997).

The mutation reporter plasmids contain *ARS-CEN* and *NatMX* markers and a *kanMX::insE-A₁₄* construct whose expression is driven by *URA3* or *LEU2* promoters. These promoters replaced the strong *TEF* promoter that drives *KanMX* expression in pFA6-KanMX; we found that the *TEF* promoter prevented the detection of large differences in reversion frequency between in and out of frame *KanMX::insE* constructs, presumably due to high rates of transcriptional slippage. In the *URA3* promoter-*kanMX::insE-A₁₄* reporter in pEAA613, a 55 bp sequence containing a +1 frameshift in the 14 bp homopolymeric A run (*insE-A₁₄*; Tran *et al.* 1997) was inserted immediately after the *URA3* ATG. This sequence was immediately followed by codons 18 to 269 of the *KanMX* open reading frame. Derivatives were constructed that

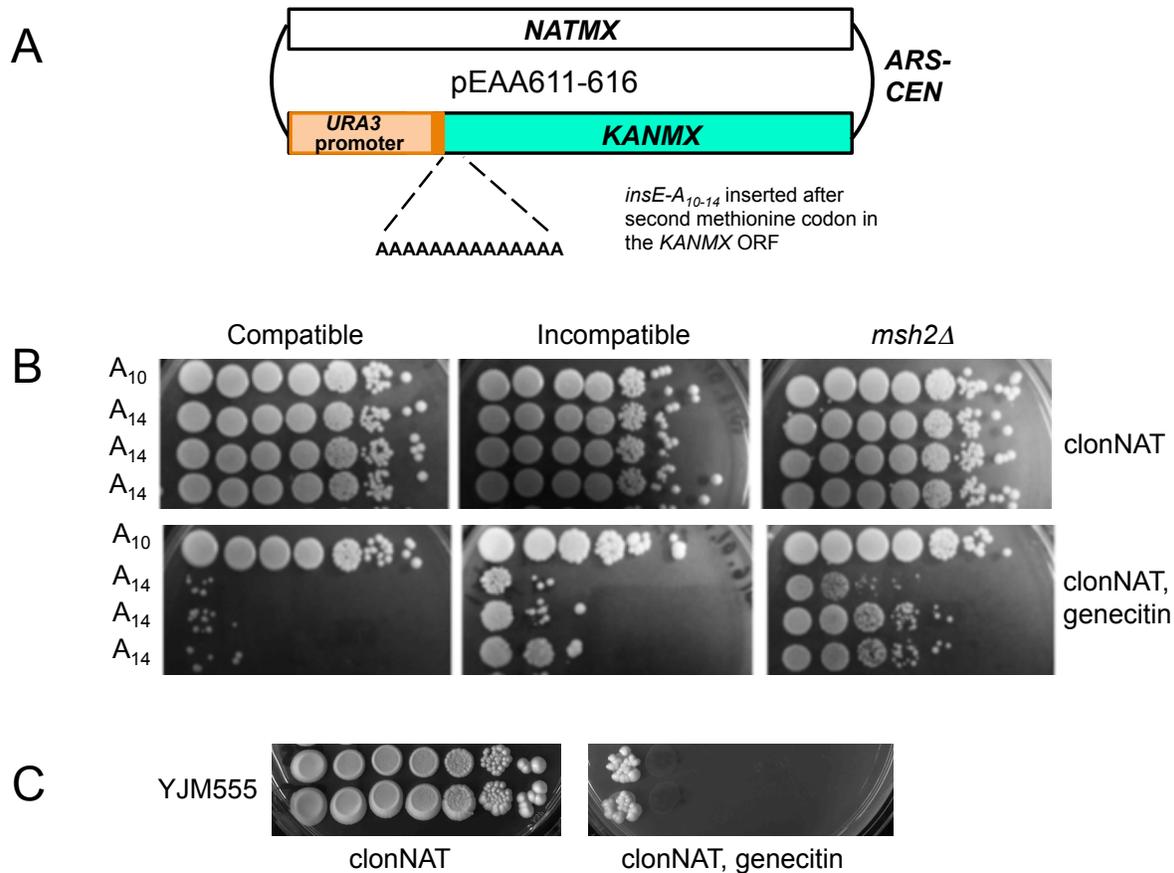


Figure 5. Use of a *URA3 promoter-kanMX::insE-A₁₄* plasmid to measure mutation rates in natural yeast isolates. **A.** The *ARS-CEN* vector pEAA613 contains a *NatMX* selectable marker and a frameshift reporter in which the *insE-A₁₄* sequence from Tran *et al.* (1997) was inserted immediately after methionine 17 in the *KanMX* open reading frame. In this reporter the *URA3* gene promoter drives expression of *KanMX*, using the methionine 17 in the *KanMX* open reading frame as the initiation codon. The resulting construct contains a +1 frameshift mutation that disrupts *KanMX* function. A reporter construct that contains the *insE-A₁₀* insertion and does not disrupt the *KanMX* open reading frame is shown as an in-frame control. Frameshift mutation events (e.g. a -1 deletion in the homopolymeric A run) are detected on YPD plates containing nourseothricin and geneticin. **B.** Examples of reversion assays performed using the *URA3 promoter-kanMX::insE-A₁₄* plasmid. EAY1369 (*cMLH1-cPMS1* compatible), EAY1370 (*cMLH1-kPMS1* incompatible) and EAY1372 (*msh2Δ*) were transformed with pEAA613 and plated in dilutions from 10 ml of a 10X concentrated overnight culture to 10 ml of 10⁰, 10⁻¹, 10⁻², 10⁻³, 10⁻⁴, 10⁻⁵ dilutions onto YPD nourseothricin (50 μg/ml), geneticin (300 μg/ml) plates. **C.** Example of the reversion assay performed for YJM555 is shown.

contained a 10 bp in-frame insertion (pEAA611), or utilized the *LEU2* promoter to drive expression (pEAA616; Table S3).

