High-Resolution Mapping of Mitotic Recombination in *Saccharomyces Cerevisiae*

by

Jordan Anne St. Charles

Department of Pharmacology and Cancer Biology
Duke University

Date:_____
Approved:

_____
Thomas D. Petes, Supervisor

_____
Sue Jinks-Robertson

_____
Daniel J. Lew

_____
David M. MacAlpine

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Pharmacology and Cancer Biology in the Graduate School
of Duke University

2012

ABSTRACT

High-Resolution Mapping of Mitotic Recombination in *Saccharomyces Cerevisiae*

by

Jordan Anne St. Charles

Department of Pharmacology and Cancer Biology
Duke University

Date:_____
Approved:

_____
Thomas D. Petes, Supervisor

_____
Sue Jinks-Robertson

_____
Daniel J. Lew

_____
David M. MacAlpine

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Pharmacology and Cancer Biology in the Graduate School
of Duke University

2012

Copyright by
Jordan Anne St. Charles
2012

Abstract

Double-stranded DNA breaks are potentially lethal lesions that can be repaired in mitotic cells by either homologous recombination (HR) or non-homologous end-joining (NHEJ) pathways. In the HR pathway, the broken DNA molecule is repaired using either the sister chromatid or the homolog as a template. Mitotic recombination events involving the homolog often result in loss of heterozygosity (LOH) of markers located distal to the crossover. In humans that are heterozygous for a mutation in a tumor suppressor gene, mitotic recombination leading to LOH can be an early step in cancer development.

In my thesis research, I analyzed mitotic recombination in the yeast *Saccharomyces cerevisiae* using oligonucleotide-containing microarrays to detect LOH of single-nucleotide polymorphisms (SNPs). In analyzing cells treated with ionizing radiation, I performed the first whole-genome analysis of LOH events done in any organism (Chapter 2). I showed that irradiated cells had between two and three unselected LOH events. I also showed that crossovers were often associated with non-reciprocal exchanges of genetic information (gene conversion events) and that these conversion events were more complex than predicted by standard models of homologous recombination.

In Chapter 3, I describe my mapping of spontaneous crossovers in a 1.1 Mb region of yeast chromosome IV. This analysis is the first high-resolution mitotic recombination map of a substantial fraction (about 10%) of a eukaryotic genome. I demonstrated the existence of recombination "hotspots" and showed that some of these hotspots were homolog-specific. Two of the strongest hotspots were formed by closely-spaced inverted repeats of retrotransposons. I demonstrated that the hotspot activity was a consequence of a secondary DNA structure formed by these repeats. Additionally, the majority of spontaneous LOH events reflect DNA lesions induced in unreplicated chromosomes during G1 of the cell cycle, indicating that G1-initiated lesions threaten genome stability more than G2-initiated lesions.

In Chapter 4, I describe mitotic crossovers associated with DNA replication stress induced by hydroxyurea (HU) treatment. Surprisingly, most HU-induced crossovers had conversion tracts indicative of DNA lesions initiated in G1. Additionally, HU-induced recombination events were very significantly associated with solo delta elements, a 330 bp sequence that is repeated several hundred times in the yeast genome.

Dedication

I would like to dedicate my thesis to my family and my partner Chris Bozzelli.

Contents

Abstract iv

List of Tables xvi

List of Supporting Tables xvii

List of Figures xix

List of Abbreviations xxiii

Acknowledgements xxv

1 Introduction 1

1.1 Systems used to select mitotic crossovers in yeast 2

1.2 Sources of DNA damage that can cause mitotic recombination 7

1.2.1 Exogenous sources of mitotic recombination-inducing DNA lesions 7

1.2.2 Endogenous sources of mitotic recombination inducing lesions 8

1.2.2.1 Inverted repeats 9

1.2.2.2 Triplet repeats 13

1.2.2.3 Stimulation of recombination related to transcription and replication 14

1.3 DSB repair 17

1.3.1 Consequences of unrepaired DSBs 17

1.3.2 Pathways of DSB repair by homologous recombination 18

1.3.2.1 Synthesis-dependent strand annealing (SDSA) pathway 20

1.3.2.2 Double-strand break repair (DSBR) pathway 22

1.3.2.3 Break-induced replication (BIR) 25

1.3.3 Mechanisms of gene conversion 25

1.3.3.1 Gene conversion as a consequence of mismatch repair in a heteroduplex 28

1.3.3.2 Gene conversion by gap filling 32

1.3.4 Cell cycle and homologous recombination 32

1.3.4.1 Homologous recombination partner choice 33

1.3.4.2 Cell cycle arrests in response to DNA damage 36

1.3.4.3 Cell cycle control of DNA resection 38

1.3.5 Non-homologous end-joining (NHEJ) 39

1.4 The association of mitotic recombination with cancer 40

2 High-resolution genome-wide analysis of irradiated (UV and γ rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events 44

2.1 Introduction 45

2.1.1 Note on collaborative work 54

2.2 Materials and methods 54

2.2.1 Strains and genetic methods 54

2.2.2 SNP microarrays: design and optimization 56

2.2.3 Methods used for microarray analysis: sample preparation, hybridization conditions, and data analysis 57

2.2.4 Generation and analysis of high-throughput DNA sequencing data 58

2.3 Results 59

2.3.1 Analysis of spontaneous selected LOH events on chromosome V by SNP arrays 63

2.3.2 Analysis of LOH events by SNP arrays in cells treated with γ rays 72

2.3.3 Analysis of LOH events by SNP arrays in cells treated with UV 77

2.3.4 HTS analysis of LOH in γ ray- and UV-treated cells 78

2.3.5 HTS analysis of mutations induced in γ ray- and UV-treated cells 81

2.4 Discussion 83

2.4.1 Lengths of gene conversion tracts 83

2.4.2 Recombinogenic DNA damage 84

2.4.3 Relationship between mitotic gene conversion and crossovers 88

2.4.4 Complex gene conversion tracts and BIR events 89

2.4.5 Relationship between the level of DNA damage and the frequency of LOH events 98

2.4.6 Mutations induced by γ rays and UV 100

2.4.7 Comparison among methods of physically mapping recombination events 103

2.4.8 Summary 104

2.5 Supporting Information: Materials and Methods 105

2.5.1 SNP microarrays: specificity of hybridization 105

2.5.2 Details of methods used for microarray analysis: sample preparation, hybridization conditions, and data analysis 106

2.5.3 Generation and analysis of high-throughput sequencing data 108

2.5.4 Mutation analysis 110

2.6 Supporting Information: Discussion 111

2.6.1 Gene conversion events that are unassociated with centromere-distal LOH 111

2.6.2 Complex gene conversion tracts 113

2.6.2.1 Complex gene conversion tracts unassociated with LOH of centromere-distal markers 113

2.6.2.2 Complex conversion tracts associated with crossovers 119

2.7 Supplemental tables 121

3 High-resolution mapping of spontaneous mitotic reciprocal crossovers on the right arm of the *Saccharomyces cerevisiae* chromosome IV 146

3.1 Introduction 147

3.2 Materials and Methods 154

3.2.1 Strains and construction 154

3.2.2 Media 155

3.2.3 Assay for crossovers 155

3.2.4 Microarray methods and analysis 156

3.2.5 Microarray design 156

3.2.6 Assay of recombination activity of a recombination hotspot containing inverted repeats 157

3.2.7 Statistical analysis 158

3.3 Results 159

3.3.1 Identification and mapping of crossovers on chromosome IV 159

3.3.2 Location of crossovers and associated conversions on the right arm of chromosome IV 167

3.3.3 Statistical analysis of the distribution of crossovers and conversions 173

3.3.4 Comparison of the physical and genetic maps of the right arm of chromosome IV 177

3.3.5 The HS4 hotspot is associated with a pair of inverted Ty elements 178

3.3.6 Types of gene conversion events: inferences about the timing of the recombinogenic DSBs during the cell cycle 184

3.3.7 Association of conversion tracts with elements of chromosome sequence or structure 188

3.4 Discussion 193

3.4.1 Frequency of reciprocal crossovers 193

3.4.2 Distribution of mitotic crossovers 194

3.4.2.1 General considerations concerning the distribution of crossovers 194

3.4.2.2 Inverted repeats as hotspots for mitotic recombination 197

3.4.2.3 Mitotic recombination events that are initiated independently of inverted repeats 200

3.4.3 Properties of mitotic gene conversion tracts 203

3.4.3.1 Conversion tract lengths 203

3.4.3.2 Conversion tract length associated with the inverted repeat hotspot 204

3.4.3.3 Complex conversion tracts 205

3.4.4 DSBs initiated during G1 versus G2 of the cell cycle 206

3.5 Summary 209

3.6 Supporting information 210

3.6.1 Supporting materials and methods 210

3.6.1.1 Strain construction 210

3.6.1.2 Design of the SNP microarrays 215

3.6.1.3 Conversion tracts associated with DSBs initiated during G1 or S/G2 of the cell cycle 216

3.6.1.4 Associations of gene conversion tracts with the chromosome elements shown in Table S 3.5. 219

3.6.1.5 Analysis of mitotic recombination hotspot activities. 222

3.6.2 Supplementary Tables 224

4 Hydroxyurea-induced crossovers 259

4.1 Introduction 259

4.2 Materials and Methods 261

4.2.1 Strains 261

4.2.2 Hydroxyurea (HU) treatment 262

4.2.3 Microarray methods and analysis 262

4.2.4 Statistical analysis 263

4.3 Results 263

4.3.1 Crossovers induced by treatment of cells with HU 263

4.3.2 Comparison of HU-induced sectors and spontaneously occurring sectors 272

4.3.3 Conversion tracts associated with HU-induced crossovers 274

4.3.4 Association of HU-induced crossovers with genomic elements 278

4.4 Discussion 281

4.4.1 Sectoring rate in the presence of HU 282

4.4.2 Comparison of the reciprocal crossover distribution of spontaneous and HU-induced sectors 282

4.4.3 Unexpected presence of G1-initiated recombination events in HU-treated cells 283

4.4.4 Enrichments of genomic elements within the HU-induced conversion tracts 285

4.5 Summary 286

4.6 Supplementary Tables 288

5 Discussion and future directions 301

5.1 High-resolution genome-wide mapping of loss-of-heterozygosity events 302

5.1.1 Complex conversion tracts associated with crossovers 303

5.1.2 Homologous repair of γ-radiation-induced DSBs 307

5.2 High-resolution mapping of spontaneous chromosome IV crossovers 308

5.2.1 DSBs repaired by HR using the sister chromatid as a template 310

5.2.2 Inverted repeats can cause mitotic recombination hotspots 311

5.2.2.1 HS4: G1-specificity and conversion tract length 312

5.2.2.2 Genetic requirements for HS4 hotspot 315

5.2.2.3 Existence of HS4 in other yeast strains 317

5.2.2.4 Recombination activity of other Ty retrotransposon repeats throughout the genome 318

5.2.3 Investigation of the spontaneous recombination hotspots, HS5 and HS7 .. 319

5.3 HU-induced crossovers on chromosome IV 321

5.3.1 Comparison of HU-induced recombination with spontaneous recombination 321

5.3.2 Timing of recombination initiation lesions in HU-induced crossovers 323

5.4 Summary 324

Appendix A- Meiotic chromosome segregation in triploid strains of *Saccharomyces cerevisiae* 326

A.1 Introduction 326

A.2 Materials and Methods 333

A.2.1 Strain construction 333

A.2.2 Genetic methods and media 333

A.2.3 CGH (Comparative Genome Hybridization) microarrays 334

A.2.4 Diagnostic PCR and restriction digests 335

A.2.5 Statistical analysis 336

A.3 Results 337

A.3.1 Description of experimental methods 337

A.3.2 Analysis of meiotic segregants of the triploid MH10 338

A.3.3 Analysis of meiotic chromosome segregation in a triploid (JSC2) 345

A.3.4 Mitotic chromosome loss 357

A.3.5 Chromosome pairing during meiosis 362

A.4 Discussion 365

A.4.1 Spore viability 365

A.4.2 Patterns of chromosome loss 366

A.4.3 Chromosome pairing and recombination 372

A.5 Acknowledgments 373

A.6 Supporting Information 373

A.6.1 Supporting Materials and Methods 373

A.6.1.1 Statistical analysis of preferential co-segregation of two different homologues into sister spores 373

A.6.1.2 Statistical analysis of preferential co-segregation of three different homologues into sister spores 374

A.6.1.3 Statistical analysis of the non-random segregation of disomic chromosomes in JSC2 based on whether the chromosomes are derived from S288c, JAY291, or YJM789 375

A.6.2 Supporting tables 377

References 392

Biography 424

List of Tables

Table 2.1 Comparison of mapping methods for four spontaneous crossovers on chromosome V. 65

Table 2.2 Summary of all crossovers diagnosed by SNP microarray 67

Table 2.3 Summary of all unselected conversion events unassociated with LOH or BIR events as diagnosed by SNP microarrays. 76

Table 2.4 Comparison between recombination events detected by high-throughput sequencing or by SNP microarrays. 81

Table 3.1 Chromosome elements that are over-represented in crossover-associated gene conversion tracts in JSC25. 191

Table 4.1 Element enrichment in HU-induced crossover-associated conversion tracts. 280

Table A.1 CGH analysis of disomy and monosomy in spores derived from the triploid MH10 342

Table A.2 Analysis of disomy and monosomy in spores derived from the triploid JSC2-1 350

List of Supporting Tables

Table S 2.1 Classes of conversion tracts associated with spontaneous recombination events and events stimulated by UV or γ rays. 121

Table S 2.2 Classes of recombination events that were observed by SNP arrays but that were altered when examined by high-throughput DNA sequencing. 129

Table S 2.3 Mapping spontaneous mitotic recombination events by SNP microarrays . 131

Table S 2.4 SNP microarray analysis of recombination events induced by γ rays 133

Table S 2.5 Analysis of γ ray-induced recombination events by high-throughput DNA sequencing 136

Table S 2.6 Analysis of UV-induced recombination events by SNP microarrays 137

Table S 2.7 Analysis of UV events by HTS 141

Table S 2.8 Mutations induced by γ ray or UV treatment 145

Table S 3.1 Strain list and strain constructions. 224

Table S 3.2 Names and sequences of primers used in strain constructions and genome analysis. 228

Table S 3.3 Depictions of conversion tracts associated with reciprocal crossovers. 230

Table S 3.4 SGD coordinates for crossover and conversion transitions on chromosome IV. 240

Table S 3.5 Chromosome elements examined for their representation in mitotic gene conversion tracts. 252

Table S 3.6 Chromosome elements that are under-represented in crossover-associated gene conversion tracts in JSC25. 254

Table S 3.7 Coordinates used to represent gene conversion tracts as single points. 255

Table S 4.1 Summary of conversion tracts associated with reciprocal crossovers. 288

Table S 4.2 SGD coordinates for HU-induced crossover and conversion transitions on chromosome IV. 294

Table S 4.3 Coordinates used to represent HU-induced crossover-associated gene conversion tracts as single points. 299

Table S A.1. Strain genotypes and constructions for triploid study 377

Table S A.2 Primers and restriction enzymes used to distinguish centromeric sequences derived from strains S288c, JAY291, and YJM789 379

Table S A.3 Primers and restriction enzymes used to distinguish telomeric sequences of chromosomes IV, V, IX, and XIII derived from strains S288c, JAY291, and YJM789 382

Table S A.4 Patterns of spore viability in the triploid strains JSC2 and MH10 384

Table S A.5 CGH analysis of disomy and monosomy in spores derived from six tetrads of the triploid JSC2 385

Table S A.6 Patterns of marker exchange on chromosomes IV, V, IX, and XIII in spores derived from the triploid JSC2 387

Table S A.7 Statistical analysis of loss of disomic chromosomes in spores derived from MH10 389

Table S A.8 Statistical analysis of loss of disomic chromosomes in spores derived from JSC2 391

List of Figures

Figure 1.1. Systems used to identify mitotic crossovers on yeast chromosome V.....6

Figure 1.2. Secondary structures formed from inverted repeats.....10

Figure 1.3. Common pathways of DSB repair by HR. 21

Figure 1.4. Conversion tracts resulting from homologous recombination. 24

Figure 1.5. Two mechanisms for gene conversion. 27

Figure 1.6. Hypothetical G1 crossover between homologs.....36

Figure 2.1. Pathways of DSB repair by homologous recombination. 46

Figure 2.2. Genetic system used to select reciprocal crossovers on the left arm of chromosome V.....49

Figure 2.3. Production of hybrid conversion tract by the repair of two broken sister chromatids. 53

Figure 2.4. Analysis of a spontaneous reciprocal crossover (PG311-2A) on the left arm of chromosome V by SNP microarrays. 60

Figure 2.5. Mapping of crossovers and associated conversion events on the left arm of chromosome V in untreated cells, and cells treated with γ rays or UV by SNP microarrays. 69

Figure 2.6. Genomic locations of unselected recombination events and *de novo* mutations in untreated cells and in cells treated with UV or γ rays as determined by SNP microarrays and HTS.....73

Figure 2.7. SNP array analysis of a gene conversion event unassociated with crossover.75

Figure 2.8. Analysis of the same recombination event by both SNP arrays and HTS. 79

Figure 2.9. Description of the Class J8 event. 91

Figure 2.10. Description of the Class J9 event. 93

Figure 2.11. One mechanism to explain Class G1.....95

Figure 2.12 Description of the L2 and L3 events (listed in Table S 2.1). 97

Figure 2.13. Generation of Class F4 by two different mechanisms. 116

Figure 2.14. Alternate mechanism for generation of Class G1. 118

Figure 3.1. Patterns of gene conversion resulting from the repair of G1- or G2-generated DSBs. 151

Figure 3.2. Crossover screening system 159

Figure 3.3. Mapping of a crossover with an associated 3:1 conversion event by SNP microarrays. 161

Figure 3.4. Mapping of a crossover with an associated 3:1/4:0/3:1 conversion event by SNP microarrays. The depiction of this event is the same as in Figure 3.3. 165

Figure 3.5. Depiction of conversion tracts and crossovers. 167

Figure 3.6. Map locations of crossovers and associated gene conversion events in JSC25 along the right arm of chromosome IV. 170

Figure 3.7. Summary of conversion tracts associated with crossovers along the right arm of chromosome IV. 172

Figure 3.8. Distribution of JSC25 crossovers into six intervals along the right arm of chromosome IV. 174

Figure 3.9. Conversion tract lengths along the right arm of chromosome IV. 176

Figure 3.10. Comparison of physical and genetic mitotic recombination maps. 177

Figure 3.11. Inverted pair of Ty elements at SGD coordinate 980 kb are present on the W303a-, but not the YJM789-derived homolog 178

Figure 3.12. Potential "hairpin" DNA structures that could be formed by inverted repeats. 181

Figure 3.13. Deletion analysis of the HS4 hotspot. 183

Figure 3.14. Generation of a complex conversion tract by "patchy" repair of mismatches in a heteroduplexes. 186

Figure 3.15. Comparison of mitotic and meiotic recombination maps on the right arm of chromosome IV. 196

Figure 3.16. Models for a double-strand break occurring at an inverted repeat during G1 of the cell cycle..... 198

Figure 3.17. Homologous recombination repair choices depending on the timing of the initiating DSB lesion. 209

Figure 4.1. Summary of crossovers induced by HU treatment. 267

Figure 4.2. Summary of conversion tracts associated with HU-induced crossovers along the right arm of chromosome IV..... 270

Figure 4.3. Number of HU-induced crossovers located in bins of the right arm of chromosome IV. 271

Figure 4.4. Comparison of the location of recombination events for HU-induced and spontaneous crossovers..... 274

Figure 4.5. Comparison of conversion tract types between the spontaneous crossovers and HU-induced crossovers. 276

Figure 4.6. Conversion tract lengths associated with spontaneous and HU-induced crossovers. 277

Figure 5.1. Proposed modifications of the YJM789-derived homolog in the region containing HS4. 314

Figure A.1. Expected meiotic segregation pattern of chromosomes of a triploid..... 329

Figure A.2. CGH microarray analysis of DNA derived from four spores of a tetrad derived from the triploid MH10 339

Figure A.3. Use of SNPs to determine the number and identity of homologous chromosome homologues within a JSC2-derived spore 346

Figure A.4. Analysis of a JSC2 tetrad by CGH microarrays 348

Figure A.5. Analysis of SNPs in colonies derived from a single meiotic spore culture (JSC2-1-8d) 353

Figure A.6. Comparison of the detection of disomes in spores derived from the triploids MH10 and JSC2 356

Figure A.7. Mitotic loss of disomic chromosomes from spore JSC2-1-8c..... 358

Figure A.8. PCR method of detecting very small amounts of JAY291-derived chromosome I in strains with S288c- or YJM789-derived chromosomes..... 360

Figure A.9. Chromosome pairing and recombination during meiosis in a triploid..... 364

List of Abbreviations

ARS: autonomously replicating sequence

BIR: break-induced replication

bp: base pair

CDK: cyclin dependent kinase

CEN: centromere

CO: crossover

dHJ: double Holliday junction

dNTPs: deoxyribnucleoside triphosphates

DSB: double-strand break

DSBR: double-strand break repair

dsDNA: double-stranded DNA

HJ: Holliday junction

HR: homologous recombination

HU: hydroxyurea

IR: ionizing radiation

kb: kilobase pair

LOH: loss of heterozygosity

MMR: mismatch repair

NCO: noncrossover

NHEJ: nonhomologous end-joining

PMS: post-meiotic segregation

PSS: post-switching segregation

RFLP: restriction-fragment length polymorphism

RNR: ribonucleotide reductase

RPA: replication protein A

RSZ: replication slow zone

SDSA: synthesis-dependent strand annealing

SSA: single strand annealing

ssDNA: single-stranded DNA

TAR: transcription associated recombination

UV: ultraviolet

Acknowledgements

I would like to thank my advisor, Tom Petes, for letting me join his lab and spending so much time training me as a scientist. Tom is a great, patient, and kind advisor who always cares what happens to his students and lab members. I would also like to thank him for spending many arduous hours reading my manuscripts (including, those, deficient, in, commas).

I would like to thank all the members of my committee: Sue Jinks-Robertson, Daniel Lew, David MacAlpine, and Sally York, for providing critical advice and support during my project. I would also like to thank Fred Dietrich for spending a significant amount of time showing me how to design a microarray. I would also like to thank the members of the SJR lab for valuable conversation and comments on my projects.

I would also like to thank all the members of the Petes Lab. Pat Greenwell, Anne Casper, and Lucas Argueso kindly spent hours demonstrating many lab techniques when I first joined the lab. Monica Hamilton and Eunice Yim were very productive and helpful during my various projects. Wei Song, Yi Yin, Sabrina Andersen, Karen O'Connell, and Jen McCulley were my partners in "monitoring the internet" and research-related conversations. Becky Pendergraft provided the best cupcakes, which fueled much of my thesis-work. The entire Petes Lab has been the most supportive

group of people, who all made it both possible and enjoyable to complete my thesis-work.

I would like to thank my whole family for being supportive and patient during my time in grad school. My partner, Chris Bozzelli, has been endlessly supportive during both the lows and highs of grad school. My sister, Whitney, has provided the most lively internet chats.

1 Introduction

Throughout cellular proliferation, DNA is subjected to various stresses that can cause DNA damage, such as double-strand breaks (DSBs). DSBs are lethal lesions if left unrepaired (FAIRHEAD and DUJON 1993). In yeast, DSBs are often repaired by homologous recombination (HR), which uses a homologous DNA molecule (such as the sister chromatid or the homolog) as a template for repair (PÂQUES and HABER 1999). HR between the broken molecule and the donor template can produce gene conversions without crossovers or gene conversions associated with crossovers. A crossover occurs when two DNA molecules exchange sequences at the location of the initiating DSB. Mitotic crossovers were first observed in *Drosophila melanogaster* by Curt Stern (STERN 1936). Despite its important role in maintaining stability of the genome, HR can also have negative consequences. If a crossover occurs between two homologs that are heterozygous for genes along their arms, the daughter cells can receive two chromosomes that are homozygous for all sequences centromere-distal to the crossover. If a cell is heterozygous for a mutant gene and a wild-type gene, loss-of-heterozygosity (LOH) can cause loss of the wild-type gene. In a human cell that is heterozygous for a mutation in a tumor suppressor gene, LOH produced by mitotic crossing over can be an early event in tumorigenesis (KNUDSON 2001).

Despite the importance of mitotic recombination, little is known about how spontaneous mitotic crossovers are initiated. In my thesis research, I mapped at high-

resolution both spontaneous and DNA-damage-induced crossovers in *Saccharomyces cerevisiae* using single-nucleotide polymorphism (SNP) microarrays. In this introductory chapter of the thesis, I describe experimental systems that are used to study mitotic crossovers (1.1), sources of DNA damage that can stimulate mitotic recombination (1.2), double-strand break repair (1.3), and the association of mitotic recombination with cancer (1.4). Chapter 2 describes experiments in which I examined genome-wide LOH induced by DNA damage from ionizing radiation. Chapter 3 describes the analysis of spontaneous reciprocal crossovers on the right arm of yeast chromosome IV, and the analysis of a mitotic recombination hotspot resulting from an inverted pair of retrotransposons. Chapter 4 describes the analysis of reciprocal crossovers induced by hydroxyurea, a drug that affects nucleotide pools. The final chapter summarizes the findings of Chapters 2 through 4 and discusses possible extensions of this research. Finally, in the Appendix of the thesis, I describe my characterization of the meiotic products of triploid yeast cells. Although this work culminated in a paper in *Genetics*, it is not germane to the main topic of my thesis.

1.1 Systems used to select mitotic crossovers in yeast

Homologous recombination can occur during both meiosis and mitosis. In most organisms, meiotic recombination is essential to ensure proper chromosome segregation (PAGE and HAWLEY 2003). In yeast, as in most other organisms, there is at least one meiotic crossover per chromosome arm (MANCERA *et al.* 2008). In comparison, mitotic

crossovers in yeast occur at a rate that is 10^4 to 10^5-fold less than the meiotic rate (BARBERA and PETES 2006). For this reason, the study of spontaneous mitotic recombination requires an experimental system to select or screen for rare events. Alternatively, some mitotic recombination studies employ recombinogenic agents such as ultraviolet light or γ rays to elevate recombination rates to a level that a selective system is unnecessary.

As described above, one consequence of mitotic crossovers is LOH of markers located centromere-distal to the crossover site. Although the two daughter cells resulting from the crossover will have reciprocal patterns of LOH, the most commonly used system to select crossovers selects for only one of these two products (Figure 1.1A). The diploid strain used in this system (HARTWELL and SMITH 1985) is heterozygous for a mutation in *CAN1*, a marker located near the left end of chromosome V. The *CAN1* gene encodes a protein required for transport of arginine (AHMAD and BUSSEY 1986); strains containing the wild-type gene are poisoned by the drug canavanine. On the homolog containing the wild-type *CAN1* gene, the diploid also has the wild-type *HOM3* gene, a gene required for production of threonine and methionine. *CAN1* and *HOM3* are located on opposite arms of chromosome V (Figure 1.1A). The heterozygous starting diploid is sensitive to canavanine (Can^S) and prototrophic for methionine (Met^+). A crossover between the centromere of chromosome V and the heterozygous *can1*/*CAN1* markers can produce one cell that is homozygous for the mutant *can1* gene and, therefore, Can^R,

as well as a Can^S cell that is homozygous for the wild-type *CAN1* gene. Thus, although one of the expected products of a crossover can be selected on solid medium containing canavanine, the other cannot. Inferences about the mechanism of the recombination event, therefore, are indirect. A Can^R derivative can also be produced by loss of the homolog with the *CAN1* gene. Such derivatives are Met$^-$ and easily distinguishable from the Can^R cells that reflect crossovers.

A system that allows selection of both daughter products of a reciprocal crossover is shown in Figure 1.1B (BARBERA and PETES 2006). The starting diploid of this system is homozygous for the ochre-suppressible *ade2-1* mutation. In the absence of an ochre suppressor, these cells are Ade$^-$ and form red colonies. The parental diploid has the ochre-suppressible *can1-100* mutation at the native *CAN1* locus on one copy of chromosome V and the ochre suppressor gene, *SUP4-o,* inserted allelically on the other homolog. The diploid, therefore, is Can^S and forms pink colonies; the colonies are pink rather than white because the suppression of the *ade2-1* mutation by the suppressor is only partial. If there is a crossover between the *can1-100*/*SUP4-o* markers and the centromere, two Can^R cells will be produced. One daughter cell will be Can^R because it lacks the *SUP4-o* marker and the other will be Can^R because it lacks the *can1-100* marker. If the crossover event occurs at the same time that the cell is plated on canavanine-containing medium, a red/white sectored colony will be formed, reflecting the absence (red sector) and presence (white sector) of the *SUP4-o* gene. The other markers shown in

Figure 1.1B (*HIS3*, *LEU2*, *HYG*, and *KAN*) are used to distinguish crossovers from chromosome losses and other types of genetic events (BARBERA and PETES 2006).

It should be pointed out that, following a crossover, there are two possible patterns of chromosome disjunction, only one of which leads to LOH events. If each of the recombined chromosomes segregates with an unrecombined chromosome (as shown in Figure 1.1), LOH events are observed. If one cell receives both recombined chromosomes and the other receives the unrecombined chromosomes, however, no LOH is observed. CHUA and JINKS-ROBERTSON (1991) showed that these two patterns are equally frequent. Thus, the frequency of crossovers is twice the frequency of sectored colonies for the systems shown in Figure 1.1B.

The system shown in Figure 1.1C is the same as that shown in Figure 1.1B except that homologs are derived from two sequence-diverged haploid strains, YJM789 and W303a (LEE *et al.* 2009; LEE and PETES 2010). There are about 55,000 single-nucleotide polymorphisms (SNPs) that are heterozygous in the resulting diploid (WEI *et al.* 2007). Following selection of Can^R red/white sectored colonies, LEE *et al.* (2009) mapped the position of the crossover by using a PCR-based method to determine whether SNPs located on the left arm of chromosome V were heterozygous or homozygous. Using a related approach (described in Chapters 2 and 3), I was able to map mitotic crossovers on the right arm of chromosome IV and unselected recombination events throughout the genome.

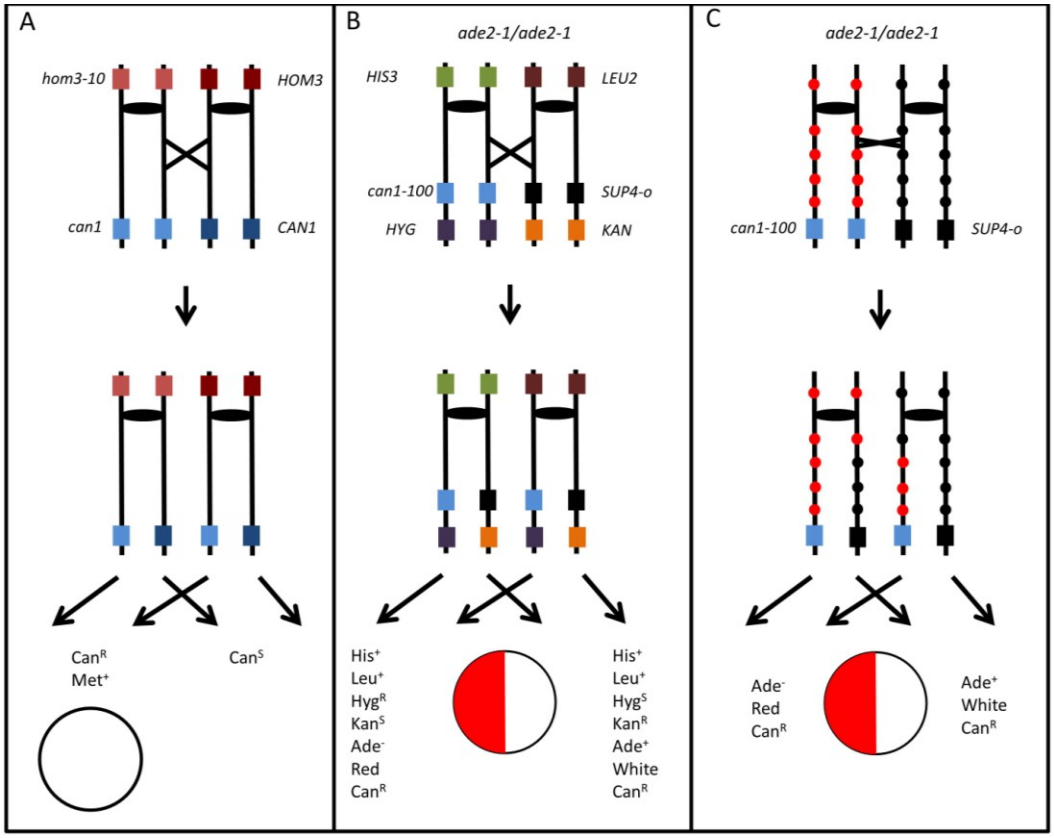


Figure 1.1. Systems used to identify mitotic crossovers on yeast chromosome V.

For all three parts of the figure, the crossover is shown as occurring between homologous chromosomes in G2. The black ovals represent the centromeres of the chromosomes and the large Xs mark the crossovers.

A. "Classic" crossover assay (HARTWELL and SMITH 1985). The starting strain is heterozygous for the recessive *can1* mutation and, therefore, sensitive to canavanine. Following a mitotic crossover between the *can1* gene and the centromere, one daughter cell will be homozygous for the mutant *can1* allele and able to grow on solid medium containing canavanine. The other daughter will be homozygous for the wild-type allele and will have the same canavanine-sensitive phenotype as the progenitor diploid.

B. System that allows for the selection of <u>both</u> daughter products of a reciprocal crossover using the *can1-100/SUP4-o* markers (BARBERA and PETES 2006). The rationale for this system is described in the text.

C. System that allows the selection of both daughter products of a crossover (as in Figure 1.1B), as well as the mapping of the position of the crossover (LEE *et al.* 2009). SNPs that distinguish the two homologs are shown as red and black circles.

1.2 Sources of DNA damage that can cause mitotic recombination

It is likely that mitotic recombination events are initiated by a DNA lesion, either a double-stranded DNA break (DSB) or a single-stranded DNA nick. Because the frequency of mitotic recombination is very low, there is little direct physical evidence about the nature of the DNA lesion that initiates spontaneous events. The assumptions about the nature of spontaneous lesions is based to a large extent on the types of exogenous agents that stimulate crossovers and the types of mutations that result in elevated recombination as described below.

1.2.1 Exogenous sources of mitotic recombination-inducing DNA lesions

Agents that produce DSBs result in elevated mitotic recombination in yeast and other organisms. Mitotic crossovers and gene conversion events are stimulated by X-rays and other types of ionizing radiation (ARGUESO *et al.* 2008; NAKAI and MORTIMER 1969; ST CHARLES *et al.* 2012; SYMINGTON 2002). The recombinogenic properties of DSBs have also been demonstrated using endonucleases that have large recognition sites, such as HO and I-*Sce*I (PÂQUES and HABER 1999). Meiotic recombination is induced by Spo11p-catalyzed DSBs throughout the genome (KEENEY *et al.* 1997; SUN *et al.* 1989). Spo11p-induced DSBs can also cause mitotic recombination in a return-to-growth assay

in which meiosis is initiated by incubating cells in sporulation media; cells that are returned to rich media before commitment to spore formation have very elevated levels of recombination (HONIGBERG and ESPOSITO 1994; SIMCHEN 2009).

The evidence that single-stranded DNA nicks can stimulate mitotic recombination is less direct than for DSBs. Ultraviolet (UV) light is a potent recombinogen in yeast (FABRE 1978). Excision of UV-induced dimers would be expected to result in single-stranded DNA nicks (FRIEDBERG 2006). It is currently unclear whether the nicked DNA engages in recombination or whether replication of nicked DNA results in recombinogenic DSBs (GALLI and SCHIESTL 1999). STRATHERN *et al.* (1991) showed that a single-stranded nick resulting from expression of the gene II from bacteriophage f1 in yeast was recombinogenic, but, as with the UV-induced damage, it is unclear whether the recombination-stimulating lesion was the nick or the DSB resulting from replication of the nicked chromosome. In summary, there is clear evidence that a DSB can stimulate mitotic, as well as meiotic recombination. The evidence concerning the recombinogenic properties of single-stranded DNA nicks is currently not definitive.

1.2.2 Endogenous sources of mitotic recombination inducing lesions

In addition to exogenous sources of recombinogenic DNA damage, the cellular environment can cause DNA damage that leads to recombination (AGUILERA and GOMEZ-GONZALEZ 2008). Certain DNA sequences are capable of forming secondary structures that are processed to recombinogenic lesions. Additionally, cellular processes

such as transcription and replication can cause DSBs. All of these mechanisms will be described in detail below.

1.2.2.1 Inverted repeats

Inverted repeats can be oriented in a head-to-head or tail-to-tail orientation such that, in addition to the Watson and Crick strands being able to pair with each other, the inverted repeats within each strand can also pair (GORDENIN and RESNICK 1998). This intrastrand pairing can produce two different types of secondary structures: "hairpins" involving intrastrand pairing of only one of the two strands (Figure 1.2A; (NAG and PETES 1991) and cruciform structures that represent double hairpins (Figure 1.2B) (LILLEY 1980). Both of these structures are thought to lead to DSBs because they are recognized as DNA damage and are processed by cellular endonucleases (GORDENIN and RESNICK 1998). Hairpins could potentially form during replication of the lagging strand and cause a replication fork impediment (GORDENIN and RESNICK 1998). Alternatively, if a single-stranded gap is formed on a DNA molecule that contains an inverted repeat, the remaining strand could form a hairpin; in Chapter 3, I will present evidence favoring this second model. If the tip of the hairpin shown in Figure 1.2A is cleaved by an endonuclease, a DSB would be produced. If the spacer between the inverted repeats is very short (<10 bp), cruciforms can form in double-stranded DNA (ALLERS and LEACH 1995; KOGO *et al.* 2007; SINDEN *et al.* 1991). Since these structures

resemble Holliday junctions (HJs), they are likely to be cleaved by HJ resolvases, resulting in a DSB (LOBACHEV *et al.* 2007).

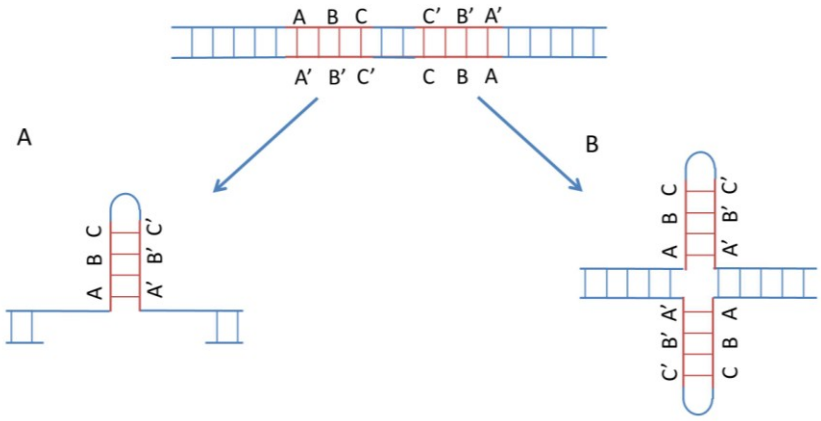


Figure 1.2. Secondary structures formed from inverted repeats.

The repeats are represented in red and single-copy sequences in blue. The blue region between the two red repeats represents the spacer.

A. If a gap on one strand includes the inverted repeats, a hairpin structure can be formed on the remaining single strand.

B. If the spacer sequences between the two repeats is small, a cruciform structure (hairpin structures on both strands) can be formed.

Both naturally-occurring and introduced inverted repeats have been shown to induce instability in yeast (CASPER *et al.* 2009; GORDENIN *et al.* 1993; LEMOINE *et al.* 2008; LEMOINE *et al.* 2005; LOBACHEV *et al.* 2002; LOBACHEV *et al.* 2000). GORDENIN *et al.* (1993) examined the properties of 1.5 kb inverted repeats (derived from the bacterial transposon Tn5) flanking a wild-type *URA3* gene. They found that Ura$^-$ derivatives of the starting strain usually contained deletions between the Tn5 inverted repeats. In wild-

type strains, the rate of loss of *URA3* was low, but it was greatly elevated in a strain that had a mutation in *POL3*, the gene encoding DNA polymerase delta. This construction also resulted in elevated frequencies of intragenic gene conversion. They observed that reducing the length of the spacer between the repeats elevated the frequency of intragenic gene conversions in wild-type strains. These results argue that the recombinogenic effects of inverted repeats are substantially influenced by the distance between the repeats.

LOBACHEV *et al.* (2000) also examined the effects of inverted repeats of the human Alu sequence in yeast. These sequences were inserted into the *LYS2* gene generating a mutation, and they determined the rate of intragenic recombination with a *lys2* gene with a mutation located elsewhere in the gene. They examined the effects of altering the distance between the inverted Alu sequences and the effects of altering sequence homology between the repeats. The inverted Alu constructs had spacers ranging from 12 – 100 bp and sequence homology between the repeats ranged from 75 – 100%. The construct with the 12 bp spacer and 100% sequence homology induced recombination in a wild-type strain 6000-fold over a construct with direct repeats. The ability to induce recombination decreased as the spacer between the repeats was increased. The ability to induce recombination also decreased as the sequence homology decreased, but all repeats with at least 86% identity were able to increase recombination relative to the control. The instability associated with increased sequence similarity

between the repeats and decreased distance between the repeats probably reflects the likelihood of secondary structure formation. Inverted repeats in *Escherichia coli* have been shown to form cruciforms when the spacer is less than 10 bp but do not form this secondary structure very often if the spacer is greater than 20 bp (ALLERS and LEACH 1995; KOGO *et al.* 2007; SINDEN *et al.* 1991).

LOBACHEV *et al.* (2002) showed that, in the absence of Mre11p, inverted repeats were incapable of inducing allelic mitotic recombination. Their results indicated that a cruciform structure formed by the inverted repeat is processed by an as yet unknown nuclease to form two nicked hairpin-capped ends. Mre11p is then required to process these ends. If the nicked hairpin-capped ends are not processed, and the nicks are ligated instead, they can cause further instability in the genome. This instability is caused by replication of the two hairpin-capped chromosome fragments, resulting in palindromic dicentric chromosomes and palindromic acentric fragments (KIM *et al.* 2008; LOBACHEV *et al.* 2002).

LEMOINE *et al.* (2005) found that Ty retrotransposons on chromosome III in a head-to-head conformation with a 283 bp spacer were a hotspot for chromosome rearrangements in strains with perturbed DNA replication. This hotspot (termed "FS2", for fragile site 2) had an elevated frequency of chromosome rearrangements in strains with low levels of DNA polymerase alpha (LEMOINE *et al.* 2005) or DNA polymerase delta (LEMOINE *et al.* 2008). Most of the chromosome rearrangements were translocations

between the Ty elements of FS2 and Ty elements on other chromosomes. In strains with low levels of DNA polymerase alpha, DSBs at FS2 were detectable (LEMOINE *et al.* 2005). Because the spacer between the Ty repeats is so large, formation of a cruciform structure at FS2 is highly unlikely, and the instability of FS2 is more likely to be due to formation of a hairpin structure. LEMOINE *et al.* (2005) suggested that, when replication is perturbed by lowering the levels of Pol1p, the intervals between Okazaki fragments are large enough to span the distance of the spacer between the inverted repeats, allowing hairpin formation.

1.2.2.2 Triplet repeats

Triplet repeats have also been associated with genome instability in yeast and other organisms, including humans. The CGG/CCG, CAG/CTG, GAA/TTC, and GCN/NGC repeats have all been shown to be associated with DSBs or elevated levels of recombination in strains with perturbed DNA replication (MIRKIN 2007). By definition, therefore, they represent fragile sites. Various triplet repeats have been shown to form secondary structures *in vitro* (MIRKIN 2007), and the instability of these repeats in yeast is affected by mismatch repair (MMR) (MOORE *et al.* 1999). For some of the repeats, their instability is dependent on their orientation relative to the nearest origin of replication (SAMADASHWILY *et al.* 1997). This result has been interpreted as indicating that the instability of repeats reflect their tendency to form secondary structures on the lagging strand during replication, followed by processing of the secondary structures to

generate the recombinogenic DNA lesion (AGUILERA and GOMEZ-GONZALEZ 2008). It should be pointed out, however, that some triplet repeats (for example, GAA/TTC) generate recombinogenic lesions by a mechanism that is independent of their orientation with respect to the replication origin (TANG *et al.* 2011).

1.2.2.3 Stimulation of recombination related to transcription and replication

In addition to secondary DNA structures causing recombination, various endogenous cellular processes can induce recombination. High levels of transcription have been observed to induce recombination, a phenomenon called "transcription-associated recombination" (TAR) (KEIL and ROEDER 1984; KIM and JINKS-ROBERTSON 2012; NEVO-CASPI and KUPIEC 1994; THOMAS and ROTHSTEIN 1989). In yeast, TAR was first observed with *HOT1,* one of the ribosomal DNA promoters (KEIL and ROEDER 1984). Mutations of the *HOT1* promoter that decreased transcription also decreased recombination (HUANG and KEIL 1995; STEWART and ROEDER 1989). The *HOT1* promoter is specific for RNA polymerase I, but elevated transcripts derived from RNA polymerase II promoters also induce mitotic recombination (SAXE *et al.* 2000; THOMAS and ROTHSTEIN 1989). The elevated rates of recombination induced by transcription require the same proteins as those induced by DSBs (GONZALEZ-BARRERA *et al.* 2002), indicating that the transcription-induced DNA lesions are likely to be DSBs.

There are a variety of possible explanations for TAR. One possibility is that TAR reflects the transcription-associated production of single-stranded DNA (ssDNA) that is

more susceptible to damaging agents than dsDNA (LINDAHL 1993). Consistent with this explanation, the non-transcribed strand during transcription is more often mutated than the transcribed strand (AGUILERA and GOMEZ-GONZALEZ 2008). DNA-RNA hybrids (R-loops) form during transcription, and if stabilized, they can extend the length of time that the non-transcribed strand is single-stranded (HUERTAS and AGUILERA 2003). R-loops have been shown to be associated with genome instability and DSB formation (LI and MANLEY 2005). Another possible mechanism of TAR is that the build-up of both positive supercoils in front of the transcription complex and negative supercoils behind the complex results in elevated frequencies of DNA lesions induced by topoisomerases or other endonucleases (KIM and JINKS-ROBERTSON 2012).

Some TAR events likely reflect DSBs generated by collisions between the transcription and replication machinery or other mechanisms that involve both transcription and replication (KIM and JINKS-ROBERTSON 2012). When the direction of transcription opposes the direction of the replication fork, it causes a replication fork pause that can induce recombination (PRADO and AGUILERA 2005). In the rDNA genes of yeast, collisions between transcription and replication forks are prevented by Fob1p, a protein that arrests replication forks in an orientation-dependent manner (BREWER and FANGMAN 1988). In addition, in many organisms, including *E. coli*, highly-transcribed genes are oriented in a way that prevents collisions between the transcription machinery and the replication forks (ELLWOOD and NOMURA 1982). It is likely that TAR events

reflect more than one type of recombination-associated DNA lesion. Replication through a DNA nick results in a DSB on one chromatid following DNA replication (CORTES-LEDESMA and AGUILERA 2006; KUZMINOV 2001). In addition, replication forks that are stalled as a consequence of a collision with the transcription machinery are likely associated with DSB formation (COURCELLE *et al.* 2003; IVESSA *et al.* 2003). It is also possible that torsional stress induced by high levels of transcription may also cause replication fork reversal and "chicken-foot" structures that can be cleaved to form DSBs by HJ-like resolvases (POSTOW *et al.* 2001; RUDOLPH *et al.* 2007).

In addition to DNA damage resulting from the interaction of replication forks with the transcription machinery, there are replication-associated DNA lesions caused by drugs or mutations that interfere with normal replication. Aphidicolin, an inhibitor of DNA polymerase alpha, causes chromosome breaks in specific regions of the mammalian genome (GLOVER *et al.* 1984). These fragile sites were originally defined as regions of the chromosome that had breaks or gaps on metaphase spreads (DURKIN and GLOVER 2007; MAGENIS *et al.* 1970). Fragile sites can be either rare and segregate in Mendelian fashion, as in the fragile site associated with the fragile X syndrome (JACOBS *et al.* 1980; KREMER *et al.* 1991; SUTHERLAND 1977), or common to all individuals (ARLT *et al.* 2003). In human cells, the stability of common fragile sites is affected by the *ATR* gene (CASPER *et al.* 2002). CASPER *et al.* (2002) showed that cells that have inactivated *ATR* and are treated with aphidicolin have 20-fold more chromosome gaps and breaks than wild-

type cells treated with aphidicolin. Cells with inactivated ATR accumulate breaks at fragile sites even in the absence of aphidicolin and other replication-perturbing drugs. The yeast *ATR* homolog, *MEC1*, affects fragile sites in yeast (CHA and KLECKNER 2002). CHA and KLECKNER (2002) showed that the absence of *MEC1* leads to genome-wide replication fork stalling and chromosome breakage at specific sites. These specific sites were shown to be replication slow zones (RSZs), however, not all RSZs show chromosome breakage in the absence of *MEC1*. They also mapped a few of these breakage regions on chromosome III. One of these regions was later shown to be the yeast fragile site, FS2, which colocalizes with inverted Ty retrotransposon repeats as mentioned above (LEMOINE *et al.* 2005).

1.3 DSB repair

As described above, there is clear evidence that DSBs can stimulate mitotic recombination. Before discussing the mechanisms of DSB repair, I will briefly discuss the consequences of failure to repair DSBs.

1.3.1 Consequences of unrepaired DSBs

If a DSB is unrepaired in a haploid cell, it eventually leads to cell death (FAIRHEAD and DUJON 1993; MALONE and ESPOSITO 1980). A DSB can be difficult to repair by homologous recombination if the break in a region that lacks an intact homologous template or because the strain is mutated for a gene (such as *RAD52*) required for homologous recombination. Haploid yeast strains with a wild-type *HO*

gene and a *rad52* mutation are inviable (MALONE and ESPOSITO 1980; WEIFFENBACH and HABER 1981), since the HO-induced DSB at the mating-type locus cannot be repaired. In diploid cells, an unrepairable DSB is not lethal, but leads to loss of the chromosome that has the DSB (KRAMER and HABER 1993; SANDELL and ZAKIAN 1993). In *rad52* diploid strains, there is progressive chromosome loss, eventually resulting in haploidy (SONG and PETES 2012).

There are likely to be several reasons why an unrepaired DSB is lethal in haploid cells. First, nucleolytic processing of the broken ends may lead to loss of essential genes (FAIRHEAD and DUJON 1993; WEIFFENBACH and HABER 1981). Second, since both products of a broken chromosome lack a telomere, and one product also lacks a centromere, these fragments will not be stably maintained. Finally, in several studies, it has been shown that DSBs in plasmids or dispensable chromosomes (yeast artificial chromosomes) can cause lethality (BENNETT *et al.* 1993; BENNETT *et al.* 1997; BENNETT *et al.* 1996). Although the reason for this lethality is not clear, it may be related to the cell cycle checkpoints.

1.3.2 Pathways of DSB repair by homologous recombination

DSBs in yeast can be repaired by either homologous recombination (HR) or by non-homologous end-joining (NHEJ) (SYMINGTON and GAUTIER 2011). Since my thesis is concerned with HR, I will discuss NHEJ only briefly following the discussion of HR. Homologous recombination is the main pathway for the repair of DSBs in yeast during

G2 of the cell cycle (BRANZEI and FOIANI 2008). HR involves the use of an intact homologous DNA template for the repair of DSBs (PÂQUES and HABER 1999). The main HR pathways are shown in Figure 1.3 including canonical double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), and break-induced replication (BIR) (PÂQUES and HABER 1999). In wild-type cells, most HR events involve either the DSBR pathway for crossovers or the SDSA pathway for gene conversions unassociated with crossovers rather than BIR (BARBERA and PETES 2006; MCMURRAY and GOTTSCHLING 2003).

The first step of DSB repair for all HR pathways shown in Figure 1.3 is 5′-to-3′ resection of the DSB ends to generate 3′ overhanging ends (MIMITOU and SYMINGTON 2011). Following DSB formation, the MRX complex (Mre11p, Rad50p, and Xrs2p) is recruited to the broken ends, and Mre11p and Sae2p catalyze the initial removal of 50-100 bp from the 5′ ends (MIMITOU and SYMINGTON 2008; SHIM *et al.* 2010; ZHU *et al.* 2008). Following this initial step, resection occurs through two different pathways: one dependent on Exo1p and a second dependent on the Sgs1p/Top3p/Rmi1p complex and Dna2p (GRAVEL *et al.* 2008; MIMITOU and SYMINGTON 2008; ZHU *et al.* 2008). Following resection, the resulting single-stranded 3′ ends are coated with replication protein A (RPA). In a Rad52p-promoted step, RPA is displaced by the Rad51p strand-transfer protein (MIYAZAKI *et al.* 2004; SUNG 1997; WANG and HABER 2004). The Rad51p-associated DNA strand, then performs a genome-wide search for homology (BARZEL

and KUPIEC 2008); the details of the mechanism of this search are not understood. When homology is detected between the Rad51p-coated end and a region of sequence homology in an unbroken template, the broken end invades the intact template, creating a D-loop (PETUKHOVA *et al.* 2000). This strand invasion step is facilitated by Rad52p, Rad59p, Rad55p, Rad57p, and Rad54p (AYLON and KUPIEC 2004). Following strand invasion, SDSA, DSBR or BIR take place (PÂQUES and HABER 1999), and each of these pathways will be discussed below.

1.3.2.1 Synthesis-dependent strand annealing (SDSA) pathway

In the SDSA pathway as diagrammed in Figure 1.3A, following processing of the broken ends, the left end invades the intact homolog. DNA synthesis primed from the 3′ invading end extends the D-loop. The D-loop is then disassembled and the single-strand of newly-synthesized DNA (shown in pale red), anneals with the right broken end. Since the newly-synthesized strand has sequences derived from the red homolog and it is annealed to a strand derived from the blue homolog, there may be mismatches within the heteroduplex (the heteroduplex region enclosed in a box in Figure 1.3A). Correction of these mismatches by the mismatch repair system to generate a duplex with the information of the blue homolog represents a gene conversion event, the non-reciprocal transfer of information between homologs (PÂQUES and HABER 1999). Other features of gene conversion will be discussed in Section 1.3.3. This pathway produces gene conversion events that are unassociated with crossovers. Since there is no crossover

associated with the conversion, markers distal to the conversion are not associated with LOH (Figure 1.4A).

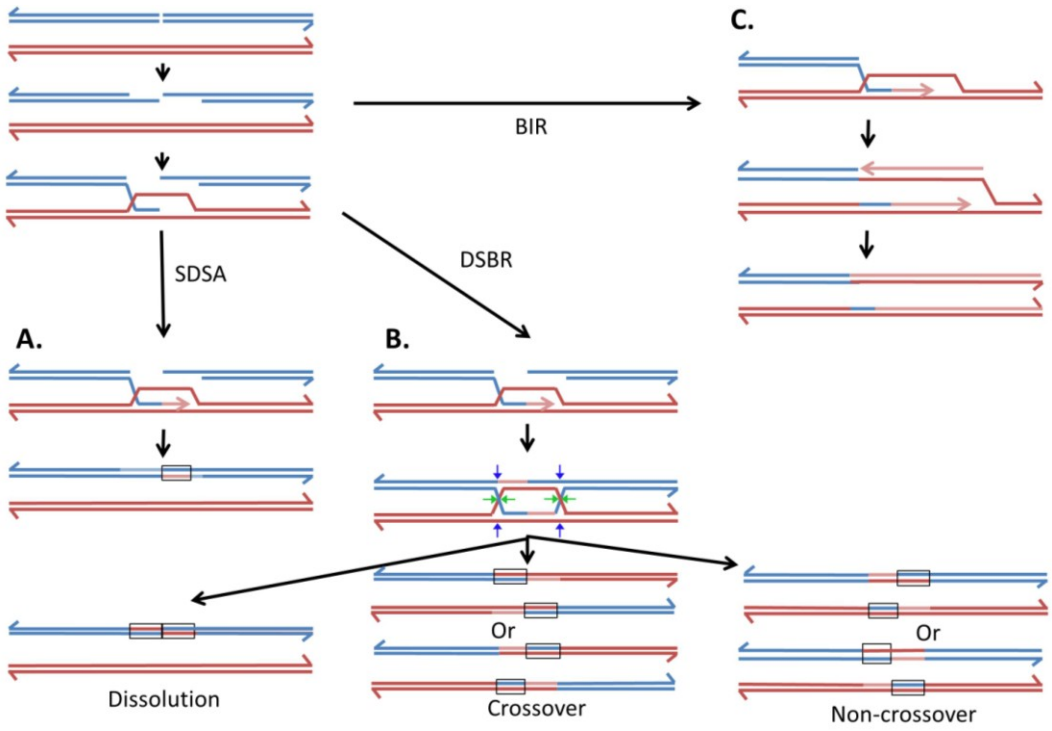


Figure 1.3. Common pathways of DSB repair by HR.

In this figure, each chromosome is shown as a double-stranded DNA molecule with arrows indicating the 3′ ends. The blue lines depict both strands of the broken chromosome and the red lines depict both strands of the intact homologous chromosome. Pale red and pale blue lines indicate DNA synthesized in the course of the repair reaction. All pathways are initiated by a DSB, followed by 5′ to 3′ resection of the broken ends. In addition, the next step in each pathway is the invasion of an intact chromosome by one of the broken ends.

A. Synthesis-dependent strand annealing (SDSA). In this pathway, following strand invasion, the 3′ end of the invading strand acts as a primer for DNA synthesis. Following synthesis, the D-loop is disassembled and the displaced newly-synthesized strand anneals with the righthand broken end, forming a heteroduplex. The heteroduplex region, which may contain mismatches, is enclosed by a black box. Finally, DNA synthesis primed by the 3′ strand of the left broken end fills in the remaining single-stranded gap (region shown as a dark blue strand and a light blue

strand). As discussed in the text, SDSA results in gene conversion events unassociated with crossovers.

B. Double-strand break repair (DSBR). This mechanism has the same early steps at the SDSA pathway. However, instead of disassembling the D-loop, the righthand broken molecule pairs with the D-loop (second-end capture), resulting in two regions of heteroduplex and formation of a double Holliday junction (dHJ). The dHJ can be processed in several ways. Dissolution (a non-crossover pathway) results in two regions of heteroduplex on one chromosome and none on the other. Cleavage of the dHJ by resolvases in opposite directions for the two junctions results in crossovers with heteroduplexes on both crossover products. Cleavage of the dHJ in the same directions for the two junctions results in non-crossover chromosomes with heteroduplex regions on both homologs.

C. Break-induced replication (BIR). In this HR mechanism, the one broken end invades the intact homolog, setting up a replication fork that copies the homolog to the end of the chromosome. In this mechanism, one side of the broken chromosome (the righthand side) is lost.

1.3.2.2 Double-strand break repair (DSBR) pathway

The classic double-strand break repair (DSBR) model was proposed by SZOSTAK *et al.* (1983). This model is depicted in Figure 1.3B. In this model, the 3′ end of the processed DSB invades a homologous template, displacing one strand, and creating a D-loop. The non-invading 3′ end of the broken chromosome then anneals to the D-loop in a step called "second-end capture". Each 3′ end can then synthesize DNA using the complementary strands as templates. In the completed dHJ intermediate, there are two patches of heteroduplex DNA.

The dHJ structure must be resolved to allow separation of the recombined DNA molecules. The dissolution pathway involves unwinding of the dHJ junction without cleavage of the junctions; this process utilizes the proteins Rmi1p, Sgs1p, and Top3p

(CEJKA *et al.* 2010). This pathway generates non-crossovers exclusively. In the alternative pathway, the junctions are cleaved by resolvases. Resolvases make concerted pairs of nicks on opposite arms of the Holliday junctions (KLEIN and SYMINGTON 2009). Possible patterns of cleavage are shown in Figure 1.3 by green and blue arrows. If the cleavages of both junctions occur with the orientations indicated by all blue arrows or all green arrows, the result is a non-crossover (NCO). Alternatively, if one junction is cleaved with the patterns shown by green arrows, and the other with the pattern shown by blue arrows, there will be a crossover associated with resolution. In mitosis, most dHJs are resolved to generate crossovers (MITCHEL *et al.* 2010). It should be emphasized that, by the DSBR model, all crossovers are associated with heteroduplex formation and, therefore, a region of potential gene conversion. In our studies, most crossovers had an associated gene conversion tract (diagrammed in Figure 1.4B).

There are a number of different protein complexes that can cleave HJs in *S. cerevisiae* including Yen1p, Mus81p-Eme1p, and Slx1p-Slx4p (KLEIN and SYMINGTON 2009). Elimination of all of these activities, however, does not substantially affect crossovers in meiotic cells. Another complex containing Exo1p, Sgs1p, and the MutL proteins Mlh1p and Mlh3p are important in resolving HJs in meiotic cells (ZAKHARYEVICH *et al.* 2012).

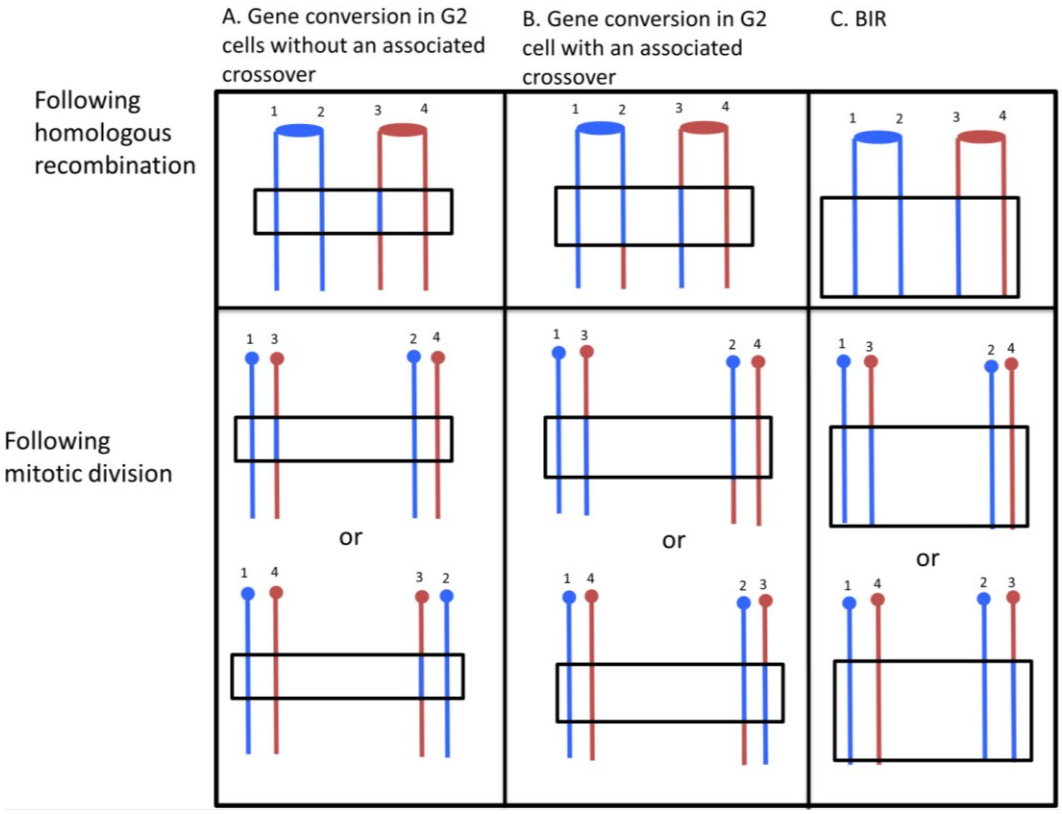


Figure 1.4. Conversion tracts resulting from homologous recombination.

In this figure, the recombination events involve duplicated chromatids (G2 events). In conversion events, the broken chromatid receives information from the unbroken chromatid (PÂQUES and HABER 1999). In this figure, each chromosome is shown as a single line and the different colors represent DNA of two different chromosome homologs. The region of conversion is enclosed by rectangles for all pathways. For all of the depicted events, there are two ways that the homologous chromatids can be segregated; both possibilities are shown in the lower portion of the figure. The blue sister chromatids are numbered 1 and 2, and the red chromatids are numbered 3 and 4.

A. Gene conversion unassociated with crossovers. As shown in Figure 1.3, conversion events of this type can be generated by SDSA or by processing of a dHJ. Note that for both segregation patterns, at least one daughter cell will have a conversion tract.

B. Gene conversion associated with a reciprocal crossover. This pattern is likely generated exclusively through processing of a dHJ (Figure 1.3B). Following the crossover, one type of segregation (chromatids 1 and 3 in one cell, and 2 and 4 in the

other) generates daughter cells that have LOH centromere-distal to the crossover. Alternatively, if chromatids 1 and 4 segregate into one daughter cell, and chromatids 2 and 3 into the other, LOH distal to the crossover is not observed, although the interstitial LOH pattern reflecting the conversion event is still evident.

C. Conversion tract resulting from BIR. Following a BIR event, for both chromosome segregation patterns, one daughter cell has no LOH and the other has an apparent very long conversion tract extending to the telomere.

1.3.2.3 Break-induced replication (BIR)

In aging yeast cells or in cells with perturbed recombination, BIR events are also observed (McMurray and Gottschling 2003; Pâques and Haber 1999). In the BIR model (Figure 1.3C), following strand invasion, the broken chromosome forms a replication fork that copies the intact homolog to the end of the chromosome (Malkova *et al.* 1996). BIR is not dependent on Rad51p (Krogh and Symington 2004), but requires Pol32p, a polymerase co-factor that is not required for other HR pathways (Lydeard *et al.* 2007). The net result of a BIR event is one recombinant and one non-recombinant chromosome (Figure 1.3C). The associated pattern of LOH for a BIR event mimics a very long gene conversion event that extends to the end of the chromosome (Figure 1.4C). If the homologous chromosome is used as the template during BIR, one daughter cell will always be homozygous for all sequences past the initiation of the replication fork, and the other daughter cell will maintain heterozygosity (Figure 1.4C).

1.3.3 Mechanisms of gene conversion

Crossovers are reciprocal events in that there is no net loss of sequences from either of the interacting homologs. In contrast, gene conversion events reflect the non-

reciprocal transfer of information from one homolog to the other (PÂQUES and HABER 1999; PETES *et al.* 1991). Conversion events can occur without crossovers or can be associated with crossovers. In meiosis, about half of conversions have an associated crossover (MANCERA *et al.* 2008; PETES *et al.* 1991). It has been argued that a smaller fraction of mitotic conversion events are associated with crossovers (PÂQUES and HABER 1999; PETES *et al.* 1991), but these arguments are based on conversion events constrained to be short. In my analysis, described in Chapter 2, about 30-40% of the conversion events are associated with crossing-over.

The properties of gene conversion were first examined in meiotic studies in fungi in which tetrad analysis was possible. In a diploid heterozygous for alleles *A* and *a*, most tetrads segregated 2 spores with the *A* allele and two with the *a* allele. At a frequency that varies for different loci between 1-20%, tetrads with 3*A*: 1*a* or 3*A*: 1*a* spores are observed. A variety of experiments showed that these events, termed "gene conversions," reflected the faithful non-reciprocal transfer of sequences between different DNA molecules (PETES *et al.* 1991). In addition, these studies indicated (as described above) that there was a non-random association between gene conversion and crossovers. These observations were critical in the formulation of molecular models of recombination by HOLLIDAY (1964) and others, including Szostak *et al.* (1983). Gene conversion events can be produced by two different types of mechanisms: 1) repair of mismatches in heteroduplex DNA (MODRICH and LAHUE 1996) and 2) repair of a

double-stranded gap in the DNA (ORR-WEAVER and SZOSTAK 1983); both are described in further detail below.

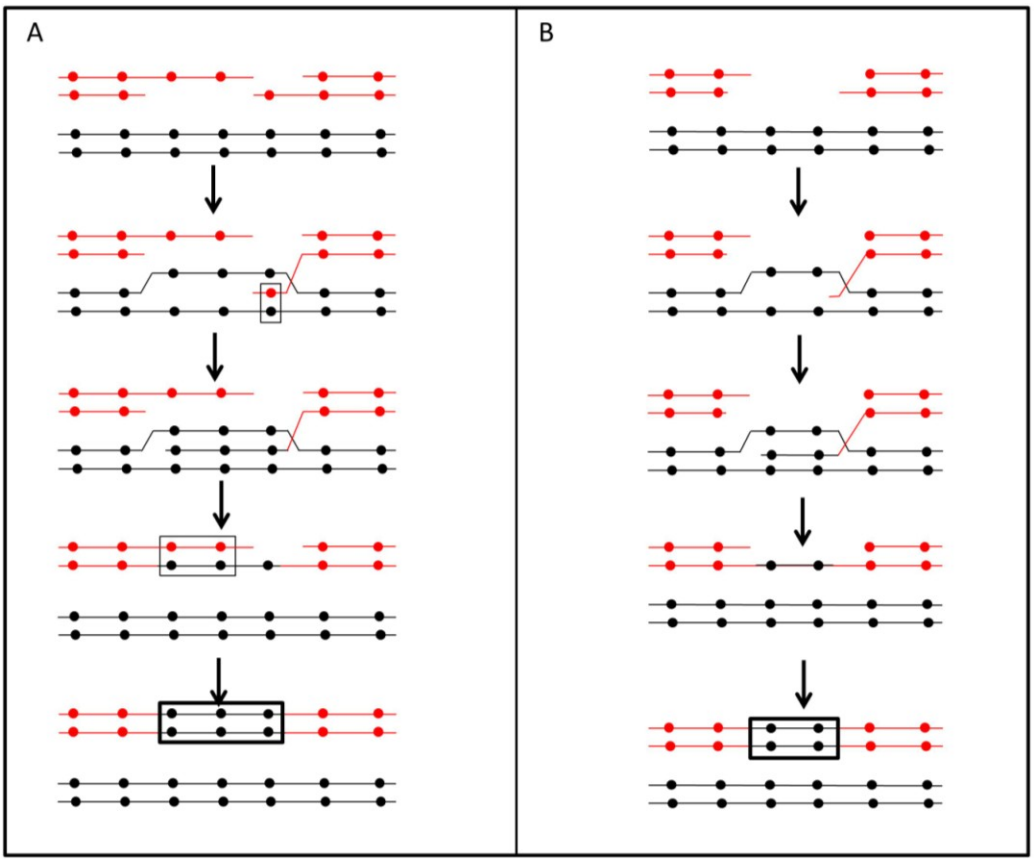


Figure 1.5. Two mechanisms for gene conversion.

This figure shows two different mechanisms of generating a gene conversion tract during repair of a DSB by the SDSA repair pathway. The chromosomes are shown as double-stranded molecules. In this depiction, the red DNA molecule is broken and the intact black duplex is used as a template for repair. The colored circles represent heterozygous SNPs on the two homologs.

A. Gene conversion as a consequence of mismatch repair in a heteroduplex. Following strand invasion, DNA synthesis is primed by the invading strand. During initial invasion a heteroduplex region is formed between the invading strand and the template, which is corrected to the black SNP. Following D-loop collapse, a heteroduplex region with two mismatches is formed. If the mismatch repair (MMR) complex removes the mismatched bases from the red strand and the resulting gap is

filled in using information from the black strand, a conversion event would be generated (as shown). Alternatively, repair in the opposite direction (removal of the SNPs on the black strand and duplication of the red SNPs) would restore 2:2 segregation. Although such restoration events are usually difficult to detect, evidence for restoration during meiotic gene conversion in yeast has been obtained (Kirkpatrick *et al.*, 1998).

B. Conversion as a consequence of gap repair. In this model, the DSB is processed into a double-stranded gap. Repair of this gap occurs by two consecutive DNA synthesis events without mismatch repair.

1.3.3.1 Gene conversion as a consequence of mismatch repair in a heteroduplex

Almost all models of HR invoke the pairing of a single-strand from the broken homolog with the unbroken homolog; this pairing ensures that the two homologs will be aligned perfectly during the recombination event (SZOSTAK *et al.* 1983). The region of the duplex that contains strands derived from two different homologs is the heteroduplex. If the intact template is the sister chromatid, there will be no differences in the sequence within the heteroduplex. However, if the intact template is the homologous chromosome, there can be mismatches within the heteroduplex. Most organisms have the enzymatic machinery to remove mismatched bases from one strand, followed by filling in the gap using the intact strand as a template (SCHOFIELD and HSIEH 2003). Although the primary function of the MMR (mismatch repair) system is likely to be the correction of errors introduced during DNA replication (MODRICH and LAHUE 1996), MMR can also generate a gene conversion event. In the example shown in Figure 1.5A, if the mismatches shown as the red SNPs are removed and the black SNPs are duplicated, a conversion event would be produced.

The evidence that most meiotic recombination in *S. cerevisiae* is a consequence of mismatch repair is based on an analysis of the regulation of gene conversion and another genetic phenomenon, post-meiotic segregation (PMS) (PETES *et al.* 1991). In general, all the cells within a single colony derived from sporulating a wild-type heterozygous *A*/*a* diploid are of one genotype, either *A* or *a*. However, at a frequency of about 1%, single spore colonies with equal numbers of *A* and *a* cells were observed (HOLLIDAY 1964; PETES *et al.* 1991). These PMS events are explicable as representing unrepaired mismatches within a heteroduplex. PMS events, like gene conversions, were shown to be non-randomly associated with crossovers. Thus, HOLLIDAY (1964) suggested that all crossovers had regions of heteroduplex. If these heteroduplex regions contained mismatches, repair of the mismatches would lead to conversion events and failure to repair mismatches would lead to PMS. The observation that gene conversion is considerably more frequent than PMS in yeast (DETLOFF *et al.* 1991) argues that MMR is efficient. One very strong argument in favor of the hypothesis that conversion events reflect MMR is that mutations that eliminate MMR in yeast result in elevated levels of PMS and reduced frequencies of gene conversion (PETES *et al.* 1991). It has been estimated that about 90% of the meiotic conversion events are a consequence of MMR in yeast. About 10% of the events could reflect gap repair (described below) or mismatch repair-utilizing systems other than the canonical MMR machinery.

Mitotic gene conversion events also can be generated by heteroduplex formation followed by mismatch repair (AYLON *et al.* 2003; HABER *et al.* 1993; PÂQUES and HABER 1999). This repair almost always converts the invading strand to the corresponding sequence of the template molecule (HABER *et al.* 1993). If the mismatch in the heteroduplex is not corrected, the resultant daughter cell can give rise to two progeny cells of different genotypes in a process called post-switching segregation (PSS) (HABER *et al.* 1993). PSS is the mitotic equivalent of PMS. As discussed below, mitotic conversion events generated by gap repair have also been observed.

Mismatches are repaired in yeast by two different complexes. The heterodimers related to MutS of *E. coli* are Msh2p-Msh6p (involved in the repair of base-base mismatches) (ALANI 1996; IACCARINO *et al.* 1996; JOHNSON *et al.* 1996), and Msh2p-Msh3p which recognizes and repairs indels (insertion/deletions) (KOLODNER and MARSISCHKY 1999). During most repair events, these complexes interact with a heterodimer of MutL-like proteins, Mlh1p and Pms1p.

In addition to their roles in the repair of mismatches, MMR proteins have other roles in recombination. Mutations affecting mismatch repair lead to increased levels of mitotic recombination between diverged sequences (ALANI *et al.* 1994; CHAMBERS *et al.* 1996; CHEN and JINKS-ROBERTSON 1998; HUNTER and BORTS 1997). The anti-recombination roles of MMR might function to prevent ectopic recombination between dispersed repeats throughout the genome that can cause genome rearrangements

(ARGUESO *et al.* 2008). Additionally, MMR proteins have been shown to influence the crossover/noncrossover ratio of DSB-initiated events. Strains with *msh2* or *msh6* mutations have increased crossovers between templates that have 1.2 kb of homology (INBAR *et al.* 2000). The Msh2p-Msh6p heterodimer is capable of recognizing Holliday junctions (MARSISCHKY *et al.* 1999). INBAR *et al.* (2000) have suggested that the Msh2p-Msh6p heterodimer might assist in the dissolution of double Holliday junctions during DSBR, resulting in a non-crossover. Finally, some MMR proteins are involved in the processing of branched DNA structures (INBAR and KUPIEC 1999). Mutations in *MSH2* and *MSH3* reduce recombination if heterologous 3′ ends are present, implying that the Msh2p-Msh3p heterodimer is necessary to remove these ends (PÂQUES and HABER 1997; SUGAWARA *et al.* 1997). The removal of these 3′ ends is also dependent on Rad1p and Rad10p (IVANOV and HABER 1995).

The amount of sequences transferred during a gene conversion event using the mechanism described above is a function of two parameters, the length of the heteroduplex and the continuity of the mismatch repair process. For meiotic conversion events in *S. cerevisiae*, most tracts are between one and two kb in size, and the conversion events associated with crossovers are longer than those unassociated with crossovers (MANCERA *et al.* 2008). As described in Chapter 3, my data show that mitotic conversion tracts in yeast are much longer than meiotic tracts.

1.3.3.2 Gene conversion by gap filling

Gene conversion can also occur as a result of repair of a double-stranded DNA gap (ORR-WEAVER and SZOSTAK 1983). DNA gaps are formed by loss of sequence information from both strands of the duplex following DSB formation (Figure 1.5B). Most of the evidence of gap repair is based on studies in which the gap is made *in vitro* using restriction enzymes in plasmids that are transformed into yeast (MITCHEL *et al.* 2010; ORR-WEAVER and SZOSTAK 1983). In such studies, repair of the plasmid, generating the conversion event, can result in either integration of the plasmid (a crossover outcome) or a failure to integrate the plasmid (conversion without crossing over) (MITCHEL *et al.* 2010; ORR-WEAVER and SZOSTAK 1983).

It is unclear at present whether gap formation *in vivo* is common enough to represent a substantial fraction of mitotic or meiotic conversion events. Most studies of the processing of DSB-generated broken ends show extensive 5′ to 3′ processing with little evidence for 3′ to 5′ processing (PÂQUES and HABER 1999; SYMINGTON and GAUTIER 2011). However, in one study ZIERHUT and DIFFLEY (2008), formation of a double-stranded DNA gap was detected in broken DNA molecules that entered the S-phase without repair.

1.3.4 Cell cycle and homologous recombination

The mechanism of DSB repair differs depending on when the DSB is formed during the cell cycle. In this section of the Introduction, I will discuss a number

of aspects of the connection between the cell cycle and DSB repair by HR including: 1) the choice of chromosome partners for the HR event, 2) DNA damage checkpoints, and 3) the cell cycle regulation of end resection.

1.3.4.1 Homologous recombination partner choice

The repair of a DSB by HR requires an intact homologous DNA molecule to act as a template for the repair event. If a DSB occurs in haploid cells in G1, no intact template exists and DSBs are repaired using NHEJ (described in Section 1.3.5). If DSBs form in diploid cells in G1, a break in one homolog can be repaired by HR using the intact homolog. In diploid cells in G2, however, there are two types of intact templates for HR, the other homolog and the sister chromatid. As described above, an HR event with the homolog is detectable since it leads to LOH centromere-distal to the position of the crossover. In contrast, a sister chromatid crossover does not result in LOH.

Most assays for sister chromatid recombination involve selection for unequal sister chromatid crossover events (KADYK and HARTWELL 1992; KLEIN 1988; SZOSTAK and WU 1980). In one assay, unequal crossover events are detected by formation of a wild-type *ADE3* gene as a consequence of recombination between tandem copies of mutant *ade3* genes, with one copy truncated at the 5′ end and the other truncated at the 3′ end (KADYK and HARTWELL 1992). If DSBs are generated in G2-synchronized yeast cells with X-rays, the majority (97%) are repaired using the sister chromatid as a template (KADYK and HARTWELL 1992). This observation implies that there must be

some cellular mechanism predisposing HR repair to use the sister chromatid as a template, because during G2 in a diploid cell, the sister chromatid only represents one-third of the homologous sequences available (one sister chromatid and two homologs). This preference is likely to reflect the closer proximity of the sister chromatid, since sister chromatids are physically linked together by sister chromatid cohesion (UHLMANN and NASMYTH 1998). DSBs during G2 of the cell cycle also trigger genome-wide formation of sister chromatid cohesion (STROM *et al.* 2007; UNAL *et al.* 2007). Mutations in genes required for sister chromatid cohesion often affect HR partner choice. Mre11p is necessary for the formation of DSB-induced sister chromatid cohesion (UNAL *et al.* 2007), and *mre11Δ* cells have decreased levels of sister chromatid recombination (BRESSAN *et al.* 1999; GONZALEZ-BARRERA *et al.* 2003). Deletions of *MRE11, RAD50,* or *XRS2* also result in increased spontaneous recombination between homologs (GOTTLIEB *et al.* 1989; IVANOV *et al.* 1992; TSUBOUCHI and OGAWA 1998), indicating that the decrease in sister chromatid recombination in *mre11Δ* is not due to an overall decrease in recombination. Another argument in favor of the preferential use for the sister chromatid over the homolog during HR is that diploid cells in G2 are more resistant to X-ray treatment than diploid cells irradiated in G1 (BRUNBORG and WILLIAMSON 1978).

Although recombination between homologs can be elevated by irradiating G2-synchronized cells, treatment of G1-synchronized cells with ionizing radiation or ultraviolet light stimulates recombination between homologs even more strongly

(ESPOSITO 1968; FABRE *et al.* 1984; KADYK and HARTWELL 1992; LEE and PETES 2010; WILDENBERG 1970). As will be discussed in Chapter 3, we suggest that chromosomes broken in G1 are replicated to produce two broken chromatids that recombine with the homologous chromosome in G2. There are several other types of previous experiments consistent with this possibility. DSBs formed during G1 of the cell cycle are not processed efficiently and are, therefore, inefficiently repaired (AYLON *et al.* 2004). Metaphase spreads of chicken cells irradiated during G1 of the cell cycle have been shown to have two chromatids broken at the same location following chromosome replication, while those irradiated during G2 only have a single broken chromatid (TAKATA *et al.* 1998). The relevant data from our lab (LEE *et al.* 2009; LEE and PETES 2010; ST CHARLES *et al.* 2012) is based on analyzing conversion tracts associated with reciprocal crossovers in sectored yeast colonies. As will be discussed in detail, in Chapter 3, these tracts demonstrate that most crossovers involve the repair of two sister chromatids broken at the same position, as expected if the broken chromatids are generated by replication of a broken chromosome in G1. It is important to note that, in our current model of mitotic recombination, most recombination events are initiated by a G1 DSB but are repaired in G2. If a DSB is made <u>and</u> repaired in G1, no sectored colony is produced (Figure 1.6).

Finally, DSBs that occur in repetitive DNA sequences, such as the dispersed Ty class of retrotranspons, can be repaired by HR events that utilize Ty elements on non-

homologous chromosomes (AYLON *et al.* 2003; FAIRHEAD and DUJON 1993; INBAR and KUPIEC 1999). Ectopic recombination resulting in crossovers between dispersed repeats causes chromosome translocations and other types of genome rearrangements (ARGUESO *et al.* 2008; DUNHAM *et al.* 2002).

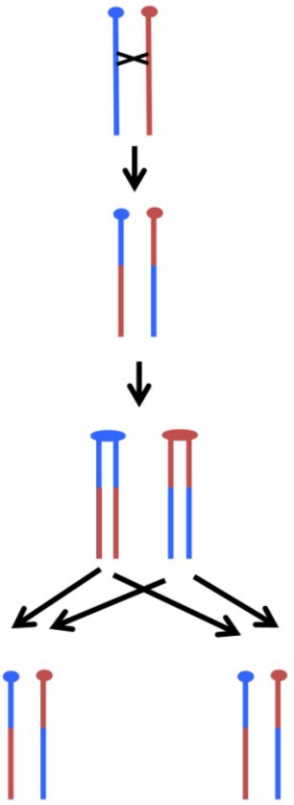


Figure 1.6. Hypothetical G1 crossover between homologs.

A crossover that occurs during G1 of the cell cycle does not generate LOH centromere-distal to the point of exchange in the daughter cells.

1.3.4.2 Cell cycle arrests in response to DNA damage

Yeast have cell cycle checkpoints that act to pause the cell cycle in response to DNA damage or replication stress (LONGHESE *et al.* 1998; WEINERT and HARTWELL 1988).

The function of the checkpoints is to allow the cells enough time to repair DNA damage before segregation of the damaged chromosomes (PUTNAM *et al.* 2009). Yeast have several checkpoints that are specific to the timing and the types of DNA damage. G1/S and G2/M DNA damage checkpoints respond to damage that occurs during G1 or G2, respectively (MICHELSON and WEINERT 2000). There is an intra-S-phase checkpoint that responds to DNA damage during replication (PAULOVICH *et al.* 1997), and a replication checkpoint that responds to aberrant replication forks (SANTOCANALE and DIFFLEY 1998). The proteins Mec1p and Tel1p are kinases that phosphorylate many proteins that function in the enzymology of DNA repair. Mec1p (along with Ddc1p) recognizes and binds to single-stranded DNA that is coated with RPA (ZOU and ELLEDGE 2003), and Tel1p is recruited to DSBs that are bound by the Mre11p-Rad50p-Xrs2p complex (NAKADA *et al.* 2003). Following their recruitment to the sites of DNA damage, Tel1p and Mec1p activate downstream pathways that result in cell cycle arrest and DNA repair (PUTNAM *et al.* 2009).

Despite the presence of a G1 DNA damage checkpoint, most repair of DSBs by homologous recombination does not occur during this cell cycle phase (AYLON *et al.* 2004). The G1 checkpoint is not very robust in yeast, and results in a transient (about 1 hour) arrest, after which the DNA damage is often unrepaired (GERALD *et al.* 2002; SIEDE *et al.* 1994). Additionally, 5′-to-3′ end resection, the initiating event of HR, is usually restricted to G2 of the cell cycle (AYLON *et al.* 2004) as described below.

1.3.4.3 Cell cycle control of DNA resection

In addition to the choices of HR pathways and templates, DSBs can be repaired by either HR or NHEJ. The details of the NHEJ pathway will be described in Section 1.3.5. A related issue that is relevant to the choice between HR and NHEJ is the regulation of DNA resection. As described below, the pattern of resection determines that NHEJ is the prominent repair pathway during G1 of the cell cycle (SHRIVASTAV *et al.* 2008) while homologous recombination is the prominent pathway during S and G2 phases (AYLON *et al.* 2004; CASPARI *et al.* 2002; IRA *et al.* 2004).

Cell cycle progression is dependent on cyclin-dependent kinases (CDKs), of which there is only one in yeast, Cdc28 (TYERS *et al.* 1993). Cdc28p forms a complex with the S and G2 cyclins (*CLB*s) that triggers DNA replication (LENGRONNE and SCHWOB 2002). Homologous recombination, and more specifically DNA resection, cannot occur without the activity of the Clb-CDK complexes (AYLON *et al.* 2004). DNA resection is initiated by the endonucleases Mre11p and Sae2p (MIMITOU and SYMINGTON 2008; SHIM *et al.* 2010; ZHU *et al.* 2008). Following this slow initiation step, processive resection occurs much more quickly (FRANK-VAILLANT and MARCAND 2002). The initial resection step that involves Sae2p is dependent on phosphorylation of Sae2p by Cdc28p and only occurs during G2 of the cell cycle (HUERTAS *et al.* 2008).

HUERTAS *et al.* (2008) also showed that the constitutive phosphorylation of Sae2p in cells that are arrested during G1 of the cell cycle results in hypersensitivity to ionizing

radiation. This phenotype probably reflects aberrant DNA resection during G1 of the cell cycle that causes many DSBs to be channeled through the HR pathway at a time during the cell cycle when the sister chromatid is not an available template for HR, resulting in cell death due to unrepaired DSBs. Cell cycle control of DSB repair pathways ensures that the most appropriate repair pathways is used during the correct phase of the cell cycle (BRANZEI and FOIANI 2008).

1.3.5 Non-homologous end-joining (NHEJ)

NHEJ involves the re-ligation of broken DNA ends by a mechanism that requires little or no sequence homology (DALEY *et al.* 2005); this mechanism occurs with minimal or no processing of the broken DNA ends (MIMITOU and SYMINGTON 2009). NHEJ requires the Ku70p, Ku80p, Lig4p, Lif1p, and Nej1p proteins (DUDASOVA *et al.* 2004), which are not involved in HR. The HR-specific step of 5′-to-3′ resection represents the commitment to HR as the repair pathway of choice, and this processing is restricted to G2 of the cell cycle (AYLON *et al.* 2004; IRA *et al.* 2004). Conversely, NHEJ is most active during G1 of the cell cycle (FRANK-VAILLANT and MARCAND 2002; SHRIVASTAV *et al.* 2008). However, many DSBs that are formed during G1 of the cell cycle are not repaired during G1 of the cell cycle, demonstrating that NHEJ is not an efficient DSB repair pathway in yeast (AYLON *et al.* 2004). Additionally, NHEJ is repressed in *MATa*/*MAT*α, but not in haploids and *mat*Δ/*MAT*α diploids, demonstrating that this pathway is under mating type regulation (ASTROM *et al.* 1999; LEE *et al.* 1999). Thus, the NHEJ pathway is

designed to be most active for haploid yeast cells in G1 that have no option for repair by HR.