We tested the sensitivity of our reporter assay by transforming pEAA613 and pEAA616 into compatible, incompatible and MMR defective (*msh2Δ*) S288c strains (Table 5; Table S6). For strains containing pEAA613 we found that the *msh2Δ* strain displayed a 2,000-fold higher rate of reversion to geneticin resistance compared to the compatible strain, and that the incompatible strain showed a ten-fold higher mutation rate compared to the compatible strain (Figure 5; Table 5; Materials and Methods). Importantly, sequencing of the *kanMX::insE-A₁₄* reporter from geneticin resistant clones showed that a -1 frameshift had occurred in the homopolymeric A sequence in all of the clones (n=14). While this range is less sensitive than was seen in the *lys2-A₁₄* reversion assay (~75 fold in an assay with a 10,000 fold range between wild-type and MMR defective; Table 4), this reporter construct provides the opportunity to estimate mutation rates in natural strains that lack genetic markers. We then transformed the *URA3 promoter-kanMX::insE-A₁₄* and *LEU2 promoter-kanMX::insE-A₁₄* plasmids into YJM555 to determine if this strain showed a mutator phenotype consistent with a MMR incompatibility. As shown in Figure 5, Table 5, and Table S6, the overall mutation rate in YJM555 was similar to that seen in the S288c strain, indicating that YJM555, a well-behaved strain derived from a YJM523 spore clone (Strope *et al.* 2015), compensated for the increased mutation rate conferred by the *MLH1-PMS1* incompatibility.

Suppression of incompatibility in YJM555 appears to involve more than one locus.

To determine if a single suppressor locus is present in YJM555 that suppressed *MLH1-PMS1* incompatibility, we mated spores of YJM555 to the incompatible strain EAY3235 (relevant

Table 5. Reversion assay using the *URA3 promoter-kanMX::insE-A₁₄* plasmid.

strain	genotype	reversion G418 ^r (10 ⁻⁷)	95% CI	relative rate	n
EAY1369	wild-type	6.1	3.1-7.0	1	17
EAY1370	incompatible	59.8 *	25.7-239	9.9	17
EAY1372	<i>msh2Δ</i>	13,451*	7,230-46,090	2,142	11
YJM555		9.8**	6.5 -10.7	1.6	14

The indicated strains (Table S1) were transformed with *ARS-CENURA3 promoter-kanMX::insE-A₁₄* plasmid pEAA613. Independent cultures (n) were examined for reversion to genecitin 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.

*Significantly different from EAY1369 (p < 0.001, Mann-Whitney test)

**Significantly different from EAY1369 (p < 0.01, Mann-Whitney test)

genotype *lys2-A₁₄*, *MLHI_{S288c}::KanMX*, *PMSI_{S288c-R818K}::HIS3*) and sporulated the resulting diploids. We used the *lys2-A₁₄* reversion assay to analyze spore clones because it is currently the most sensitive assay to measure DNA mismatch repair incompatibility (Heck *et al.* 2006). While it is possible that lysine auxotrophy could alter a complex phenotype, our initial goal was to determine if suppression involved the contributions of more than one locus.

We picked 42 *lys2-A₁₄* spore clones and tested them for reversion to Lys⁺ using a colony reversion patch assay (Materials and Methods). As shown in Figure 6 and Table S7, the compatible control strain EAY3234 displayed a median reversion of 0 colonies (n = 10) and the incompatible control strain EAY3235 showed a median reversion of 91 (n = 10). These values fit with the ~two orders of magnitude difference in mutation rate between the two strains (Table 4; Heck *et al.* 2006). It is important to note that YJM555 displayed a mutation rate that was slightly higher than observed in S288c strains (Table 5), and so it is likely that the median colony number would be slightly higher than seen in EAY3234 if YJM555 bearing a *lys2-A₁₄* allele was tested for Lys⁺ reversion.

As shown in Figure 6, the mutator phenotype of the 42 spore clones distributed into roughly four groups, with one group showing Lys⁺ colony numbers close to EAY3234 (3 spore clones, 1-6 median colony number). A second group displayed Lys⁺ colony numbers that fell between those found for the compatible and incompatible controls (18 spore clones, 18-48 median colony number). A third group displayed phenotypes similar to EAY3235 (8 spore clones, 68-107 median colony number), and a fourth displayed colony numbers that were higher than EAY3235 (13 spore clones, >120 median colony number). If a single locus was responsible for a suppressor phenotype, and suppression was specific to *PMSI_{YJM555}* (the *MLHI* ORF sequence is the same in *MLHI_{S288c}* and *MLHI_{YJM555}*; Figure 4), then 25% of the spore clones

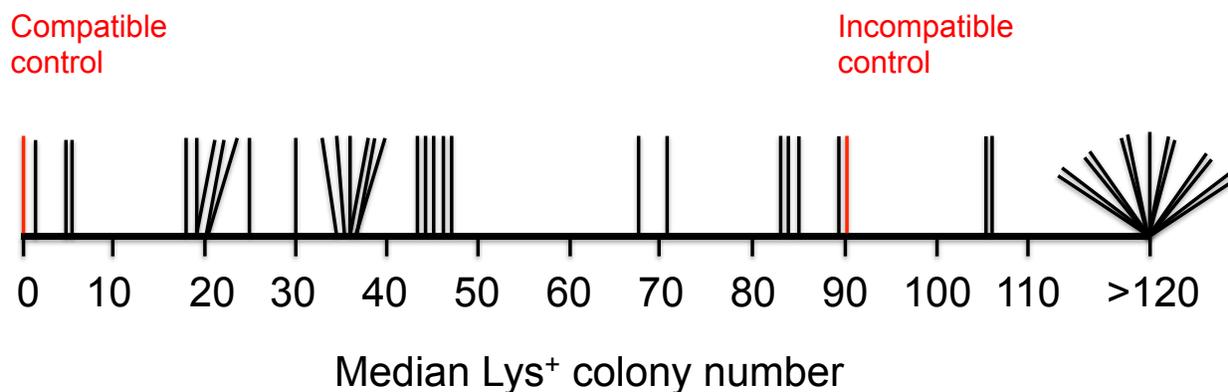


Figure 6. Mutator phenotype of 42 spore clones derived from YJM555 x EAY3235 cross. 42 *lys2-A₁₄* spore clones were tested for reversion to Lys⁺ in a colony patch assay (Materials and Methods). The median number of colonies from each spore clone assay is shown, from lowest to highest. Assays performed on compatible (EAY3234) and incompatible (EAY3235) control strains are shown in red.

would have been expected to show a suppressor phenotype, with the remainder showing phenotypes similar to that seen in EAY3235 (showing a bimodal distribution). 50% of spore clones would show a suppressed mutator phenotype if a single locus suppressor could interact with either *PMSI_{YJM555}* or *PMSI_{S288c-R818K}* (also showing a bimodal distribution). The fact that we see a range of mutator phenotypes, rather than a bimodal distribution, with only 3/42 spore clones showing phenotypes similar to compatible, and 13/42 showing phenotypes more severe than the incompatible, suggests that the suppression of mutator phenotype in YJM555 is unlikely to involve a single suppressor locus. A reasonable way to explain how the suppressor in YJM555 arose is that incompatible and compatible ancestral strains had mated, yielding spores that were incompatible and displayed an initial adaptive advantage. These spore clones then incurred a large number of mutations, some of which were beneficial, and in combination, suppressed the incompatible mutator phenotype, ultimately yielding YJM555. We cannot exclude the possibility that the spectrum of mutator phenotypes seen in the progeny results from suppressor loci interactions in combination with incompatibilities involving other genome stability loci (see Demogines *et al.* 2008).