1.4 The association of mitotic recombination with cancer

Paradoxically, although homologous recombination acts primarily to maintain genome integrity and prevent cancer (MOYNAHAN and JASIN 2010; NEGRINI *et al.* 2010), HR can also promote carcinogenesis and other diseases. The importance of HR in maintaining genome stability is based on several types of experiments. Mice lacking *Rad51* fail to fully develop and die early during embryonic development (LIM and HASTY 1996); this lethality is thought to reflect the failure to repair spontaneous DSBs. Mutations in the *BRCA1* and *BRCA2* genes greatly increase the chance of developing breast and ovarian cancers (NATHANSON *et al.* 2001). The protein products of these genes have been shown to interact with known HR proteins and signaling proteins (SCULLY *et al.* 1997; WONG *et al.* 1997). Mutations in these genes result in high levels of chromosome breakage, deletions, and aneuploidy (KRAAKMAN-VAN DER ZWET *et al.* 2002; TUTT *et al.* 1999; VENKITARAMAN 2001). Additionally, loss of checkpoint genes required for efficient HR, such as ataxia-telangiectasia-mutated gene (ATM, the human ortholog of the yeast gene *TEL1*), leads to a very high rate of cancer development (FRIEDENSON 2007).

The cancer-promoting role of HR is related to its property of generating LOH of heterozygous markers distal to the crossover. The importance of LOH in cancer development was first demonstrated for retinoblastoma, a cancer in which fetal

retinoblasts fail to differentiate into post-mitotic retinal photoreceptors and neurons, but instead continue to divide and form tumors (KNUDSON 2001). KNUDSON (1971) observed that 40% of cases of retinoblastoma occur in individuals that are heterozygous for a germ-line mutation in the *Rb* gene. However, not all carriers of mutant *Rb* genes developed the disease. He suggested a "Two-hit hypothesis" to account for this discrepancy. He argued that retinoblastoma requires two events: the first hit is a germline mutation of the *Rb* gene, and the second hit is an event that inactivates the remaining wild-type copy of *Rb*. This second hit was later shown to be associated with LOH of the *Rb* gene (CAVENEE *et al.* 1983).

LOH in human tumor samples can be analyzed by comparing tumor and normal tissue samples by two-allele RFLP and Southern analysis (WEISSENBACH *et al.* 1992) or by comparing tumor and normal tissue samples with SNP microarrays (ZHENG *et al.* 2005). In addition to retinoblastoma, LOH for the *Rb* gene has been observed in bladder, lung, and breast cancers (BORG *et al.* 1992; GOUYER *et al.* 1994; XU *et al.* 1993; YOO *et al.* 1994). LOH for *Rb* correlates with reduced expression of the *Rb* gene, indicating that LOH of this gene and the subsequent reduction in expression offers a growth advantage to cancer cells (BORG *et al.* 1992; GOUYER *et al.* 1994; XU *et al.* 1993; YOO *et al.* 1994). The *Rb* gene encodes for a protein that has a cell cycle control function for the G1-S transition (SHERR 1994).

In addition to LOH for the *Rb* gene, LOH has been detected in tumor samples for regions containing other tumor suppressor genes including *p53, BRCA1,* and *BRCA2* (BAKER *et al.* 1990; COLLINS *et al.* 1995; SMITH *et al.* 1992; WELCSH and KING 2001). For this reason, a number of researchers have examined tumor samples to determine if frequently observed regions of LOH might identify unknown tumor suppressor genes. However, since mutations in some genes observed in cancers lead to increased rates of global LOH (RELIENE *et al.* 2007), it has been difficult to distinguish cause from effect in these studies (SHEN *et al.* 2000). The observance of regions of LOH that are common to a subtype of cancer as opposed to regions that are rarely homozygous within a different subtype might be helpful in distinguishing causal LOH from symptomatic LOH (WANG *et al.* 2004).

It should be pointed out that LOH events such as those described above can be generated by mechanisms other than mitotic crossovers. LOH can also result from chromosome loss, chromosome loss followed by re-duplication, mitotic gene conversion events, and deletions (CAVENEE *et al.* 1983; TUNA *et al.* 2009). Although the relative numbers of these events appears to be quite variable, depending on chromosome context and the type of tumor examined (TUNA *et al.* 2009), the evidence that a substantial fraction of LOH events represents mitotic crossovers is strong (SHAO *et al.* 1999). For example, about 70% of the LOH events observed for high-grade gliomas result from chromosome loss, and 30% reflect mitotic recombination(LO *et al.* 2008).

Although HR can be cancer-promoting, as well as cancer-repressing, the observation that all free-living organisms share the enzymatic machinery required to repair DSBs by HR argues strongly that this system is of great benefit to the organism. In the following chapters of my thesis, I describe my analysis of spontaneous and damage-induced mitotic crossovers. This study reveals important new information about the mechanism of initiating mitotic crossovers. In the course of this analysis, I developed a high-resolution genetic map for mitotic recombination that includes about 10% of the yeast genome.

2 High-resolution genome-wide analysis of irradiated (UV and γ rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events

In diploid eukaryotes, repair of double-stranded DNA breaks (DSBs) by homologous recombination often leads to loss of heterozygosity (LOH). Most previous studies of mitotic recombination in *S. cerevisiae* have focused on a single chromosome or a single region of one chromosome at which LOH events can be selected. In this study, we used two techniques (single-nucleotide polymorphism [SNP] microarrays and high-throughput DNA sequencing [HTS]) to examine genome-wide LOH in a diploid yeast strain at a resolution averaging one kb. We examined both selected LOH events on chromosome V and unselected events throughout the genome in untreated cells, and cells treated with either γ radiation or ultraviolet radiation (UV). Our analysis shows: 1) spontaneous and damage-induced mitotic gene conversion tracts are more than three times larger than meiotic conversion tracts, and conversion tracts associated with crossovers are usually longer and more complex than those unassociated with crossovers, 2) most of the crossovers and conversions reflect the repair of two sister chromatids broken at the same position, and 3) both UV and γ radiation efficiently induce LOH at doses of radiation that cause no significant loss of viability. Using HTS, we also detected new mutations induced by γ-rays and UV. To our knowledge, our

study represents the first high-resolution genome-wide analysis of DNA damage-induced LOH events performed in any eukaryote.

2.1 Introduction

All organisms experience DNA damage from both exogenous and endogenous sources. Endogenous DNA damage includes spontaneous deamination of nucleotides, depurination/depyrimidination, oxidative damage, and double-stranded DNA breaks (DSBs) (FRIEDBERG 2006). DSBs are likely to be particularly deleterious, since unrepaired DSBs can lead to chromosome rearrangements or chromosome loss. Although the sources of endogenous DSBs have not been completely determined, some DSBs appear to reflect nuclease processing of secondary DNA structures (such as DNA "hairpins") or head-on collisions between the replication and transcription machineries (AGUILERA 2002). Below, in addition to examining spontaneous recombination events that presumably reflect the repair of endogenous DNA damage, we also analyze recombination events induced by two exogenous sources: γ rays and ultraviolet light (UV).

Both γ rays and UV cause a variety of different types of DNA damage. γ rays cause DSBs, single-stranded DNA nicks, and base damage (FRIEDBERG 2006; WARD 1990). UV results in pyrimidine dimers (FRANKLIN *et al.* 1985; SETLOW 1966), DNA-DNA or DNA-protein crosslinks (PEAK and PEAK 1986), and single-stranded DNA nicks resulting from the dimer excision (BREEN and MURPHY 1995).

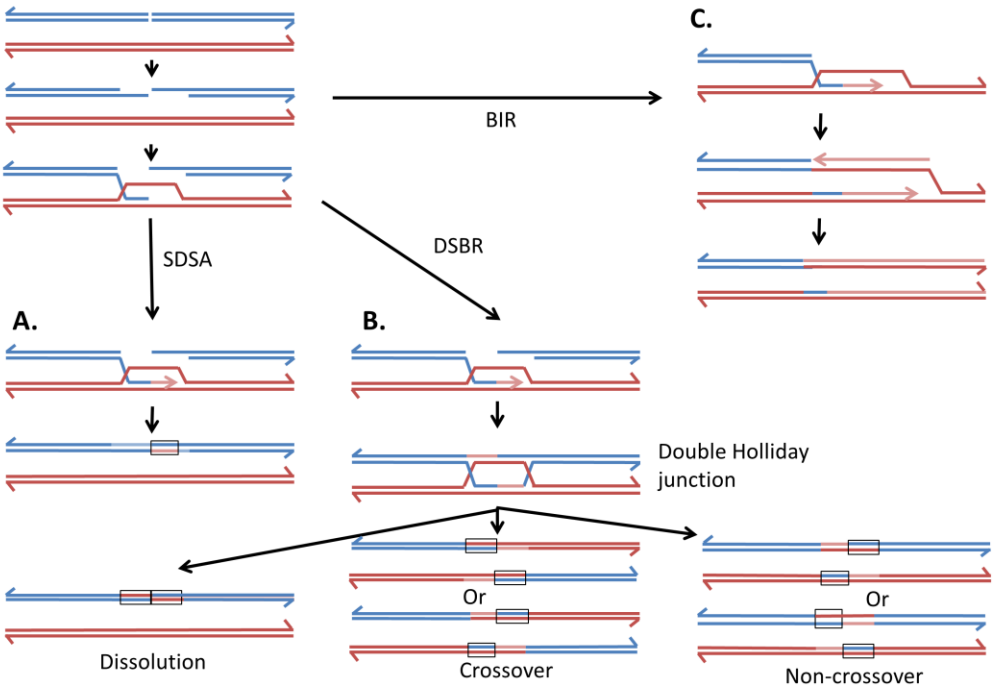


Figure 2.1. Pathways of DSB repair by homologous recombination.

In this figure, we show accepted models of DSB repair by homologous recombination. DNA strands from two different homologues are shown in red and blue; light red and blue lines indicate newly-synthesized DNA. Regions of the duplex that have strands of different colors represent heteroduplexes. These pathways are described in detail in the text.

A. Synthesis-dependent strand annealing (SDSA) pathway. Following processing of the DSB, the 3′ end of the left end of the broken DNA molecule invades the other duplex. Following DNA synthesis, the invading strand is displaced, and hybridizes to the right end of the broken chromosome. This pathway results in conversion events unassociated with crossovers.

B. Double-strand break repair (DSBR) pathway. In this pathway, a double Holliday junction (dHJ) is formed. Resolution of these junctions by resolvase cleavage can result in two different crossover products (middle panel) and two different non-crossover products (right panel). These products have two regions of heteroduplex located in *trans*. Alternatively, the dHJ can be dissolved by the action of

topoisomerases/helicases resulting in a non-crossover product with heteroduplexes located in *cis*.

C. Break-induced replication (BIR) pathway. One of the broken ends invades the homologous chromosome, and duplicates sequences from the point of invasion to the telomere. The net result of BIR events is an apparent long terminal gene conversion event.

In yeast, as in most eukaryotes, there are two recombination pathways that are used to repair DSBs: non-homologous end-joining (NHEJ) and homologous recombination (HR). In NHEJ events, as the name implies, broken DNA molecules are re-joined by a mechanism that requires little or no homology (DALEY *et al.* 2005). This mechanism is most active in haploid yeast cells during G1 of the cell cycle (SHRIVASTAV *et al.* 2008). In diploid cells, the dominant pathway is HR. HR uses an intact homologous DNA molecule, either the sister chromatid or the homologous chromosome, as a template for repair of the broken chromosome.

DSBs can be repaired by several different HR pathways (HEYER *et al.* 2010). The repair of a DSB by gene conversion unassociated with a crossover is shown in Figure 2.1A. This process involves the non-reciprocal transfer of sequences from the intact donor molecule to the broken chromosome in several steps: 1) invasion of one broken end into the intact template molecule, followed by DNA synthesis primed by the invading 3′ strand, 2) removal of the invading end and reannealing of this end back to the other broken end, forming a heteroduplex with mismatches, and 3) repair of the mismatches. This mechanism (synthesis-dependent strand-annealing or SDSA) was first

suggested to explain some features of meiotic recombination in yeast (ALLERS and LICHTEN 2001). In the second pathway (Figure 2.1B), gene conversion may be associated with crossovers. In this pathway, a double Holliday junction (dHJ) is formed that can be resolved to yield a crossover or non-crossover. In this pathway, heteroduplexes flank the original position of the DSB. Although the heteroduplex regions have the same size in Figure 2.1B, in both meiosis (JESSOP *et al.* 2005; MERKER *et al.* 2003) and mitosis (MITCHEL *et al.* 2010; TANG *et al.* 2011), the conversion tracts flanking the DSB are often of different lengths. The dHJ can also be dissolved without nucleolytic cleavage of DNA strands to yield non-crossover products with heteroduplexes located in *cis* on one of the two interacting chromosomes (HEYER *et al.* 2010). In the third pathway (Figure 2.1C), one part of the broken DNA molecule is lost and a complete chromosome is then reconstructed by break-induced replication (BIR). In this mechanism, one of the broken ends invades the intact template molecule and a replication fork is set up that duplicates the template from the site of invasion to the telomere.

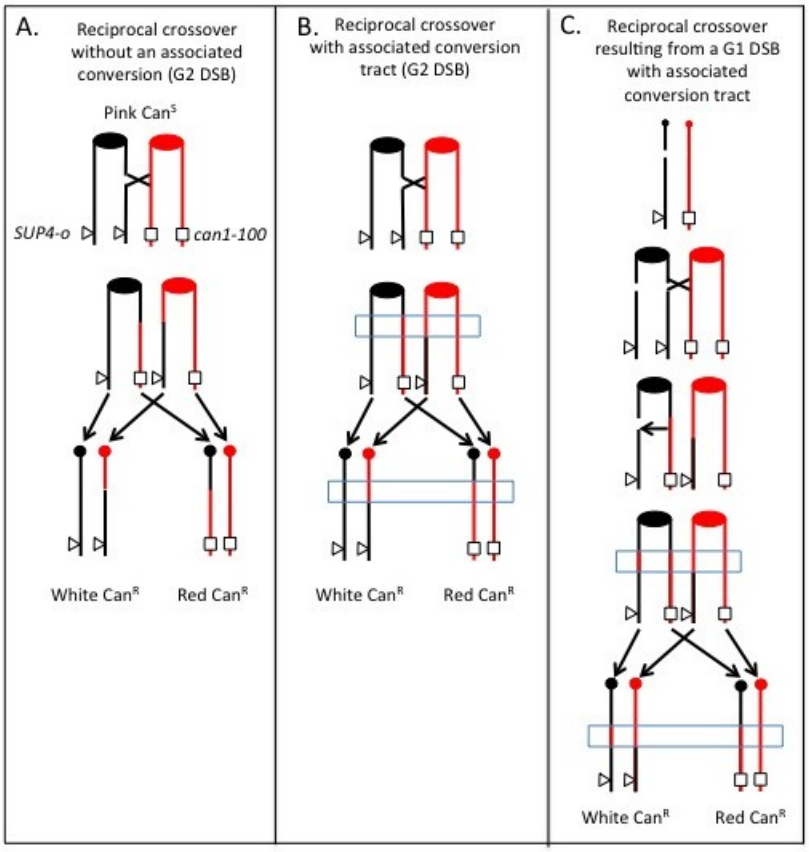


Figure 2.2. Genetic system used to select reciprocal crossovers on the left arm of chromosome V.

The starting diploid strain PG311 has the ochre-suppressible *can1-100* on one copy of chromosome V (shown in red) and the *SUP4-o* gene (encoding an ochre suppressor tRNA gene) at an allelic position on the other homologue (shown in black). The strain is homozygous for the ochre-suppressible *ade2-1* allele. Strains with an unsuppressed *ade2*-1 mutation form red colonies. The starting diploid strain is canavanine-sensitive and forms pink colonies.

A. Reciprocal crossover without an associated gene conversion initiated by a single DSB in G2. This type of event produces a canavanine-resistant red/white sectored colony (BARBERA and PETES 2006). The transition from heterozygous markers to LOH is identical in the two sectors.

B. Reciprocal crossover with an associated conversion event initiated by a single DSB in G2. If a DSB forms on one of the black chromatids, a conversion associated with the crossover may occur. This event will also result in a canavanine-

resistant red/white sectored colony in which the transitions between heterozygous markers and LOH are different in the two sectors. The region of conversion is indicated by the blue rectangle.

C. Reciprocal crossover and conversion resulting from a DSB formed in G1. A black chromosome with a DSB is replicated to form two sister chromatids that are broken at the same place. One chromatid is repaired to yield a reciprocal crossover and an associated conversion; the second is repaired to yield a conversion without a crossover. The resulting red and black sectors will have a 4:0 conversion event, a region in which both sectors are homozygous for SNPs derived from the red chromatid (included within the blue rectangle).

If HR involves an interaction between two homologues that can be distinguished by single-nucleotide polymorphisms (SNPs), conversions without crossovers will produce a small patch of LOH within a chromosome that is otherwise heterozygous, whereas both crossovers and BIR result in LOH that extends from the site at which the event initiates to the end of the chromosome. Repair of a DSB by HR in which the homologous chromosome is used as a template will result in LOH, but repair events involving the sister-chromatid will not. Although most sister-chromatid recombination events are genetically silent, unequal sister-chromatid exchanges can be detected in yeast by a variety of different systems (PETES and HILL 1988). Using one of these systems, KADYK and HARTWELL (1992) showed that, in diploid cells, sister chromatids are the preferred substrate for the repair of DSBs generated by X-rays. Despite this preference, it is clear that ionizing radiation and UV strongly stimulate HR events (both mitotic crossovers and gene conversions) between homologous chromosomes (FABRE 1978; NAKAI and MORTIMER 1969).

One problem with studying spontaneous mitotic recombination is that most analytic systems do not allow the selection of both daughter cells that contain the recombinant chromosomes. Several years ago, we developed a method of selecting reciprocal crossovers on chromosome V that surmounts this difficulty (BARBERA and PETES 2006; LEE *et al.* 2009). One copy of chromosome V has the *can1-100* allele (an ochre mutation) and, in the other copy, the *CAN1* gene is replaced by *SUP4-o*, a gene encoding an ochre-suppressing tRNA (Figure 2.2A). In the absence of the suppressor, strains with the *can1-100* allele are resistant to canavanine, but because of the suppressor, the diploid used in our experiments is canavanine-sensitive. In addition, the diploid is homozygous for the *ade2-1* mutation (an ochre mutation). Strains with this mutation, in the absence of the *SUP4-o* gene form red colonies, but form pink colonies if one copy of the *SUP4-o* gene is present. Thus, the diploid strain is canavanine-sensitive and forms pink colonies. If a crossover occurs between the centromere of chromosome V and the *can1-100*/*SUP4-o* markers (a distance of about 120 kb), a canavanine-resistant red/white colony is formed (Figure 2.2A).

Although this method was first used in diploid strains lacking polymorphisms, subsequent studies were done in which a diploid was constructed using two haploid strains that had about 0.5% sequence divergence (LEE *et al.* 2009; LEE and PETES 2010), resulting in about 55,000 SNPs distributed throughout the genome. Crossovers and associated gene conversions on chromosome V were mapped to a resolution of about 4

kb by using PCR, restriction digests, and gel electrophoresis to look for LOH. Although a few of the crossovers had no associated conversion, most of the crossovers were associated with an adjacent conversion event (boxed regions in Figure 2.2). In the 3:1 class of events (Figure 2.2B), in the boxed region, three of the four chromosomes had one of the two forms of the SNP and one of the chromosomes had the other form (one sector being homozygous for a SNP with the other sector being heterozygous). In addition, about 40% of the crossovers were associated with a conversion event in which the same SNP was homozygous in both sectors (Figure 2.2C); we term these events "4:0 conversions." The observation of the 4:0 events argues that about half of mitotic crossovers result from the repair of two sister chromatids that are broken at approximately the same positions. One simple mechanism for generating this intermediate is to have the DSB occur in G1, and the broken chromosome replicate to form two broken chromatids (LEE *et al.* 2009). This proposed mechanism was confirmed by analysis of the types of conversion events stimulated by γ rays in synchronized G1 and G2 cells (LEE and PETES 2010). In addition to 3:1 and 4:0 conversion events, 3:1/4:0 hybrid tracts are also observed (LEE *et al.* 2009; LEE and PETES 2010). Such tracts can also be explained as a consequence of repair of two broken sister chromatids (Figure 2.3).

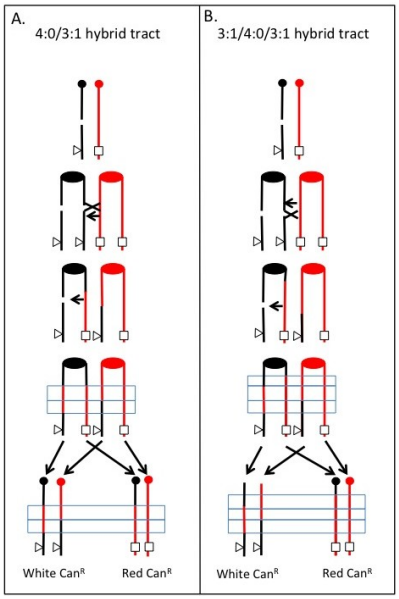


Figure 2.3. Production of hybrid conversion tract by the repair of two broken sister chromatids.

The black chromosome is broken in G1 and replicated to yield two broken sister chromatids.

A. Production of a 3:1/4:0 hybrid tract. If the two DSBs are repaired to yield conversion tracts that have the same centromere-proximal boundary, but different centromere-distal boundaries, a 3:1/4:0 hybrid will be generated (shown in the blue rectangle).

B. Production of a 3:1/4:0/3:1 hybrid tract. If one conversion event is extended beyond the other at both the centromere-proximal and centromere-distal boundaries, a 3:1/4:0/3:1 tract will be formed.

In the current study, we use SNP microarrays, and HTS to map selected events on chromosome V as well as unselected events throughout the genome at a resolution of about 1 kb. To our knowledge, these studies are the first to measure the numbers and types of LOH events throughout the genome induced by doses of ionizing radiation (100

Gy) and UV (10-15 J/m^2) that have no significant effect on cell viability. We also determined the number of mutations induced in the genomes by these treatments.

2.1.1 Note on collaborative work

This chapter of my thesis is the result of a collaborative paper. I designed the SNP microarrays that were used for analysis within the paper, and analyzed the γ radiation-induced sectors. Sabrina Andersen completed the analysis of the spontaneous sectors. Yi Yin completed the analysis of the UV radiation-induced sectors. Einat Hazkani-Covo analyzed all high-throughput sequencing (HTS) data. Ewa Malc and Piotr Mieczkowski performed the high-throughput sequencing of all samples sequenced. Patricia Greenwell confirmed the *de novo* mutations detected within by HTS. Fred Dietrich helped with the design of the SNP microarrays.

2.2 Materials and methods

2.2.1 Strains and genetic methods

All experiments were conducted with the diploid strain PG311 (LEE *et al.* 2009). The relevant genotype of PG311 is *MATα::NAT/MAT***a** *URA3/ura3-1 ade2-1/ade2-1 TRP1/trp1-1 HIS3/his3-11,15 GAL2/gal2 SUP4-o/can1-100 V9229/V9229::HYG V261553/V261553::LEU2*. This diploid was generated by crossing the haploid strains PSL2 and PSL5 which are isogenic with strains W303a and YJM789, respectively, except for alterations introduced by transformation (LEE *et al.* 2009). Below, we will refer to the haploid parents of PG311 as W303a and YJM789. In general, PG311 has the SNPs

predicted from the haploid parents. The disruption of *MATα* in PG311 allows synchronization of this diploid by alpha factor. Although diploids that lack *MATα* do not sporulate under normal conditions, such strains can be sporulated on plates containing 5 mM nicotinamide (J. Rine, personal communication). For experiments in which we analyzed meiotic products of PG311, the stains were pre-grown on YPD plates with 5 mM of nicotinamide, and then transferred to sporulation plates containing 5 mM nicotinamide. Plates were incubated at 25° for 2-4 days before tetrad dissection.

Standard media were used (GUTHRIE and FINK 1991) unless noted. To detect spontaneous crossovers, we first isolated single colonies of PG311 grown on rich growth medium (YPD) at 30° for two days. Individual colonies were suspended in 400 μl of dH_2O, and 100 μl of this mixture was plated on canavanine-containing plates (SD-arg + 120 μg/ml canavanine). The plates were incubated four days at room temperature, followed by incubation for 16 hours at 4°; the 4° incubation allows better visualization of the red sectors. We purified cells from the red and white sectors for subsequent analysis.

In the experiments in which cells were irradiated, we synchronized cells in G1 using alpha factor (LEE and PETES 2010). The synchronized cells were treated with either γ radiation in a Shepherd Mark 1 Cesium-137 irradiator at 100 Gy or with UV (254 nm) derived from a TL-2000 Ultraviolet Translinker at a dosage of 10 or 15 J/m^2. Following radiation, the cells were plated either on non-selective plates (SD- arg) or plates that lacked arginine and contained 120 μg/ml canavanine. The subsequent growth of the cells

and the analysis of sectors were done as described above for the spontaneous selection with the exception that sectored colonies for the UV-treated samples were isolated from non-selective plates grown at 30° instead of room temperature.

2.2.2 SNP microarrays: design and optimization

We designed the SNP arrays based on genomic sequence information available from SGD for S288c (very closely related to W303a) (WINZELER *et al.* 2003) and YJM789 (WEI *et al.* 2007). Microarrays that were capable of detecting LOH for SNPs in PG311 were designed based on principles outlined by GRESHAM *et al.* (2010). For each SNP represented on the array, four 25-base oligonucleotides were used: one for each strand of the W303a SNP, and one for each strand of the YJM789 SNP. The SNP was located in the middle of the 25-base oligonucleotide. Although there are approximately 55,000 SNPs in PG311, about three-quarters of these SNPs were not used for our analysis. We excluded most of the SNPs found in repeated genes. We also screened out oligonucleotides that had a melting temperature for the perfectly-matched 25 base pair duplex that was less than 55° or greater than 59°. The remaining oligonucleotides (representing about 15,000 SNPs) that were printed on the microarray are listed in Table S3 of ST CHARLES *et al.* (2012). We also included on the array about 120 oligonucleotides that were not different between W303a and YJM789; these are listed in Table S4 of ST CHARLES *et al.* (2012).

Oligonucleotides were printed onto the microarrays by Agilent Technologies on slides containing about 105,000 oligonucleotides. Many of the oligonucleotides are

represented more than once in the microarrays. Following experiments to determine which oligonucleotides resulted in the most specific hybridization signals (described in Supporting Information), we reduced the number of SNPs used in our analysis to 13,000. This final set of oligonucleotides is presented in Table S5 of ST CHARLES *et al.* (2012).

2.2.3 Methods used for microarray analysis: sample preparation, hybridization conditions, and data analysis

The methods used for sample preparation, hybridization conditions, and data analysis are similar to those described previously (LEMOINE *et al.* 2005; MCCULLEY and PETES 2010). A detailed description of these procedures is presented in Supporting Information. In brief, genomic DNA from the experimental strain was labeled with Cy5-dUTP, DNA from the control strain (PG311) was labeled with Cy3-dUTP, and the two labeled samples were competitively hybridized to the microarrays. The arrays were then scanned at wavelengths of 635 and 532 nm using a GenePix scanner and GenePix Pro software using settings recommended by the manufacturer. The ratio of the medians (635 nm/532 nm; R_M) for each probe was used for analysis, and replicate R_Ms were averaged. The data were centered around a value of one by dividing each probe R_M by the average of all of the probe R_Ms in order to normalize for differences in the hybridization levels for the control and experimental strain samples.

We calculated 95% confidence intervals on the median sizes of conversion tracts using Table B11 of ALTMAN (1990). Comparisons of conversion tract lengths under

different experimental conditions were done using the Mann-Whitney test on the VassarStats Website (http://faculty.vassar.edu/lowry/VassarStats.html).

2.2.4 Generation and analysis of high-throughput DNA sequencing data

Samples were prepared for HTS as described above for the SNP microarray sample preparation with the exception that genomic DNA was sonicated to 300-700 bp fragments. The DNA was then prepared for sequencing using the protocol recommended by Illumina for the Genome Analyzer IIx. The samples were sequenced using an Illumina GAIIx machine, generating 67- to 75-bp paired end reads. For the eight sequenced samples, coverage varied from 90- to 180-fold.

The details of the HTS data analysis are presented in Supporting Information. In brief, we detected regions of LOH by identifying SNPs in the experimental strains in which at least 90% of the "reads" that were originally heterozygous were now identical to one of the original alleles. We identified new mutations by finding bases that were identical in the original diploid, but had a novel base in at least 40% of the "reads" in the irradiated diploid; we use the 40% criterion because we expect that any new mutation will be heterozygous. Mutations that appeared in more than one independent isolate were not counted as *de novo* mutations, since such mutations presumably arose in the strain before treatment with the DNA damaging agent.

2.3 Results

As described in the Introduction, we previously selected spontaneous mitotic crossovers, as well as crossovers induced by γ rays, that occurred between the centromere of chromosome V and the *can1-100/SUP4-o* markers, an interval of about 120 kb (Figure 2.2) (LEE *et al.* 2009; LEE and PETES 2010). The diploid used in these studies (PG311) was constructed by a cross between two haploids that are isogenic with W303a and YJM789, and is heterozygous for about 55,000 SNPs. In our previous analysis, the positions of the crossovers and associated gene conversion events were mapped to a resolution of about 4 kb using a PCR-based strategy that determined whether the SNPs were heterozygous or homozygous. This procedure is impractical for genome-wide mapping of recombination events. Below, we describe the use of SNP arrays to map spontaneous, UV-induced and γ ray-induced crossovers selected on chromosome V, as well as unselected crossovers and gene conversion events throughout the genome.

SNP arrays have been used previously to map LOH events in tumor cells (LINDBLAD-TOH *et al.* 2000), to map meiotic recombination events in *S. cerevisiae* (MANCERA *et al.* 2008), to characterize chromosome rearrangements and chromosome loss in *C. albicans* (ABBEY *et al.* 2011), and in a variety of other experiments. Using principles outlined by Gresham *et al.* (GRESHAM *et al.* 2010), we developed a SNP array to examine LOH events throughout the genome in the diploid PG311. This array has oligonucleotides that distinguish over 13,000 SNPs, resulting in an average density of

one oligonucleotide per kb of genomic DNA. The details of the array design and the specific sequences of the probes are in Supporting Materials and Methods and in Tables S3-S5 of ST CHARLES *et al.* (2012).

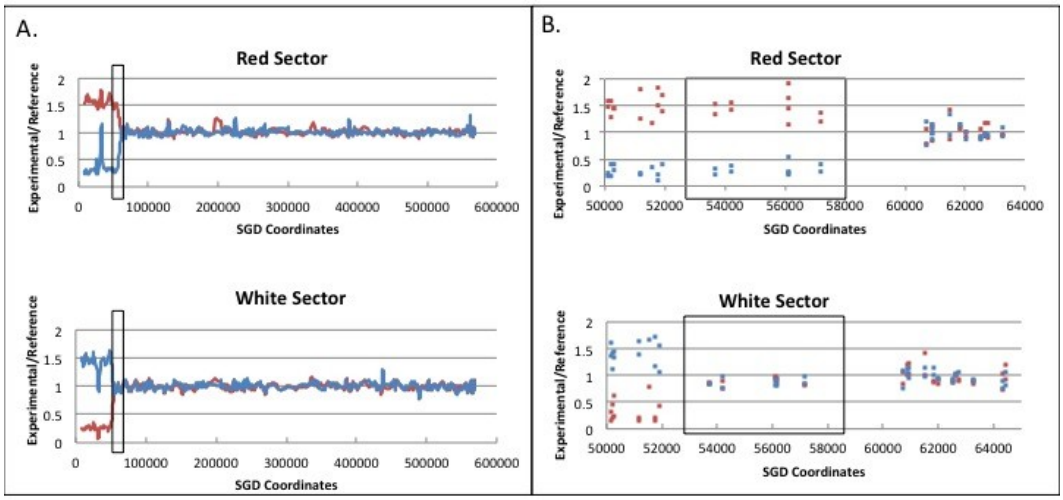


Figure 2.4. Analysis of a spontaneous reciprocal crossover (PG311-2A) on the left arm of chromosome V by SNP microarrays.

Most of the details concerning this figure are explained in the text. In brief, DNA samples isolated from the red and white sectors were labeled with one fluorescent nucleotide and DNA from a heterozygous control strain was labeled with a different fluorescent nucleotide. The samples were competitively hybridized to the SNP array and we measured the ratio of hybridization of the probes to SNP-specific oligonucleotides. The red and blue colors indicate hybridization to the W303a- and YJM789-specific oligonucleotides, respectively. *CEN5* is located approximately at SGD coordinate 152000.

A. Low-resolution depiction of the samples derived from the red and white sectors. In the boxed region, the red sector has a region of LOH whereas the white sector is heterozygous at the same position. This pattern indicates a 3:1 conversion associated with the crossover. Centromere-distal to the conversion event, the red and white sectors are homozygous for the W303a- and YJM789-specific SNPs, respectively.

B. High-resolution depiction of the samples derived from the red and white sectors. Each blue and red square represents hybridization to a single oligonucleotide on the array; the converted region is boxed.

In our experiments, we labeled genomic DNA from a diploid with a recombination event with one fluorescent nucleotide and DNA from the control diploid PG311 with a different fluorescent nucleotide. The samples were mixed and competitively hybridized to the SNP arrays. If SNPs retained heterozygosity in the experimental strain, then the hybridization signal for the oligonucleotides representing that SNP was similar to that of the control strain. LOH was detected by an increase in hybridization to oligonucleotides that had one form of the SNP (for example, the W303a form) and a decrease in hybridization to oligonucleotides that had the other form (for example, the YJM789 form).

As described above, crossover events between *CEN5* and the *can1-100/SUP4-o* markers produce canavanine-resistant red/white sectored colonies. Thus, all of the samples analyzed had a selected recombination event on chromosome V, and we found that samples treated with UV or γ rays also had unselected events on other chromosomes. We isolated genomic DNA from both the red and white sides of the sectors and examined the DNA by SNP arrays for diploids untreated with recombinogenic agents, as well as for diploids treated with UV or γ rays. An example of the analysis for chromosome V for a selected spontaneous crossover (PG311-2A) is shown on Figure 2.4. In Figure 2.4A, a low-resolution depiction of the hybridization levels is shown for both the red (top) and white (bottom) sectors. In both sectors, the hybridization pattern indicates that the transition from heterozygous SNPs to

homozygous SNPs is at approximately SGD coordinate 55000. As expected, the DNA that is centromere-distal to the crossover junction from the red sector hybridizes well to the W303a-specific probes (red line) and poorly to the YJM789-specific probes (blue line), since the red sector is generated by LOH events that include the *can1-100* marker that is derived from the W303a-related homologue (Figure 2.2). Genomic DNA from the white sector shows the reciprocal pattern of hybridization. The "spike" of increased hybridization in the red sector for YJM789 SNPs near SGD coordinate 30000 is an artifact resulting from a deletion of YJM789 sequences during the insertion of the *SUP4-o* gene into the YJM789-derived chromosome.

In Figure 2.4B, we show the same crossover event at higher resolution. Each square in this figure shows the hybridization ratio to a specific oligonucleotide. In the red sector, the transition between the homozygous SNPs and the heterozygous SNPs is between SGD coordinates 57170 and 60701. In the white sector, the transition occurs between 51915 and 53692. Thus, there is a region (boxed in Figure 2.4B) in which one sector is homozygous for SNPs and the other is heterozygous. This region is a 3W:1Y gene conversion tract (W and Y indicating W303a-derived and YJM789-derived SNPs, respectively), equivalent to the boxed region in Figure 2.2B; in our subsequent discussions, a 3:1 conversion event indicates 3W:1Y conversion and 1:3 indicates 1W:3Y conversion. We estimate the length of the gene conversion tract by averaging the maximal length of the tract (the distance between markers that are not within the gene

conversion tract, 8.8 kb) and the minimal length of the tract (the distance between the converted markers, 3.5 kb). For the tract shown in Figure 2.4B, this length is about 6.2 kb. It is important to emphasize that the presence and extent of gene conversion tracts can only be identified when the patterns of LOH are analyzed in genomic DNA from both sectors of the sectored colony.

With the genome-wide SNP analysis, we found that the parental diploid PG311 and all of its subsequent derivatives had two LOH events that were unexpected from the sequence of the parental haploids. All strains were homozygous for W303a-derived SNPs centromere-distal to SGD coordinate 685 kb on chromosome XIII, and were homozygous for W303a-derived SNPs between coordinates 412715 and 414085 on chromosome X. Since these events, presumably generated during sub-culturing of PG311, were present in all strains, they were excluded from our analysis.

2.3.1 Analysis of spontaneous selected LOH events on chromosome V by SNP arrays

We examined by SNP arrays genomic DNA from both the red and white sectors from fourteen independent canavanine-resistant colonies. The crossover events in five of the isolates had been mapped previously by PCR-amplifying regions along the left arm of chromosome V that contained polymorphic restriction enzyme sites and testing for homozygosity or heterozygosity by a restriction enzyme digest analysis (LEE *et al.* 2009). We also mapped an additional nine events solely by SNP arrays. In the previous study,

we used 34 markers in the 120 kb *CEN5-can1-100/SUP4-o* interval. We monitored 172 markers in this same interval by SNP arrays

Table 2.1 Comparison of mapping methods for four spontaneous crossovers on chromosome V.

	PCR-based method			SNP microarrays		
Strain name	Cen-prox. coordinate	Cen-dist. coordinate	Event description	Cen-prox. coordinate	Cen-dist. coordinate	Event description
PG311-1.4	133080	125754	CO, no conversion	130096	127038	CO + 3:1 tract
PG311-1.7	151440	146855	CO, no conversion	151419	150291	CO, no conversion
PG311-4.1	99267	60163	CO + hybrid tract	98763	62494	CO + complex tract
PG311-4.11	133080	94329	CO + hybrid tract	129511	97792	CO + hybrid tract

In our previous study, we found that most of the spontaneous crossovers were crossovers without conversions (Figure 2.2A), crossovers with associated 3:1 conversions (Figure 2.2B), crossovers with 4:0 conversions (Figure 2.2C), or crossovers with hybrid conversions (3:1/4:0 or 3:1/4:0/3:1 tracts; Figure 2.3) (LEE *et al.* 2009). A comparison of the mapping of recombination events by the PCR-based method and SNP arrays for four of the sectored colonies is shown in Table 2.1. In this table, we define the position of the crossovers with only two SGD coordinates: the position of the centromere-proximal heterozygous marker that is closest to the crossover/conversion event and the position of the centromere-distal homozygous markers representing the crossover. Although the agreement between the two methods was reasonably good, as expected, the SNP array mapped events with better resolution and also revealed that some of the conversion events were more complex than previously determined. For example, in PG311-1.4, we previously mapped a crossover between SGD coordinates 125754 and 133080 that appeared to be unassociated with gene conversion. With the SNP arrays, we mapped the transition at higher resolution between SGD coordinates 127038 and 130096, and we detected a SNP in this region that had undergone gene conversion. The complete description of all spontaneous crossovers and associated gene conversion events is given in Table S 2.3.

Table 2.2 Summary of all crossovers diagnosed by SNP microarray

In this table, we summarize data from selected crossovers and associated conversion events on the left arm of chromosome V as well as unselected crossovers and associated conversions on other chromosomes. For this table, the data obtained with high-throughput DNA sequencing was not used.

Types of reciprocal crossovers	Spontaneous		γ-ray		UV	
	Selected	Unselected	Selected	Unselected	Selected	Unselected
No detectable conversions	2	0	0	0	0	0
3:1 or 1:3 conversions	2	0	0	1	0	2
4:0 or 0:4 conversions	1	0	1	0	0	0
Hybrid conversions[1]	1	0	2	3	2	0
Complex conversions	7	0	4	0	1	5
Total crossovers	**13**	**0**	**7**	**4**	**3**	**7**

One exception to the generally good agreement between the two mapping methods is isolate PG311-1.6. This event was originally classified as crossover associated with a conversion tract that extended from SGD coordinate 31694 to 63936. SNP analysis demonstrated that the white sector had a terminal deletion on chromosome V, beginning near coordinate 62000. The same sector also had a large terminal duplication on chromosome VII. Although this rearrangement has not been fully characterized, since there are a cluster of delta elements near the breakpoint on chromosome V and Ty elements at the breakpoint on chromosome VII, it is possible that the strain has a chromosome V-VII translocation, similar to those that we have characterized previously (ARGUESO *et al.* 2008). No alterations were detected on either chromosome V or VII in the

[1] 3:1/4:0, 3:1/4:0/3:1, 1:3/0:4, or 1:3/0:4/1:3 conversion events.

red sector. Sequence analysis indicated that the red sector retained the *SUP4*-o gene. It is possible that the cell that gave rise to the red sector lost the prion *PSI*, that affects the efficiency of ochre suppressors (SHKUNDINA and TER-AVANESYAN 2007), although other possibilities cannot be excluded. Whatever the details of the genetic alterations in PG311-1.6, the event does not represent a conventional allelic crossover on chromosome V and, therefore, is excluded from our analysis.

Of the 13 colonies with spontaneous reciprocal recombination events analyzed by SNP arrays, the numbers of colonies of various classes were: 1) two crossovers without detectable conversions, 2) two crossovers with 3:1 conversion events, 3) one crossover with a 0:4 conversion, 4) one crossover associated with a hybrid tract (1:3 and 0:4 segments), and 5) seven crossovers with complex conversion tracts (Table 2.2). The complex conversion tracts will be discussed further below.

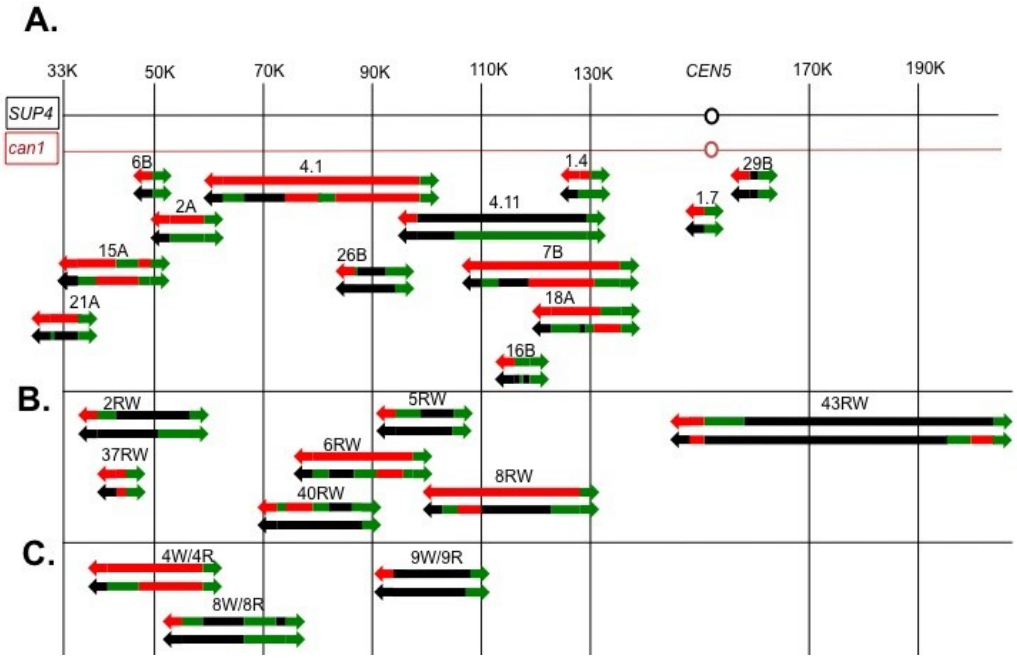


Figure 2.5. Mapping of crossovers and associated conversion events on the left arm of chromosome V in untreated cells, and cells treated with γ rays or UV by SNP microarrays.

Each red/white sectored canavanine-resistant colony is given a number and is depicted as a pair of lines with the upper line representing the red sector and the lower line showing the white sector. The colored segments signify heterozygosity for the markers (green), homozygosity for the YJM789-derived SNPs (black), or homozygosity for the W303a-derived SNPs (red). The green arrows show that the markers are heterozygous from the position that the green segment begins to the end of the chromosome, and the red and black arrows indicate that the markers are homozygous for the W303a- or the YJM789-derived SNPs, respectively, from the point that the segment begins to the telomere of the chromosome. Internal regions of heterozygosity and homozygosity are shown as line segments without arrows and are drawn approximately to scale. The numbers at the top of the figure are SGD coordinates and the region between *CEN5* and the *can1-100*/*SUP4-o* markers is about 120 kb in length.

A. Spontaneous crossovers and conversions.

B. γ ray-induced crossovers and conversions.

C. UV-induced crossovers and conversions.

The locations of the spontaneous crossovers and associated conversion tracts are shown in Figure 2.5A. Each sectored colony is depicted as a pair of lines with the upper line representing the red sector and the lower line representing the white sector. The red and black line segments indicate that the sector is homozygous for the W303a-associated SNPs and the YJM789-related SNPs, respectively. The green line segment indicates that the sector is heterozygous for the SNPs. We show the two chromosomes within each sector as a single line because our analysis does not allow us to determine the coupling relationships for heterozygous SNPs between the two homologues. The median length of all crossover-associated conversion tracts was 6.1 kb, similar to the median observed in our previous study of spontaneous conversion tracts (6.5 kb, LEE *et al.* (2009)). Only one unselected LOH was observed in unirradiated cells. Both sectors in PG311-7B had a gene conversion event on chromosome VIII unassociated with a crossover (Table S 2.3). Thus, the frequency of spontaneous unselected LOH events/cell is very low (about 0.08) as expected.

For the events shown in Figure 2.5A, 3:1 conversion events could reflect an initiating DNA lesion occurring anywhere within the tract, since events can be propagated bidirectionally from the DSB (TANG *et al.* 2011). For 4:0 or 3:1/4:0 hybrid tracts, the initiating lesion presumably occurs within the 4:0 region of the tract (Figure 2.3). Although we do not see any strong hotspots for spontaneous events with this

limited dataset, in a larger sample, we found that the region between SGD coordinates 41,000 and 60,000 had a significantly elevated level of crossovers and the region near *CEN5* had a significantly reduced level of events (LEE *et al.* 2009).

As shown in Figure 2.5A, many of the recombination events are associated with multiple transitions between heterozygosity and homozygosity. In Table S 2.3, for each sectored colony, we assigned a letter to represent each transition point; for each transition point, we also show two SGD coordinates, indicating the positions of the closest SNPs on the array that flank the transition points. The simplest events (crossovers without gene conversion, 6B and 1.7 in Figure 2.5A) have a single transition point. The recombination event shown in Figure 2.4 (which corresponds to event 2A in Figure 2.5A) has two transition points at different positions, one in the red sector and one in the white sector. In contrast to these relatively simple events, the sectored colony 18A (Figure 2.5A) has six transitions, one in the red sector and five in the white sector. In our analysis, if the transition point is identical in both sectors, it is only counted once. In Table S 2.3, we also assign a class (A to L) for all events. In Table S 2.1 and Table S 2.2, each event class is diagrammed using the same approach employed in Figure 2.5A. In these figures, we also indicate the number of events in each class, and which supplementary figure (Figs. S1-S40 of ST CHARLES *et al.* (2012)) shows the pattern of DNA repair events consistent with the specific conversion event. The same methods are used to describe the recombination events induced by DNA damage as were used to

describe the spontaneous events. Multiple transitions within conversion tracts could reflect "patchy" repair of mismatches within long heteroduplexes (discussed further below) or template switching between homologues. For the spontaneous conversion events in Figure 2.5A, we did not find a correlation between SNP density and the number of transitions within the tract (r^2 =0).

2.3.2 Analysis of LOH events by SNP arrays in cells treated with γ rays

We analyzed PG311 sectors that were generated in a previous study (LEE and PETES 2010) by treatment of G1-arrested cells with 100 Gy of γ radiation, followed by selection of red/white sectored colonies on canavanine-containing plates; this dose of γ rays elevated the frequency of sectoring about 26-fold. All of the colonies examined had a crossover on chromosome V. We analyzed seven of these sectored colonies with SNP microarrays, and two of these were also examined by HTS. The SNP array data are shown in Table S 2.4 with depiction of the recombination events in Table S 2.1 and Table S 2.2.

The positions of the selected crossovers and associated gene conversion events in the *CEN5-can1-100/SUP4-o* interval are shown in Figure 2.5B. Our mapping of these events by SNP arrays is in reasonably good agreement with our PCR-based mapping method (LEE and PETES 2010). All of the conversion events had at least one SNP that was homozygous on both sides of the sector (4:0 conversion) as expected if the

recombination events were a consequence of repair of two sister chromatids broken at the same position (Figure 2.2 and Figure 2.3).

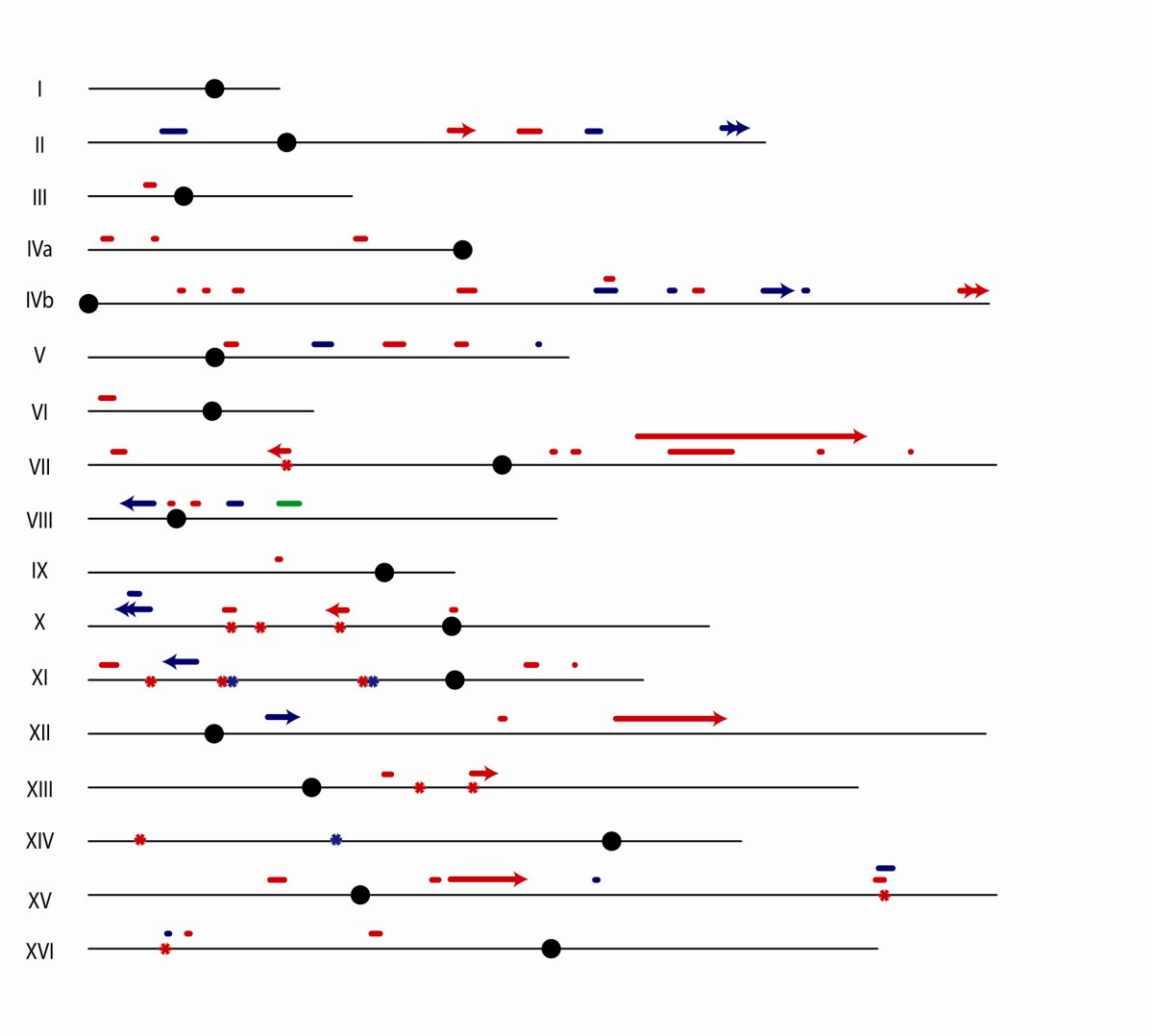


Figure 2.6. Genomic locations of unselected recombination events and *de novo* mutations in untreated cells and in cells treated with UV or γ rays as determined by SNP microarrays and HTS.

The horizontal black bars depict each chromosome and are proportional to the chromosome length except for chromosome XII. The black circles depict the centromere of each chromosome. Short horizontal bars above each chromosome depict conversion events unassociated with crossovers and the length of each bar is approximately proportional to the length of the conversion tract. All conversion tracts

are shown as single solid lines without regard to the complexity of the event (for example, transitions between 4:0 and 3:1). Single arrowheads depict reciprocal crossovers and double arrowheads depict BIR events. Asterisks located on the chromosome indicate the approximate positions of mutations induced by UV or γ rays; two of the mutations (located at SGD coordinates 171529 and 301552 on X) are in regions of LOH. Events observed in untreated cells, cells treated with UV, and cells treated with γ rays are shown in green, red, and blue, respectively. None of the events selected on the left arm of V are shown in this figure.

In addition to the selected events, from our genome-wide analysis, we observed 17 unselected events on other chromosomes among the seven colonies: four crossovers associated with conversion (Table 2.2), eleven conversions that were not associated with crossovers, and two break-induced replication (BIR) events (Table 2.3). Since the frequency of unselected crossovers in unirradiated samples is very low (less than 0.1/cell), it is likely that all of the events in the irradiated cells reflect the repair of γ ray-induced DNA damage. The locations of these unselected events are shown as blue symbols in Figure 2.6. The events appear randomly distributed in this small data set. The SNP arrays for radiation-induced unselected crossovers and associated conversions have patterns similar to the selected crossover shown in Figure 2.4. In addition, we observed many conversion events unassociated with crossovers (Figure 2.7). This figure shows at low and high resolution a 0:4 conversion event in which both sectors have gained YJM789 SNPs and lost W303a SNPs. This pattern could represent an event that occurred prior to radiation. However, since such events were observed commonly in irradiated cells but not in control diploids, we assume that most (or all) were induced by γ rays.

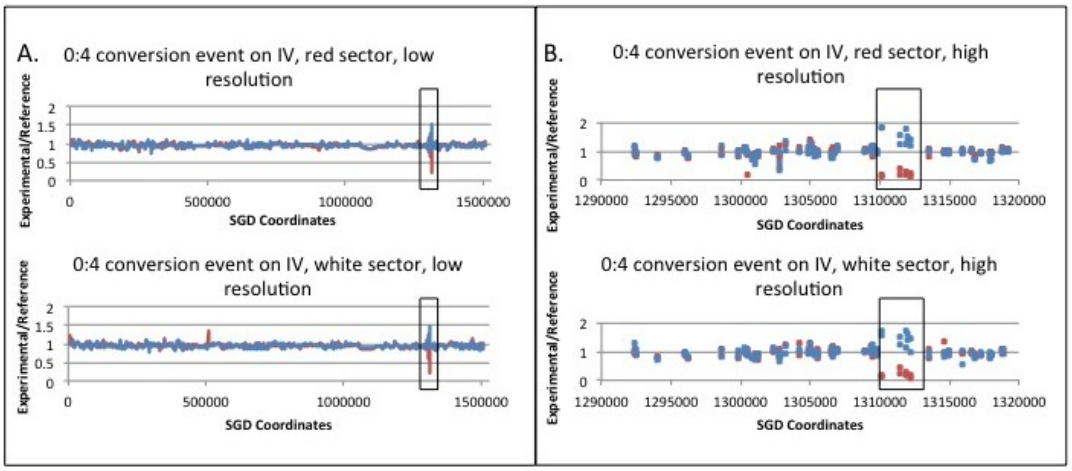


Figure 2.7. SNP array analysis of a gene conversion event unassociated with crossover.

In cells treated with γ rays, one of the canavanine-resistant red/white sectored colonies (43RW) had an unselected gene conversion event on chromosome IV. As shown at low- (Figure 2.7A) and high- (Figure 2.7B) resolution, both sectors had an LOH region in which YJM789-derived SNPs became homozygous (0:4 conversion event). The depiction of the SNP array data is the same as in Figure 2.4. The length of the conversion tract is about 3 kb. *CEN4* is located approximately at SGD coordinate 450000.

Since the red and white sectors are produced by the two daughter cells resulting from the division of radiation treatment of a G1-synchronized cell, the analysis of genomic DNA from both sectors gives valuable mechanistic information even for unselected events. For example, if we observed an interstitial LOH event by examining only the white sector, we would not know whether this event was a consequence of a 3:1 conversion, a 4:0 conversion, or a two-strand double crossover. This ambiguity can be resolved by examining genomic DNA from the red sector.

Table 2.3 Summary of all unselected conversion events unassociated with LOH or BIR events as diagnosed by SNP microarrays.

Type of event	Spontaneous	γ-ray	UV
3:1 or 1:3 conversions	0	4	9
4:0 or 0:4 conversions	1	4	12
Hybrid conversions[2]	0	2	9
Complex conversions	0	1	3
BIR	0	2	1
Totals	**1**	**13**	**34**

We observed two sectored colonies that had BIR events. In single BIR events (such as Class L2 in Table S 2.1), one sector has an LOH event that extends from an internal site on the chromosome to the telomere, whereas the other sector is heterozygous for the same SNPs. In double BIR events, both sectors have LOH events extending from an internal site to the telomere (Class L1 in Table S 2.1). Interestingly, in the colony with the single BIR event, there is a conversion event on the chromosome that was originally the sister chromatid of the one involved in BIR. This result argues that both sister chromatids had DSBs at approximately the same position. The molecular interactions required to produce Classes L2 and L3 are shown in Figure 2.12.

All of the selected and unselected reciprocal crossovers induced by γ rays were associated with conversion tracts. The median length of all conversion tracts (both associated and unassociated with crossover) was 12.9 kb (95% confidence limits of 5.2-20.4 kb). The median lengths of conversion tracts associated and unassociated with

[2] 3:1/4:0, 3:1/4:0/3:1, 1:3/0:4, or 1:3/0:4/1:3 conversion events.

crossovers were 18.4 kb (10.8-25.3) and 8.4 kb (2.6-13.3), respectively. By the Mann-Whitney test, the median lengths of crossover-associated and crossover-unassociated γ-ray-induced conversion tracts were significantly different (p=0.01).

2.3.3 Analysis of LOH events by SNP arrays in cells treated with UV

G1-synchronized PG311 cells were treated with a UV dose of 10-15 J/m^2. This dose resulted in no significant loss of viability but stimulated the frequency of sectors by about 1000-fold. We examined three sectored colonies by SNP arrays and two of these colonies were also analyzed by HTS. In addition to the selected crossover on chromosome V, among the three colonies, there were seven unselected crossovers, 33 unselected conversion events, and one BIR event (Table 2.2 and Table 2.3). Thus, there were about fourteen unselected LOH events per UV-treated cell. The locations of selected chromosome V events and the unselected LOH events are shown in Figure 2.5 and Figure 2.6, respectively. The UV-induced LOH events are distributed fairly evenly throughout the genome (Figure 2.6). As observed for the γ ray-induced BIR events, the UV-induced BIR event is located close to the telomere. The detailed information about breakpoints in UV-treated cells is shown in Table S 2.6 and Table S 2.1 and Table S 2.2.

All of the crossover events had an associated conversion. In most of the conversion events, there was at least one SNP with the 4:0 or 0:4 pattern, suggesting that UV-induced damage in G1 may result in DSBs (Figure 2.2). The median length of all UV-induced conversion tracts was 9.2 kb (6.5-10.3), whereas the median lengths of

conversion tracts associated and unassociated with crossovers were 10.3 (7.0-18.9) and 7.5 (4.5-10.2) kb, respectively. Although the crossover-associated conversions are longer than the crossover-unassociated conversions, this difference is not significant (p=0.08 by Mann-Whitney test).

2.3.4 HTS analysis of LOH in γ ray- and UV-treated cells

Since G1-arrested yeast cells treated with 100 Gy of ionizing radiation have about 35 DSBs per cell (LEE and PETES 2010), the average number of unselected LOH events per cell (two to three) indicates most events must be repaired by mechanisms that do not generate LOH. An alternative possibility is that a substantial fraction of the events have short conversion tracts that are not detectable by the SNP arrays. Since HTS can detect LOH events for all of the 55,000 SNPs existing in the diploid strain, rather than the 13,000 SNPs represented on the SNP array, we sequenced genomic DNA samples from both red and white sectors of two sectored colonies of γ ray-treated samples (PG311-GR-37R/W and PG311-GR-40R/W) and two colonies of UV-treated samples (PG311-UV-8R/W and PG311-UV-9R/W). The details of this analysis are described in Supporting Information. All of the LOH events detected by SNP arrays were also found by HTS. LOH events that had not been previously detected by the SNP arrays were confirmed by re-sequencing the relevant PCR fragment. The patterns of LOH as identified by HTS in the γ ray- and UV-treated samples are in Table S 2.5 and Table S 2.7.

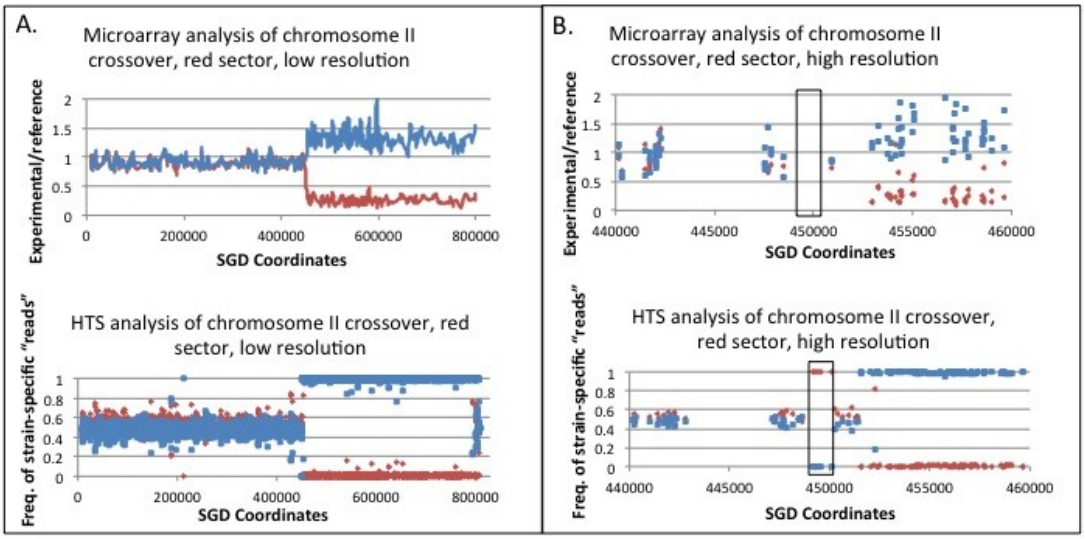


Figure 2.8. Analysis of the same recombination event by both SNP arrays and HTS.

This figure shows the analysis of the unselected recombination event on chromosome II in the red sector of the UV-induced sectored colony 8. Our standard SNP array analysis (top parts of Figure 2.8A and Figure 2.8B) showed a single transition between heterozygosity and homozygosity at about SGD coordinate 452000. In the bottom panels of the figure, we show HTS data for the same genomic sample. For the HTS data, the Y axis represents the frequency of YJM789 SNP (blue) or W303a SNP (red) "reads" for the experimental sample when assembled to the PSL2 genome. For heterozygous regions, there should be approximately equal frequencies of the two types of SNPs. It is clear from the high-resolution depictions of the HTS data that there is a short LOH region (boxed) located near SGD coordinate 450000 that was not detected by the SNP arrays. This region was not detected because oligonucleotides containing these SNPs were not present on the array. In the low-resolution depiction of the HTS data, within the LOH region, there is a small region near SGD coordinate 800000 in which SNPs appear to be heterozygous. These signals are artifacts based on "reads" from the repeated diverged *MAL* and *MPH* genes that were incorrectly mapped by the genome analysis software to chromosome II. *CEN2* is located near SGD coordinate 238000.

Figure 2.8 shows a comparison between SNP microarray and HTS data for an LOH event on chromosome II in a UV-treated sample (PG311-UV-8R). The SNP microarray indicates that a transition between heterozygous and homozygous SNPs

occurs somewhere between SGD coordinates 450919 and 452926, whereas the HTS data refine the mapping of this transition between SGD coordinates 451337 and 451581. Additionally, the HTS data showed that the recombination event was more complex than indicated by the microarray data. By HTS, we found that the region between coordinates 448628 and 450279 had undergone LOH; this event was not detected by microarrays because there were no oligonucleotides between 448488 and 450919 on the microarray. A summary of the comparison of data from SNP arrays and HTS for the same samples is given in Table 2.4.

Although more LOH events were observed with HTS than with SNP arrays, the difference was not large. For example, in the two UV-treated samples, we observed 32 LOH events by SNP arrays and 35 events by sequencing. In the γ ray-treated samples, we found five events by SNP arrays and six events by sequencing. Since more than 80% of the events detected by HTS were also detected by microarrays, it is unlikely that our estimates of LOH events are substantially affected by a high frequency of gene conversion events with short conversion tracts. We cannot rule out, however, the possibility of gene conversion events with very short (<100 bp) tracts. As shown in Table 2.4, a number of the gene conversion tracts analyzed by HTS were more complex than the same tracts examined by the SNP arrays. The frequencies of complex tracts, as determined by HTS and SNP microarrays, were 0.37 and 0.22, respectively. Despite the differences in the numbers and types of LOH events detected by HTS and the SNP

microarrays, it is clear that most of the LOH events are detectable by the SNP microarrays.

Table 2.4 Comparison between recombination events detected by high-throughput sequencing or by SNP microarrays.

Type of event	γ -ray		UV	
	Array	HTS	Array	HTS
Crossovers with no detectable conversion tracts	0	0	0	0
Crossovers with 3:1 or 1:3 conversion tracts	0	0	2	0
Crossovers with 4:0 or 0:4 conversion tracts	1	0	0	0
Crossovers with hybrid conversion tracts	0	1	1	1
Crossovers with complex conversion tracts	1	1	6	8
3:1 or 1:3 conversion tracts without crossovers	1	0	8	9
4:0 or 0:4 conversion tracts without crossovers	1	1	5	4
Hybrid conversion tracts without crossovers	0	1	8	7
Complex conversion tracts without crossovers	1	2	1	5
BIR	0	0	1	1
Total recombination events	**5**	**6**	**32**	**35**[3]

2.3.5 HTS analysis of mutations induced in γ ray- and UV-treated cells

About 99.5% of the bases between the two homologues of PG311 are identical. The HTS data generated for the same colonies examined for LOH events were analyzed for radiation-induced mutations. We analyzed both sectors of two sectored colonies induced by UV and two sectored colonies induced by γ rays for new mutations (Table S

[3] This table includes data from two UV-induced sectored colonies and two sectored colonies induced by gamma radiation that were analyzed by both SNP arrays and HTS. The table includes both selected recombination events on chromosome V and unselected events on other chromosomes. All conversion events unassociated with LOH were unselected.

2.8). There were three and twelve *de novo* point mutations detected in the γ ray- and UV-treated samples, respectively.

All three of the mutations induced by γ rays and six of the twelve mutations induced by UV were in both red and white sectors of the colony. The presence of the mutation in both sectors indicates that the mutation induced in G1 by the radiation was represented in both strands of the duplex prior to replication. The mutations in the UV-treated cells that were present in only one sector could reflect a mutant base in only one of the two strands. These two types of events have been observed previously in UV-treated cells (ECKARDT and HAYNES 1977; JAMES and KILBEY 1977). Nine of twelve of the UV-induced mutations and two of three of the γ ray-induced mutations were transitions. In much more extensive study of spontaneous and UV-induced mutations at the *SUP4-o* locus (KUNZ *et al.* 1987), spontaneous mutations had a ratio of transitions:transversions of 4:6, whereas UV-induced mutations were biased toward transitions (4:1).

Most (11 of 15) of the induced point mutations were located within genes rather than between genes (Table S 2.8). By Chi-square analysis, the distribution of mutations throughout the genome is non-random ($p = 0.002$). Five of the fifteen mutations are located on the left arm of XI and, remarkably, two mutations (one induced by γ rays and one induced by UV) are within the *NUP120* gene. By similar methods used to detect new base substitutions, we failed to detect any insertion/deletion mutations (in/dels) in

the eight sequenced samples. It should be pointed out, however, that detection of in/dels in HTS data with short-paired reads is challenging, particularly in a diploid that is heterozygous for many pre-existing in/dels.

2.4 Discussion

In this study, we mapped both selected and unselected mitotic recombination events in a genome-wide analysis. Most of the events were mapped using only SNP microarrays, but four events were examined by both SNP microarrays and HTS. To our knowledge, this study is the first to examine spontaneous and DNA damage-induced LOH events throughout the yeast genome. The conclusions from this study are: 1) the gene conversion tracts analyzed by SNP arrays and HTS were often more complex than inferred from our earlier lower-resolution mapping studies (LEE *et al.* 2009; LEE and PETES 2010), 2) doses of radiation that result in little or no loss of cell viability in G1-synchronized diploid cells resulted in multiple unselected LOH events, and 3) the same doses of γ rays and UV used in the LOH study result in very low levels of *de novo* mutations. In addition, we conclude that, although HTS has four-fold better resolution than SNP microarrays, the SNP arrays detect most of the same events identified by HTS. These conclusions will be further discussed below.

2.4.1 Lengths of gene conversion tracts

The median lengths (95% confidence limits shown in parentheses), as measured by SNP arrays, of gene conversion tracts associated with crossovers for spontaneous,

UV-induced, and γ ray-induced events were 6.1, 10.3, and 18.5 kb, respectively. As we have observed previously (LEE *et al.* 2009), mitotic gene conversion tracts are substantially longer than meiotic conversion tracts (MANCERA *et al.* 2008). In addition, for the γ ray-induced conversion events, the conversion tracts associated with crossovers were significantly longer than those unassociated with crossovers as has been observed previously (AGUILERA and KLEIN 1989; MANCERA *et al.* 2008). One simple explanation of this observation is that conversion events unassociated with crossovers usually are a consequence of SDSA, and such events might involve limited processing of the broken chromosome ends. In contrast, crossovers likely proceed through formation of a double Holliday junction. Formation of this dHJ intermediate may require more extensive processing of broken DNA ends and/or more extensive DNA synthesis primed from the invading end. It is also possible that branch migration of the dHJ intermediate could extend the length of the heteroduplex associated with the crossover; this possibility will be further discussed below.

2.4.2 Recombinogenic DNA damage

Although it is clear from a variety of studies that DSBs stimulate mitotic recombination, the DNA lesion responsible for initiating spontaneous recombination events is not certain. We previously showed that about half of mitotic crossovers on chromosome V are associated with a gene conversion tracts that are exclusively 4:0 or 0:4 or that have a region of 4:0 or 0:4 (hybrid tracts). Such conversion tracts indicate that

both sister chromatids have breaks at approximately the same position and one simple mechanism consistent with this property is that these spontaneous conversion events reflect a DSB formed in G1 of the cell cycle. Supporting this conclusion, many (44%) of the conversion events induced by γ rays in G1 of the cell cycle have regions of 4:0 or 0:4 segregation, whereas none of the conversion events induced by γ rays in G2 had this pattern (LEE and PETES 2010). Among the mechanisms that could produce the spontaneous lesions required to initiate recombination are: 1) closely-opposed nicks generated during removal of adducts caused by oxidative DNA damage, 2) DSBs caused by Top2p or other cellular endonucleases, 3) lesions resulting from collisions between replication forks and the transcription machinery, 4) DSBs resulting from the collision of converging replication forks, and 5) nuclease-dependent processing of secondary DNA structures ("hairpins" and cruciforms). Thus far, we have been unable to associate the spontaneous recombination events with any of these mechanisms. For example, the two positions that represent the convergence of replication forks on the left arm of chromosome V (FACHINETTI *et al.* 2010) are not hotspots for recombination in our limited dataset. If there are several different mechanisms that can produce recombinogenic DNA lesions, we will need to map many events to detect significant associations.

In our current analysis of γ ray-induced gene conversion events in G1-synchronized cells by SNP microarrays, we found that ten of the eleven conversion tracts associated with crossovers had a 4:0 or 0:4 segment, and eight of the eleven

conversion tracts that were unassociated with crossovers had such a segment. This observation is consistent with the possibility that most of the observed recombination events in IR-treated cells reflected a DSB introduced by γ rays in G1.

The recombinogenic effect of UV-induced DNA damage is less clear. One possibility is that small gaps resulting from the removal of UV-induced dimers are the recombinogenic lesion. GALLI and SCHIESTL (1999) found that UV did not stimulate mitotic recombination between direct repeats (single-strand annealing) in G1-arrested cells unless the cells were allowed to enter the S-period after the UV treatment. In contrast, when G1-arrested cells were treated with IR, stimulation of single-strand annealing was observed without requiring the cells to enter S. If the recombinogenic DNA lesion is a DSB, the likely explanation of the different results is that IR directly creates DSBs whereas the repair of UV lesions results in nicks that result in DSBs when the nicked molecule is replicated (GALLI and SCHIESTL 1999). By this explanation, it is surprising that many of the conversion events induced by UV in G1 in our experiments had 4:0 or 0:4 segments, suggesting that these exchanges were a consequence of a G1-stimulated DSB. Such a DSB could be generated in G1 if the removal of dimers on opposite DNA strands resulted in a very short (<10 bp) duplex region separating the 30 bp gaps. Based on previous estimates of the number of dimers induced by 20 J/m^2 of UV (DAIGAKU *et al.* 2010), we calculate that there are about 7500 dimers/diploid genome induced by a UV dose of 15 J/m^2 and, based on a Poisson distribution, there would be

about 35 regions per genome in which two dimers are on opposite strands within 75 base pairs of each other. Since the DSB would require two closely-opposed gaps rather than two closely-opposed dimers, the kinetics of gap formation and gap repair affect the probability of DSB formation by this mechanism. Another complicating factor is that the frequency of closely-opposed dimers is higher than expected if dimer formation is random (REYNOLDS 1987).

A second possibility is that DSB formation is initiated by gaps on opposite strands that are relatively close together ($\leq$ 500 bp apart), but too far apart to generate a G1 DSB. If a DNA molecule with such gaps is replicated, the product would be two sister chromatids with DSBs located 500 bp or less apart. Processing of the broken ends to yield single-stranded DNA regions 500 bp or greater would preclude formation of a dHJ involving the two sister chromatids. Thus, such molecules would likely be repaired using the intact homologue as the template, generating a 4:0 conversion. This model is consistent with the Galli and Schiestl interpretation. We calculate that cells irradiated with 15 J/m^2 would have about 234 dimers on opposite strands within 500 bp of each other.

An alternative possibility is that recombination events are a consequence of DSBs occurring at replication forks stalled at unexcised pyrimidine dimers. Unrepaired UV-induced damage has been demonstrated to block replication forks and replication of such damaged templates promotes sister-chromatid recombination (BRANZEI and FOIANI

2010). Although we cannot exclude this model, the observed UV-induced 4:0 events would require that the replication fork stalled at the dimer result in two broken sister chromatids, perhaps by increasing the probability of a replication fork collision. It should be emphasized that although UV very efficiently stimulates crossovers between homologues, most of the UV-induced recombination events are likely to represent sister-chromatid interactions (KADYK and HARTWELL 1992).

2.4.3 Relationship between mitotic gene conversion and crossovers

In meiosis in yeast, about half of conversion events are associated with crossovers (MANCERA *et al.* 2008; PETES *et al.* 1991). In our previous mitotic studies, we selected crossovers and found that most (>80%) of these events were associated with an adjacent tract of gene conversion (LEE *et al.* 2009; LEE and PETES 2010); conversion events unassociated with crossovers could not be selected with our system. In the current study, for unselected events stimulated by radiation, we can estimate the fraction of conversion events that are associated with crossovers.

For IR-treated cells, including all non-selected events except BIR, we found four conversions associated with crossovers and eleven conversions unassociated with crossovers (Table 2.2 and Table 2.3). In these experiments, we detect the associated crossover because it generates LOH from the conversion tract to the end of the chromosome. As shown in Figure 2.2, depending on the pattern of chromosome segregation, we expect that only half of the crossovers will lead to LOH of markers distal

to the point of exchange, and this expectation has experimental support (CHUA and JINKS-ROBERTSON 1991). In addition, as discussed in the Supporting Information, we found preliminary evidence in our experiments for conversion events associated with crossovers that did not result in LOH.

Thus, we calculate that of the fifteen conversion events induced by γ rays, it is likely that eight were associated with crossovers (53% association). Similarly, among UV-induced recombinants, since there were seven unselected conversions associated with crossovers and 33 unselected conversions not associated with LOH (Table 2.2 and Table 2.3), we calculate that about 35% of the UV-induced conversion events are associated with crossovers. Our conclusion that the frequency of crossovers associated with conversions is not very different for mitotic and meiotic conversion events is consistent with other recent studies (HO *et al.* 2010). In yeast studies in which conversion events are limited in size, the association between conversion and crossovers is weaker (PÂQUES and HABER 1999). Also, in Drosophila and mammalian cells, conversion events are only rarely associated with crossovers (ANDERSEN and SEKELSKY 2010).

2.4.4 Complex gene conversion tracts and BIR events

Previously, we classified conversion tracts as "simple" if the markers within the tract had one of the following patterns: 1) all markers were 3:1 or 1:3 (not mixtures of 3:1 and 1:3 in same tract), 2) all markers were 4:0 or 0:4, or 3) hybrid tracts of the form 3:1/4:0, 1:3/0:4, 3:1/4:0/3:1, or 1:3/0:4/1:3. All such tracts can be explained as the

consequence of the repair of one or two broken chromatids by one of the HR pathways shown in Figure 2.1. There were, however, conversion tracts that were more complicated (termed "complex tracts"), which will be described below. In the UV-treated samples, six of ten of the crossover-associated conversion tracts were complex, although only three of thirty-three tracts were complex in conversions unassociated with crossovers (Table 2.2 and Table 2.3); this difference is significant ($p<0.01$) by the Fisher exact test. In the IR-treated samples, the conversion events associated with crossovers were usually more complex than those that were not (Table 2.2), although the difference was not significant. MANCERA *et al.* (2008) reported that 11% of meiotic crossovers had complex conversion tracts, whereas the frequency of complex tracts among conversions unassociated with crossovers was 3%. One explanation of this difference could be that crossovers that proceed through the pathway shown in Figure 2.1B are associated with two regions of heteroduplex, while conversions resulting from SDSA or dHJ dissolution have only a single region of heteroduplex (Figure 2.1A and C). Second, because gene conversion tracts associated with crossovers are usually longer than those unassociated with crossovers, there may be a greater chance to observe patchy repair of mismatches (as defined below) in tracts associated with crossovers.

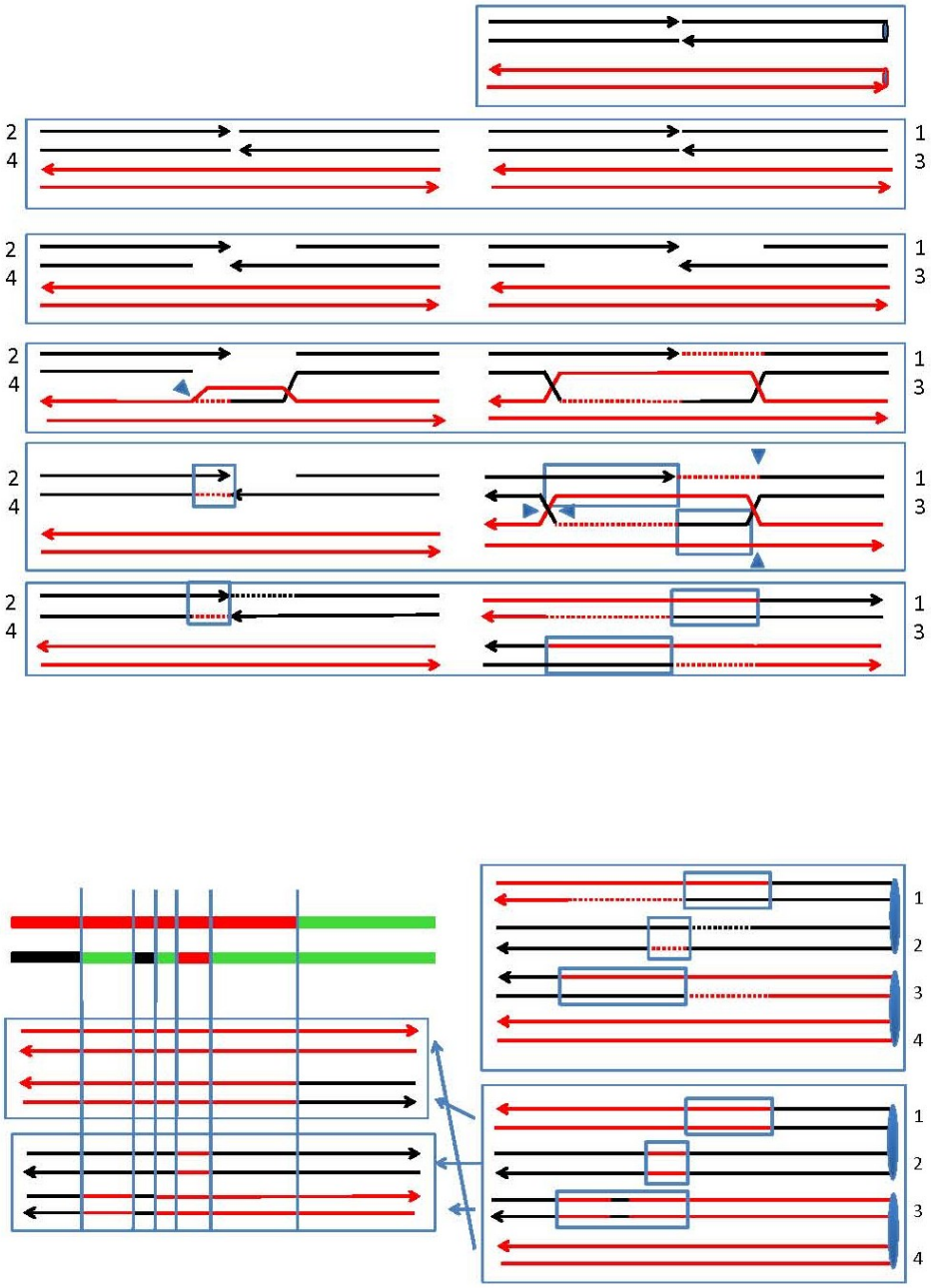


Figure 2.9. Description of the Class J8 event.

In this event, the crossover is associated with a complex conversion tract. This event can be explained as a consequence of the repair of two DSBs, one by the DSBR pathway and one by the SDSA pathway. In addition, two of the heteroduplexes have "patchy" repair of mismatches.

Diagrams of all recombination events in our study are shown in Table S 2.1 and Table S 2.2, and figures showing the patterns of DSB repair required to produce the recombination events are shown in Figs. S1-S40 of ST CHARLES *et al.* (2012). Most of the complex conversion tracts could be divided into two categories: those that had multiple transitions between 3:1, 4:0, and heterozygosity within the tract and those tracts in which 3:1 and 1:3 or 4:0 and 0:4 segments occurred within one tract. Examples of conversion tracts with multiple transitions are strains 18A (Class J9, Table S 2.1) and 4.1 (Class J8, Table S 2.1); both 18A and 4.1 are also depicted in Figure 2.5A. The complex tract in 4.1 is consistent with the repair of two DSBs with "patchy" repair of mismatches in two of the resulting heteroduplexes (Figure 2.9). Heteroduplexes will often contain multiple mismatches that can be repaired to produce either a conversion event or a restoration event (KIRKPATRICK *et al.* 1998). For example, in Figure 2.1A, repair of the heteroduplex resulting in a duplex with two "red" strands would represent conversion-type repair, since this pattern produces 3:1 segregation; repair of the mismatch to produce a duplex with two "blue" strands represents restoration-type repair, since this pattern generates two cells that retain heterozygosity at the position of the original heteroduplex. Although multiple mismatches within one heteroduplex are generally converted in a concerted manner yielding a continuous conversion tract, tracts with mixtures of conversion-type and restoration-type repair have been detected in both

meiosis (MANCERA *et al.* 2008; SYMINGTON and PETES 1988) and mitosis (MITCHEL *et al.* 2010; NICKOLOFF *et al.* 1999).

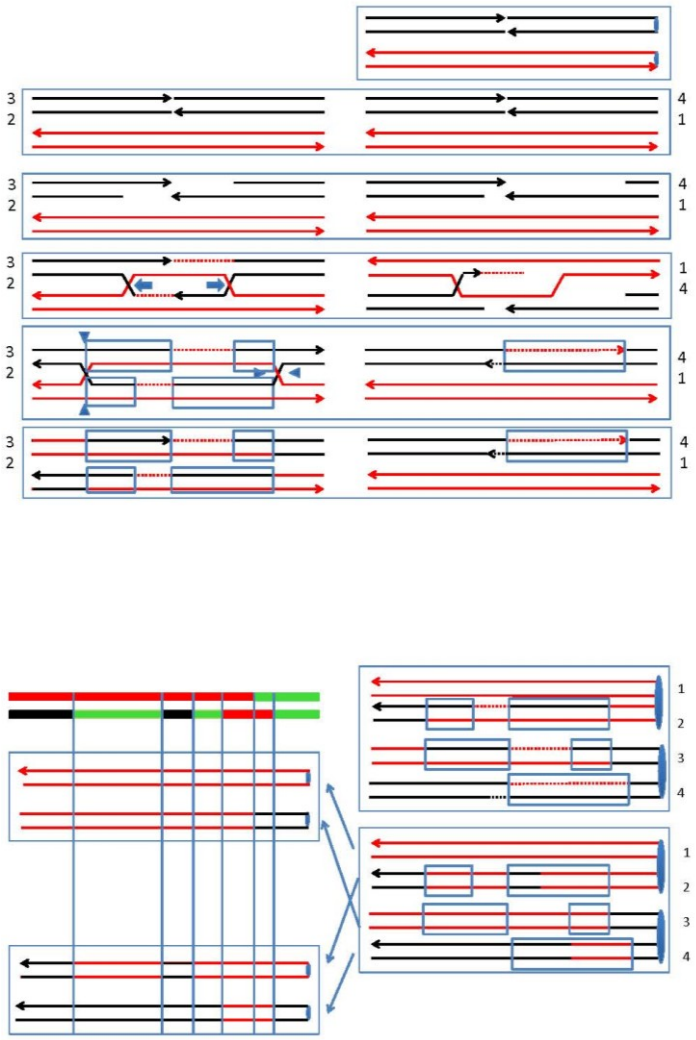


Figure 2.10. Description of the Class J9 event.

In this event, the crossover is associated with a complex conversion tract. This event can be explained as a consequence of the repair of two DSBs, one by the DSBR intermediate, and one by the SDSA pathway. In addition, we postulate branch migration of the dHJ intermediate, and that two of the heteroduplexes have "patchy" repair of mismatches.

The pathway of DSB repair shown to explain the pattern of markers in the strain 18A conversion event (Figure 2.10) invokes patchy repair and branch migration of the dHJ. During recombination in *E. coli*, a Holliday junction can be translocated by branch migration, resulting in symmetric heteroduplexes (WEST 1997). Although genetic evidence argues against the formation of symmetric heteroduplexes during meiotic recombination in *S. cerevisiae* (PETES *et al.* 1991), symmetric heteroduplexes have been invoked previously to explain certain classes of mitotic gene conversions (ESPOSITO 1978; NICKOLOFF *et al.* 1999; ROITGRUND *et al.* 1993). Branch migration can also generate patterns of repair in which a single DSB can produce both 3:1 and 1:3, or 4:0 and 0:4 events as shown in Figure 2.11.