Discussion

MMR Incompatibility is effective only as a short-term strategy for adaptation to new environments

We performed competitions involving populations of yeast that contained different combinations of evolved, unevolved, MMR compatible, and MMR incompatible genotypes (Figure 2 and Figure 3). In these repetitive transfer experiments we observed long-term fitness costs for MMR incompatibility, suggesting that there is an elevated mutation rate associated with MMR

incompatibility that could not be stably maintained in natural populations. We proposed that following a transient adaptive advantage provided by MMR incompatibility, incompatible strains can mate to compatible strains to maintain beneficial mutations, and thus return to MMR compatibility. Such a strategy would avoid the long-term fitness cost of being a mutator. This model is consistent with our survey of 1,010 yeast isolates showing that while there are a significant number (18 isolates; Table 1) heterozygous for MMR incompatibility, which is recessive, only one (YJM523) is homozygous, and is not maintained as a mutator (Table 1; Table 5; Bui *et al.* 2015; Wielgoss *et al.* 2013; Heck *et al.* 2006; Demogines *et al.* 2008; Argueso *et al.* 2003).

As indicated above we identified a single yeast isolate from 1,010 sampled (YJM523) that is homozygous for the incompatible genotype. Importantly, our data suggest that there are unknown mutations elsewhere in the genome of YJM523 that contribute to suppressing the mutator phenotype associated with incompatibility (see below). The molecular evolution experiments presented in Figure 3, our genotyping analysis of 1,010 yeast isolates, and the associated contingency test, suggest that once an adaptive mutation reached fixation a mutator state would no longer be favorable and would in fact become a liability. Consistent with this idea are observations made by Taddei *et al.* (1997), who proposed that a modestly elevated rate in an *E. coli* population would facilitate adaptation, but once adaptive mutations were fixed, the mutation load present in individuals in the population, through the accumulation of deleterious mutations throughout their genomes, would lead to a decrease in their frequency. Genomic analyses (Figure 4B; Strobe *et al.* 2015) and direct mutation rate measurements indicate that YJM523, the parent of YJM555, is not a mutator (Table 5; Table S6; Heck *et al.* 2006; Demogines *et al.* 2008). YJM555 displays an overall mutation rate similar to that seen for S288c

compatible strains, and whole genome neighbor joining analysis suggests that YJM555 is not an outlier compared to other genomes (Figure 4B; Table 5; Table S6; Strobe *et al.* 2015). As outlined below, we hypothesize that the *MLH1-PMS1* incompatibility seen in YJM523 was formed through a single mating between S288c and SK1 groups (Figure 1), and that these strains rapidly acquired extragenic suppressors that lowered their cellular mutation rates. In this scenario incompatible strains containing a beneficial mutation for adaptation continue to rapidly acquire mutations; among these are suppressors of the mutator phenotype that reduce genetic load (Wielgoss *et al.* 2013). Such a scenario has been observed in single cell organisms that have relatively large population sizes (e.g. *E. coli*; Wielgoss *et al.* 2013).

There are likely to be other strategies to suppress the mutator phenotype associated with incompatibility. For example, among the 1,010 isolates analyzed, the frequency of a known intragenic suppressor of the mutator phenotype (*MLH1-L271P*, see Table 3 and Demogines *et al.* 2008) is very high. In total, 812 of 937 isolates (73 could not be genotyped due to ambiguity at the 271 position, possibly due to heterozygosity) contained this polymorphism, but YJM523 does not contain it. The fact that the *MLH1-P271* suppressor is so common supports the idea that there is an abundant availability of this apparently non-deleterious suppressor polymorphism that can allow strains to minimize mutation rates.

We recently learned that Helen Murphy and colleagues (Skelly *et al.* 2016) analyzed strains presented in Strobe *et al.* (2015) for incompatible *MLH1-PMS1* combinations. They also found that YJM555, which contains the incompatible *MLH1-PMS1* combination, displays a much lower mutation rate than predicted based on its incompatible genotype, and their genomic analysis does not suggest evidence for an elevated mutator phenotype in YJM555. They postulated, as we demonstrated in our work, that *MLH1-PMS1* incompatibility is suppressed by

background genetic modifiers.

Finally, in this study we focused on a haploid model because we can uncover recessive phenotypes and monitor genotypes rapidly. However it is clear that most yeast in nature are diploid (1002 Yeast Genomes Project - <http://1002genomes.u-strasbg.fr/>). Our identification of isolates heterozygous for incompatibility indicate that in stress conditions that result in sporulation haploid progeny will be produced which display incompatible *MLH1-PMS1* genotypes. At present we cannot determine if the incompatible haploids will diploidize before their contribution of adaptation, but we plan to further characterize the heterozygous diploids (Table 1), with the hope that future analyses will better address these concerns. For example, we are intrigued about the possibility that loss of heterozygosity (LOH), which can result in the rapid fixation of recessive mutations (See Gerstein *et al.* 2014), provides a means to generate a homozygous incompatible genotype in the absence of sporulation.

How was the *MLH1-PMS1* incompatible in YJM523 generated?