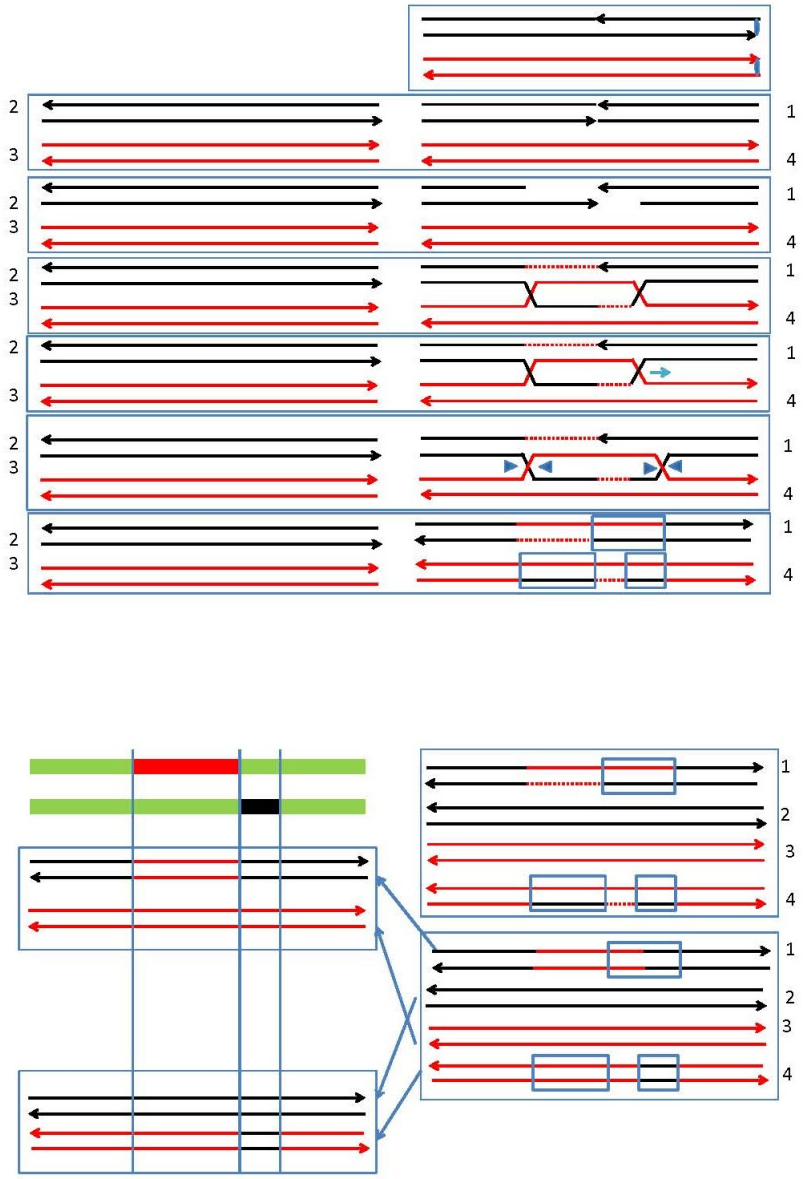


Figure 2.11. One mechanism to explain Class G1.

The Class G1 event has adjacent conversion tracts of 3:1 and 1:3 unassociated with a crossover. In this model, the event is produced by repair of mismatches in symmetric heteroduplexes produced by branch migration (DSBR pathway).

Our data do not allow us to determine unambiguously the pathways required to generate the observed conversion tracts. However, we can state that many of the

complex tracts are inconsistent with the simplest form of the recombination models shown in Figure 2.1. In particular, it is likely that patchy repair of mismatches is a relatively common feature of mitotic gene conversion tracts. A detailed discussion of all of the conversion tracts in our studies is given in the Supporting Information.

There were three unselected BIR events observed in our study (Class L of Table S 2.1). For two of the three events, we observed a region of conversion associated with the BIR event. This pattern is consistent with the repair of two DSBs, one by SDSA and one by BIR (Figure 2.12). The BIR events were about three-fold less frequent than unselected crossovers, as expected from previous studies (HO *et al.* 2010; MCMURRAY and GOTTSCHLING 2003).

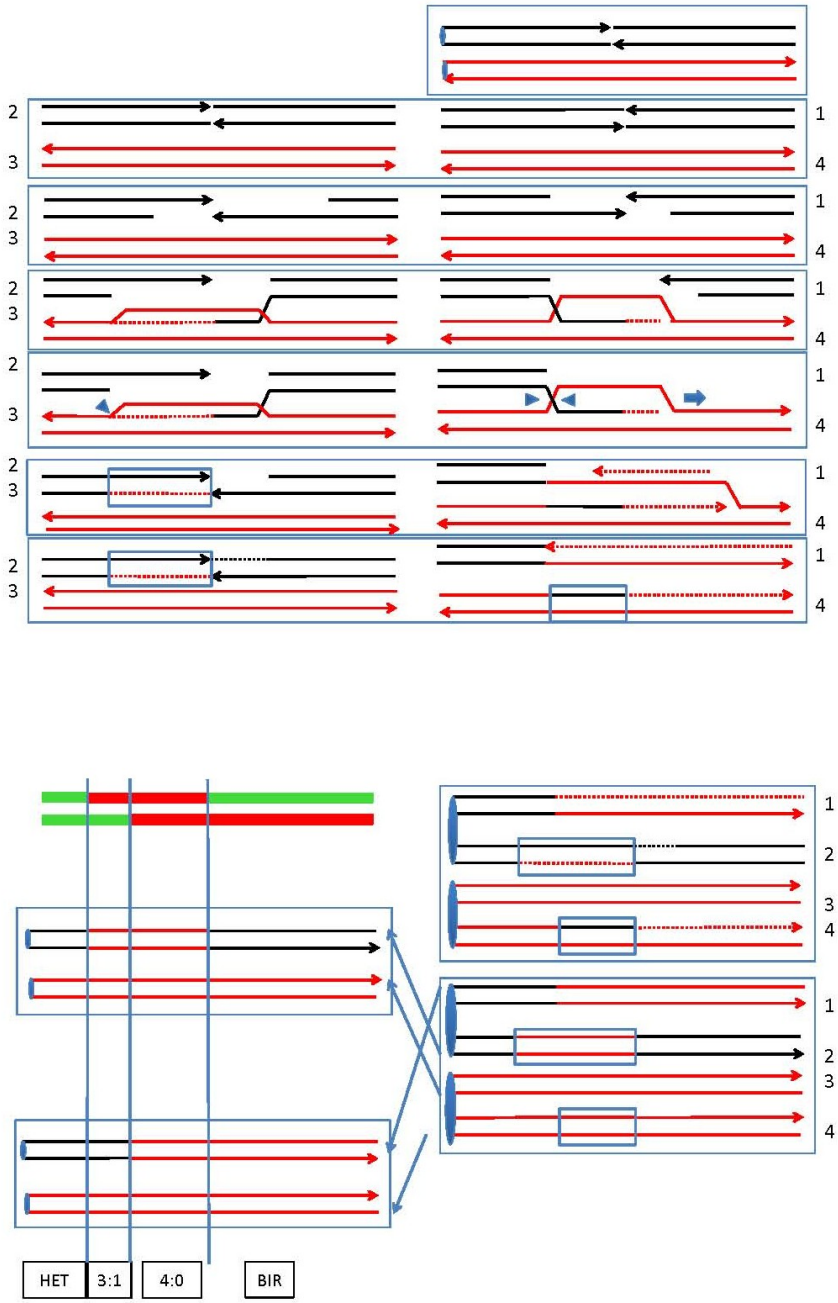


Figure 2.12 Description of the L2 and L3 events (listed in Table S 2.1).

In these events, there are hybrid conversion tracts associated with a BIR event. This event can be explained as a consequence of the repair of two DSBs, one by the SDSA pathway and the second by BIR. There are heteroduplexes associated with both repair events, and these heteroduplexes have different lengths.

2.4.5 Relationship between the level of DNA damage and the frequency of LOH events

The 100 Gy dose of IR used in our experiments is expected to produce about 35 DSBs/diploid genome (LEE and PETES 2010). Since we observed only 2.4 LOH events/irradiated cell, most of these DSBs must be repaired by a mechanism that does not produce a detectable LOH event. Since the cells in our experiments were irradiated in G1, the DSBs must have been repaired either by an interaction with the homologous chromosome or by NHEJ. We suggest several possible explanations. First, it is possible that the repair of the DSB frequently involves an interaction with the homologue that is associated with a very short conversion tract. Tracts shorter than 50 bp would be rarely detected, even by HTS. Such a repair event would likely involve very limited processing of broken DNA ends as well as short excision repair tracts. A system of short-patch (often less than 12 base pairs) mismatch repair that is independent of the classical mismatch repair system in *S. cerevisiae* was described by COIC *et al.* (2000), although the genes involved in this type of repair have not been identified. In addition, conversion tracts less than 53 bp have been detected among HO-induced events (PALMER *et al.* 2003). A related possibility is that gene conversion events occur non-randomly in regions of the genome that are not represented on our microarrays (regions that are identical between W303a and YJM789 or regions with repeated genes). A third possibility is that the repair of the DSB is associated with restoration-type repair of mismatches within the heteroduplexes. Since most of the crossovers in our study are

associated with detectable gene conversion tracts, we would have to hypothesize that conversion events that are not associated with crossovers are much more prone to restoration-type repair than conversion events that are associated with crossovers. A fourth possibility is that the IR-induced DSBs are frequently (DALEY *et al.* 2005) repaired by NHEJ events. Although NHEJ events are repressed in *MAT***a**/*MATα* diploids, since PG311 lacks the *MATα* locus, NHEJ events will occur. Although NHEJ events will not produce LOH, depending on the nature of DNA ends (compatible single-strand overhangs or blunt), some NHEJ events would be expected to result in loss or gain of a few base pairs . Although we did not observe in/dels in our HTS analysis, this observation does not rule out the possibility that some repair events reflect NHEJ. It is also possible, of course, that all four possibilities described above are partly responsible for the "missing" LOH events.

One explanation that we can exclude as a major contributor to the discrepancy between the number of lesions and the number of LOH events is chromosome loss. Chromosome loss can be readily detected by the SNP microarrays, and no losses were observed in cells treated with γ rays or UV. In experiments in which eight-fold higher doses of γ rays were used, about 10% of the treated yeast cells had chromosome loss (ARGUESO *et al.* 2008).

Although we detected more than 50 unselected LOH events in cells treated with γ rays and UV, no duplications or deletions were detected. Thus, SNP arrays that can

detect both LOH and changes in copy number are a much more efficient method of detecting recombinogenic DNA lesions than comparative genomic hybridization (CGH) arrays. In our previous analysis of γ ray-treated diploid cells by CGH (ARGUESO *et al.* 2008), we found that most of the irradiated cells had one or more chromosome rearrangements, usually non-reciprocal translocations with retrotransposons at the breakpoints. In these experiments, we treated G2-synchronized cells with doses of radiation that were eight-fold higher than the doses used in our current study.

2.4.6 Mutations induced by γ rays and UV

We found only a few mutations induced by γ rays and UV, averages of 1.5 and 6 mutations/irradiated cell, respectively. Although there are no genome-wide studies of the frequencies of mutations induced by γ rays, extrapolating from the frequency of induction of X-ray-induced mutations at the *CAN1* locus (GOCKE and MANNEY 1979) and the rate of spontaneous mutations per base pair at *CAN1* (LANG and MURRAY 2008), we calculate that the expected frequency of mutations per genome is about two/diploid cell, close to our observed number. The most direct comparison for the UV-induced mutations is with data obtained from HTS of UV-treated stationary-phase haploid yeast strains (BURCH *et al.* 2011). These strains in these studies had a temperature-sensitive mutation in *CDC13*. However, three of the sequenced isolates were treated at the permissive temperature. By extrapolating their data to our UV dose, we would expect about 14 mutations/diploid cell, only two-fold different from our observed frequencies.

In summary, our HTS data detected roughly the expected number of mutations per irradiated strain.

As described in the Results, the mutations induced by UV and γ rays are non-randomly distributed among the yeast chromosomes. Although this non-random distribution needs to be verified with a large dataset, it is possible that the mutagenic DNA damage is distributed non-randomly because of the specific position of different chromosomes within the nucleus or chromosome-specific chromatin domains. Since the UV-irradiated strains have about 7500 DNA lesions (as discussed above), the vast majority of these lesions must be repaired by nucleotide excision repair in a manner that does not result in LOH or mutations.

We assume that most of the UV-induced mutations reflect errors introduced during the bypass of pyrimidine dimers by Rev1p and Polζ, since 90% of UV-induced mutations require these activities (LAWRENCE 2002). The source of the mutations in the γ ray-treated samples is less clear. Since the mutations are not associated with regions of LOH, the mutations probably do not reflect errors introduced by DSB repair. It is possible that bases damaged by γ rays are bypassed by error-prone polymerases by a mechanism similar to that associated with UV-induced DNA damage.

All three of the mutations introduced by γ irradiation and about half of the mutations caused by UV were found in both sectors of sectored colonies. This result indicates that the introduced mutation was placed into both strands of the duplex before

DNA replication. Such events, which have been observed previously for UV-induced DNA damage (ECKARDT and HAYNES 1977; JAMES and KILBEY 1977), have been termed "two-strand" mutations (ABDULOVIC *et al.* 2006). One model for such events is that they reflect the repair of two closely-opposed DNA lesions by nucleotide excision repair (NER). During the repair of one lesion, a mutation is introduced. The repair of the second lesion on the opposite strand results in a gap that includes the mutant substitution, and filling-in of the gap results in mutant substitutions in both strands of the duplex. Whatever the explanation of two-strand events, both UV and γ rays efficiently produce this type of mutation.

The repair of DSBs is associated with a 100-fold elevation in the frequency of reversion of a closely-linked mutation (STRATHERN *et al.* 1995), and approximately 1000-fold elevated rates of mutation have been observed during BIR (DEEM *et al.* 2011) and other gene conversion events that result in two newly-synthesized strands (HICKS *et al.* 2010). In addition, the frequency of UV-induced mutagenesis is elevated more than 100-fold in regions of single-stranded DNA next to DSBs or abnormal telomeres (YANG *et al.* 2008). Based on these observations, we checked whether the *de novo* mutations were non-randomly associated with LOH regions associated with gene conversion or BIR. There were 15 base substitutions observed among four sectored colonies resulting from the irradiation of G1 cells. The total lengths of the unselected gene conversion and BIR events among these strains were 163 kb (PG311-UV-8R/W), 271 kb (PG311-UV-9R/W), 22

kb (PG311-IR-37R/W), and 50 kb (PG311-IR-40R/W). The fraction of the genome with these LOH regions was about 0.01. Two of the fifteen (0.13) mutations were in regions of LOH. Although this calculation suggests that the LOH regions may have a significantly elevated frequency of mutations, most of the induced mutations are located outside of the LOH regions.

2.4.7 Comparison among methods of physically mapping recombination events

In our previous studies, we mapped recombination events by a PCR-based technique (described in the Introduction). As employed in our analysis of chromosome V events, this approach was time-consuming and expensive, and mapped events with relatively poor resolution (about 4 kb). More importantly, this method could not be easily used to map events throughout the genome. In addition, the PCR-based approach did not allow us to examine changes in gene dosage (deletions or duplications). For example, we found that an event classified as a crossover on chromosome V by the PCR-based method was actually a terminal deletion on V when examined by SNP arrays.

In contrast, both SNP arrays and HTS allow analysis of events throughout the genome. The advantages of SNP arrays compared to HTS are: 1) relatively low cost (about $100/sample), 2) speed of analysis (about four hours for SNP arrays versus a week for HTS), and 3) relative ease in detecting changes in gene dosage. The major advantages of HTS are: 1) higher resolution (1 kb for SNP arrays versus 250 bp for HTS), and 2) the ability to detect *de novo* mutations. In addition, diagnosis of LOH by HTS can be done

with any diploid in which the progenitor haploid strains have been sequenced, whereas diagnosis of LOH by SNP arrays requires the construction of strain-specific microarrays. Although SNP arrays are probably a more cost-effective and faster approach for mapping large numbers of recombination events at present, as HTS becomes cheaper and analysis of HTS data becomes faster, HTS is likely to be the method of choice in the future. Neither SNP microarrays nor HTS, however, can map recombination events that do not involve LOH (for example, sister chromatid exchanges).

2.4.8 Summary

In conclusion, we have used SNP microarrays and HTS to map crossovers and gene conversion events at high resolution throughout the yeast genome. These studies represent the first genome-wide measurement of the number and types of unselected LOH events induced by UV and γ rays. In G1-synchronized cells treated with either UV or γ rays, 4:0 conversion events are common, suggesting that many of the LOH events reflect the repair of two sister chromatids broken at approximately the same position. In addition, the high-resolution analysis of recombination events by SNP arrays and HTS reveals that gene conversion tracts, particularly those associated with crossovers, are more complex than was previously recognized by low-resolution studies.

2.5 *Supporting Information: Materials and Methods*

2.5.1 SNP microarrays: specificity of hybridization

To confirm the specificity of the pattern of hybridization to the oligonucleotides, we isolated genomic DNA from the haploid strains PLS2/W303a and PSL5/YJM789. The DNA was labeled using either Cy3-dUTP (YJM789) or Cy5-dUTP (W303a) (LEMOINE *et al.* 2005). The samples were mixed and competitively hybridized to the microarrays (details below). Our first experiments were done with arrays containing about 60,000 oligonucleotides (Table S3 in ST CHARLES *et al.* (2012)), representing about 15,000 SNP positions. Following scanning of the arrays, we measured the ratio of the signals at wavelengths specific for the Cy3- and Cy5-labeled samples (532 and 635 nm, respectively). The 635 nm/532 nm ratio was analyzed for each oligonucleotide. The average value of the median ratios of the control probes (those that were identical in the two haploid strains listed in Table S4 of ST CHARLES *et al.* (2012)) was calculated and used to normalize the ratios of the experimental probes to a value of 1 by dividing each probe ratio by the average control probe ratio. All W303a probes (designated by an "SF" or "SR" as the last two characters of the ProbeID in Table S3 of ST CHARLES *et al.* (2012)) that had a centered ratio less than 1.5 were discarded. All YJM789 probes (designated by an "YF" or "YR" as the last two characters of the ProbeID in Table S3 of ST CHARLES *et al.* (2012)) that had a centered ratio greater than 0.67 were discarded. These criteria require at least a 50% difference in signal between the two strains for any given probe. We then

repeated this experiment, switching the dyes that were used to label the genomic DNA samples. Only probes that satisfied the criteria described above in both experiments were used in our analysis. The final probe set used for analysis is listed in Table S5 of ST CHARLES *et al.* (2012).

2.5.2 Details of methods used for microarray analysis: sample preparation, hybridization conditions, and data analysis

Five ml YPD yeast cultures were grown at 30° overnight with agitation. Approximately 55 mg of the pelleted culture was resuspended in 500 ml of a melted agarose solution (0.5% low melt agarose, 100 mM EDTA pH 7.5 at approximately 42°), and then 20 μl of a 25 mg/ml Zymolase solution was added. This cell-agarose mixture was distributed into five plug molds and allowed to solidify. The plugs were then incubated in 700 μl of an EDTA/Tris solution (500 mM EDTA, 10 mM Tris, pH 7.5) overnight at 37°. The next day, 400 μl of a sarcosyl/proteinase K solution (5% sarcosyl, 5 mg/ml proteinase K, 500 mM EDTA pH 7.5) was added to the tubes containing the plugs, and they were incubated at 50° for five hours to overnight. DNA was isolated from three plugs per strain analyzed using the Fermentas Life Sciences GeneJet Gel Extraction Kit (#K0692). Following the addition of the "Binding" buffer from the kit, the samples were incubated at room temperature (~25°) until the agarose had melted and were then incubated on ice for 5 minutes. These samples were sonicated using a BioRupter sonicater for two 15-minute sessions of 30-second high pulses, followed by 30

seconds without sonication, resulting in DNA fragments of about 200-300 bp. Following sonication, the samples were treated according to the kit protocol.

We labeled two reactions for each experimental strain and a single reaction for the reference PG311 strain using the Invitrogen Bioprime Array CGH Genome Labeling Module using the kit protocol for all except for the last step (stop buffer). Experimental strains were labeled with Cy5-dUTP and PG311 was labeled with Cy3-dUTP. After incubation at 37°, all of the reactions intended for a single microarray (two reactions of the experimental strain and one reaction of the reference strain) were combined into a single tube and purified using the Fermentas Life Sciences GeneJet PCR Purification kit (#K0702) following the kit protocol but eluting with 79 µl of dH_2O.

The hybridization reactions were prepared using an Agilent Oligo aCGH/ChIP-on-Chip Hybridization kit (5188-5220) following kit instructions. Arrays were incubated for 48 hours at 62°. Following hybridization, the arrays were washed for 5 minutes in Oligo aCGH/ChIP-on-Chip Wash Buffer 1 (Agilent 5188-5221) and 1 minute in Oligo aCGH/ChIP-on-Chip Wash Buffer 2 (Agilent 5188-5222) that was pre-warmed to 37°. The arrays were then scanned at wavelengths of 635 and 532 nm using the GenePix scanner and the GenePix Pro software using settings recommended by the manufacturer.

Microarrays could be re-used approximately 4-6 times by removing the hybridized labeled DNA sequences from the oligonucleotides. Microarrays and gasket

slides were stripped separately in 1x stripping buffer (10 mM potassium phosphate, pH6.6). The slides were slowly heated to the boiling point in the stripping buffer for 30-45 minutes. After stripping, they were transferred to deionized water, and then slowly removed and stored in a nitrogen cabinet. The gasket slides were centrifuged at 500 rpm to remove excess liquid. Labels on microarrays were removed prior to stripping.

The data generated by GenePix Pro were exported to text files and analyzed with Microsoft Excel. Probes that were flagged by the software were deleted from the analysis. The ratio of the medians (635 nm/532 nm) for each probe was used for analysis, and replicate probe medians were averaged. The data was centered around one by dividing each probe median by the average of all of the probe medians in order to normalize for differences in the hybridization levels for the reference and experimental strain samples. The data were plotted separately for each haploid parental strain and, in plots showing whole chromosomes, the medians were "smoothed" by averaging over nine consecutive probes.

2.5.3 Generation and analysis of high-throughput sequencing data

Prior to analysis of the experimental strains, a reference genome was compiled from the sequences of the two parental haploid strains of PG311; this process required several steps. First, from S288c-YJM789 contig alignments (WEI *et al.* 2007), using in-house PERL scripts, we extracted the S288c and YJM789 sequences separately. The PSL2/W303a "reads" were then assembled using these S288c contigs and the

PSL5/YJM789 "reads" were assembled using the YJM789 contigs; this analysis was performed using CLC Genomics software with the "random" flag option. From these assemblies, we used CLC Genomic software to generate the first-stage reference genomic sequence for PSL2/W303a and PSL5/YJM789. Second, we re-assembled the PSL2 and PSL5 "reads" using the first-stage reference sequences; CLC Genomics software was used for this process with the "unique" flag option. The resulting assemblies were used to generate a second-stage reference sequence for PSL2/W303a and PSL5/YJM789. In the construction of the second-stage reference sequences for PSL2/W303a and PSL5/YJM789, we altered the base from the first-stage reference sequence if ≥75% of the "reads" in the new alignments had a base that was different from the first-stage reference sequence.

We then aligned the second-stage PSL2/W303a and PSL5/YJM789 contigs with the previously aligned S288c-YJM789 contig using the MAFFT program with the Sequence-to-skeleton alignment with the '-add' option (KATOH *et al.* 2009). The resulting alignments allowed us to translate between positions in the assemblies of the two paternal haploids using in-house PERL scripts. In addition, we used in-house PERL scripts to locate the positions of all of the approximately 55,000 SNPs that distinguished the two haploid strains.

Each of the irradiated diploid strains was independently assembled to the consensus sequences of both of PSL2 and PSL5 using CLC Genomics software and, in

addition, the Burrows-Wheeler Alignment Tool (BWA, (LI *et al.* 2009)). SAMtools was then used to extract "pileup" files out of the assembled files (LI *et al.* 2009), creating files that show the number of bases supported at each position. For each experimental strain, from the "pileup" files, we determined the frequency of each SNP allele at each heterozygous SNP position. Positions where the frequency of one of the two original alleles averaged ≥90% of the reads (average of frequencies from PSL2/W303a and PSL5/YJM789 assemblies) were candidates for LOH events. To be regarded as confirmed LOH events, the candidate events had to pass the same criterion for the assemblies obtained with both the CLC Genomics and BWA software. In order to translate the sequence coordinates of each putative LOH event from our analysis of contigs to SGD coordinates, we aligned SGD genomic sequences from the 16 chromosomes with the S288c-YJM789-PSL2-PSL5 alignment by using the MAFFT program. In-house PERL scripts were used to identify the exact SGD coordinates for each putative LOH event and to exclude putative LOH events that were located in repetitive regions of the genome; the coordinates of repetitive regions are given in SGD.

2.5.4 Mutation analysis

In the irradiated experimental diploids, we looked for induced mutations in non-repetitive regions of the genome that were originally identical in the two parental haploids. The mutations were identified using the SNP-calling software of CLC Genomics based on both second-stage parental haploid assemblies. *De* novo mutations

were identified if they were supported by ≥ 40% of the reads in one experimental strain, and less than 15% of the reads in the other experimental strains. We validated each of the identified mutations by a manual comparison to the BWA assembly using Integrative Genomic Viewer (ROBINSON *et al.* 2011). Finally, we confirmed the existence of the mutations by DNA sequence analysis of PCR fragments containing the relevant region.

2.6 Supporting Information: Discussion

Below, we expand our discussion of two topics introduced in the main text: 1) evidence that some of the unselected gene conversion events that are not associated with centromere-distal LOH are associated with crossovers, and 2) the mechanisms involved in forming complex gene conversion events.

2.6.1 Gene conversion events that are unassociated with centromere-distal LOH

In order for a mitotic crossover to produce LOH, the two daughter cells must receive one recombinant chromosome and one non-recombinant chromosome (Figure 2.2). If one daughter cell receives both recombinant chromosomes and the other receives both non-recombinant chromosomes, LOH will not be observed (CHUA and JINKS-ROBERTSON 1991). Thus, about half of the gene conversions that <u>are</u> associated with crossovers will not be detectable by our analysis. To look for crossover-associated gene conversion events that did not result in LOH, we looked for changes in linkage of

markers flanking the conversion by analyzing the meiotic products that resulted from sporulating diploids derived from sectored colonies.

We examined the meiotic products of eleven sectored colonies (three from γ ray-treated cells and eight from UV-treated cells) that had unselected 4:0 gene conversion events unassociated with LOH. The meiotic segregation patterns of the heterozygous SNPs flanking the 4:0 conversion tract were analyzed using the PCR-based method described previously (LEE *et al.* 2009). Since the physical distances separating the flanking markers were relatively small (< 20 kb for most intervals), we expected that most of the detected crossovers would reflect mitotic rather than meiotic events in the tested interval.

If no mitotic crossover was associated with the 4:0 tract, we expected that, in the tetrads derived from both sectors, most tetrads would have two spores with the YJM789 form of SNPs flanking the conversion tract and two spores with the W303a form of SNPs flanking the conversion tracts. If a mitotic crossover occurred in G2, we expect that one sector would produce tetrads in which all four spores have the recombinant configuration of the flanking markers, and the other sector would have tetrads in which the spores have flanking markers in the original parental configurations. Finally, if both a gene conversion event <u>and</u> a crossover occur in G1, we would expect to find the recombinant configuration of markers in tetrads derived from both sectors. Of the eleven sectored colonies examined, eight had the pattern expected for gene conversion

unassociated with the crossover; data is presented in Table S12 of (ST CHARLES *et al.* 2012). One had the pattern expected for a crossover in G2 with the segregation of two recombinant chromosomes in one cell and two non-recombinant chromosomes in the other. Interestingly, two sectored colonies had the pattern consistent with a conversion event and a crossover occurring in G1. This pattern, however, could also be a consequence of the repair of two broken chromosomes in G2, both associated with a crossover.

2.6.2 Complex gene conversion tracts

As discussed in the text, many of the conversion tracts associated with crossovers were complex, involving multiple transitions between 3:1, 4:0, and 2:2 regions within the tract or tracts in which 3:1 and 1:3 segments occurred within one tract. Below, we will first discuss gene conversion events that are not associated with LOH of centromere-distal markers, followed by a discussion of LOH-associated conversions. Diagrams of all recombination events in our study are shown in Table S 2.1 and Table S 2.2, and figures showing the patterns of DSB repair required to produce the recombination events are shown in Figs. S1-S40 of ST CHARLES *et al.* (2012).

2.6.2.1 Complex gene conversion tracts unassociated with LOH of centromere-distal markers

It is important to note that gene conversion events unassociated with crossovers could occur by three pathways (Figure 2.1): 1) SDSA in which there is a single region of heteroduplex, 2) resolution of a double Holliday junction (dHJ) which results in two

regions of heteroduplex located in *trans*, and 3) dissolution of the dHJ by topoisomerase leading to two regions of heteroduplex located in *cis* on the chromosome. In an analysis of plasmid-chromosome recombination in yeast, MITCHEL *et al.* (2010) concluded that most mitotic gene conversion events were a consequence of SDSA.

As discussed below, most (about 80%) of the conversion events unassociated with LOH of distal markers are explicable as a consequence of repair of one or two DSBs by the SDSA pathway. These conversion events were divided into Classes A-G. The defining attributes of each class and the mechanism(s) required to produce each class are: 1) Class A (3:1 or 1:3 conversion, SDSA repair of single G2 DSB), 2) Class B (4:0 or 0:4 conversion, SDSA repair of two DSBs located at the same position on sister chromatids), 3) Class C (3:1/4:0 or 1:3/0:4 hybrid conversions, SDSA repair of two DSBs located at the same position on sister chromatids with different length of conversion tracts), 4) Class D (3:1 tract in which the homozygous region is split between the two sectors, repair of two DSBs by SDSA with the conversion tracts on opposite sides of the DSB), 5) Class E (3:1/4:0/3:1 or 1:3/0:4/1:3 conversions, repair of two DSBs, one by SDSA and one by the DSB repair pathway as shown in Figure 2.1), 6) Class F (3:1, 1:3 or hybrid conversion tracts are disrupted by segment of 2:2 segregation, repair of DSBs in which there is "patchy" repair (defined below), and 7) Class G (two closely-linked conversions with different donors (3:1 and 1:3 or 4:0 and 0:4); models to explain Class G will be described below.

In summary, Classes A-E, which account for more than 80% of the conversion events, can be simply explained by the repair of one or two DSBs by the mechanisms shown in Figure 2.1. Concerning Classes F and G, several additional points need to be discussed. First, although yeast cells can efficiently repair a double-stranded DNA gap (MITCHEL *et al.* 2010; ORR-WEAVER and SZOSTAK 1983), it is likely that most mitotic gene conversion events reflect heteroduplex formation followed by the repair of the resulting mismatches (WENG and NICKOLOFF 1998). Mismatch repair can result in a detectable conversion event or a restoration event. For example, in Figure 2.1A, repair of the heteroduplex resulting in a duplex with two "red" strands would represent conversion-type repair, since this pattern produces 3:1 segregation; repair of the mismatch to produce a duplex with two "blue" strands represents restoration-type repair, since this pattern generates two cells that retain heterozygosity at the position of the original heteroduplex. Although multiple mismatches within one heteroduplex are generally converted in a concerted manner yielding a continuous conversion tract, tracts with mixtures of conversion-type and restoration-type repair have been detected in both meiosis (MANCERA *et al.* 2008; SYMINGTON and PETES 1988) and mitosis (MITCHEL *et al.* 2010; NICKOLOFF *et al.* 1999). Second, we point out that some of the events have more than one plausible interpretation. For example, the F4 recombination event (Table S 2.1) can be explained by patchy repair of two DSBs in G2 (Figure 2.13A) or patchy repair of a single DSB in G1 (Figure 2.13B).

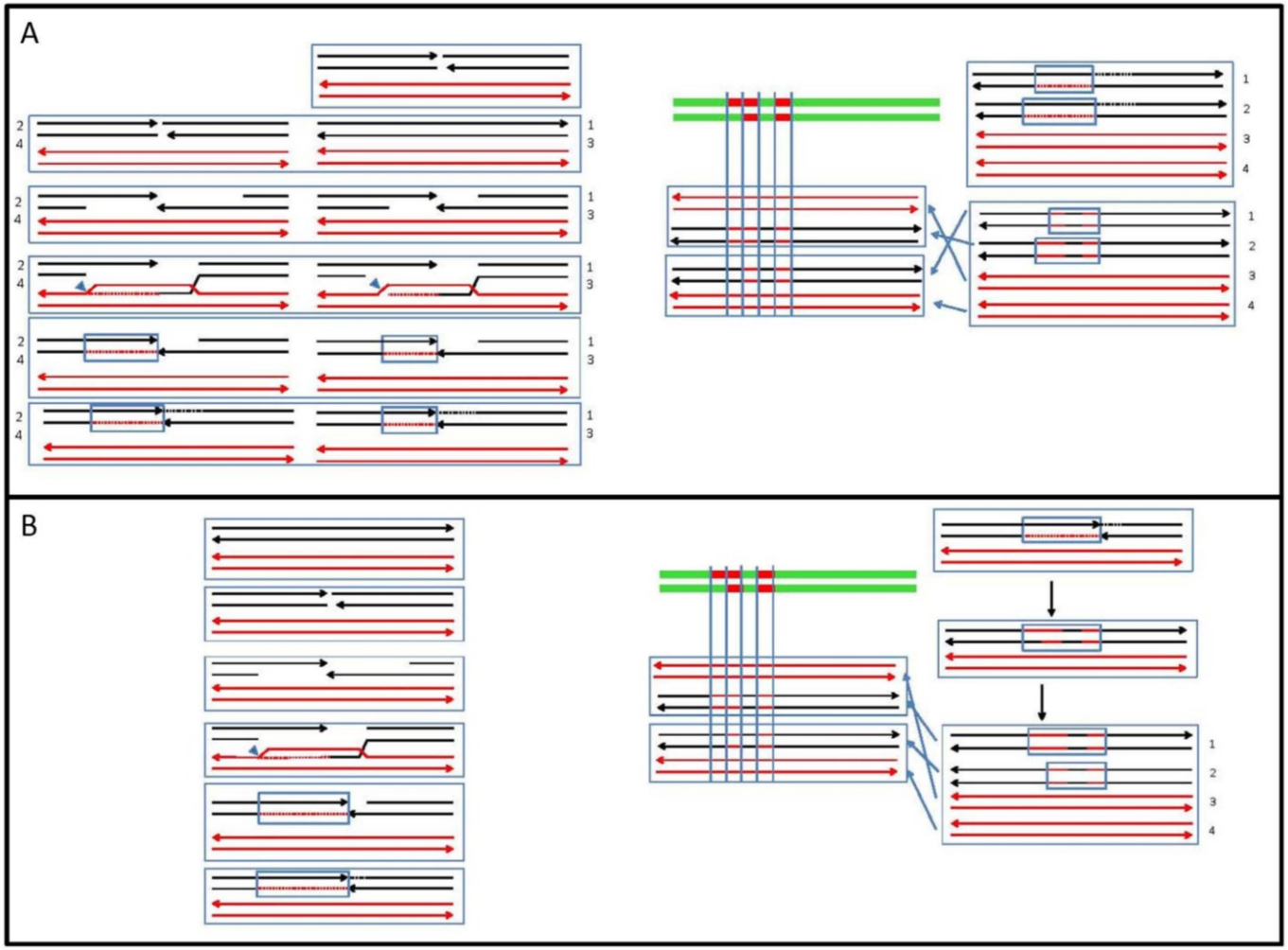


Figure 2.13. Generation of Class F4 by two different mechanisms.

In Class F4, there is a 3:1/4:0 conversion tract separated from a second 4:0 tract by a heterozygous segment. In Figure 2.13A, we show this pattern generated by the repair of two DSBs using the SDSA pathway. The heteroduplex tracts are of different lengths and are repaired in a "patchy" manner. In Figure 2.13B, we show Class F4 as generated by repair of a single G1 DSB. Mismatches in the resulting heteroduplex are repaired in a "patchy" manner in G1 with one segment containing unrepaired mismatches. Replication of this molecule would produce the F4 pattern.

In the Class G events, two conversions in opposite directions are observed (3:1 and 1:3 or 4:0 and 0:4). Although such events could be explained as a consequence of two independent DSBs, their frequency and the close linkage of the two types of conversion tracts indicate that they likely reflect the repair of a single DNA lesion. There are two different modifications of the models shown in Figure 2.1 that can explain the conversion events with two different donors. In one model (Figure 2.14), there are two different rounds of mismatch repair associated with an SDSA event. In the first round, conversion occurs in which information from the invading strand is transferred to the invaded strand. Following the reversal of the strand invasion, a second round of mismatch repair occurs to produce the 3:1/1:3 conversion pattern. In an alternative pathway, a dHJ intermediate is formed, followed by branch migration of one of the junctions (Figure 2.11). During recombination in *E. coli*, a Holliday junction can be translocated by branch migration, resulting in symmetric heteroduplexes (WEST 1997). Although genetic evidence argues against the formation of symmetric heteroduplexes during meiotic recombination in *S. cerevisiae* (PETES *et al.* 1991), symmetric

heteroduplexes have been invoked previously to explain certain classes of mitotic gene conversions (ESPOSITO 1978; NICKOLOFF *et al.* 1999; ROITGRUND *et al.* 1993).

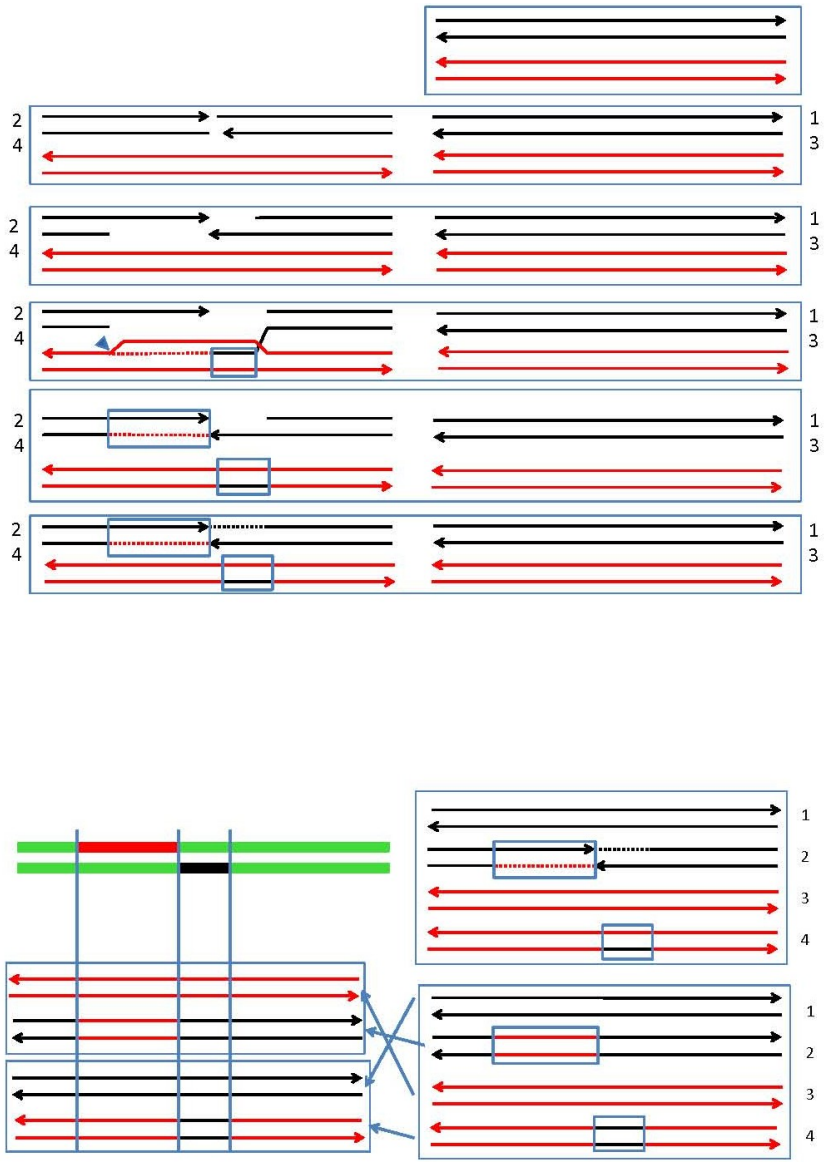


Figure 2.14. Alternate mechanism for generation of Class G1.

In this model, the G1 event is produced by two rounds of mismatch repair during SDSA, one associated with the invading strand, and a second after strand displacement.

As described above, because of the patterns of chromosome segregation following mitotic crossovers, only half of the events lead to LOH of centromere-distal markers. Thus, some of the gene conversion events that are unassociated with LOH for centromere-distal markers could reflect the crossover-associated conversions. Thus, there are likely to be additional pathways of repair other than those shown in the supplementary figures. It is important to stress, however, that about 80% of the conversion events are simply explained as a consequence of the SDSA pathway shown in Figure 2.1, and that most of the events (about three-quarters) are most simply explained as a consequence of the repair of two sister chromatids that are broken at approximately the same position.

2.6.2.2 Complex conversion tracts associated with crossovers

We divided the gene conversion tracts associated with crossovers into the following classes: 1) Class H (no detectable conversion tract or 3:1 conversion associated with repair of single DSB, 2) Class I (repair of two DSBs by canonical repair pathways), 3) Class J (repair of one or two DSBs involving either patchy repair or branch migration), and 4) Class K (recombination events that involve more than one independent DSB on one chromosome arm). Of the crossover-associated conversions, about half were relatively simple (Classes H and I), and half required patchy repair and/or branch migration.

As discussed above, conversion tracts associated with crossovers are significantly more complex than those unassociated with crossovers. MANCERA *et al.* (2008) reported that 11% of meiotic crossovers had complex conversion tracts, whereas the frequency of complex tracts among conversions unassociated with crossovers was 3%. One explanation of this difference could be that crossovers that proceed through the pathway shown in Figure 2.1 are associated with two regions of heteroduplex, while conversions resulting from SDSA have only a single region of heteroduplex. Second, since gene conversion tracts associated with crossovers are usually longer than those unassociated with crossovers (AGUILERA and KLEIN 1989; MANCERA *et al.* 2008), there may be a greater chance to observe patchy repair in tracts associated with crossovers.

2.7 *Supplemental tables*

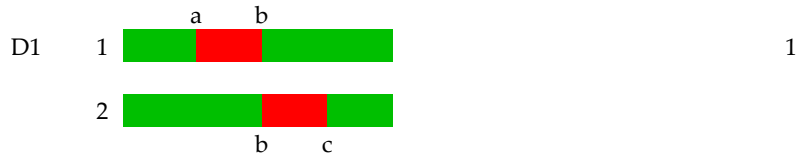
Table S 2.1 Classes of conversion tracts associated with spontaneous recombination events and events stimulated by UV or γ rays.

In this table, we depict the various classes of crossovers (with associated gene conversion events), gene conversion events unassociated with crossovers, and break-induced replication (BIR) events detected in our study of spontaneous mitotic recombination, and DNA damage-induced recombination. Each class contains a pair of lines, with one line representing each sector of a sectored colony. Green, red, and black line segments indicate that SNPs are heterozygous, homozygous for W303a-derived SNPs, and homozygous for YJM789-derived SNPs, respectively. Each transition between different colored segments is given a letter, and the SGD coordinates of these transitions are given in Table S2.3-2.7. Mechanisms that explain these patterns are presented in Figs. S1-S40 of ST CHARLES *et al.* (2012). In column T, we show the number of events in each class.

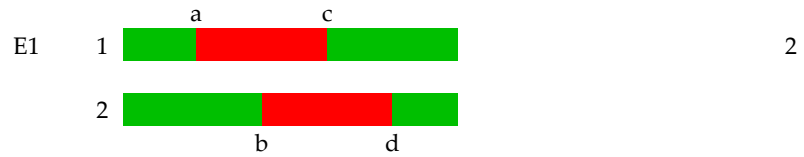
Class		Conversion tract	Number of events
Class A		a b	
A1	1		5
	2		
A2	1		7
	2		
		a b	
Class B			
		a b	
B1	1		10
	2		
		a b	
		a b	
B2	1		6
	2		
		a b	

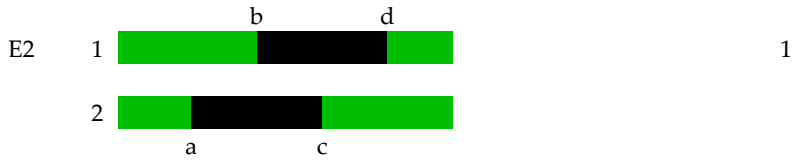
Class C

Class D

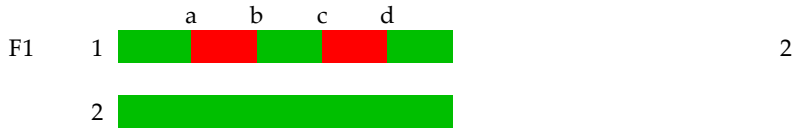


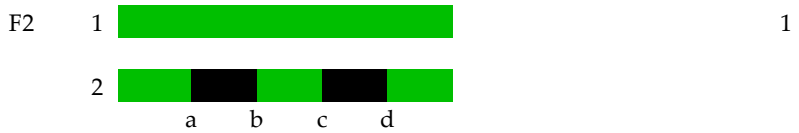
Class E

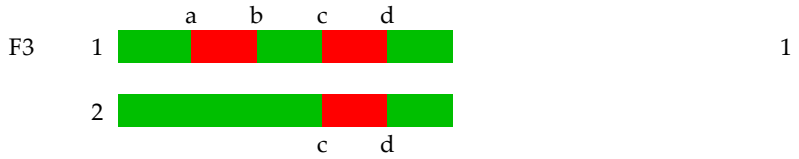


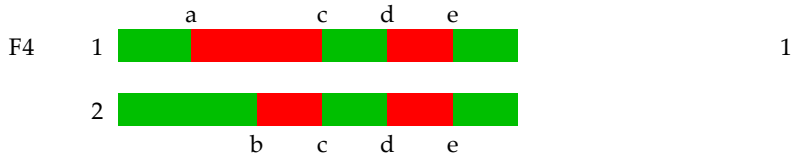


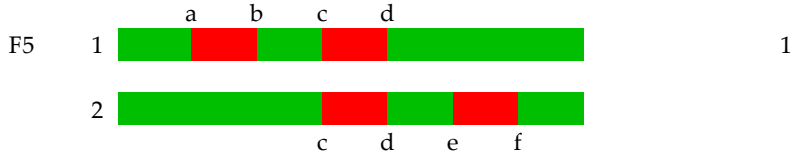
Class F



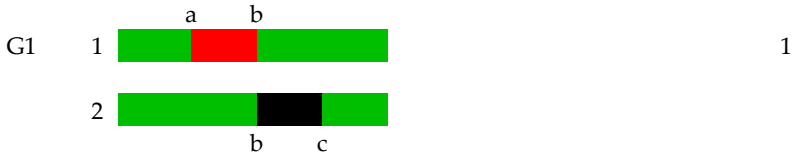


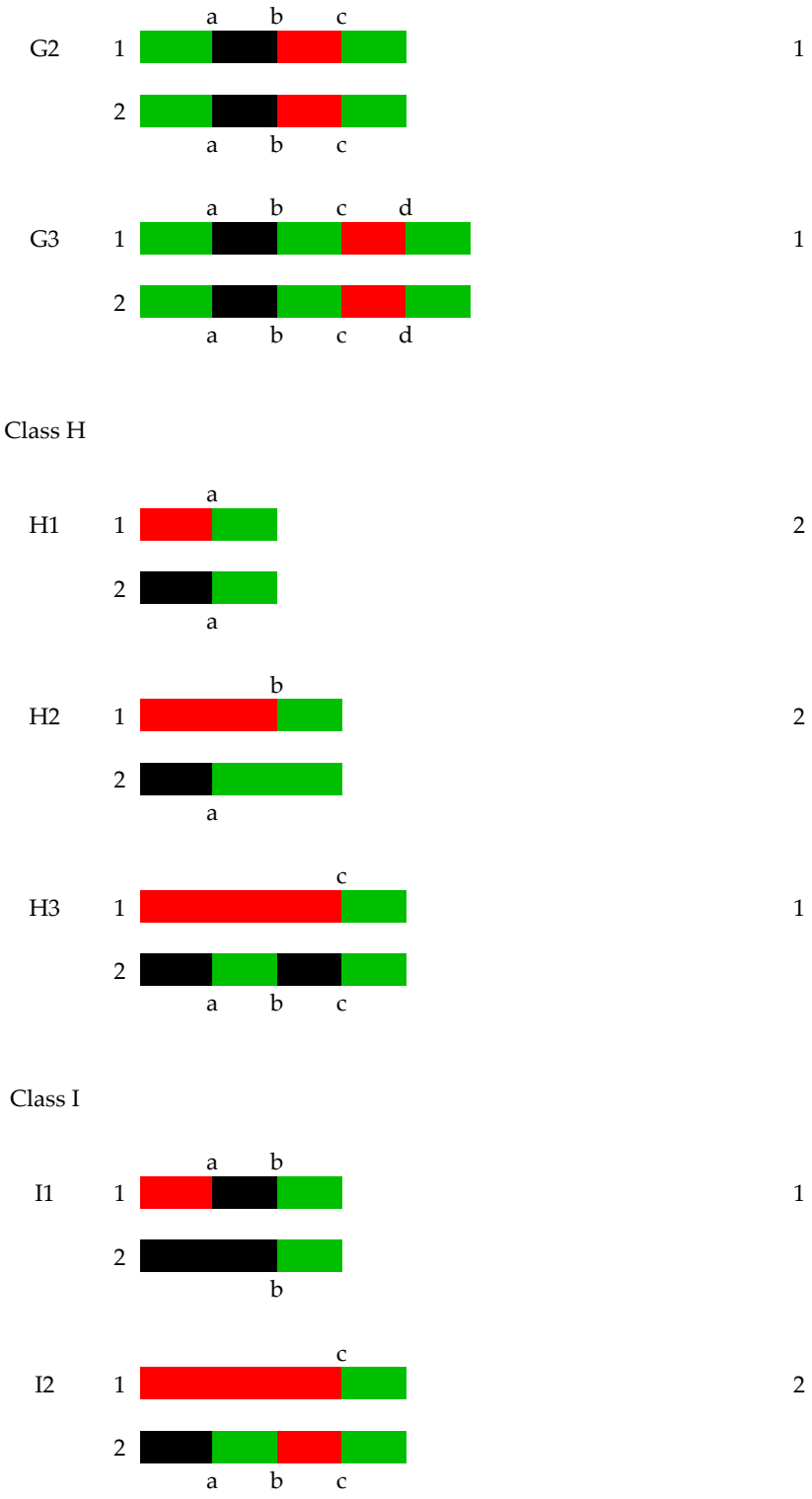


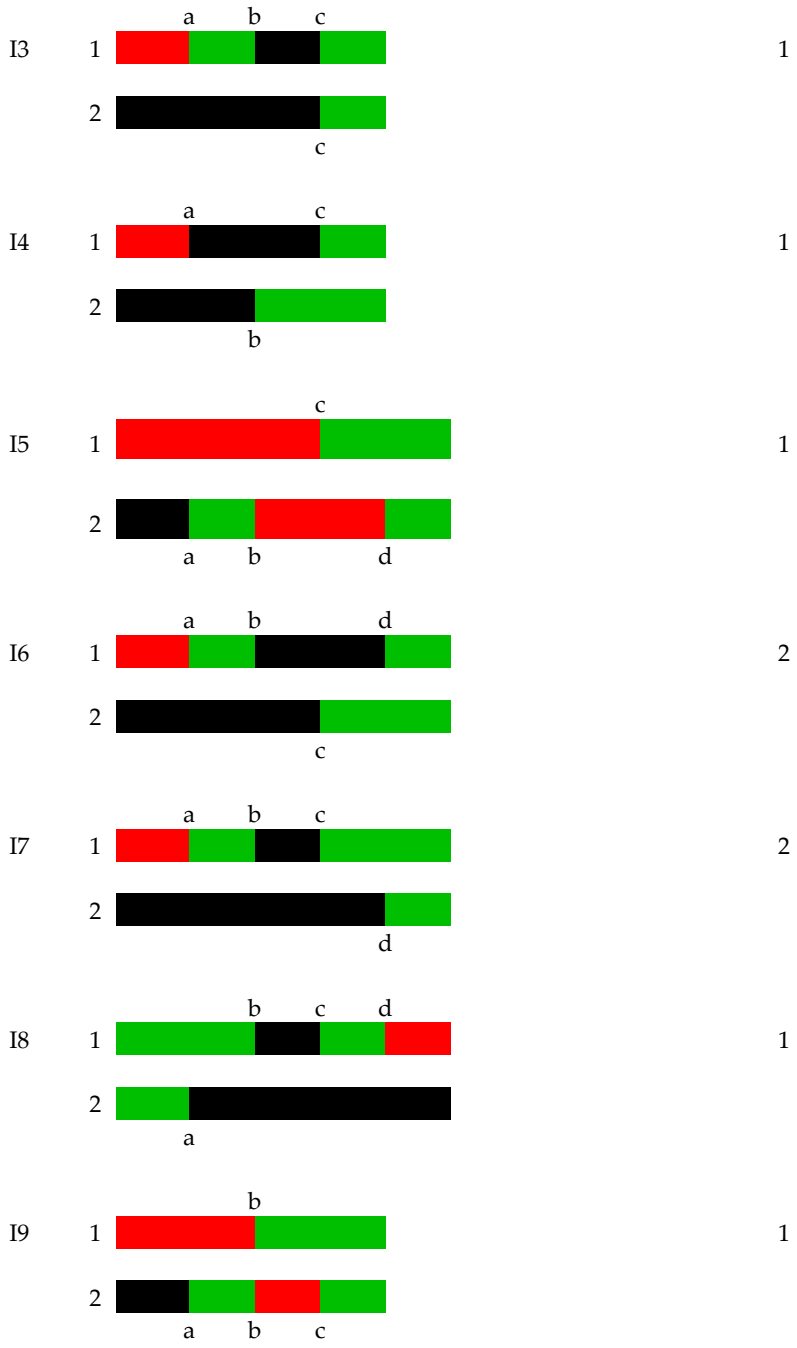




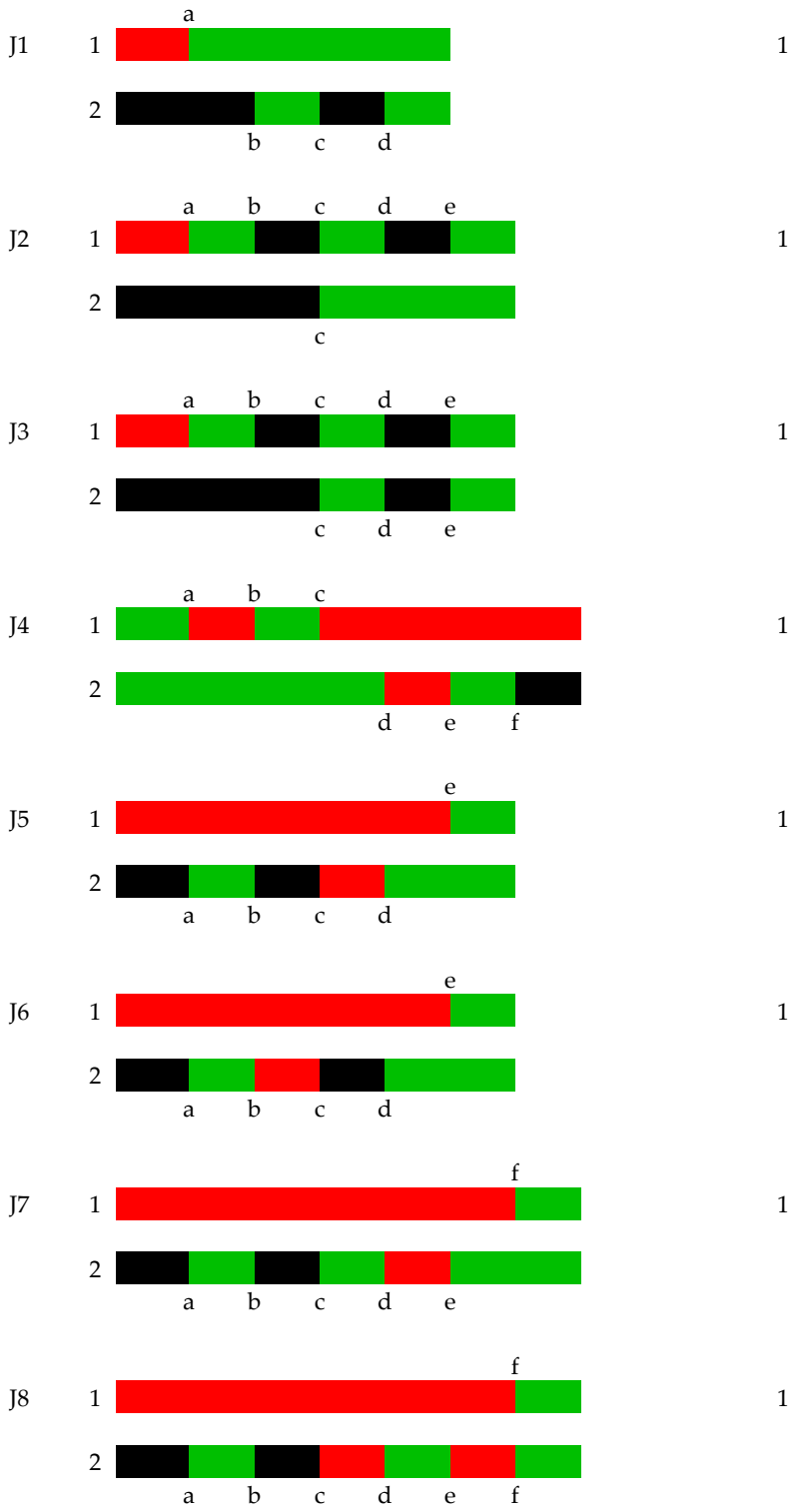
Class G

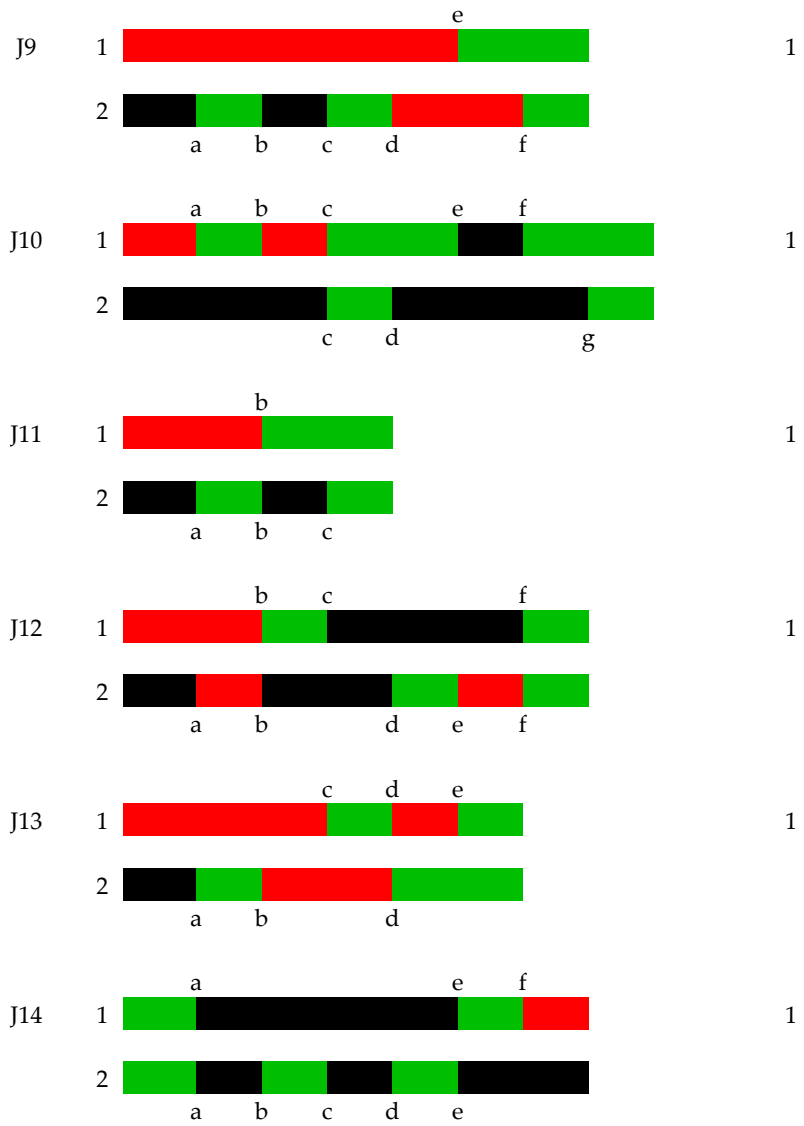



G2
1
2
a b c
1
G3
1
2
a b c d
1
Class H
H1
1
2
a
2
H2
1
2
b
a
2
H3
1
2
c
a b c
1
Class I
I1
1
2
a b
b
1
I2
1
2
c
a b c
2

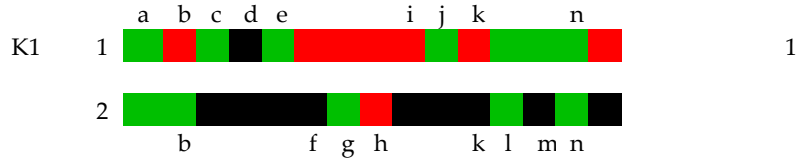

I3
1
a b c
2
c
1
I4
1
a c
2
b
1
I5
1
c
2
a b d
1
I6
1
a b d
2
c
2
I7
1
a b c
2
d
2
I8
1
b c d
2
a
1
I9
1
b
2
a b c
1

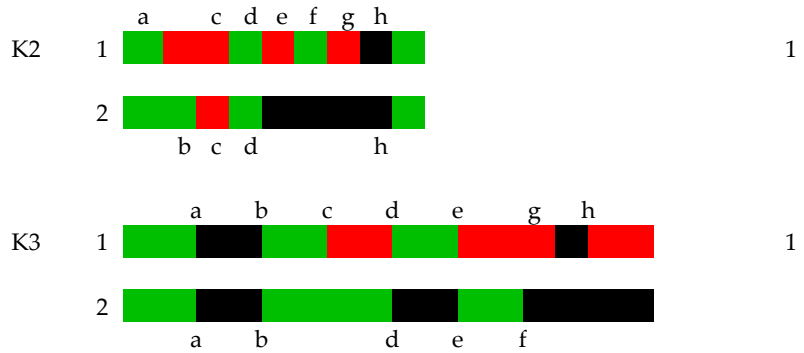
Class J


J1
1
2
a
b c d
1
J2
1
2
a b c d e
c
1
J3
1
2
a b c d e
c d e
1
J4
1
2
a b c
d e f
1
J5
1
2
e
a b c d
1
J6
1
2
e
a b c d
1
J7
1
2
f
a b c d e
1
J8
1
2
f
a b c d e f
1


J9
1
2
e
a b c d f
1
J10
1
2
a b c e f
c d g
1
J11
1
2
b
a b c
1
J12
1
2
b c f
a b d e f
1
J13
1
2
c d e
a b d
1
J14
1
2
a e f
a b c d e
1

Class K


K1
1
2
a b c d e i j k n
b f g h k l m n
1


K2
1
2
a c d e f g h
b c d h
1
K3
1
2
a b c d e g h
a b d e f
1

Class L

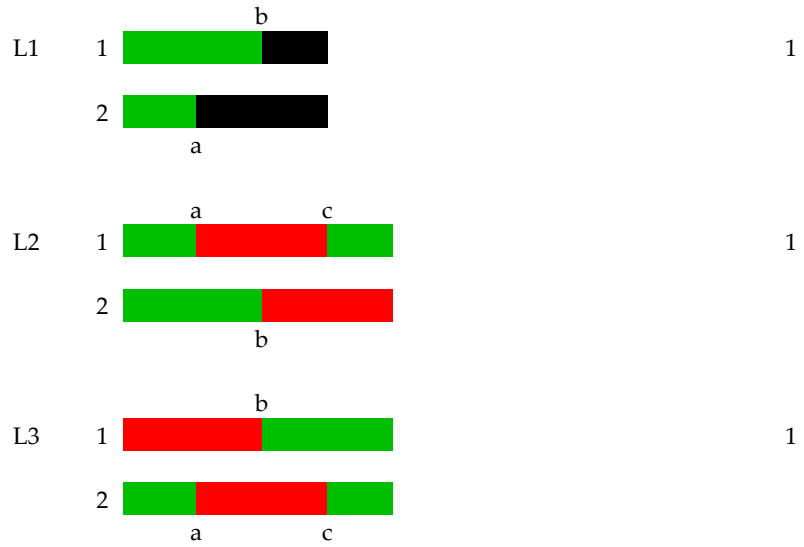

L1
1
2
b
a
1
L2
1
2
a c
b
1
L3
1
2
b
a c
1

Table S 2.2 Classes of recombination events that were observed by SNP arrays but that were altered when examined by high-throughput DNA sequencing.

Some of the sectored colonies that were analyzed by both SNP microarrays and by high-throughput DNA sequencing (HTS) became more complex in the HTS data. If the event detected by the microarray was not already represented by a class in Table S2.1, it was assigned to Classes M1-M7. Since these classes, based only on microarrays, were less accurate than those based on HTS, we did not depict Classes M1-M7 in Figs. S1-S40 of ST CHARLES *et al.* (2012).

Class		Conversion tract	Number of events
M1	1		1
	2		
M2	1		2
	2		
M3	1		1
	2		

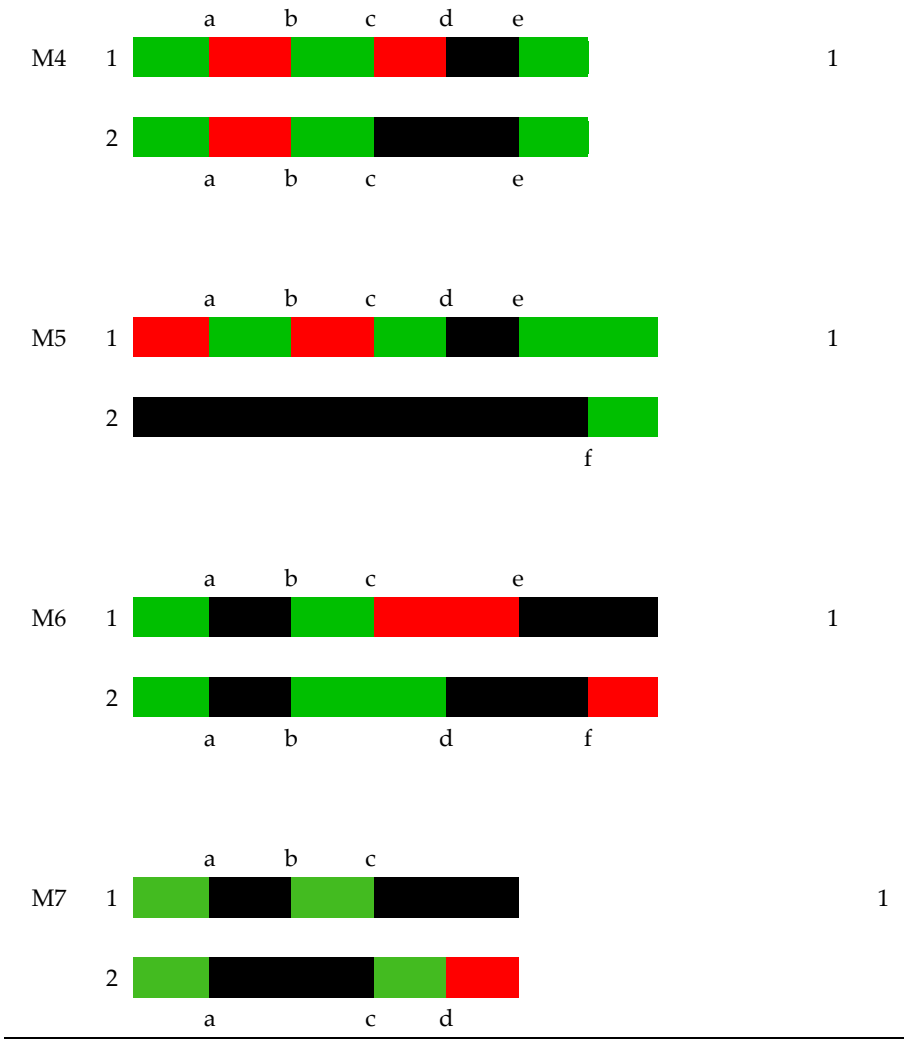

a b c d e
M4 1
1
2
a b c e
a b c d e
M5 1
1
2
f
a b c e
M6 1
1
2
a b d f
a b c
M7 1
1
2
a c d

Table S 2.3 Mapping spontaneous mitotic recombination events by SNP microarrays

We selected chromosome V crossovers as red/white sectored colonies. DNA was isolated from each sector and examined for LOH by SNP arrays. In most of these sectors, there were one or more transitions between heterozygous SNPs and SNPs that were homozygous for either the W303a or YJM789 form of the SNP. The "Sector" column shows the name of the sectored colony, and the "Chrom." shows the chromosome location; for this table, all events except one were on chromosome V, since we were selecting for events on V. The event on chromosome VIII in 7BRW was unselected. The "Event class" column shows the event class as defined in Table S2.1-S2.2. The numbers in the "Line number" columns relate to the depictions of the events in Table S2.1-S2.2 and define which line drawing (numbered 1 or 2) corresponds to which sector (red or white). The letters in the "Transistion label" column correspond to transitions between heterozygous and homozygous segments within each sector as shown in Table S2.1-S2.2. The "Markers flanking transitions" columns show the SGD coordinates of the SNPs that most closely flank the transition points.

Sector	Chrom.	Event class	Line number		Transition label	Markers flanking transitions	
			Red sector	White sector		Left	Right
1.4RW	V	H2	1	2	a	127038	128941
					b	129616	130096
1.7RW	V	H1	1	2	a	150291	151419
4.1RW	V	J8	1	2	a	62494	62635
					b	65685	67302
					c	72490	75719
					d	80069	81264
					e	82770	84997
					f	98369	98763
4.11RW	V	I4	1	2	a	97792	98369
					b	104636	105299
					c	129310	129511
2ARW	V	H2	1	2	a	51915	53692
					b	57170	60701
6BRW	V	H1	1	2	a	49385	49896
7BRW	V	J5	1	2	a	109029	110820
					b	112452	113907
					c	118129	119505
					d	130096	131261

					e	134783	136012
	VIII	B1	1	2	a	228989	231822
					b	231822	253188
15ARW	V	J13	1	2	a	34481	37492
					b	38068	41126
					c	41533	44506
					d	45587	48837
					e	49044	49385
16BRW	V	J1	1	2	a	115078	117001
					b	117001	117276
					c	117276	118129
					d	118137	119505
18ARW	V	J9	1	2	a	121397	124190
					b	127038	128941
					c	128941	129100
					d	130096	131261
					e	131261	132281
					f	134832	136012
21ARW	V	H3	1	2	a	30652	31251
					b	31251	32231
					c	34481	37492
26BRW	V	I7	1	2	a	86758	86861
					b	86861	87888
					c	91739	93200
					d	94077	94399
29BRW	V	I1	1	2	a	158692	159409
					b	159409	161995

Table S 2.4 SNP microarray analysis of recombination events induced by γ rays

Seven sectored colonies induced by γ rays were analyzed by SNP arrays. As expected, all colonies had a crossover event on chromosome V. In addition, most of the sectored colonies had unselected events on other chromosomes. The description of the information concerning the various columns in this table is given in Table S2.3. The depictions of the M classes of events are given in Table S2.2 with all other classes in Table S2.1.

Sector	Chrom.	Event class	Line number		Transition label	Markers flanking transitions	
			Red sector	White sector		Left	Right
2RW	II	D1	2	1	a	599438	599687
					b	604780	605756
					c	610185	614979
	V	I6	1	2	a	38068	41126
					b	41533	44506
					c	50295	51183
					d	56137	57170
	VIII	C3	1	2	a	168743	170576
					b	173252	178613
					c	181063	183463
	XI	I9	2	1	a	105100	112922
					b	117632	119303
					c	121099	122536
5RW	IV	B1	1	2	a	1148058	1148418
					b	1152992	1153949
	IV	I8	1	2	a	1265332	1267763
					b	1272145	1273200
					c	1275710	1277206
					d	1286930	1287823
	V	I3	1	2	a	94077	94339
					b	98736	98985
					c	104636	105299
	V	C3	2	1	a	271363	272155
					b	276195	278498
					c	278498	291700
6RW	V	J7	1	2	a	78080	80069
					b	81264	82754
					c	86732	86758

					d	90200	91730
					e	95684	95906
					f	97404	97743
	V	B1	1	2	a	540335	541057
					b	541472	541633
	XII	I7	2	1	a	221657	226427
					b	236112	236226
					c	237597	237803
					d	244367	244661
8RW	II	L1	2	1	a	769886	771821
					b	773533	775053
	V	J6	1	2	a	102173	103260
					b	105481	105830
					c	109050	110820
					d	121397	124190
					e	127038	128941
	XV	A1	2	1	a	608670	608946
					b	611151	611700
37RW	V	M3	1	2	a	41533	44506
					b	44506	45284
40 R/W	II	A2	2	1	a	88072	89532
					b	102421	116691
	V	M5	1	2	a	72004	72490
					b	72490	75719
					c	78080	80069
					d	81264	82754
					e	85566	86732
					f	87888	88109
	X	M1	2	1	a	51711	52024
					b	52024	52381
					c	57132	57722
					d	57722	58117
					e	60412	61667
					f	62052	64080
	XVI	B1	1	2	a	93747	95738
					b	96463	98869
43RW	IV	B2	1	2	a	1309753	1310149
					b	1312257	1313530
	V	J12	1	2	a	145732	150291
					b	150291	151419

					c	157756	158692
					d	192712	197339
					e	197480	201569
					f	201569	205990
	VIII	I5	1	2	a	50240	53891
					b	55920	57162
					c	67614	70935
					d	75202	77201
	X	L3	2	1	a	59483	59933
					b	68077	70059
					c	73999	75330
	XV	A1	1	2	a	949471	958063
					b	958063	966210

Table S 2.5 Analysis of γ ray-induced recombination events by high-throughput DNA sequencing

Sector	Chromo.	Event class	Line number		Transition label	Markers flanking transitions	
			Red sector	White sector		Left	Right
37RW	IV	F2	2	1	a	1060007	1065083
					b	1071768	1075945
					c	1077848	1079627
					d	1079627	1082848
	V	I2	1	2	a	41687	42238
					b	42490	42959
					c	44999	45229
40 R/W	II	C4	2	1	a	88493	89545
					b	103654	104643
					c	104850	115777
	V	J10	1	2	a	72326	72502
					b	72990	73131
					c	78835	79231
					d	79231	79309
					e	81276	81713
					f	85693	86742
					g	87900	87938
	X	F5	2	1	a	51723	52036
					b	52036	52231
					c	57448	57612
					d	57849	57860
					e	61049	61448
					f	62856	63404
	XVI	B1	1	2	a	94414	94922
					b	96934	97162

Table S 2.6 Analysis of UV-induced recombination events by SNP microarrays

Three sectored colonies derived from UV-treated cells were examined by SNP microarrays as described previously. In addition to the expected crossovers/conversions on chromosome V, we observed many unselected events on other chromosomes. The information shown in Columns A to H is comparable to that shown in Tables S2.4 and S2.5.

Sector	Chrom.	Event class	Line number		Transition label	Markers flanking transitions	
			Red sector	White sector		Left	Right
4WR	II	F1	1	2	a	517739	518366
					b	534963	536407
					c	536711	537814
					d	541372	542341
	III	A1	2	1	a	68840	72774
					b	77887	78730
	IV	B2	1	2	a	589427	590872
					b	591064	593197
	IV	B1	1	2	a	895029	909432
					b	911910	913818
	V	I2	1	2	a	41126	41533
					b	45587	48837
					c	57170	60701
	V	B2	1	2	a	442725	450243
					b	452240	453723
	VII	B2	1	2	a	29386	32836
					b	33882	43381
	VII	B1	1	2	a	557235	558489
					b	559492	560550
	VII	G2	1	2	a	699064	730597
					b	737182	747266
					c	759626	773191
	VII	B1	1	2	a	878454	878865
					b	880297	881258
	VIII	B2	1	2	a	97859	99073
					b	100063	101004
	XII	C2	1	2	a	494992	495797
					b	496424	498208
					c	499664	500129

8WR	II	M2	2	1	a	442290	447517
					b	450919	452926
	IV	E2	2	1	a	15561	18870
					b	21769	21917
					c	24733	27078
					d	27078	27666
	IV	G1	2	1	a	78170	78934
					b	79589	79790
					c	80845	81765
	IV	B1	1	2	a	559223	560565
					b	563067	565309
	IV	A2	1	2	a	624148	628588
					b	632639	634749
	IV	B1	1	2	a	1065071	1075942
					b	1075942	1143034
	V	J2	1	2	a	54198	56117
					b	57170	60701
					c	65685	67302
					d	72004	72490
					e	72490	75719
	V	E1	2	1	a	356821	357717
					b	368841	370202
					c	373720	374799
					d	377174	380708
	VII	J3	1	2	a	231705	233805
					b	234547	236270
					c	236523	236846
					d	237107	237661
					e	237690	239200
	VII	C1	2	1	a	985827	991002
					b	991889	992929
					c	996609	1000337
	X	J11	2	1	a	297498	298815
					b	306935	307352
					c	307352	309462
	X	B1	1	2	a	436583	437739
					b	439631	441601
	XV	C4	1	2	a	218129	221061
					b	223649	226687
					c	234269	235360
	XV	A2	2	1	a	412646	412925

					b	420322	420877
	XV	B2	1	2	a	945646	946473
					b	946473	948727
	XVI	A1	1	2	a	118240	119094
					b	121545	122461
	XVI	C3	1	2	a	339827	341256
					b	342256	342732
					c	344770	347422
9WR	IV	A2	1	2	a	317781	325063
					b	332254	332733
	IV	B1	1	2	a	1178332	1181398
					b	1184854	1187255
	IV	L2	2	1	a	1503756	1504918
					b	1505988	1507520
					c	1510998	1515286
	V	I6	1	2	a	93407	94077
					b	95906	96550
					c	106469	107912
					d	107912	108369
	V	C1	1	2	a	165534	167628
					b	167628	169096
					c	176440	178678
	VII	C3	1	2	a	580170	582792
					b	585672	589039
					c	589039	592008
	VII	M2	2	1	a	663893	670102
					b	670102	682631
	IX	A1	2	1	a	227036	227449
					b	229500	230665
	X	A2	2	1	a	163640	164139
					b	172148	175014
	XI	E1	2	1	a	15265	19930
					b	23922	26696
					c	29975	30912
					d	30912	34433
	XI	A2	1	2	a	523522	527412
					b	534897	538280
	XII	M4	1	2	a	633203	640474
					b	642140	645790
					c	645790	649852
					d	751132	751818

					e	751963	755543
	XIII	A2	2	1	a	352162	356136
					b	363039	365184
	XIII	M7	2	1	a	469500	470798
					b	474925	476981
					c	476981	479133
					d	479133	480314
	XV	M6	2	1	a	438251	440736
					b	441476	442416
					c	443020	443590
					d	446473	448742
					e	508790	509957
					f	513325	514332

Table S 2.7 Analysis of UV events by HTS

Two of the sectored colonies previously examined by microarrays were also examined by HTS. Genomic DNA isolated from both sectors of the two colonies were sequenced. If the recombination event defined by HTS was more complex than that previously observed by SNP arrays, the event was assigned a different class. Columns A-H have the information comparable to that in Tables S2.3-S2.6.

Sector	Chrom.	Event class	Line number		Transition label	Markers flanking transitions	
			Red sector	White sector		Left	Right
8WR	II	J4	2	1	a	440187	440214
					b	440214	440343
					c	442813	447152
					d	448628	449146
					e	450124	450279
					f	451448	451581
	IV	E2	2	1	a	17461	17898
					b	21807	21893
					c	24937	25924
					d	27210	27498
	IV	G1	2	1	a	78265	78579
					b	79622	79688
					c	81258	81586
	IV	B1	1	2	a	559235	559934
					b	563217	563467
	IV	A2	1	2	a	625248	628600
					b	632651	633880
	IV	C1	2	1	a	1071768	1075945
					b	1075945	1077715
					c	1077848	1079627
	V	J2	1	2	a	54912	55841
					b	57449	57885
					c	66297	67311
					d	72326	72502
					e	73131	75594
	V	E1	2	1	a	356833	357729
					b	369253	369440
					c	373732	375013
					d	377186	378582

	VI	A1	2	1	a	14770	25898
					b	29907	30187
	VII	J3	1	2	a	231717	232620
					b	234558	235232
					c	236715	236858
					d	237223	237461
					e	237744	238566
	VII	C1	2	1	a	988076	990250
					b	992453	992804
					c	997417	998195
	X	J11	2	1	a	298226	298676
					b	306947	307281
					c	307364	308210
	X	B1	1	2	a	436595	437751
					b	439643	440839
	XV	C4	1	2	a	218142	218547
					b	223662	225904
					c	234931	235232
	XV	A2	2	1	a	412659	412752
					b	420368	420716
	XV	G3	1	2	a	945871	946487
					b	946487	948117
					c	949629	954763
					d	955483	956016
	XVI	A1	1	2	a	118260	119014
					b	121557	121751
	XVI	F4	1	2	a	339839	340193
					b	342268	342371
					c	346500	347398
					d	350087	350218
					e	350218	350248
9WR	IV	A2	1	2	a	321142	325075
					b	332266	332745
	IV	B1	1	2	a	1178344	1181373
					b	1186623	1187267
	IV	L2	2	1	a	1503985	1504011
					b	1506000	1506534
					c	1514249	1514361
	V	I6	1	2	a	93419	93878
					b	95918	96109
					c	106822	107341

					d	107929	108381
	V	C1	1	2	a	165445	166837
					b	167640	168778
					c	177478	177570
	VII	F3	1	2	a	582439	582804
					b	585684	586091
					c	586091	589051
					d	589641	590531
	VII	K1	1	2	a	665927	670084
					b	670378	671690
					c	676446	676844
					d	677488	678458
					e	679031	679475
					f	701128	703861
					g	715098	715157
					h	717276	729400
					i	807421	807501
					j	807518	807579
					k	919536	920585
					l	921250	921313
					m	921313	921358
					n	921655	922011
	VIII	A2	1	2	a	125460	128730
					b	131716	131893
	IX	F1	2	1	a	227050	227148
					b	227148	227298
					c	227306	227358
					d	230144	230506
	X	A2	2	1	a	163652	163826
					b	172656	175026
	XI	E1	2	1	a	15916	18015
					b	24894	25338
					c	30575	30717
					d	31755	33674
	XI	A2	1	2	a	526070	527065
					b	537771	538292
	XI	B2	1	2	a	584267	584460
					b	584460	584586
	XII	K2	1	2	a	633784	634129
					b	634191	634579
					c	642559	643162

					d	646908	649864
					e	656914	657024
					f	657047	657236
					g	751400	751829
					h	752105	755555
	XIII	A2	2	1	a	355164	355412
					b	363615	363819
	XIII	J14	1	2	a	469570	470696
					b	475854	476481
					c	477050	478086
					d	478637	478672
					e	478672	479145
					f	479745	480326
	XV	K3	2	1	a	439049	439789
					b	441489	442429
					c	443073	443157
					d	444291	445218
					e	445218	445650
					f	446486	447364
					g	509319	509621
					h	513338	513608

Table S 2.8 Mutations induced by γ ray or UV treatment

Eight strains derived from four sectored colonies (two each from UV- and γ ray-treated cells) were sequenced and de novo mutations were identified. If the mutations were observed in both sectors of a treated colony, in the "Strain" column, the letters R (red) and W (white) are written. The chromosome number and SGD coordinate of the altered base are shown in "Chr" and "SGD pos" columns, respectively. For this table only, we use the SGD coordinates dated December, 2011; most of the SGD coordinates of February, 2010 differ from those of December, 2011 by less than 100 bp. The L431X nonsense mutation ("Mutated gene" column) is in the essential gene SMC3. As expected, when the diploid containing this mutation was sporulated, two of the four spores of each tetrad failed to grow.

Treatment	Strain	Chr	SGD pos	Reference	Mutation	Mutated gene	A. a. change
X-ray	37W/37R	11	333864	C	T	*NUP120*	G36R
X-ray	40W/40R	11	168580	C	T	*DBR1*	G84S
X-ray	40W/40W	14	298475	G	A	*YNL181W*	W47X
UV	8W/8R	7	238320	C	T		
UV	8W	10	206437	C	T	*MDV1*	P379L
UV	8W	11	74661	A	G	*tN(GUU)K*	
UV	8W/8R	11	159842	G	A	*RSM22*	P501S
UV	8W/8R	11	332479	C	A	*NUP120*	M497I
UV	8W/8R	13	396919	A	T		
UV	8W/8R	15	956654	C	T		
UV	8W/8R	16	93515	A	G	*IQG1*	L532S
UV	8R	10	301558	A	T	*SMC3*	L431X
UV	9W	10	171529	A	G	*URA2*	L280P
UV	9W	13	463981	A	G	*YMR099C*	Y283H
UV	9R	14	62590	G	A		

3 High-resolution mapping of spontaneous mitotic reciprocal crossovers on the right arm of the *Saccharomyces cerevisiae* chromosome IV

Homologous recombination (HR) is an important pathway for the repair of double strand breaks (DSBs) in the DNA of mitotically-dividing eukaryotic cells. Recombination between the homologs in diploid organisms, however, can also have negative consequences. For example, in a human cell that is heterozygous for a mutation in a tumor suppressor gene, a crossover event centromere-proximal to the mutation can result in the mutation becoming homozygous in one of the daughter cells. Such loss of heterozygosity (LOH) events are likely to be an early step in tumorigenesis for many solid tumors (FISCHER and STRINGER 2008).

To investigate the properties of mitotic recombination, we mapped about 140 spontaneous reciprocal crossovers (RCOs) on the right arm of the yeast chromosome IV using SNP microarrays. This 1.1 Mb region contains about 10% of the yeast genome. Our mapping and subsequent experiments demonstrate that inverted repeats of Ty retrotransposable elements can be homolog-specific mitotic recombination hotspots in the absence of replication perturbation, and that the hotspot activity is quenched if the inverted repeat structure is disrupted. Additionally, about 70% of the DNA lesions that result in LOH are generated during G1 of the cell cycle and repaired during S or G2. We also show that different genetic elements are associated with reciprocal crossover conversion tracts depending on the cell cycle timing of the initiated DSB. G-quadruplex

motifs are associated with crossovers that are initiated by lesions during G1 of the cell cycle, and replication-termination regions are associated with crossovers initiated by lesions in S or G2.

3.1 Introduction

A DNA double-strand break (DSB) is a potentially lethal DNA lesion that can lead to genomic instability and chromosome rearrangements if not promptly repaired. DSBs and other forms of DNA damage (for example, single-stranded nicks or base damage) can result from exogenous sources such as γ or UV radiation. They can also form as a result of endogenous DNA damage. Possible sources of endogenous DNA damage resulting in DSBs include head-on collisions between the replication and transcription machinery AGUILERA (2002), and processing of DNA secondary structures such as those formed by certain triplet repeats (FREUDENREICH *et al.* 1998; TANG *et al.* 2011) or by inverted repeats (GORDENIN *et al.* 1993; LEMOINE *et al.* 2005; LOBACHEV *et al.* 2000).

Inverted repeats have been called "at-risk motifs" because of their ability to form recombinogenic secondary structures (GORDENIN and RESNICK 1998). Inverted repeats are capable of forming either hairpin structures when single-stranded (intrastrand pairing between the inverted repeats) or cruciform structures (both strands forming hairpins at the same location) if the DNA is double stranded (LILLEY 1980;

PANAYOTATOS and WELLS 1981). Hairpin structures can cause replication fork pausing and lead to DSBs (VOINEAGU *et al.* 2008).

Increasing either the size of a spacer between IRs or the amount of sequence divergence between the repeats results in less frequent formation of the cruciform structure (VOINEAGU *et al.* 2008). A cruciform structure is thought to promote genome instability because this structure resembles a Holliday junction that can be cleaved by Holliday junction resolvases resulting in hairpin-capped ends. The resolvase responsible for cleavage of a cruciform is unknown, but Mre11p is necessary for repair of the hairpin-capped ends (LOBACHEV *et al.* 2002). A pair of inverted repeats of Ty elements located on chromosome III, known as FS2, have been previously shown to be a source of genomic instability in yeast cells with low levels of alpha or delta polymerase (LEMOINE *et al.* 2008; LEMOINE *et al.* 2005). There are at least five locations in the yeast genome, including FS2, in which Ty repeats are present as inverted repeats with less than 1 kb between the repeats. These locations (distance between the repeats shown in parentheses) are: chromosome III at coordinate 167 kb (FS2, 283 bp spacer), chromosome IV from 872 to 884 kb (524 bp spacer), chromosome IV from 980 to 992 kb (25 bp spacer), chromosome VII from 811 to 823 kb (these Ty repeats overlap by 4 bp), and chromosome XVI from 844 to 856 kb (189 bp spacer). Since FS2 is composed of two Ty1 repeats that are not present in the Saccharomyces Genome Database (SGD), the location coordinate indicates the site of insertion in strains that have FS2.