YJM523 is a clinical isolate from Stanford University hospital; it appears to have an unmapped mutator suppressor(s) elsewhere in its genome because *MLH1-PMS1* combinations from this isolate conferred a higher mutation rate in the S288c background compared to its native backgrounds (compare data in Tables 4 and 5). How it acquired extragenic suppressors remains unclear. The SNP polymorphism data presented in Figure 4 appear consistent with YJM523 and YJM320 being created through a single mating between S288c and SK1 group strains followed by sporulation and the segregation of markers. This model is consistent with the presence of *PMS1* SNP variants that are identical in YJM523 (incompatible) and YJM320 (compatible; Figure 4). Furthermore, there are four *PMS1* SNPs (Figure 4B) and two *PMS1* amino acid

polymorphisms (V392, V564; Table 3) in these two isolates that are not found in 1,007 of the 1,010 yeast isolates (they were found in the closely related strain YJM554 (Strope *et al.* 2015), and YJM320 is on the same branch as YJM555 (Figure 4B; Strope *et al.* 2015). Taken together, this pattern suggests that SK1 and S288c group isolates mated to yield compatible and incompatible *MLH1-PMS1* genotypes. This suggests that YJM523 experienced an elevated mutation rate, presumably providing an adaptive advantage to a stress environment, which was subsequently suppressed by extragenic suppressor(s) (Table 4; Table 5). However, we cannot exclude the possibility that YJM523 never encountered situations where it had entered a mutator state in response to a changing environment, and was instead created in a background where the initial mutation rate was low and could tolerate an incompatible genotype. Unfortunately, the data are not available to further test this idea. We do not favor such a hypothesis based on the complexity of the suppression phenotype seen in YJM555 (Figure 6), and previous work showing that genomic mutation rates tend to vary by less than an order of magnitude in organisms such as *E. coli*, *Neurospora crassa*, and *S. cerevisiae* (Zeyl and Devisser 2001; Drake *et al.* 1998), suggesting that deviations from a typical rate would be selected against.

Development of a reporter construct to sensitively detect mutation rates in natural yeast isolates

The *ARS-CEN KanMX::insE-A₁₀₋₁₄* reporter construct described in the paper provides a new approach to measure mutation rates in a large number of isolates that lack markers that are typically used for genetic manipulation and whose chromosome content can vary with respect to ploidy and chromosome copy number (Strope *et al.* 2015; Storchova 2014). In these constructs the *NatMX* antibiotic resistance marker is used to select for the plasmid, and the *KanMX* marker

is used to identify frameshift reversion events (Figure 5). The pEAA613 construct displayed a 2000-fold difference in mutation rate between wild-type and MMR null. This difference is only five-fold lower than that seen using an integrated *lys2-A₁₄* reporter, which to our knowledge is the most sensitive assay developed to measure differences in frameshift reversion frequency between wild-type and MMR mutants (Table 5; Tran *et al.* 1997). Importantly, our reporter plasmid assay was capable of distinguishing between mutation rates in wild-type and MMR incompatible strains. In the *lys2-A₁₄* assay this difference was 75 to 100-fold in a 10,000-fold range; in the *KanMX::insE-A₁₀₋₁₄* assay it was 10-fold in a 2000 fold range (Tables 4, 5). We believe that this sensitive plasmid-based assay, in conjunction with whole-genome sequencing data and classical mutation accumulation analyses (Nishant *et al.* 2009, 2010; Lujan *et al.* 2014), will allow groups to accurately measure variation in mutation rate in natural isolates and determine if this rate varies under different environmental conditions. Furthermore, one can modify the sequences present in the *insE* part of the reporter to examine reversion to genecitin resistance due to base substitution and other types of frameshift events. pEAA613 and derivatives can also be easily modified to integrate the *KanMX::insE-A₁₀₋₁₄* reporter at a specific chromosomal location.

Acknowledgments

We thank Claudia Caradec, members of the Alani lab, and Jae Young Choi for helpful comments and technical advice, John McCusker for advice and providing us with the YJM521 and YJM523 natural isolates, and Helen Murphy for sharing unpublished data. J.S. is a member of the Institut Universitaire de France. D.T.B., N.A.A., and E.E.A. were supported by NIH GM53085. C.F.A. was supported by NIH GM095793. The 1002 yeast genomes project (J.S., G.L., 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. is 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.

References

Argueso, J. L., A. W. Kijas, S. Sarin, J. Heck, M. Waase, and E. Alani, 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: 873-886.

Boe, L., M. Danielsen, S. Knudsen, J. B., Petersen, J. Maymann, *et al.*, 2000 The frequency of mutators in populations of *Escherichia coli*. *Mutat. Res.* 448: 47–55. doi:10.1016/S0027-5107(99)00239-0

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

<http://doi.org/10.1371/journal.pgen.1005407>

Chao, L. and E. C. Cox, 1983 Competition between high and low mutating strains of *Escherichia coli*. *Evolution* 37: 125-134.

Demogines, A., E. Smith, L. Kruglyak, and E. Alani, 2008 Identification and dissection of a complex DNA repair sensitivity phenotype in Baker's yeast. *PLoS Genet.* 4: e1000123. doi: 10.1371/journal.pgen.1000123

Demogines, A., A. Wong, C. Aquadro, and E. Alani, 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

Denamur, E., G. Lecointre, P. Darlu, O. Tenaillon, C. Acquaviva, *et al.*, 2000 Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. Cell 103: 711-721.

[doi:10.1016/S0092-8674\(00\)00175-6](https://doi.org/10.1016/S0092-8674(00)00175-6)

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

Dobzhansky, T. 1936 Studies on Hybrid Sterility. II. Localization of Sterility Factors in *Drosophila Pseudoobscura* Hybrids. Genetics 21: 113–135. PMID: 17246786

Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow, 1998 Rates of spontaneous mutation. Genetics 148: 1667-1686.

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

Gerstein A. C., A. Kuzmin, and S. P. Otto, 2014 Loss-of-heterozygosity facilitates passage through Haldanes sieve for *Saccharomyces cerevisiae* undergoing adaptation Nature Communications 5: Article number 3819. doi:10.1038/ncomms4819

Gietz, R. D., and R. H. Schiestl, 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

Giraud, A., I. Matic, O. Tenaillon, A. Clara, M. Radman, *et al.*, 2001 Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. Science 291: 2606-2608.

doi:10.1126/science.1056421

Goldstein, A. L., and J. H. McCusker, 1999 Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast 15: 1541-1553. doi:10.1002/(SICI)1097-

0061(199910)15:14<1541::AID-YEA476>3.0.CO;2-K

Hartl, D., and A. Clark, 2007 Principles of Population Genetics. 4th Edition. Sinauer Associates. doi:10.1002/ajpa.1330800314

Heck, J. A., J. L. Argueso, Z. Gemici, R. G. Reeves, A. Bernard, *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. USA 103: 3256–3261.

doi:10.1073/pnas.0510998103

Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease, 1989 Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77: 51-59.

Hoffman, C. S. and F. Winston, 1987 A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene 57: 267-272.

Holm, C., D. W. Meeks-Wagner, W. L. Fangman, and D. Botstein, 1986 A rapid, efficient method for isolating DNA from yeast. Gene 42: 169-173.

Kimura, M. 1967 On the evolutionary adjustment of spontaneous mutation rates. Genetical Res. 9: 23-34.

LeClerc, J. W., B. Li, W. L. Payne, and T. A. Cebula, 1996 High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science 274: 1208-1211.

doi:10.1126/science.274.5290.1208

Liti, G., D. M. Carter, A. M. Moses, J. Warringer, L. Parts, *et al.*, 2009 Population genomics of domestic and wild yeasts. Nature 458: 337-441. doi: 10.1038/nature07743.

Liti, G. and E. J. Louis, 2005 Yeast evolution and comparative genomics. Annu. Rev. Microbiol. 59: 135-153. doi:10.1146/annurev.micro.59.030804.121400

Lujan, S. A., A. R. Clausen, A. B. Clark, H. K. MacAlpine, D. M., MacAlpine, *et al.*, 2014

Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res.* 24: 1751-1764. doi:10.1101/gr.178335.114.

Magwene, P. M., Ö. Kayıkçı, J. A. Granek, J. M. Reininga, Z. Scholl, *et al.*, 2011 Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 108: 1987-1992. doi:10.1073/pnas.1012544108

Mann, H. B. and D. R. Whitney, 1947 On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 18: 50–60.

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

Muller, H. J., 1939 Reversibility in evolution considered from the standpoint of genetics. *Biol. Rev. Camb. Philos. Soc.* 14: 261–280.

Muller, H. J., and G. Pontecorvo, 1940 Recombinants between *Drosophila* Species the F1 Hybrids of which are sterile. *Nature* 146: 199–200.

Murphy, H. A., and C. W. Zeyl, 2010 Yeast Sex: Surprisingly high rates of outcrossing between asci. *PLoS One* 5: e10461. doi:10.1371/journal.pone.0010461

Nishant, K. T., N. D. Singh, and E. Alani, 2009 Genomic mutation rates: what high-throughput methods can tell us. *Bioessays* 31: 912-920. doi:10.1002/bies.200900017.

Nishant, K. T., W. Wei, E. Mancera, J. L. Argueso, A. Schlattl, *et al.*, 2010 The baker's yeast diploid genome is remarkably stable in vegetative growth and meiosis. *PLoS Genet.* 6: e1001109. doi:10.1371/journal.pgen.1001109.

Orr, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* 139: 1805–1813. PMID: 7789779

Raynes, Y., M. R. Gazzara, and P. D. Sniegowski, 2012 Contrasting dynamics of a mutator allele

in asexual populations of differing size. *Evolution* 66: 2329-2334. doi:10.1111/j.1558-5646.2011.01577.x

Reuter, M., G. Bell, and D. Greig, 2007 Increased outbreeding in yeast in response to dispersal by an insect vector. *Curr. Biol.* 17: R81-83. doi:10.1016/j.cub.2006.11.059

Rose, M. D., F. Winston, and P. Hieter, 1990 *Methods in yeast genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. doi:10.1016/0307-4412(91)90039-B

Ruderfer, D. M., S. C. Pratt, H. S. Seidel, and L. Kruglyak, 2006 Population genomic analysis of outcrossing and recombination in yeast. *Nat. Genet.* 38: 1077-1081. doi:10.1038/ng1859

Schacherer, J., J. A. Shapiro, D. M. Ruderfer, and L. Kruglyak, 2009 Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* 458: 342-345. doi: 10.1038/nature07670.

Skelly D.A., P.M. Magwene, B. Meeks and H.A. Murphy, 2016 The effect of known mutator alleles in a cluster of clinical *Saccharomyces cerevisiae* strains. bioRxiv 090498; doi:<https://doi.org/10.1101/090498>

Steinmetz, L. M., H. Sinha, D. R. Richards, J. I. Spiegelman, P. J. Oefner, *et al.*, 2002 Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 416: 326-330.

Storchova, Z., 2014 Ploidy changes and genome stability in yeast. *Yeast* 31: 421-430. doi:10.1002/yea.3037.

Strope, P. K., D. A. Skelly, S. G. Kozmin, G. Mahadevan, E. A. Stone, *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: 762-774.

Taddei, F., M. Radman, J. Maynard-Smith, B. Toupance, P. H. Gouyon, *et al.*, 1997 Role of

mutator alleles in adaptive evolution. *Nature* 387: 700-702.

Tanaka, M. M., C. T. Bergstrom, and B. R. Levin, 2003 The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. *Genetics* 164: 843–854.

Thompson, D. A., M. M. Desai, and A. W. Murray, 2006 Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. *Curr. Biol.* 16: 1581-1590.

doi:10.1016/j.cub.2006.06.070

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

Tran, H. T., J. D. Keen, M. Krickler, M. A. Resnick, and D. A. Gordenin, 1997 Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants.

Mol. Cell. Biol. 17: 2859-2865.

Wielgoss, S., J. E. Barrick, O. Tenaillon, M. J. Wiser, W. J. Dittmar, *et al.*, 2013 Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc. Natl. Acad. Sci. USA* 110: 222-227. doi: 10.1073/pnas.1219574110

doi:10.1073/pnas.1219574110

Wilcoxon, F., 1945 Individual comparisons by ranking methods. *Biometrics Bulletin* 1: 80–83.

Winston, F., C. Dollard, and S. L. Ricupero-Hovasse, 1995 Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. *Yeast* 11: 53-55.

doi:10.1002/yea.320110107

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

Table S1. Strains used in this study.

EAY235	<i>MATa, ura3-52, leu2Δ1, trp1Δ63</i>
EAY280	<i>MATa, ura3-52, leu2Δ1, trp1Δ63, msh2Δ::hisG</i>
EAY1365	<i>MATa, ura3-52, leu2Δ1, trp1Δ63, his3 Δ 200, lys2::insE-A₁₄, mlh1Δ::KanMX4, pms1Δ::KanMX4</i>
EAY1369	<i>MATalpha, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2::insE-A₁₄, cPMS1::HIS3, cMLH1</i>
EAY1370	<i>MATalpha, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2::insE-A₁₄, kPMS1::HIS3, cMLH1</i>
EAY1372	<i>MATa, ura3-52, leu2Δ1, trp1Δ63, lys2::insE-A₁₄, msh2Δ::hisG</i>
EAY3191	<i>MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, kPMS1::HIS3, kMLH1::KanMX</i>
EAY3234	<i>MATa, lys2::insE-A₁₄, ura3-52, trp1Δ63, leu2Δ1, cMLH1::KanMX, cPMS1::HIS3</i>
EAY3235	<i>MATa, lys2::insE-A₁₄, ura3-52, trp1Δ63, leu2Δ1, cMLH1::KanMX, kPMS1::HIS3</i>
EAY3236	<i>MATalpha, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, cMLH1::KanMX, kPMS1::HIS3</i>
EAY3241	<i>MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, kPMS1::HIS3, kMLH1::NatMX</i>
EAY3242	<i>MATalpha. ura3-52, leu2Δ1, trp1Δ63, his3Δ200, kPMS1::HIS3, kMLH1::NatMX</i>
EAY3688	EAY3236, <i>pmr1-T459A</i>
EAY3689	EAY3236, <i>pmr1-T412C</i>
EAY3690	EAY3236, <i>pmr1-T2G</i>
YJM521	Natural isolate obtained from John McCusker, Duke University
YJM523	Natural isolate obtained from John McCusker, Duke University
YJM555	Homozygous diploid strain derived from a spore clone of YJM523 (Strope <i>et al.</i> 2015)