In *S. cerevisiae*, DSBs can be repaired by the relatively inefficient non-homologous end joining (NHEJ) pathway or by homologous repair (HR) (SYMINGTON and GAUTIER 2011). HR uses an intact donor DNA molecule (usually either a sister chromatid or a homolog) to repair the broken recipient molecule. In addition, if the DSB is located within a repeated gene, ectopically-located repeats on other chromosomes can act as a substrate for homologous recombination (ARGUESO *et al.* 2008). If the repair event uses the homolog as a template, it can result in conversions (the non-reciprocal transfer of DNA sequence) that can be visualized as LOH events. Conversion events can be associated or unassociated with reciprocal crossovers (RCOs) (swapping of the donor and recipient chromosomes from the site of the break to the end of the molecule); in meiosis, about half of the gene conversion events are associated with crossovers (PETES *et al.* 1991). Conversion events result from either correction of mismatches in heteroduplex DNA by the cellular mismatch repair (MMR) system (MODRICH and LAHUE 1996), or repair of a double-stranded DNA gap located near the site of the DSB (ORR-WEAVER and SZOSTAK 1983).

Previously in the Petes lab, diploid yeast strains heterozygous for many (55,000) single-nucleotide polymorphisms (SNPs) markers distributed throughout the genome were used to map conversion tracts associated with spontaneous RCOs on chromosome V (LEE *et al.* 2009), and crossovers induced by ionizing and UV radiation (LEE and PETES 2010; ST CHARLES *et al.* 2012). The diploids used in the mapping experiments were

generated by crosses of the sequence-diverged haploids, YJM789 and W303a (WEI *et al.* 2007). As described in detail in Chapter 2, we used oligonucleotide-containing microarrays to detect regions of LOH throughout the genome in cells treated with UV or γ rays (ST CHARLES *et al.* 2012).

In this chapter, I describe our mapping of crossovers and associated gene conversion events on the right arm of chromosome IV, an interval of about 1.1 Mb. Unlike the analysis described in Chapter 2, we concentrate on spontaneous events rather than those induced by DNA damaging agents. The previous study of spontaneous crossovers on chromosome V (LEE *et al.* 2009) was restricted to a small region (about 100 kb) of the genome that lacked a number of common chromosome elements such as retrotransposons. In addition, the mapping of the crossover events was done at relatively low (about 3 kb) resolution. In my analysis of events on chromosome IV, I use single-nucleotide polymorphism (SNP) microarrays to map events at approximately 1 kb resolution.

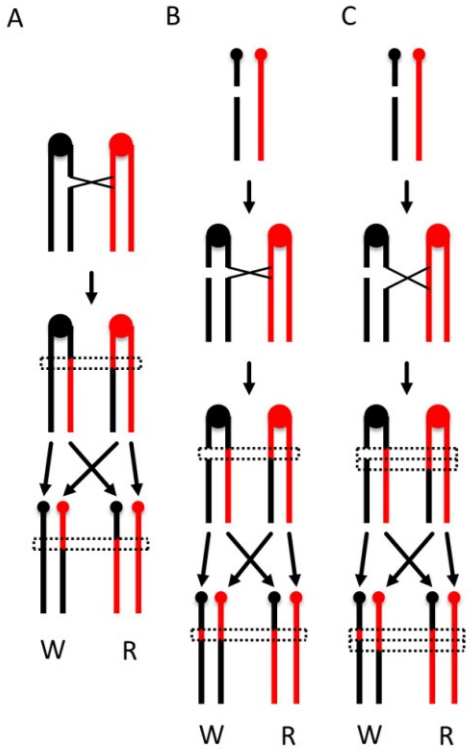


Figure 3.1. Patterns of gene conversion resulting from the repair of G1- or G2-generated DSBs.

Chromatids that are depicted in black represent YJM789-derived chromatids, and those in red represent W303a-derived chromatids. "R" and "W" indicate the homologs present in the red and white sectors, respectively, of a sectored colony. The centromeres are shown as circles.

A. 3:1 conversion event. Repair of a G2-associated DSB results in a 3:1 gene conversion tracts (enclosed in dotted lines) associated with the LOH event. Note that LOH distal to the conversion tract is observed if the daughter cells contain one non-recombinant and one recombinant chromatid. If one daughter has two recombinant chromosomes and the other two non-recombinant chromatids, no LOH is observed. Thus, about half of the crossovers result in detectable LOH. Note also that the chromatid that receives the DSB acts as a recipient of information from the unbroken chromatid of the other homolog.

B. 4:0 conversion event. A chromosome in a G1 cell is replicated to form two sister chromatids that are broken at approximately the same position. Repair of these two broken chromatids in G2 can result in a region in which all four chromatids have identical SNPs, a 4:0 conversion tract. For this event, both conversion tracts are of the same length.

C. 4:0/3:1 hybrid tract. As in Figure 3.1B, a broken chromosome is replicated to form to two broken sister chromatids. The conversion tracts associated with the repair of the two DSBs, however, are of different lengths, resulting in a 4:0/3:1 hybrid tract.

Previous studies in the lab have indicated that differences in the cell cycle timing of the DSBs that initiate recombination result in different types of conversion tracts (LEE *et al.* 2009; LEE and PETES 2010). Figure 3.1 shows a comparison of conversion tracts resulting from repair of a DSB formed during S or G2 of the cell cycle (Figure 3.1A), and conversion tracts resulting from repair of a DSB formed in G1 (Figure 3.1B and C). Figure 3.1 depicts reciprocal crossovers in a color-sectoring system in which one daughter cell will give rise to a while colony and the other daughter cell will give rise to a red colony (designated by the "W" and "R" letters). This system will be described in further detail in the Results section. Figure 3.1A shows a DSB on the YJM789 homolog during G2 of the cell cycle that is repaired by homologous recombination, resulting in a crossover and associated 3:1 conversion tract. If the chromatids segregate to generate distal LOH, there will be a short conversion tract near the site of the crossover, such that the red sector (as depicted in the figure) is homozygous for the W303a background and the white sector is heterozygous at the same location. This type of conversion is called a 3:1 conversion event because within the conversion tract, three of the chromatids have red SNPs and only one chromatid has black SNPs. Figure 3.1B shows a DSB on the

YJM789 homolog during G1 of the cell cycle that has been replicated to generate two broken chromatids. Since both chromatids are broken in the same location, only the homolog is available as a template for homologous recombination. If one chromatid is repaired resulting in a crossover and associated conversion event, and the second chromatid is repaired as a conversion event unassociated with crossovers, following chromosome segregation to generate LOH, there will be a 4:0 conversion tract (indicated in the dotted line box in Figure 3.1B). Figure 3.1C also shows a DSB on the YJM789 homolog during G1 of the cell cycle. Following replication, repair occurs as in Figure 3.1B, but one broken chromatid has been processed to a greater extent than the other, resulting in unequal-sized conversion tracts. Following chromosome segregation to generate LOH, the resulting conversion tract is a 4:0/3:1 hybrid tract.

We found that mitotic recombination can occur throughout the chromosome arm, although the distribution of certain types of exchanges is non-random. In particular, the inverted repeat formed by two Ty retrotransposons, oriented in tail-to-tail configuration, located around SGD coordinates 980 kb is a hotspot for reciprocal crossovers initiated by DSBs during G1 of the cell cycle in the W303a-derived homolog. Our results also demonstrate that LOH events are more frequently a consequence of G1-initiated DSBs than G2-initiated DSBs, and that conversion tracts associated with the G1 DSBs are longer than those associated with G2-initiated DSBs. I also found that different

genetic elements are associated with reciprocal crossovers initiated during G1 of the cell cycle in comparison to those initiated during G2 of the cell cycle.

3.2 Materials and Methods

3.2.1 Strains and construction

A list of all strains used in this chapter are present in Table S 3.1, and a list of all primers used in this study are listed in Table S 3.2, unless otherwise specified. Mapping of crossovers was performed with the strain JSC25 (*MATa/MATα::HYG leu2-3,112/LEU2 his3-11,15/HIS3 ura3-1 /ura3 GAL2/gal2 ade2-1/ade2-1 trp1-1/TRP can1-100::NAT/ CAN1::NAT RAD5/RAD5 IV1510386::KANMX-can1-100/IVI1510386::SUP4-o*) which is a hybrid diploid generated by mating the sequenced strains W303a and YJM789 (WEI *et al.* 2007). A detailed description of the strain construction is presented in the Supplementary Information. In brief, the *KANMX* gene (which renders cells resistant to geneticin) was inserted at position 1510386 on the W303a-derived chromosome IV homolog, and the *SUP4-o* gene was inserted at the allelic position on the YJM789 homolog. These two haploids were mated. As discussed in the Introduction, since JSC25 is homozygous for the nonsense-suppressible *ade2-1* mutation, crossovers result in red/white sectored colonies. As in previous studies (LEE *et al.* 2009), we replaced the *MATα* locus with the *HYG* gene. This manipulation prevents the diploid from sporulating to generate artifactual red/white sectors.

We also used strains that had modifications of the HS4 recombination hotspot (inverted pair of Ty elements) located near SGD coordinate 980 kb on chromosome IV. The strain construction is also presented in detail in the Supporting Information. In brief, in the control diploid JSC71-1, in the W303a-derived homolog, the *HYG* gene was inserted at about position 957 kb, which is upstream of the inverted repeat hotspot; in the same homolog, the *URA3* gene was inserted at about position 1013 kb, which is downstream of the hotspot. We then constructed several derivatives of this strain: 1) a strain in which the centromere-proximal Ty2 element was replaced by *KANMX* (JSC73-2), 2) a strain in which the delta element associated with the centromere-proximal Ty2 was replaced by *KANMX* (JSC74-1), and 3) a strain in which *KANMX* was inserted between the two Ty elements of HS2 (JSC77-1).

3.2.2 Media

All media were prepared using standard recipes (Guthrie and Fink 1991), except that the SD-Arg plates contained only 10 μg/mL of adenine, as previously described (BARBERA and PETES 2006).

3.2.3 Assay for crossovers

To detect crossovers, we first streaked JSC25 for single colonies onto rich solid medium (YPD plates). The cells were then grown for three days at 30°. A single (pink) colony was then resuspended in water, and plated onto SD-Arg plates at a density of about 2000-5000 cells per plate. Plates were incubated at 25° for three days. The plates

were then put at 4° for an additional day to let the red pigmentation develop. Plates were examined using a dissecting microscope to look for red/white sectors. In colonies in which the red sector was at least one-eighth of the colony, we purified colonies from both the red and white sides of the sector. We confirmed that the red and white colonies were Ade^- and Gen^S, respectively. Single red and single white colonies from each sectored colony were subsequently analyzed.

3.2.4 Microarray methods and analysis

Microarray samples were prepared, hybridized, and analyzed as previously described (ST CHARLES *et al.* 2012). Briefly, agarose plugs containing genomic DNA were prepared from strains derived from the sectored colonies and from the control diploid reference strain. DNA from both the experimental and reference strains was sonicated to yield DNA fragments of about 200-400 base pairs (bp); these samples were labeled with Cy5-dUTP (experimental) and Cy3-dUTP (control). DNA samples from the experimental and control strains were mixed and competitively hybridized to SNP microarrays. Following washing and scanning of the arrays, probe signals were analyzed as ratios of hybridization of the experimental and control samples for each SNP-specific oligonucleotide represented on the arrays.

3.2.5 Microarray design

The chromosome IV SNP microarray was designed and optimized following the principles described previously (ST CHARLES *et al.* 2012). In short, most of the SNPs

analyzed on the microarray are represented by four 25-bp oligonucleotides: two specific to the Watson and Crick strands of the W303a form of the SNP, and two specific to the Watson and Crick strands of the YJM789 form of the SNP. For the majority of the probes, the SNP is located in the center of the oligonucleotide, and the melting temperature for the oligonucleotide/genomic DNA hybrid is between 55 and 59° C. Minor deviations from these principles are described in the Supporting Information. To map events on chromosome IV, we used microarrays containing 9200 oligonucleotides on IV, allowing us to assay about 2300 SNPs. The location and sequence of each SNP on the microarray will be included in the Supplementary Information online, but are not included in the thesis.

3.2.6 Assay of recombination activity of a recombination hotspot containing inverted repeats

In order to assay the frequency of recombination at the hotspot containing the inverted pair of Ty elements, we inserted *HYG* and *URA3* centromere-proximal and centromere-distal to the hotspot, respectively. Both markers were inserted on the W303a-derived homolog. A crossover between the markers generates a 5-fluoro-orotate-resistant (5-FOA^R) derivative that is also resistant to hygromycin. A crossover centromere-proximal the *HYG* gene results in a 5-FOA^R Hyg^S strain. Consequently, we examined the ratio of Hyg^R 5-FOA^R colonies to the total amount of 5-FOA^R colonies derived from the experimental strains containing the *HYG* and *URA3* markers (JSC71-1, JSC73-2, JSC74-1, and JSC77-1). Each strain was struck for single colonies on YPD plates.

After incubating the cultures at 30° for 30 hours, we made patches of individual colonies on plates containing medium with 5-FOA. The plates were then incubated for two days at 30°. A single 5-FOA^R colony from each patch was then transferred as a patch to YPD plates. After one day of growth at 30°, the samples were replica-plated to YPD plates containing hygromycin. The Hyg^R 5-FOA^R / total 5-FOA^R ratio was then determined by counting the number of patches that grew on the YPD+Hyg plates and the total number of 5-FOA^R colonies examined. About 400 5-FOA^R colonies were analyzed for each strain.

3.2.7 Statistical analysis

We used chi-square goodness-of-fit tests to compare expected values with observed values. VassarStats (http://vassarstats.net/) and Microsoft Excel were used for this statistical test. We used Table B11 of ALTMAN (1991) to calculate 95% confidence intervals on median estimates of conversion tract length. The Mann-Whitney test from the VassarStats website was used to compare conversion tract lengths for events initiated during G1 versus G2 of the cell cycle.

3.3 Results

3.3.1 Identification and mapping of crossovers on chromosome IV.

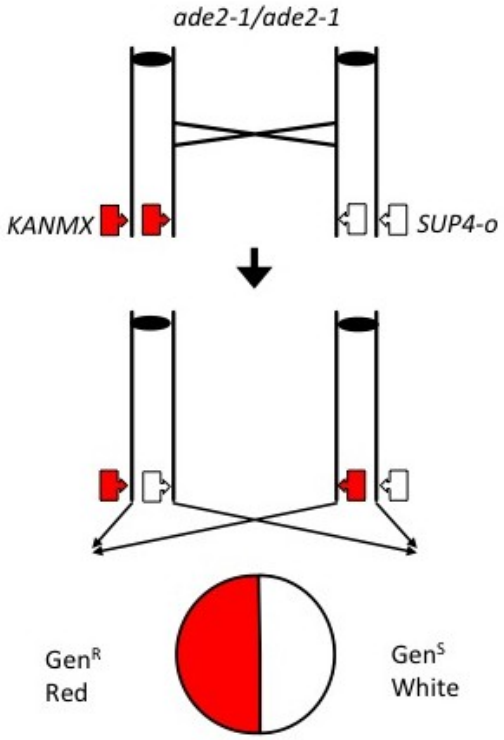


Figure 3.2. Crossover screening system

We used a colony-color screening assay to identify reciprocal crossovers on chromosome IV, similar to the system that our lab has previously used to study mitotic recombination on chromosome V (BARBERA and PETES 2006; LEE *et al.* 2009)(Figure 3.2). Our experimental diploid, JSC25, is a hybrid generated by mating two sequenced haploid strains, W303a and YJM789. We inserted the geneticin-resistance gene (*KANMX*) near the telomere on the right arm of the W303a-derived copy of chromosome IV. At the allelic position, on the YJM789-derived homolog, we inserted the *SUP4-o* gene, which

encodes an ochre-suppressing tRNA. JSC25 is also homozygous for the ochre-suppressible *ade2-1* mutation. Diploid strains of this genotype that lack *SUP4-o* form red colonies, strains with one copy of the suppressor form pink colonies, and those with two copies of *SUP4-o* form white colonies.

If there is a crossover between the centromere and the *SUP4-o* locus, there is a 50% chance, following chromosome segregation that the two daughter cells will be homozygous for the region centromere distal of the crossover (CHUA and JINKS-ROBERTSON 1991). If the crossover occurs at the time that the diploid JSC25 is plated, the daughter cell that is homozygous for *SUP4-o* will produce a white Gen^S sector (Figure 3.2). The other daughter cell will produce a red Gen^R sector. Although one would expect that the red and white sectors would be of equal sizes if the crossover occurred in the cell that was initially plated, we observed at least two-fold variation in the sizes of the red and white sectors. In addition, we also observed single colonies that had pink, white, and red sectors. Such colonies presumably reflect crossovers that occurred after the plated cells had started to divide. In our experiments, we did not analyze sectored colonies unless the red portion of the colony was more than one-eighth of the colony. One potential artifact that could generate a red sector in a diploid that is heterozygous for *SUP4-o* is sporulation, resulting in haploid spore colonies that lack *SUP4-o.* To prevent this complication, we deleted *MATα* in constructing JSC25. Diploid strains usually require both mating types in order to sporulate (HOPPER and HALL 1975).

As discussed above, the frequency of "pure" red/white sectored colonies should reflect the rate of crossovers between *CEN4* and the *KANMX*/*SUP4-o* markers. We observed a frequency of red/white sectored colonies of 3.1×10^{-5}/division. Since only one half of reciprocal crossovers on the right arm of chromosome IV result in a sectored colony, the crossover frequency is twice the sector frequency or 6.2×10^{-5}/division.

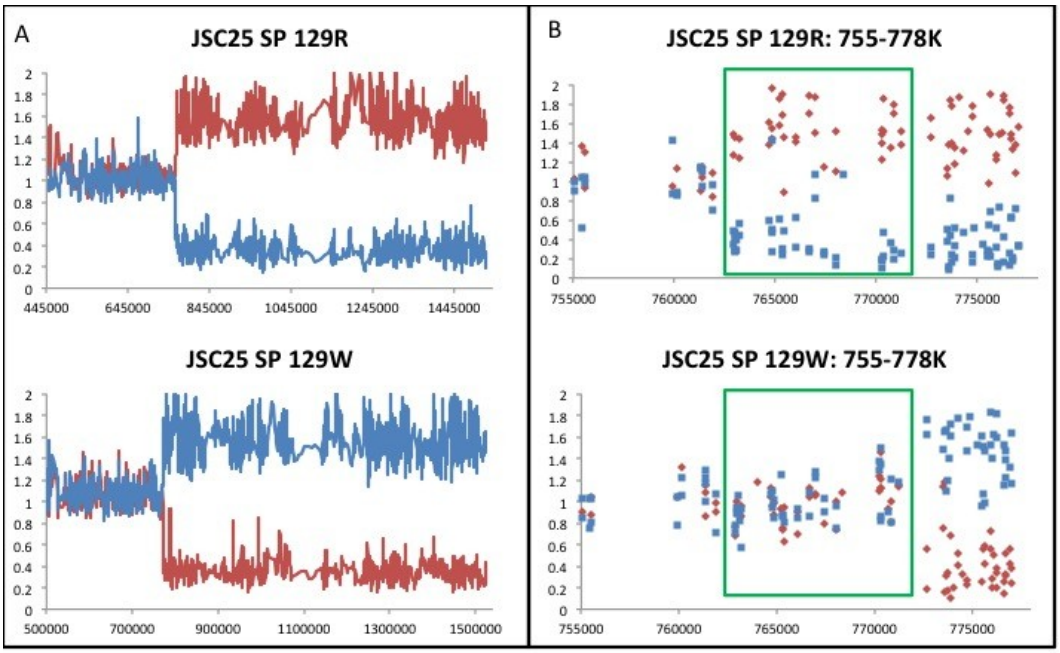


Figure 3.3. Mapping of a crossover with an associated 3:1 conversion event by SNP microarrays.

The values on the Y-axis show the experimental/reference hybridization ratio of genomic DNA to oligonucleotides that are specific to SNPs from the W303a and YJM789 backgrounds. The values on the X-axis indicate the SGD coordinates of the SNPs along chromosome IV. The red and blue lines depict the hybridization levels of probes specific for the W303a and YJM789 backgrounds, respectively. Where both the red and blue lines have a value of 1, the diploid is heterozygous for the SNPs. If the strain is homozygous for the W303a variant, the red signal increases to about 1.6 and the blue signal diminishes to about 0.3. If the strain is homozygous for the YJM789 variant, we observe the reverse pattern.

A. Low-resolution depiction of a reciprocal crossover analyzed by the SNP microarrays. In both the red (top plot; JSC25 SP 129R) and white (bottom plot; JSC25 SP 129W) sectors, there is a transition between heterozygosity and homozygosity at approximately SGD coordinate 770 kb.

B. High-resolution depiction of a reciprocal crossover. These two graphs show the hybridization to strain-specific SNPs at high resolution. Each square and diamond indicates the hybridization level to a specific oligonucleotide. The genomic DNA from the red sector has a transition between heterozygosity and homozygosity at approximately coordinate 762.5 kb (top panel), whereas genomic DNA from the white sector has a transition at about coordinate 772 kb. Thus, the regions shown in green rectangles represent the 3:1 gene conversion event.

We mapped crossovers in about 140 sectored colonies of JSC25. Based on the sequences of W303a and YJM789 (the two haploids used to generate the diploid), this strain is heterozygous for about 55,000 SNPs. In order to characterize a mitotic recombination event on chromosome IV, one needs to determine whether SNPs are heterozygous (as in the original progenitor diploid) or homozygous in the region between *CEN4* (SGD coordinate of about 450 kb) and the *KANMX*/*SUP4-o* markers (SGD coordinate about 1,510 kb). We examined about 2300 SNPs in this 1.1 Mb region using oligonucleotide-containing microarrays, resulting in a mapping resolution of about 500 bp. As in our previous study (St Charles *et al.* 2012), for each SNP, we designed four 25-base oligonucleotides: two representing the Watson and Crick strands of the W303a form of the SNP and two representing the Watson and Crick strands of the YJM789 form of the SNP. If the hybridization of genomic DNA to these oligonucleotides is done under stringent conditions, it is possible to determine whether a particular strain is

homozygous for the W303a- or YJM789-derived SNPs, or is heterozygous (GRESHAM *et al.* 2010; ST CHARLES *et al.* 2012). Genomic DNA samples were prepared from both the red and white sectors. In separate experiments, these samples were labeled with Cy5-dUTP and hybridized in competition with the control diploid strain labeled with Cy3-dUTP. The normalized ratio of hybridization of these samples to each oligonucleotide allowed us to determine whether the experimental strains were heterozygous or homozygous for a specific SNP. In previous experiments (St Charles *et al.* 2012), we used high-throughput DNA sequencing to confirm the validity of results obtained with the SNP array.

Figure 3.3 shows an example of the analysis of genomic DNA from a sectored colony (JSC25 SP 129). In this figure, the hybridization of genomic DNA to the W303a-derived SNPs is shown by a red line and the hybridization of genomic DNA to the YJM789-derived SNPs is shown by a blue line. In the low resolution array analysis for the red side of the sector (top part of Figure 3.3A), genomic DNA from the red sector hybridizes equally efficiently to both W303a- and YJM789-derived SNPs (normalized ratio of 1) from SGD coordinates 445 kb to 770 kb. At approximately SGD coordinate 770 kb, there is a transition from heterozygosity to homozygosity for the W303a form of the SNPs. In the white sector (bottom part of Figure 3.3A), the genomic DNA has a heterozygous region of approximately the same size as the red sector and a transition to homozygosity for the YJM789 form of the SNPs at approximately the same position. This

LOH pattern is expected for a reciprocal crossover occurring near SGD coordinate 770 kb. Based on the higher resolution depiction (Figure 3.3B), it is clear that the transition points between heterozygosity and homozygosity are not identical between the two sides of the sector. This transition occurs at SGD coordinate 762 kb in the red sector and 772 kb in the white sector. The region boxed in this figure represents a 3:1 gene conversion tract. The "3:1" notation indicates that within this boxed region, considering both sectors, three of the chromosomes have W303a-derived SNPs and only one has YJM789-derived SNPs. This pattern indicates that the gene conversion tract was initiated by a DNA lesion on the YJM789-derived chromosome, since the chromosome with the lesion acts as a recipient during the conversion event (PÂQUES and HABER 1999). Of 121 crossover-associated conversions, 29 (23%) were associated with a simple 3:1 conversion event.

In Figure 3.4, we show a crossover associated with a 3:1/4:0/3:1 hybrid conversion tract. In the crossover event shown in this figure, the red sector has a hybridization pattern with a single transition, whereas the white sector has three transitions. In the region marked with the red box, all four chromosomes have the SNP derived from W303a (the 4:0 portion of the tract). This 4:0 region is flanked by two 3:1 regions (marked with green boxes). This type of hybrid conversion tract is explicable as reflecting the repair of two broken sister chromatids, with one repair tract extending beyond the other (LEE and PETES 2010; ST CHARLES *et al.* 2012). The region in which both

conversion tracts overlap make up the 4:0 part of the conversion tract, and the region in which one chromatid was processed to a greater extent becomes the flanking 3:1 parts of the conversion tract.

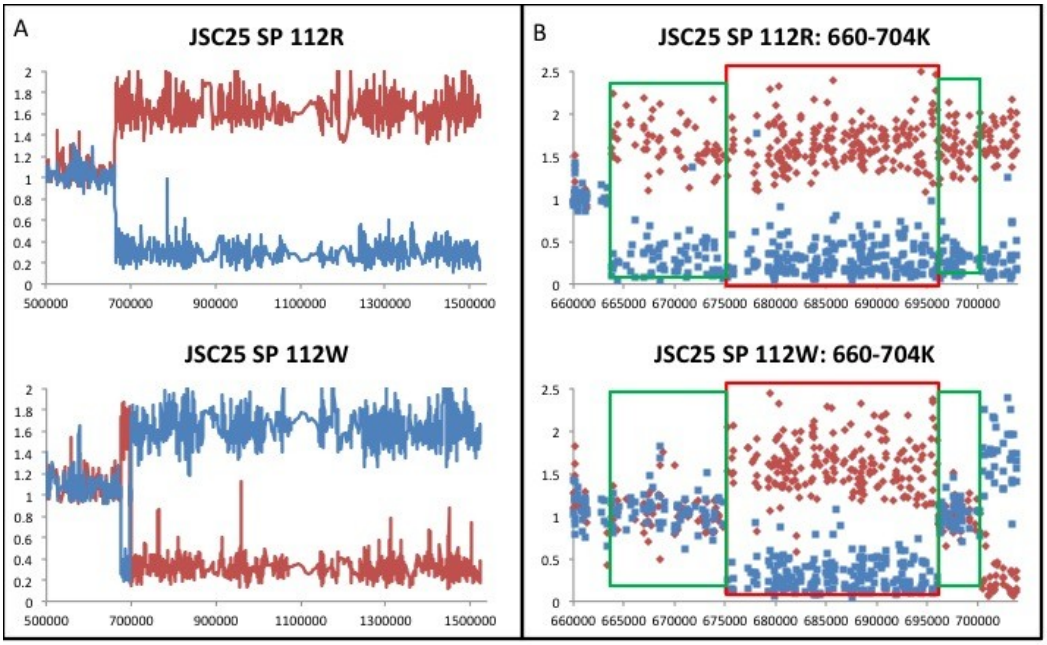


Figure 3.4. Mapping of a crossover with an associated 3:1/4:0/3:1 conversion event by SNP microarrays. The depiction of this event is the same as in Figure 3.3.

A. Low-resolution depiction of the crossover. In both the red (top plot; JSC25 SP 112R) and white (bottom plot; JSC25 SP 112W) sectors, there is a transition between heterozygosity and homozygosity at approximately SGD coordinate 700 kb, although it is evident in JSC25 SP 112W that there are at least two transitions.

B. High-resolution depiction of 3:1/4:0/3:0 conversion associated with a reciprocal crossover. A comparison between the patterns of SNP heterozygosity and homozygosity demonstrate that the crossover in this sectored colony was associated with a 3:1/4:0/3:1 conversion tract. The 3:1 segments of the tract are included within the green rectangles and the 4:0 portion is outlined in the red rectangles.

Of 139 crossovers analyzed, 121 (87%) were associated with a contiguous gene conversion event. Of the 121 conversion tracts, there were 29 simple 3:1 conversions, 7

simple 4:0 conversions, and 46 simple 4:0/3:1 or 3:1/4:0/3:1 hybrid conversions; 39 tracts had more complicated patterns of conversion (discussed below). In order to detect and classify gene conversion events, the patterns of heterozygous and homozygous SNPs in both sectors must be considered. In Figure 3.5, we show the method that we used to depict gene conversion events. The lines on top and bottom represent information from the red and white sectors, respectively. A green line indicates that the sector is heterozygous, a red line indicates that the sector is homozygous for the W303a form of the SNPs, and the black line indicates that the sector is homozygous for the YJM789 form of the SNPs. Each transition between heterozygosity and homozygosity is indicated with a lower case letter. Figure 3.5 shows crossovers with no associated conversions (Figure 3.5A), simple 3:1 conversions (Figure 3.5B), simple 4:0 conversions (Figure 3.5C), simple 3:1/4:0/3:1 tracts (Figure 3.5D), and one example of a complex tract (Figure 3.5E). A summary of all conversion tracts observed in our study is presented in Table S 3.3 with the SGD coordinates of each transition point given in Table S 3.4.

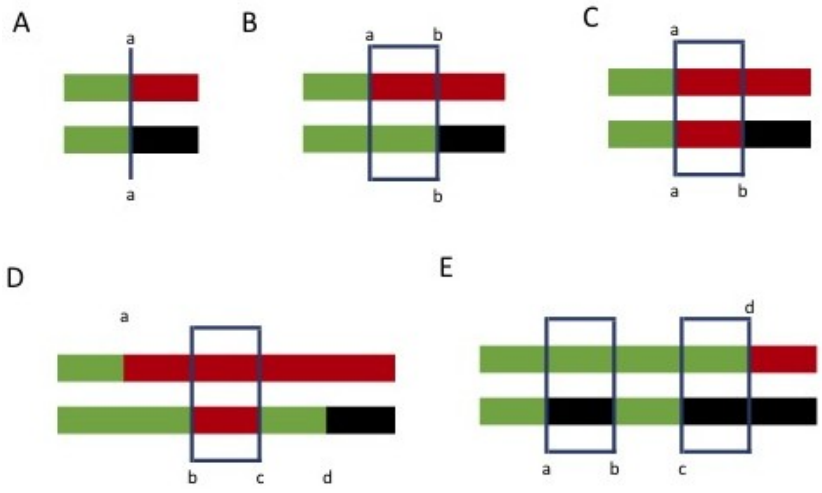


Figure 3.5. Depiction of conversion tracts and crossovers.

In order to diagnose a conversion event within a sectored colony, one must analyze LOH patterns for both the red and white sectors. In the images, green, black, and red colors indicate SNP heterozygosity, homozygosity for the YJM789-derived SNPs, and homozygosity of the W303a-derived SNPs, respectively. Each sectored colony is represented by two lines, with the top line showing the pattern of heterozygosity and homozygosity for the red sector and the bottom line showing the pattern for the white sector. The small letters mark the transition points (Table S 3.3 and Table S 3.4).

A. Crossover without an associated conversion.

B. Crossover with an associated 3:1 conversion.

C. Crossover with simple 4:0 conversion.

D. Crossover with an associated 3:1/4:0/3:1 hybrid conversion tract.

E. Crossover with a complex conversion tract.

3.3.2 Location of crossovers and associated conversions on the right arm of chromosome IV

We collected and analyzed 138 sectors and 139 reciprocal crossovers in the strain JSC25; one sectored strain had a double crossover. The locations of crossovers and

associated conversion tracts are shown in Figure 3.6. The crossovers associated with conversion tracts were separated into two categories: those in which the initiating double strand break occurred in the W303a homolog and those that occurred in the YJM789 homolog. The separation is based on which strain is the donor in the conversion event (Figure 3.1 legend). For example, if W303a-derived chromosome is the donor in a conversion tract, then the YJM789-derived homolog had the originating DSB; conversion events of this type are shown on the bottom part of Figure 3.6. Conversion events in which W303a-derived homolog had the initiating DSB are shown at the top part of Figure 3.6. Although this analysis was usually straightforward, about 2% (3/121) of the conversion tracts involved more than one donor (segments of red and black within one tract in Figure 3.6). For such tracts, we considered the homolog that had donated the smallest amount to the conversion tract to be the chromosome with the initiating DSB. Crossovers unassociated with conversion tracts are not shown in Figure 3.6, because it is impossible to determine which homolog had the originating DSB.

Some interesting conclusions can be made from the data in Figure 3.6. First, although the number of crossovers initiated from DSBs on the W303a- and YJM789-derived homologs is approximately equal (61 versus 59, respectively), the distribution of crossovers along each homolog appears to be different. For instance, there are 16 conversion tracts that start between 600 and 800 kb on the YJM789-derived homolog but only 5 conversion tracts initiated on the W303a-derived homolog within the same

region. In contrast, more conversion events in the region between 800 and 1000 kb are initiated on the W303a-derived homolog than the YJM789-derived homolog. In Figure 3.6, the lengths of the conversion tracts are indicated by the lengths of the pairs of lines. The conversion tract lengths in the 800-1000 kb region on the W303a-derived homolog are longer than those initiated on the YJM789-derived homolog. The reasons for these differences will be discussed further below.

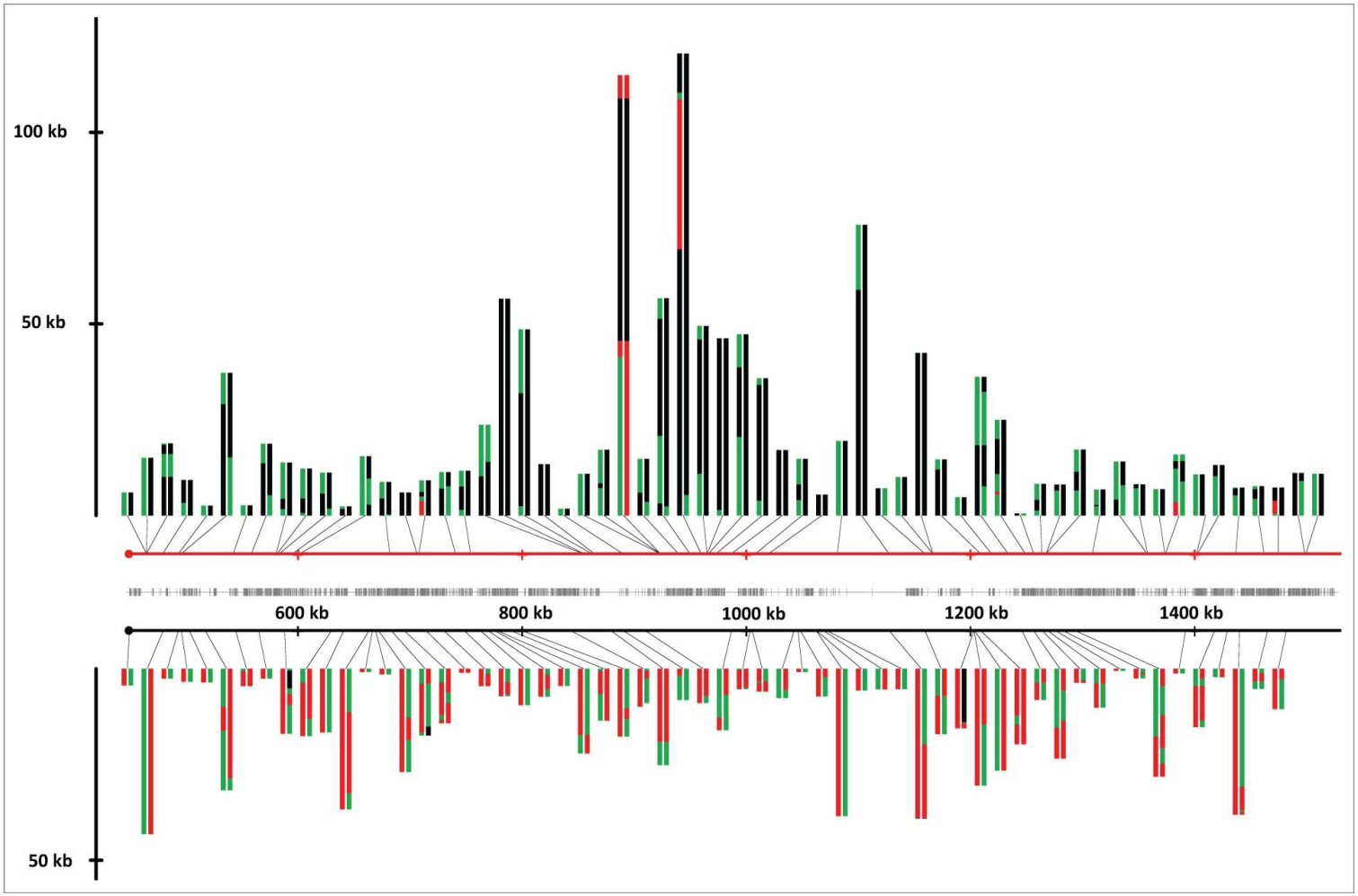


Figure 3.6. Map locations of crossovers and associated gene conversion events in JSC25 along the right arm of chromosome IV.

The X-axis represents the SGD coordinates (*CEN4* at 449 kb), and the Y-axis represents the length of the individual conversion tract. The red and black horizontal lines in the center of the figure depict the W303a- and YJM789-derived homologs, respectively. Individual conversion tracts are depicted as two horizontal vertical lines that are attached to either homolog line by a black connector line. Where the thin black lines join the horizontal red and black lines indicate the centromere-proximal margin of the conversion tract. The left line of each pair of lines shows the conversion tract of the red side of the sector, and the right line shows the white side of the sector. Green, red, and black show heterozygosity, homozygosity for W303a-derived SNPs, and homozygosity for YJM789-derived SNPs, respectively. Each segment of the depicted conversion tract is proportional to the size of that segment within the sector. The hash marks between the two homologs indicate the density of the SNPs represented on the microarray.

Figure 3.7 summarizes the location of crossover-associated gene conversion events in a different way. In this figure, the X-axis shows the SGD coordinates and the Y-axis shows the number of conversion events that include the SNP located at that coordinate summed over all mapped conversion events. Thus, the heights of the peaks reflect both the frequency of DNA lesions that initiate gene conversions, and the lengths of conversion tracts emanating from the sites of the DNA lesion. All crossover-associated conversion events are shown in Figure 3.7A. The median number of times individual SNPs were involved in a conversion event was about two. In Figure 3.7A, we label regions where multiple SNPs were involved multiple times in a gene conversion event as potential hotspots (HS1-HS7). In Figure 3.7B, we show separately those events initiated on the W303a- and YJM789-derived homologs. These distributions are strikingly different. It is also clear that the distributions of G1-initiated conversions (tracts with 4:0 segments or other features that indicate a G1-associated DSB) and G2-

associated conversions (3:1 conversions without a 4:0 segment) are different (Figure 3.7C). We will discuss the statistical analysis of these distributions below.

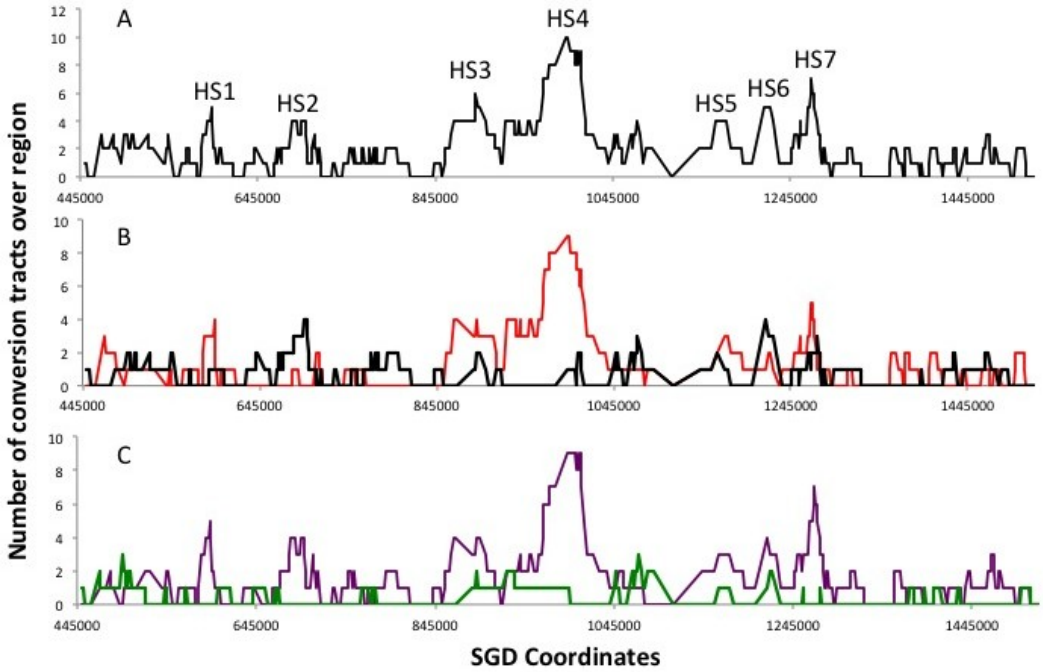


Figure 3.7. Summary of conversion tracts associated with crossovers along the right arm of chromosome IV.

These plots represent the number of times a SNP represented on the microarray is involved in a conversion tract. A SNP was considered involved if it was between the start and end of the conversion tract.

A. This portion of the figure includes all conversion tracts. The labels HS1-HS7 represent potential hotspots for recombination.

B. This plot is the same analysis as Figure 3.7A except that we show the conversion events initiated on the W303a-derived homolog (red) separately from those initiated from the YJM789-derived homolog (black).

C. This plot is the same analysis as Figure 3.7A, except we show the conversion events initiated in G1 (those with a 4:0 tract, indicated in purple), separately from those initiated in G2 (those with a 3:1 tract, indicated in green).

3.3.3 Statistical analysis of the distribution of crossovers and conversions

In Figure 3.7A, the crossover-associated gene conversion events appear clustered into the middle third of the right arm of chromosome IV. As discussed above, however, the depiction of events shown in Figure 3.7A is somewhat misleading since the number of times a SNP is involved in a conversion event is a function of both the frequency of initiating events near the SNP and the length of the conversion tracts. For example, if the SNPs are 1 kb apart, a chromosome region in which conversion tracts are 10 kb in length will result in a 10-fold elevation in the number of times a SNP is included within the tract compared to a region in which the conversion tracts are only 1 kb in size, even if the frequency of recombinogenic DNA lesions is identical for the two chromosome regions. One reason that this issue is of concern is that many of the conversion tracts near the middle of the chromosome arm are longer than the average (Figure 3.6). Consequently, as described below, we developed methods of performing the statistical analysis that eliminated the effect of the size of the conversion tract.

We used several different statistical methods to determine if the distribution of crossovers along the chromosome arm is non-random. In our first statistical analysis, we divided the right arm of chromosome IV into five bins of 200 kb (starting at the centromere at coordinate 445 kb) and a sixth bin of about 80 kb. We counted the number of crossovers that occurred within each region (Figure 3.8A). For those events in which

the crossover was associated with a conversion tract that spanned two bins, we assigned the crossover to a bin based on whether the mid-point of the conversion tract was within the bin (additional details discussed in the Supporting Materials and Methods at the end of this chapter). As can be seen in Figure 3.8A, if all events are examined, the number of crossovers per bin is approximately proportional to the size of each bin (p=0.96 by chi-square goodness-of-fit). However, when we examined specific classes of events within each bind, we observed significant departures from a random distribution (Figure 3.8B).