Table S2. List of S288c derived and SK1 derived *MLH1-PMS1* isolates.

S288c derived isolates (182)

YJM280_b, YJM326_b, YJM428_1b, YJM436_1b, CBS3093_1b, DBVPG4651_1b, EM93_3, TL229S2.2, CLIB413_1b, CLIB483_1b, Y10_1b, YJM269_1b, DBVPG1339_1b, UC8_1b, WE372_1b, CLIB208_1, CBS7765, CBS7962, CBS7964, CBS2973, CBS4454, CBS1252, CBS4456, CBS7959, CBS1464, CBS1385, CBS7961, CBS1544, CBS6333, CBS7963, CBS3081, CBS3093, CBS1487, E4, MC8, MC10, CLQCA_02-003, CLQCA_10-027, CLQCA_17-060, SJ5L16, SJ5L18, DBVPG6696, A-6, A-18, CBS1250, CBS4903, CBS5378, CBS420, CBS1257, CBS2087, CBS2421, YJM308, CECT10497, CECT10266, CH02, CH06, CH10, CH11, CH14, CH13, 2282, 2644, MAJ_G, CLIB650, CLIB651, CLIB653, CLIB654, EXF-5042, EXF-5046, EXF-7145, CLIB484, CLIB485, CLIB552, CLIB655, CLIB1060, CLIB1077, DBVPG1076, DBVPG1080, DBVPG1105, DBVPG1122, DBVPG1123, DBVPG1125, DBVPG1126, DBVPG1145, DBVPG1163, DBVPG1339, DBVPG1490, DBVPG1523, DBVPG1584, DBVPG1594, DBVPG1608, DBVPG1975, 368428S, 269521J, 2698, DBVPG1841, DBVPG1843, DBVPG1848, DBVPG1880, DBVPG1895, DBVPG4317, DBVPG4460, DBVPG4651, DBVPG6302, DBVPG6350, M.9.1, RP.10.14, SM.9.1.AL1, SM.9.1.BL7, SM.9.2.BR3(L), 4LBI-3, DBVPG3051, 1, CECT10331:9-1, CECT10351-1B, CECT10171-5B, DBVPG1620-1A, N34:2-4(a), N26-1-10:7-5(a), N134:7-1(a), ho_SB, UCD_40-49, UCD_05-780, UCD_09-448, UCD_11-601, Lib61, SRC 120, SRC 147, SRC 213, CLI_1, CLI_2, CLI_3, CLI_4, CLI_5, CLI_6, CLI_7, CLI_8, CLI_9, CLI_10, CLI_11, CLI_12, CLI_13, CLI_14, UL G84F88I90, ultra lev. 1127, CLI_15, CLI_17, CLI_18, CLI_20, CLI_21, CLI_24, SM.9.4.BR2, SM.8.7.BR1, SM.8.8.BL1, SA.10.1.VL1, RP11.4.1, RP11.4.11, RP11.4.14, 1560, FTPW4, FTPW3, FTPW32, FTPW35, RB3-7Sc2, SC2-37, W303, 273614N, YJM451, YJM682, YJM683, YJM1199, YJM1242, YJM1252, YJM1311, YJM1386, YJM1433, YJM1450, YJM1463, YJM1477, YJM1527, YJM1549, YJM1078

SK1 derived isolates (28)

CBS2165a, CBS1388a, YJM320b, YJM653_1b, CBS2441, CBS1428, CBS1509, CBS1463, 21-4-114, CBS1239, M9, Y1, 2281, CLQCA_10-386, CLIB1085, CLIB1083, B9b-1, B9c-1, Ponton 12, Ponton 14, CLQCA_17-111, SK1, YJM681, YJM1307, YJM1381, YJM1385, YJM1419, YJM1478

Incompatible isolate (1)

YJM523

Table S3. Plasmids.

	Relevant genotype
pRS413	<i>ARS-CEN HIS3</i>
pRS415	<i>ARS-CEN, LEU2</i>
pRS416	<i>ARS-CEN, URA3</i>
pLZ259	<i>ARS-CEN, NATMX</i>
pEAA611	<i>ARS-CEN, NATMX,URA3 promoter-KANMX::insE-A₁₀</i>
pEAA612	<i>ARS-CEN, NATMX,URA3 promoter-kanMX::insE-A₁₁</i>
pEAA613	<i>ARS-CEN, NATMX,URA3 promoter-kanMX::insE-A₁₄</i>
pEAA614	<i>ARS-CEN, NATMX,LEU2 promoter-KANMX::insE-A₁₀</i>
pEAA615	<i>ARS-CEN, NATMX,LEU2 promoter-kanMX::insE-A₁₁</i>
pEAA616	<i>ARS-CEN, NATMX,LEU2 promoter-kanMX::insE-A₁₄</i>
pEAA609	<i>ARS-CEN, LEU2, MLH1-YJM523/YJM555</i>
pEAA213	<i>ARS-CEN, LEU2, MLH1-S288c</i>
pEAA214	<i>ARS-CEN, LEU2, MLH1-SK1</i>
pEAA610	<i>ARS-CEN, HIS3, PMS1-YJM523/YJM555</i>
pEAA238	<i>ARS-CEN, HIS3, PMS1-S288c</i>
pEAA239	<i>ARS-CEN, HIS3, PMS1-SK1</i>

Table S4. DNA primers used to sequence *MLH1* and *PMS1* from YJM521 and YJM523.

Primers used to sequence *MLH1*

AO324 5' ATAGTGTAGGAGGCGCTG

AO325 5' CATGTGGCAACAGTCACAGTAACG

AO327 5' ATGTAAGCCACTCCCAAGAGGCAG

AO328 5' GACGAGTTAAATGACGATGCTTCC

AO821 5' AACTTTGCGGCCGCGGATCCAGCCAAAACGTTTTAAAGTTA

AO3567 5' CAGGAAATAAACAAAAAACTTTGG

AO3568 5' AACAAAAAACTTTGGTATTACAGC

AO3569 5' CTTTGGTATTACAGCCAAAACG

Primers used to sequence *PMS1*

AO480 5' CCCAAAAGAATGTGCTCACAATCAGAAC

AO481 5' CCACGTTCATATTCTTAATGGCTAAGC

AO483 5' GAGAAAGAATATATCCTCAG

AO485 5' AAAGTA TCTGACGTTAACAGTTTC

AO548 5' CGATTCTAATACAGATTTTAATGACC

AO3564 5' TCATATTTTCGTAATCCTTCGAAAATG

AO3565 5' TTCGTAATCCTTCGAAAATGAGCTCC

AO3566 5' ATCCTTCGAAAATGAGCTCCAATCACG

Ten primer-pairs used to make *MLH1* PCR products, with the expected sizes

AO324 (205 bp upstream of ATG), AO821 (19 bp downstream of stop codon): 2.5 kb

AO325 (315 bp downstream of ATG), AO3567 (48 bp downstream of stop codon): 2.0 kb

AO325 (315 bp downstream of ATG), AO3568 (39 bp downstream of stop codon): 2.0 kb

AO325 (315 bp downstream of ATG), AO3569 (30 bp downstream of stop codon): 2.0 kb

AO327 (1309 bp downstream of ATG), AO3567 (48 bp downstream of stop codon): 1.0 kb

AO327 (1309 bp downstream of ATG), AO3568 (39 bp downstream of stop codon): 1.0 kb

AO327 (1309 bp downstream of ATG), AO3569 (30 bp downstream of stop codon): 1.0 kb

AO324 (205 bp upstream of ATG), AO3567 (48 bp downstream of stop codon): 2.6 kb

Table S4 (continued)

AO324 (205 bp upstream of ATG), AO3568 (39 bp downstream of stop codon):	2.6 kb
AO324 (205 bp upstream of ATG), AO3569 (30 bp downstream of stop codon):	2.5 kb
Ten primer-pairs used to make <i>PMSI</i> PCR products, with the expected sizes	
AO548 (125 bp upstream of ATG), AO481 (146 bp downstream of stop codon):	2.9 kb
AO480 (1090 bp downstream of ATG), AO3564 (at stop codon):	1.5 kb
AO480 (1090 bp downstream of ATG), AO3565 (6 bp upstream of stop codon):	1.5 kb
AO480 (1090 bp downstream of ATG), AO3566 (12 bp upstream of stop codon):	1.5 kb
AO483 (633 bp downstream of ATG), AO3564 (at stop codon):	2.0 kb
AO483 (633 bp downstream of ATG), AO3565 (6 bp upstream of stop codon):	2.0 kb
AO483 (633 bp downstream of ATG), AO3566 (12 bp upstream of stop codon):	2.0 kb
AO548 (125 bp upstream of ATG), AO3564 (at stop codon):	2.7 kb
AO548 (125 bp upstream of ATG), AO3565 (6 bp upstream of stop codon):	2.7 kb
AO548 (125 bp upstream of ATG), AO3566 (12 bp upstream of stop codon):	2.7 kb

Table S5. Fitness of evolved incompatible and compatible strains as a function of transfer in YPD and YPD + NaCl media.

Transfer	<i>pmr1</i> genotype	Fitness, $w \pm$ SEM		n
		YPD	YPD+NaCl	
3	<i>pmr1-T459A</i>	0.99±0.01	0.97±0.01	4
	<i>pmr1-T412C</i>	0.96±0.02	0.97±0.03	4
	<i>pmr1 -T2G</i>	0.97±0.01	1.00±0.01	2
4	<i>pmr1-T459A</i>	0.99±0.01	0.98±0.01	4
	<i>pmr1-T412C</i>	0.92±0.01	0.88±0.02	2
8	<i>pmr1-T459A</i>	0.97±0.01	0.97±0.03	4
	<i>pmr1-T412C</i>	0.93±0.03	0.92±0.08	4
	<i>pmr1 -T2G</i>	0.92±0.01	0.88±0.05	2
12	<i>pmr1-T459A</i>	0.97±0.02	1.04±0.03	4
	<i>pmr1-T412C</i>	0.84±0.01	1.09±0.32	2
14	<i>pmr1-T459A</i>	0.98±0.02	1.12±0.04	4
	<i>pmr1-T412C</i>	0.86±0.03	1.04±0.13	4
	<i>pmr1 -T2G</i>	0.93±0.04	0.96±0.01	2
16	<i>pmr1-T459A</i>	0.96±0.02	1.11±0.05	4
	<i>pmr1-T412C</i>	0.87±0.04	1.03±0.15	4
	<i>pmr1 -T2G</i>	0.93±0.02	0.97±0.02	2
20	<i>pmr1-T459A</i>	0.94±0.02	1.09±0.05	4
	<i>pmr1-T412C</i>	0.88±0.06	1.04±0.16	4
	<i>pmr1 -T2G</i>	0.89±0.03	1.01±0.07	2
24	<i>pmr1-T459A</i>	0.94±0.02	1.11±0.06	4
	<i>pmr1-T412C</i>	0.83±0.04	1.08±0.13	4
	<i>pmr1 -T2G</i>	0.95±0.07	1.05±0.12	2

The proportion of compatible and incompatible genotypes were determined after the indicated number of transfers for spore clone competitions performed in YPD and YPD+ 1.2 M NaCl (Fig 2; Fig 3). Fitness (w) values were calculated as $w = ((p_t/q_t)/(p_0/q_0))^{1/t}$, where t equals the number of generations after T transfers (7 generations per transfer; Hartl and Clark 2007), p_0 and q_0 are the number of incompatible and compatible cells, respectively at Transfer 0, and p_t and q_t are the number of incompatible and compatible cells, respectively, at the indicated Transfer. n is the number of unique competitions performed for each *pmr1* genotype.

Table S6. Reversion assay using the *LEU2 promoter-kanMX::insE-A₁₄* plasmid.

strain	genotype	reversion G418 ^r (10 ⁻⁷)	95% CI	relative rate	n
EAY1369	wild-type	6.9	5.3-8.9	1	36
EAY1370	incompatible	32.6*	3.3-38.5	4.7	29
EAY1372	<i>msh2Δ</i>	6,070*	1,843-7,778	878	29
YJM555		7.2	4.9-10.4	1.04	35

The indicated strains (Table S1) were transformed with *ARS-CEN LEU2 promoter-kanMX::insE-A₁₄* plasmid pEAA616. Independent cultures (n) were examined for reversion to genecitin 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. *Significantly different from EAY1369 ($p < 0.001$, Mann-Whitney test)

Table S7. Phenotype of spore clones obtained by crossing spores of YJM555 with EAY3235.

	<i>PMS1</i> genotype	<i>MLH1</i> genotype: G418 resistance	Lys+ colonies/patch	Lys+ median	Leu+/-	Ura+/-	Trp+/-
Controls							
	EAY3234 <i>PMS1-S288c::HIS3</i>	resistant	0, 0, 0, 0, 0, 1, 1, 1, 17	0	-	-	-
	EAY3235 <i>PMS1-S288c-R818K::HIS3</i>	resistant	34, 38, 66, 73, 90, 91, 94, 107, >120, >120	91	-	-	-
	YJM555 <i>PMS1-YJM555</i>	sensitive	not tested		+	+	+
Spore clones							
2_1	<i>PMS1-YJM555</i>	resistant	5, 4, 8, 13, 5	5	-	-	-
2_2	<i>PMS1-S288c-R818K::HIS3</i>	resistant	38, 17, 35, 22, 42	35	-	-	-
2_3	<i>PMS1-YJM555</i>	resistant	2, 9, 19, 63, 22	19	-	-	+
2_4	<i>PMS1-YJM555</i>	resistant	17, 19, 23, 24, 15	19	-	+	-
2_5	<i>PMS1-S288c-R818K::HIS3</i>	resistant	>120, 83, >120, 79, 81	83	-	+	-
2_6	<i>PMS1-YJM555</i>	sensitive	>120, >120, >120, >120, >120	>120	-	-	-
2_7	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	66, >120, >120, >120, 117	>120	+	-	-
2_8	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	>120, >120, 115, >120, >120	>120	-	+	+
2_9	<i>PMS1-S288c-R818K::HIS3</i>	resistant	7, 13, 42, 40, 34	34	-	-	+
2_10	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	44, 37, 48, 25, 25	37	+	-	-
2_11	<i>PMS1-S288c-R818K::HIS3</i>	resistant	72, >120, >120, >120, 76	>120	+	-	+
3_1	<i>PMS1-YJM555</i>	sensitive	107, >120, 88, 96, >120	107	+	+	-
3_2	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	>120, 47, 37, 106, 45	47	-	-	-
3_3	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	45, 71, 77, 64, 105	71	+	+	+
3_4	<i>PMS1-S288c-R818K::HIS3</i>	resistant	10, 20, 13, 51, 42	20	-	+	+
3_5	<i>PMS1-YJM555</i>	sensitive	106, >120, 104, 69, >120	106	+	-	+
3_6	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	31, 82, 113, 39, 68	68	+	+	-
3_7	<i>PMS1-YJM555</i>	resistant	>120, >120, >120, >120, >120	>120	+	+	-
3_8	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	50, 27, 46, 17, 82	46	-	+	-
3_9	<i>PMS1-S288c-R818K::HIS3</i>	resistant	>120, >120, 54, >120, 86	>120	+	+	+
3_10	<i>PMS1-YJM555</i>	resistant	>120, >120, >120, >120, >120	>120	-	+	+
3_11	<i>PMS1-S288c-R818K::HIS3</i>	resistant	20, 41, 38, 9, 8	20	+	-	+
4_1	<i>PMS1-S288c-R818K::HIS3</i>	resistant	46, >120, 29, 43, 3	43	+	-	-
4_2	<i>PMS1-S288c-R818K::HIS3</i>	resistant	14, 28, 6, 5, 1	6	+	-	-
4_3	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	86, >120, 80, 89, >120	89	-	+	+
4_4	<i>PMS1-YJM555</i>	resistant	>120, >120, >120, >120, >120	>120	+	-	-
4_5	<i>PMS1-S288c-R818K::HIS3</i>	resistant	17, 36, 5, 18, 46	18	-	-	+
4_6	<i>PMS1-S288c-R818K::HIS3</i>	resistant	36, 69, 23, 28, 99	36	-	+	-
4_7	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	>120, 37, 30, 35, 67	37	+	-	+
4_8	<i>PMS1-YJM555</i>	sensitive	>120, >120, >120, >120, >120	>120	+	+	+
4_9	<i>PMS1-YJM555</i>	sensitive	115, 26, >120, 48, 37	48	+	-	-
4_10	<i>PMS1-YJM555</i>	sensitive	>120, >120, >120, >120, >120	>120	-	+	+
4_11	<i>PMS1-YJM555</i>	resistant	112, 36, 8, 38, 3	36	-	-	+
5_1	<i>PMS1-YJM555</i>	sensitive	>120, >120, >120, >120, >120	>120	-	+	+
5_2	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	>120, 85, 30, 47, >120	85	+	-	+
5_3	<i>PMS1-YJM555</i>	sensitive	91, >120, >120, >120, >120	>120	+	+	-
5_4	<i>PMS1-YJM555</i>	resistant	1, 0, 1, 18, 63	1	-	-	+
5_5	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	65, 0, 30, 30, 20	30	-	+	+
5_6	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	58, 44, 5, 64, 0	44	-	+	-
5_7	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	25, 24, 13, 56, >120	25	+	-	+
5_8	<i>PMS1-YJM555</i>	resistant	85, 113, 71, >120, 25	85	+	-	+
5_9	<i>PMS1-S288c-R818K::HIS3</i>	sensitive	>120, >120, >120, >120, >120	>120	+	-	-

See Materials and Methods for details. EAY3234 is compatible, EAY3235 is incompatible.

The proportion of compatible and incompatible genotypes was determined after the indicated number of transfers for spore clone competitions performed in YPD and YPD+ 1.2 M NaCl (Figure 2; Figure 3). Fitness (w) values were calculated as $w = ((p_t/q_t)/(p_0/q_0))^{1/t}$, where t equals the number of generations after T transfers (7 generations per transfer; Hartl and Clark 2007), p_0 and q_0 are the number of incompatible and compatible cells, respectively at Transfer 0, and p_t and q_t are the number of incompatible and compatible cells, respectively, at the indicated Transfer. n is the number of unique competitions performed for each genotype in each transfer.