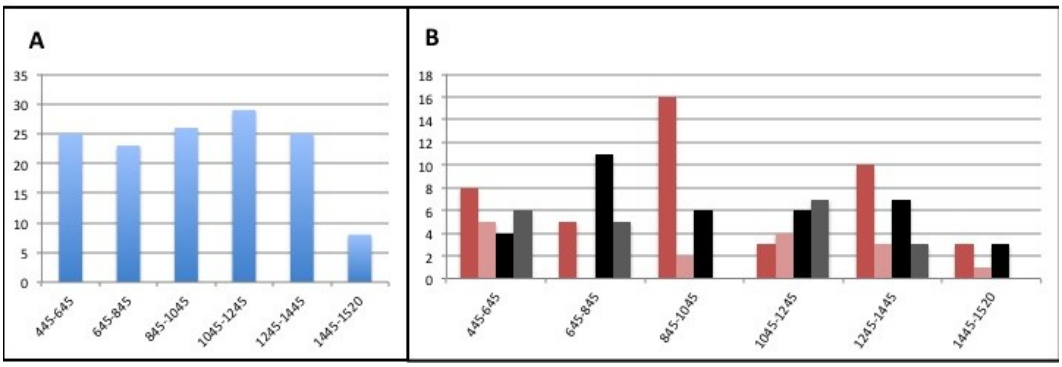


Figure 3.8. Distribution of JSC25 crossovers into six intervals along the right arm of chromosome IV.

As described in the text, we grouped the observed crossovers into six bins along the right arm of chromosome IV. The Y-axis shows the number of events/bin, and the X-axis shows the SGD coordinates of each bin.

A. For this analysis, all crossovers were examined without regard to which homolog had the initiating DNA lesion or whether the conversion event was initiated by a G1- or G2/S-associated DSB.

B. For this analysis, we classified the crossovers in each bin into four groups: crossovers initiated on the W303a-derived chromosome in G1 (red), crossovers initiated on the W303a-derived chromosome in G2 (pink), crossovers initiated on the YJM789-derived chromosome in G1 (black), and crossovers initiated on the YJM789-derived chromosome in G2 (gray).

The data depicted in Figure 3.8B were examined in various ways using a contingency chi-square test. By this test, we determined that the distribution of crossovers initiated on the W303a-derived chromosome was significantly different from the distribution of crossovers initiated on the YJM789-derived chromosome ($p=0.01$). This result suggests that the W303a and YJM789 homologs have different distributions of recombinogenic DNA lesions. Additionally, the distributions of DSBs initiated during G1 or G2 of the cell cycle were significantly different from each other ($p=0.007$). Both of these characteristics can also be seen in Figure 3.7B and C. We also note that the bin from 845-1045 kb has a very significant enrichment of crossovers initiated on the W303a-derived homolog during G1 of the cell cycle compared to a random distribution of this class of event among all of the bins ($p=2.3 \times 10^{-5}$).

The conversion tract lengths in different regions of chromosome IV, as defined by the bins, are also non-random. We calculated the conversion tract length by taking the midpoint of the two markers flanking the transitions at each end of the tract (the first transition from heterozygosity to homozygosity in either sector and the last transition in which both sectors become homozygous). Figure 3.9 shows both the average conversion tract length in each 200 kb bin. The median conversion tract length for all events examined is 10.6 kb, and the average conversion length is 16.7 kb. For these calculations, we estimated conversion tract length for crossovers with no detectable conversion event as half the distance between the last heterozygous probe and the first homozygous

probe. With the exception of the two bins from 845 to 1045 kb and from 1045 to 1245 kb, all of the bins have average conversion tract lengths of about 11 kb.

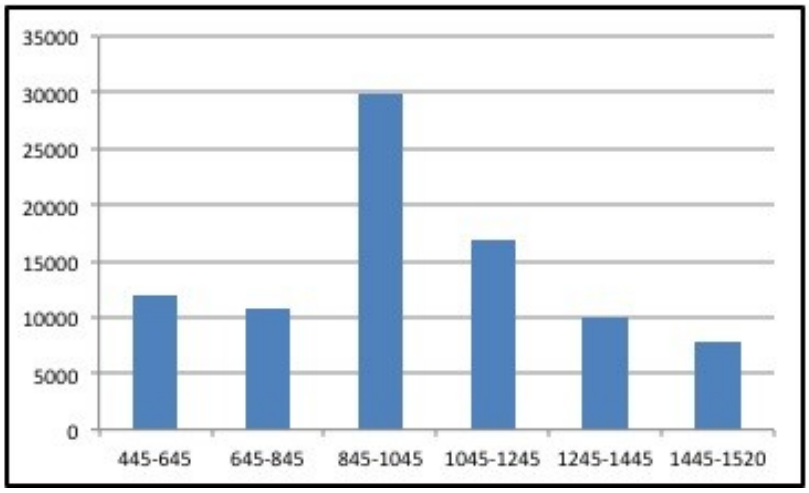


Figure 3.9. Conversion tract lengths along the right arm of chromosome IV.

Average conversion tract length in six windows along the right arm of chromosome IV. The Y-axis shows the length of the conversion tract in kb and the X-axis shows the SGD coordinates of the bins which are the same as those used in Figure 3.8.

3.3.4 Comparison of the physical and genetic maps of the right arm of chromosome IV

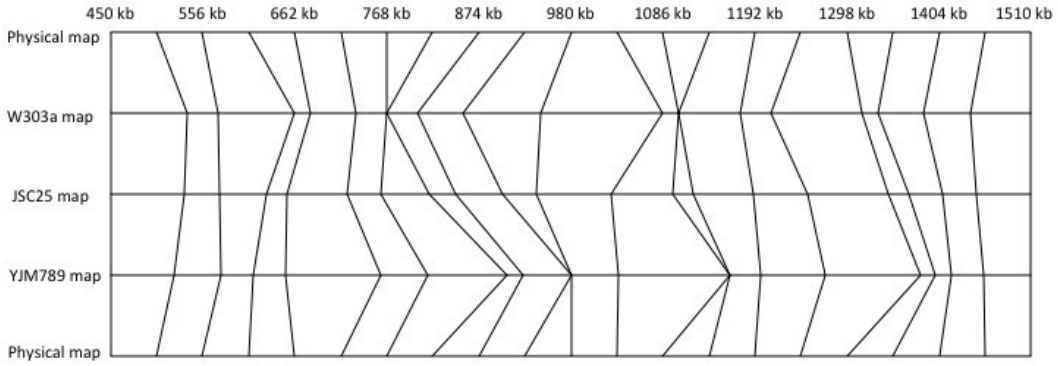


Figure 3.10. Comparison of physical and genetic mitotic recombination maps.

This figure shows a comparison of the physical map of the 1.1 Mb right arm of chromosome IV with a genetic map of mitotic crossovers. We show three genetic maps: one that includes the crossovers initiated from both homologs (JSC25 map), one including the crossovers initiated from the W303a-derived homolog (W303a map), and one including the crossovers initiated from the YJM789-derived crossover (YJM789 map). The chromosome arm was divided into 20 regions of about 53 kb. The genetic distances are based on the proportion of events within each 53 kb interval, using the "point-analysis" method employed for Figure 3.8. and described in the Supporting Materials and Methods section. The top and bottom horizontal lines show the physical map. The vertical lines allow comparisons of the genetic maps with the physical maps, and of the different genetic maps with each other.

In Figure 3.10, we show a comparison between the physical map of the right arm of chromosome IV with the genetic maps based on recombination data from JSC25. In addition, we show separately the data from the W303a- and YJM789-derived homologs. This figure emphasizes the differences in recombination activity between the two different homologs. For example, the region between 927 kb and 1,033 kb in the map of W303a is larger than expected from the physical map, but the same region is smaller than expected in the YJM789 map. Conversely, the segment between 768 kb and 821 kb

has no events in the W303a map, but a greater than average number of events in the W303a map. Such differences demonstrate the difficulty and, perhaps, the futility of generating a universal mitotic recombination map for *S. cerevisiae*. As shown below, however, differences in the recombination activities of the two homologs can lead to mechanistic insights into mitotic recombination.

3.3.5 The HS4 hotspot is associated with a pair of inverted Ty elements

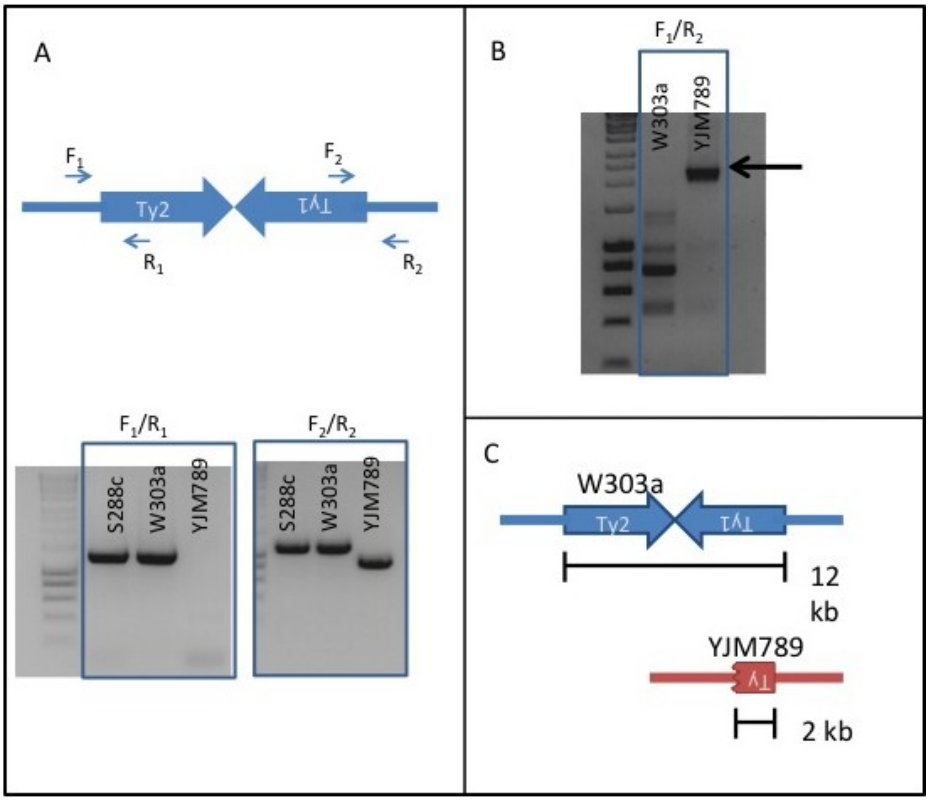


Figure 3.11. Inverted pair of Ty elements at SGD coordinate 980 kb are present on the W303a-, but not the YJM789-derived homolog

A. Top: Structure of the Ty inverted repeat located near 980 kb on chromosome IV. F_1, F_2, R_1, and R_2 are primers used to diagnose the existence of these repeats on the W303a and YJM789 homologs. Primers F_1, R_1, F_2, and R_2 are listed in Table S 3.2 as IV 980403 F, Ty2 R, Ty1.2 R, and IV 993256 R, respectively. Bottom: Gel analysis of the PCR products using F_1/R_1 and F_2/R_2 primer pairs and the indicated template DNA; S288c is very similar in DNA sequence to W303a (WINZELER *et al.* 2003).

B. Gel analysis of the PCR reaction using primer pair F_1/R_2. This analysis indicates the existence of a partial Ty element in the YJM789 genome.

C. Summary of the structure of the inverted Ty repeats at HS4 in the W303a-derived homolog and the absence of this structure in the YJM789-derived homolog. These results are also consistent with the genomic sequencing of the two strains.

The HS4 hotspot, located between coordinates 970 kb and 1,000 kb, has the highest level of recombination as measured by the number of conversion events (Figure 3.7A). This hotspot is specific to the W303a-derived homolog (Figure 3.7B), and is associated with G1-initiated gene conversion events (Figure 3.7C). Analysis of the DNA sequences in the hotspot region shows that the peak of recombination activity of HS4 overlaps with a closely-spaced inverted pair of Ty elements*YDRWTy2-3* and *YDRCTy1-3* located between coordinates 981 and 992 kb; Ty elements are 6 kb retrotransposons present in about 40 copies per haploid (KIM *et al.* 1998). Artificially-constructed inverted repeats stimulate deletions and certain types of recombination in yeast (GORDENIN *et al.* 1993; LOBACHEV *et al.* 2002), and inverted Ty elements have been shown to increase the frequency of chromosome aberrations in yeast strains with low levels of DNA polymerase alpha (LEMOINE *et al.* 2005).

To determine whether the hotspot activity of HS4 was related to the inverted pair of Ty elements, we first examined whether both of the haploid progenitor strains

contained the inverted pair of elements. Using oligonucleotide primers specific for Ty1, Ty2, and the regions flanking these repeats, we found that W303a, but not YJM789, had the inverted Ty pair (Figure 3.11A). This result is also consistent with the DNA sequences of W303a (SGD database) and YJM789 (WEI *et al.* 2007), showing that YJM789 is missing the centromere-proximal Ty element and has only a 2 kb fragment of the distal Ty element. Using primers that flank the Ty element(s) (Figure 3.11B), we found that the YJM789 strain had a 2 kb band as expected, whereas the W303a strain had multiple small non-specific bands. This result is not unexpected, since the pair of Ty elements represents a 12 kb palindromic fragment that would be difficult to amplify. These results are summarized in Figure 3.11C. The observation that the HS4 hotspot is specific to the W303a-derived homolog that contains the inverted pair of Ty elements implicates this structure in the hotspot activity.

Figure 3.12 shows the sequence and potential secondary structure of the central 98 bp of the inverted repeat hotspot if this region was single-stranded. Figure 3.12A shows the linear sequence with regions of complementary shown in brackets. Figure 3.12B and C show two possible secondary structures that can be formed by this sequence. The spacer between the pairs of inverted repeats is 25–66 bases, depending on whether the secondary structure is tolerant of mismatches. In addition to forming a hairpin, inverted repeats are capable of forming cruciform structures in which two hairpin structures are formed on opposite strands (LOBACHEV *et al.* 2002).

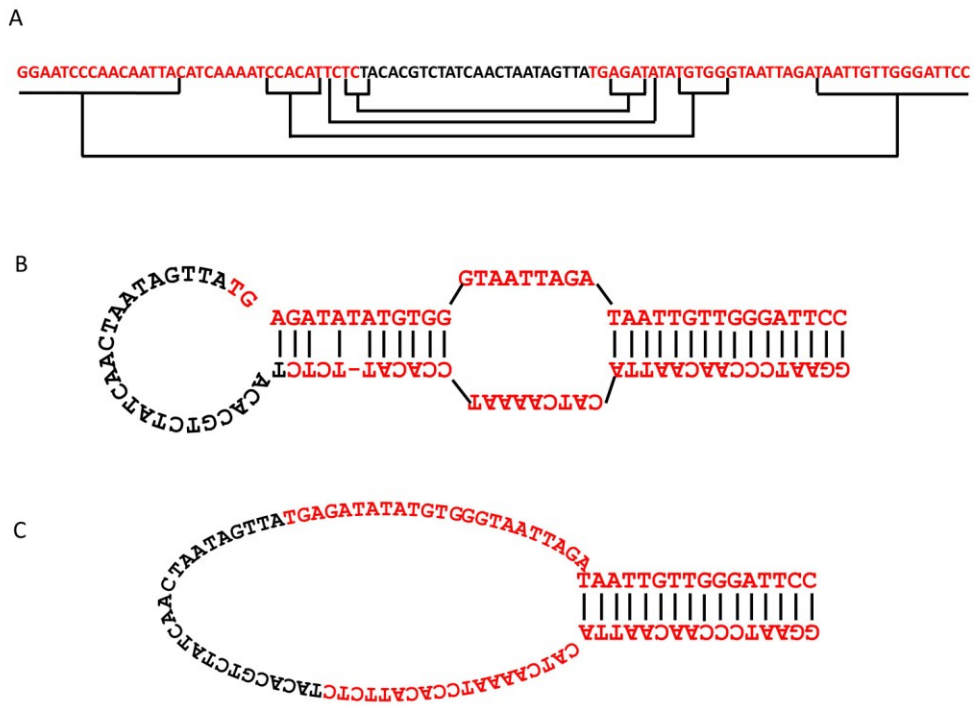


Figure 3.12. Potential "hairpin" DNA structures that could be formed by inverted repeats.

Bases with homology to Ty or delta sequences are shown in red; bases with no homology to Ty are shown in black.

A. Sequence of 98 bases from one strand of the DNA at the center of the inverted Ty pair (SGD coordinates 987091-987188) at HS4. Regions of possible intrastrand pairing are indicated by brackets.

B. Secondary structure formed by HS4 with a terminal 25 bp spacer and an unpaired 9 base loop.

C. Secondary structure formed by HS4 with a 66 bp spacer.

To directly test whether the inverted Ty elements were responsible for the hotspot activity of HS4, we measured hotspot activity in four isogenic stains: one with the wild-type HS4 sequence, one with a deletion of the centromere-proximal Ty element,

one with an insertion of a marker that expanded the spacer by about two kb, and one with a deletion of a delta element of the centromere-proximal hotspot. For all four strains, we inserted *URA3* about 23 kb centromere-distal to HS4 (957 kb) and the hygromycin-resistance gene *HYG* about 20 kb centromere-proximal to HS4 (1,013 kb). These strains are depicted in Figure 3.13A and details of the strain construction are in the Supporting Materials and Methods.

A crossover initiated at HS4 will result in a daughter cell that is Hyg^R and 5-FOA^R 50% of the time. A crossover initiated between *CEN4* and *HYG* will result in a daughter cell that is Hyg^S and 5-FOA^R 50% of the time. By monitoring the ratio of Hyg^R 5-FOA^R recombinants to all 5-FOA^R recombinants, we measured the recombination rate between the *HYG* and *URA3* markers. For this analysis, we examined about 400 independent 5-FOA^R derivatives of each strain. The distance between *CEN4* and *URA3* is about 568 kb and the distance between *HYG* and *URA3* is about 56 kb. Thus, if the numbers of crossovers are proportional to the size of the physical intervals, we expect that about 10% of the 5-FOA^R strains will be Hgr^R. The observed frequency of derivatives of this class in the strain with the wild-type hotspot (JSC71-1) was 0.17 (Figure 3.13C). The difference in the number of observed crossovers and expected crossovers in the *HYG-URA3* interval is statistically significant ($p<0.001$), as expected since this region contains the HS4 hotspot. In a strain with the deletion of the Ty2 element (JSC73-2) or a strain in which the spacer between the Ty elements was expanded

by insertion of the *KANMX* gene (JSC74-1), the hotspot activity is significantly reduced. In contrast, in a strain in which the centromere-proximal delta element of Ty2 is deleted (JSC77-1), no significant loss of hotspot activity is observed.

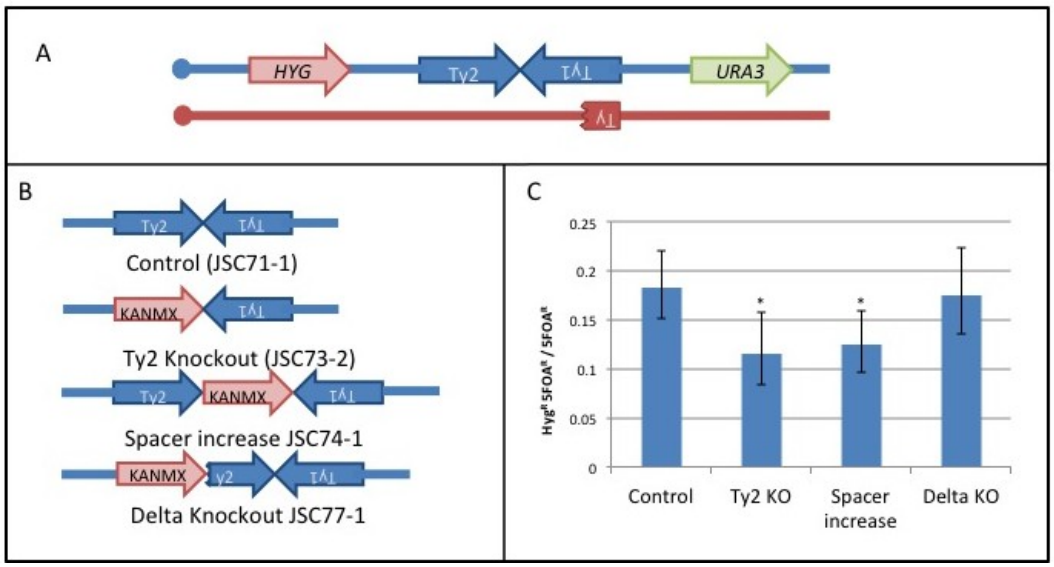


Figure 3.13. Deletion analysis of the HS4 hotspot.

A. Experimental system used to activity of the wild-type HS4 hotspot and various deletion derivatives. The *URA3* and *HYG* genes were inserted centromere-distal and centromere-proximal to the inverted Ty elements, respectively, on the W303a-derived homolog (in blue). The YJM789-derived homolog (in red) was unaltered.

B. Modifications of the inverted repeat structure. In the control strain JSC71-1, the unmodified hotspot consists of a Ty2 element on the Watson strand, 25 bp of non-Ty sequence, and a Ty1 element on the Crick strand. In the JSC73-2 strain, the Ty2 element is replaced with *KANMX*. In the JSC74-1 strain, the space between the two Ty elements, is increased by about 2 kb by insertion of the *KANMX* marker. In the strain JSC77-1, the centromere-proximal delta element of Ty2 was replaced by the *KANMX* gene.

C. Frequency of recombination between *HYG* and *URA3* in the strain with wild-type HS4 activity and in various deletion derivatives. Asterisks indicate a

significant ($p < 0.05$) reduction in recombination activity compared to the control strain.

Deletion of the Ty2 element would eliminate the possibility of hairpin formation at HS4 and an increase in the spacer between the repeated elements would be expected to substantially reduce the probability of hairpin formation (GORDENIN *et al.* 1993; LEMOINE *et al.* 2005). The 330 bp delta elements flank Ty elements (KIM *et al.* 1998) and the 5′ delta element acts as a transcriptional promoter (MORILLON *et al.* 2002). In the strain containing the deletion of the 5′ delta, Ty2 should no longer be transcriptionally active. The results summarized in Figure 3.13, therefore, indicate that the formation of a secondary structure between the two Ty elements is important for hotspot activity at HS4, but the transcriptional activities of both of the Ty elements are not required for hotspot activity.

3.3.6 Types of gene conversion events: inferences about the timing of the recombinogenic DSBs during the cell cycle

Most (121/139; 87%) of the crossovers examined in our study were associated with gene conversion events. Of these conversion tracts, about two-thirds (82/121) are simple 3:1 conversion, simple 4:0 conversions, or hybrid 3:1/4:0 or 3:1/4/0/3:1 tracts as diagrammed in Figure 3.1. As discussed previously, the simple 3:1 tracts are most simply interpreted as events initiated in S or G2, and the other two classes are likely to reflect events initiated in G1. Thus, if we count only the simple conversions as illustrated in Figure 3.1, two-thirds of the crossovers have gene conversions indicative of a G1-

initiated event. These data support the surprising conclusion that most recombinogenic events that lead to crossovers between homologs are initiated in G1 rather than in S or G2 (LEE *et al.* 2009). The numbers of conversion events of all types, as well as their schematic depictions, are given Table S 3.3.

The simple conversion events shown in Figure 3.5B-D can be simply explained as resulting from the repair of one (Figure 3.5B) or two (Figure 3.5C and D) DSBs by the canonical DSB repair model (LEE *et al.* 2009; ST CHARLES *et al.* 2012; SZOSTAK *et al.* 1983). In this model, DSB formation is followed by processing of the broken ends to yield ends with 3′ overhangs. These single-stranded DNA molecules "invade" the homolog forming regions of heteroduplex. Gene conversion events reflect repair of mismatches within the heteroduplex; repair of mismatches during meiotic conversion events is usually continuous, with mismatches deleted from one strand and corrected using the other strand as a template (PETES *et al.* 1991).

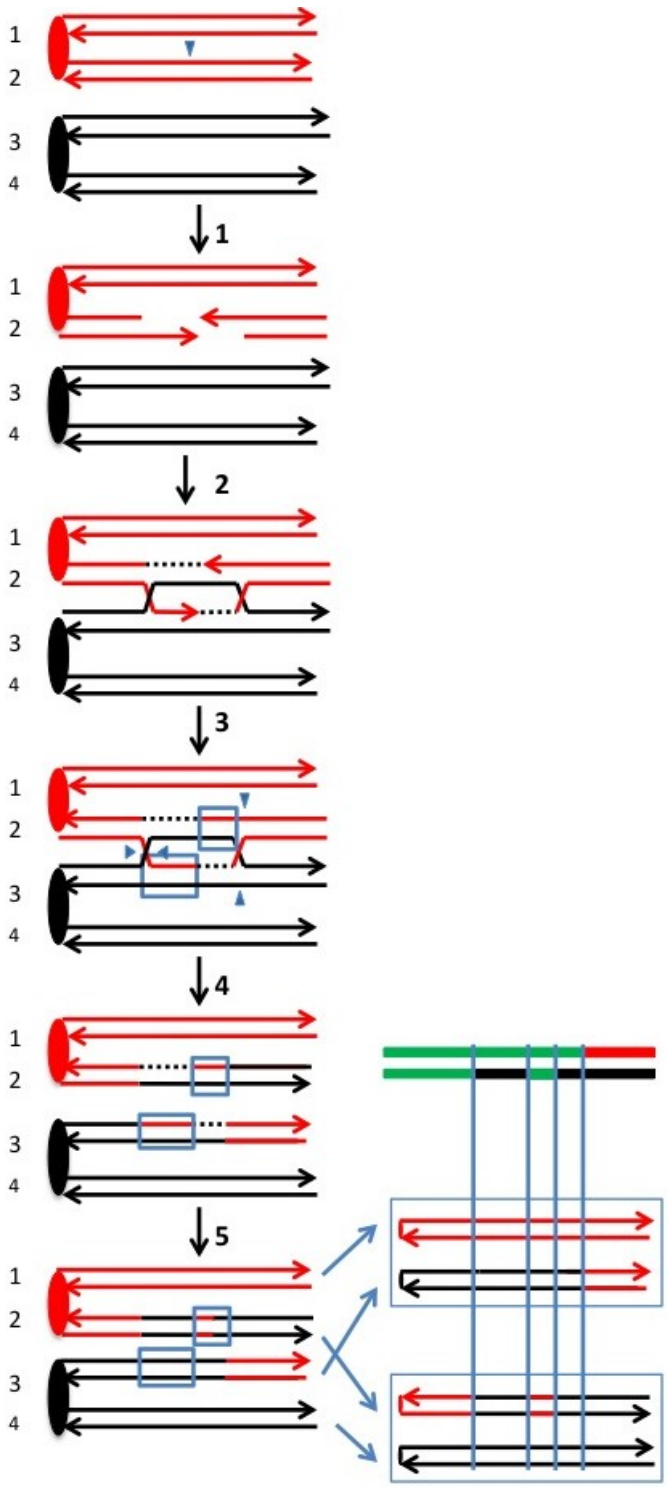


Figure 3.14. Generation of a complex conversion tract by "patchy" repair of mismatches in a heteroduplexes.

This figure shows one mechanism for generating the complex gene conversion tract of Figure 3.5E. The homologs are depicted as double-stranded DNA molecules with the W303a-derived homolog shown as red lines and the YJM789-derived homolog shown as black lines. Recombination is initiated by a DSB on chromatid 2 in G2. The broken chromatid undergoes 5′ to 3′ resection (step 1), followed by strand invasion of the left end into chromatid 3 (step 2). DNA synthesis initiated by the broken end (shown as dotted lines), displaces one of the strands of chromatid 3, allowing second-end capture and DNA synthesis by the right hand broken end. The resulting double Holliday junction has two regions of heteroduplex (shown in blue rectangles), and is processed by cleaving the left and right junctions (cleavage sites shown by triangles) (step 3). Following junction resolution, the chromosome regions flanking the heteroduplexes are in the recombined configuration (step 4). In step 5, the mismatches within the two heteroduplexes are repaired. All of the mismatches in the heteroduplex on chromatid 3 are repaired in the same direction (duplicating YJM789-derived SNPs), whereas the heteroduplex on chromatid 2 undergoes "patchy" repair. Segregation of chromatids 1 and 3 into one cell and chromatids 2 and 4 into the other result in a crossover associated with the complex conversion tract depicted in Figure 3.5E.

About one-third of the conversion tracts had more complex patterns that require modifications of the canonical model. For example, in Figure 3.5E, a 3:1 conversion tract is interrupted by a region that is heterozygous in both sectors. This pattern is consistent with repair of a DSB formed during S or G2, followed by heteroduplex formation and "patchy" repair of mismatches (Figure 3.14). Since many of the classes of complex conversion events are similar or identical to complex tracts that we observed previously in other studies (ST CHARLES *et al.* 2012), we will not discuss models for their formation in any detail.

We also examined the conversion tract lengths for the G2- and G1-associated conversions. The median tract lengths for the simple conversions were 4.5 kb (95% confidence limits of 2.6-9.5) for the simple 3:1 conversions (G2-initiated) and 12.8 (8.7-

17.1) kb for the simple 4:0 and 4:0/3:1 or 3:1/4:0/3:1 hybrid conversions. If we include all of the conversion tracts, interpreting G2- and G1-associated events as shown in Table S 3.3, the median tract lengths of G2- and G1-associated crossovers were 4.7 kb (2.6-9.5 kb) and 14.8 kb (11.7-17.5 kb), respectively. By the Mann-Whitney test, the G1 and G2 conversion tract lengths are significantly different ($p<0.0001$). The median tract length for all conversion events was 10.6 kb (95% confidence limits of 8.2-13.6 kb).

3.3.7 Association of conversion tracts with elements of chromosome sequence or structure

From many genetic studies in yeast, it has been shown that the site at which recombination initiates is within or adjacent to the conversion tract (PÂQUES and HABER 1999). For example, in our previous study, we found that almost all of the crossovers stimulated by the trinucleotide repeats $(GAA/TTC)_{230}$ include these repeats in the conversion tract (TANG *et al.* 2011). For HS4, the relevant recombinogenic DNA sequences are the inverted pairs of Ty elements previously discussed. The position of HS3 colocalizes with another inverted pair of Ty elements located about 100 bp apart (SGD coordinates 870-885 kb), and it is likely that the hotspot activity of HS3 is a consequence of the same mechanism as for HS4. The other putative hotspots are not associated with inverted repeats. There are a number of elements of chromosome structure in yeast that have been identified as potential chromosome fragile sites or potential hotspots of recombination including G4-forming sequences, ARS elements, highly-transcribed genes, replication-termination regions, peaks of γH2-AX, and other

chromosome elements. We compared our mapping of gene conversion events with 18 elements listed in Table S 3.5.

To examine associations with these elements, we determined the frequency of these elements within the conversion tracts in our experiments compared with their expected frequency assuming that there is no significant association of the elements with the conversion tracts; we corrected p-values for multiple comparisons (Hochberg and Benjamini 1990). We examined these associations using four related methods as explained in detail in the Supplementary Information. In brief, we looked at associations using windows that included: 1) the whole conversion tract regardless of length (Method 1), 2) the whole conversion tract excluding lengths over 20 kb (Method 2), 3) the portion of conversion tracts likely to have the initiating DSBs (for example, the 4:0 region of a 4:0/3:1 hybrid tract) (Method 3), and 4) the portion of the conversion tracts likely to have the initiating DSB, excluding windows greater than 20 kb (Method 4). We also examined different sets of recombination data with these methods including: 1) all of the conversion events, 2) all of the conversion events initiated on the W303a-derived chromosome, 3) all of the conversion events initiated on the YJM789-derived chromosome, 4) all of the G1 conversion events, and 5) all of the G2 conversion events. The significant associations from these analyses are in Table 3.1.

In the analysis in which all conversions are included, the only significant positive associations were with Ty elements and long terminal repeats (delta elements) that flank

Ty elements. These same associations were also significant in the G1-associated conversions (Methods 1 and 3) and the conversion events of the W303a-derived chromosome (Methods 1 and 3). These associations are expected since the HS4 hotspot is G1-specific and specific for the W303a-derived chromosome (Figure 3.7B and C). No significant associations were observed between conversion tracts and delta elements that were unassociated with Ty elements. Using Method 3, we also found significant associations involving tRNA genes with the complete dataset of JSC25, the G1-initiated events, and the events initiated on the W303a-derived chromosome. The non-random linkage between Ty elements and tRNA genes in the genome (KIM *et al.* 1998) may explain this association.

Table 3.1 Chromosome elements that are over-represented in crossover-associated gene conversion tracts in JSC25.[1]

Element	Data[2]	Method of analysis	p-value
Long terminal repeats (deltas)	JSC25 all	1	1.55E-07
Long terminal repeats (deltas)	JSC25 all	3	1.04E-11
Long terminal repeats (deltas)	JSC25 G1	1	6.06E-09
Long terminal repeats (deltas)	JSC25 G1	3	1.44E-15
Long terminal repeats (deltas)	W303a all	1	1.28E-06
Long terminal repeats (deltas)	W303a all	3	7.31E-14
Ty elements	JSC25 all	1	1.39E-05
Ty elements	JSC25 all	3	2.13E-06
Ty elements	JSC25 G1	1	1.70E-06
Ty elements	JSC25 G1	3	4.45E-08
Ty elements	W303a all	1	3.44E-06
Ty elements	W303a all	3	1.28E-07
tRNA genes	JSC25 all	3	0.011
tRNA genes	JSC25 G1	3	0.0019
tRNA genes	W303a all	3	0.002
G4 motifs	JSC25 G1	4	3.72E-04
G4 motifs	YJM789 all	4	4.27E-04
Replication termination region	JSC25 G2	1	2.77E-04
Replication termination region	JSC25 G2	2	3.82E-04
Replication termination region	JSC25 G2	3	1.67E-04
Replication termination region	JSC25 G2	4	1.98E-04
Rrm3p pause sites[3]	YJM789 all	2	0.0024

[1] As described in the text, we used four related methods of analysis to determine whether various structure/sequence chromosome motifs (listed in Supplementary Table 3.5) were over-represented in gene conversion tracts associated with crossovers on chromosome IV.

[2] The data examined by Methods 1-4 were: JSC25 all (G1 and G2 events initiated on either W303a- and YJM789-derived homologs), JSC25 G1 (G1 events initiated on either W303a- and YJM789-derived homologs), JSC25 G2 (G2 events initiated on either W303a- and YJM789-derived homologs), W303a all (G1 and G2 events initiated on the W303-derived homologs), and YJM789 all (G1 and G2 events initiated on the YJM789-derived homologs).

[3] Rrm3p is a helicase that promotes replication fork progression through regions at which the forks are paused (Azvolinsky *et al.*, 2009).

Replication forks do not proceed at a constant rate in the genome and replication-pause sites have been associated with a number of chromosome elements including the ribosomal RNA fork barrier, centromeres, tRNA genes, convergent replication forks, G-quadruplex motifs, and highly-transcribed RNA polymerase II genes (reviewed by AZVOLINSKY *et al.* (2009)). We found several similar motifs associated with conversion events. G2, but not G1, conversion events were very significantly associated with chromosome termination regions (*TER*s). We also found a significant association between G-quadruplex motifs (CAPRA *et al.* 2010) and G1 conversion events, as well as an association between G-quadruplex motifs and conversion events initiated on the YJM789-derived chromosome; G-quadruplex structures slow replication fork progression (PAESCHKE *et al.* 2011). Finally, the conversion events that are initiated on the YJM789-derived chromosome have an over-representation of sites of stalling of the Rrm3p helicase; this helicase is thought to aid replication through genomic regions at which DNA polymerase is paused (AZVOLINSKY *et al.* 2009).

In summary, we found two types of significant associations between chromosome elements and the location of conversion events. Ty elements were significantly associated with G1-initiated events and conversions that were initiated on the W303a-derived homolog, as expected from other types of analysis. Conversion events in S/G2 were significantly associated with regions in which replication forks are slowed or stalled. In addition to positive associations between the location of conversion

events and elements of chromosome structure, we found a number of negative associations (Table S 3.6).

3.4 Discussion

In this study, we developed a high-resolution map of spontaneous crossovers on the right arm of chromosome IV, a region that includes about 10% of the yeast genome. Below, we will discuss: 1) the frequency of mitotic crossovers, 2) the distribution of crossover events, 3) the properties of mitotic gene conversion tracts, and 4) the timing of recombination events during the cell cycle.

3.4.1 Frequency of reciprocal crossovers

We mapped a total of 139 reciprocal crossovers between *CEN4* and the *SUP4-o/KANMX* markers. The sectoring frequency for this chromosome arm was 3.1×10^{-5}. Since only half of the crossovers result in a sectored colony, the frequency of crossovers in this interval was about 6.2×10^{-5}/division or about 6.2×10^{-8} crossovers per kb. As a convenient unit of mitotic crossovers, we suggest that 10^{-6} crossovers/division be defined as one micro Stern (μS), named after Curt Stern, the discover of mitotic recombination (Stern 1936). Thus, the right arm of chromosome IV has a genetic length of about 62 μS.

In our previous analysis of the left arm of chromosome V, we observed a frequency of crossovers of 6.6×10^{-6}/division (LEE *et al.* 2009). Since the physical distance for the examined interval on V is about one-tenth of the chromosome IV interval, the numbers of events per kb on chromosomes IV and V are similar. If the remainder of the

yeast genome has the same frequency of crossovers per kb as chromosomes IV and V, we calculate that the yeast genome has a spontaneous crossover rate of about 6.2×10^{-4}/cell division and a genetic map length of about 620 μS. In contrast, the yeast genome has a genetic length of about 4100 cM (SGD), representing about 81 meiotic crossovers/division. Thus, the rate of crossovers in meiosis is about 10^5 higher than the mitotic rate. Although the rate of mitotic crossovers/division is low relative to the meiotic rate, since the number of mitotic divisions is likely to greatly exceed the number of meiotic divisions, mitotic crossovers are likely to be a potent mechanism for generating novel combinations of alleles.

3.4.2 Distribution of mitotic crossovers

3.4.2.1 General considerations concerning the distribution of crossovers

The distribution of mitotic crossovers and their associated gene conversion tracts are shown in Figure 3.6 and Figure 3.7. Although the distribution of events appears strikingly non-random in Figure 3.7, several points concerning the distribution should be emphasized. First, by our statistical analysis, not all of the peaks labeled as hotspots have significantly elevated levels of recombination. Second, in Figure 3.7, we show the number of times each SNP was involved in a gene conversion tract; this parameter is affected by both the frequency of recombination events near the SNP and the lengths of conversion tracts in the region. Third, for conversion-associated crossovers, the position of the initiating DNA lesion can be anywhere within the conversion event. Previously,

we showed that 230 repeats of the GAA trinucleotide was a hotspot for DSB formation (TANG *et al.* 2011). In individual conversion events associated with this hotspot, the conversion tracts were propagated either toward the centromere, away from the centromere or bidirectionally from the tract. Thus, our mapping of DNA lesions and the resolution of the mitotic recombination map, therefore, are limited by the size of the conversion tracts rather than the distribution of SNPs.

An important conclusion from our analysis is that two homologs can differ significantly in their distribution of recombination events (Figure 3.7B). From the data shown in Figure 3.7B, it appears that HS1, HS3, and HS4 are hotspots on the W303a-derived chromosome, HS2 and HS6 are hotspots on the YJM789-derived chromosomes, and HS5 and HS7 are hotspots on both homologs. Thus, construction of a universal genetic map of mitotic events is difficult.

Meiotic recombination events have been mapped throughout the genome using various types of microarrays (BLITZBLAU *et al.* 2007; BORDE *et al.* 2004; BUHLER *et al.* 2007; GERTON *et al.* 2000). Meiotic crossover hotspots are associated with intergenic GC-rich regions with a preference for intergenic regions located at the 5′ ends of genes, although there are also regional effects on DSB formation (PAN *et al.* 2011; PETES 2001). The patterns of meiotic and mitotic recombination events on the right arm of chromosome IV show no evident similarity (Figure 3.15). Since meiotic recombination are initiated by meiosis-specific cleavage of the genome by Spo11p and the DNA lesions that initiate

mitotic recombination are likely to have a variety of sources (discussed further below), this difference in meiotic and mitotic patterns of recombination is expected.

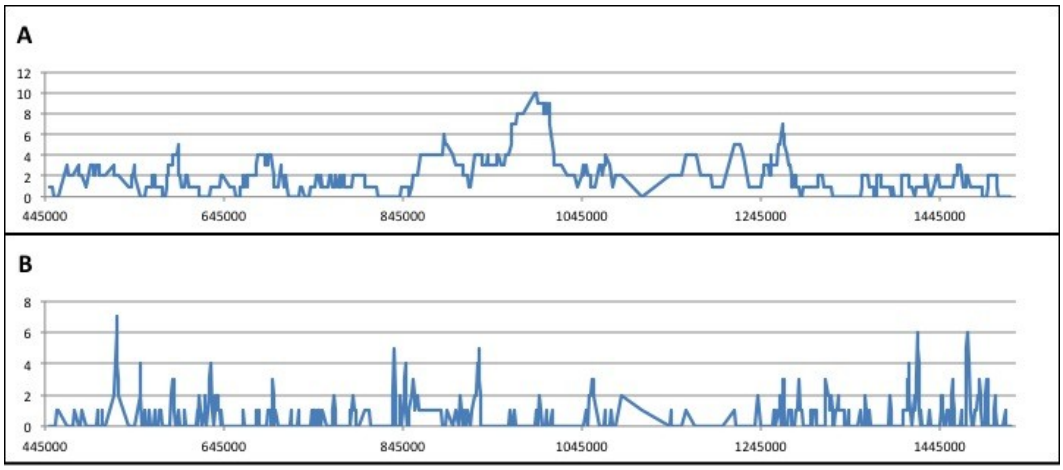


Figure 3.15. Comparison of mitotic and meiotic recombination maps on the right arm of chromosome IV.

The Y-axis in the figure shows the number of times individual SNPs are included in a crossover-associated gene conversion. The Y-axis shows SGD coordinates on chromosome IV.

A. Summary of our mitotic mapping of events in JSC25.

B. Summary of the mapping of meiotic events in a closely-related diploid strain by Mancera *et al.* (2008).

The levels of local meiosis-specific DSBs at different places in the genome vary at least 400-fold (PAN *et al.* 2011). In contrast, the difference between the average number of times a conversion event included a SNP in our mitotic experiments (twice) was only five-fold less than the maximum number of times a SNP was included (ten). Thus, our data suggest that hotspots and coldspots are more pronounced for meiotic recombination than for mitotic recombination.

3.4.2.2 Inverted repeats as hotspots for mitotic recombination

Both HS3 (near SGD coordinates 880 kb) and HS4 (at SGD coordinates 980 kb) colocalize with inverted Ty retrotransposon repeats. Inverted repeats have been previously shown to be hotspots for certain types of mitotic recombination and translocations in yeast strains with impaired replication (LEMOINE *et al.* 2008; LEMOINE *et al.* 2005; LOBACHEV *et al.* 2002; LOBACHEV *et al.* 1998; LOBACHEV *et al.* 2000). Most of these studies have involved non-yeast DNA sequences inserted into the genome by transformation. For example, LOBACHEV *et al.* (2000) showed that ectopic heteroallelic recombination was stimulated about 1000-fold by inverted copies of the human Alu sequence that were separated by 12 bp. Sequence divergence between the repeats and increased spacer distance between the repeats substantially reduced the recombinogenic effects of the inverted repeats (LOBACHEV *et al.* 2000). In our previous study (LEMOINE *et al.* 2005), a hotspot for chromosome rearrangements on chromosome III was an inverted pair of Ty elements separated by about 280 bp. This pair of elements is a hotspot for chromosome rearrangement only when cells have low levels of DNA polymerase alpha. In contrast, the HS3 and HS4 pairs of repeats stimulate mitotic recombination in cells with wild-type levels of DNA polymerase.

The conversion tracts associated with inverted repeats at HS3 and HS4 are primarily 4:0 or hybrid events, arguing that the recombinogenic lesion is formed prior to replication. Two types of secondary structures have been associated with inverted

repeats: cruciforms (Figure 3.16A) and hairpins (Figure 3.16B). Evidence for cruciform processing (presumably by a Holliday junction-like resolvase) resulting in two hairpin-capped ends has been obtained (LOBACHEV *et al.* 2002). However, spacers greater than 20 bases substantially inhibit cruciform formation in *E. coli* (ALLERS and LEACH 1995; KOGO *et al.* 2007; SINDEN *et al.* 1991). It is possible that various stresses associated with HS3 or HS4 (such as supercoiling or single stranded regions from R-loops forming during transcription of Ty repeats) increase the likelihood of cruciform formation.

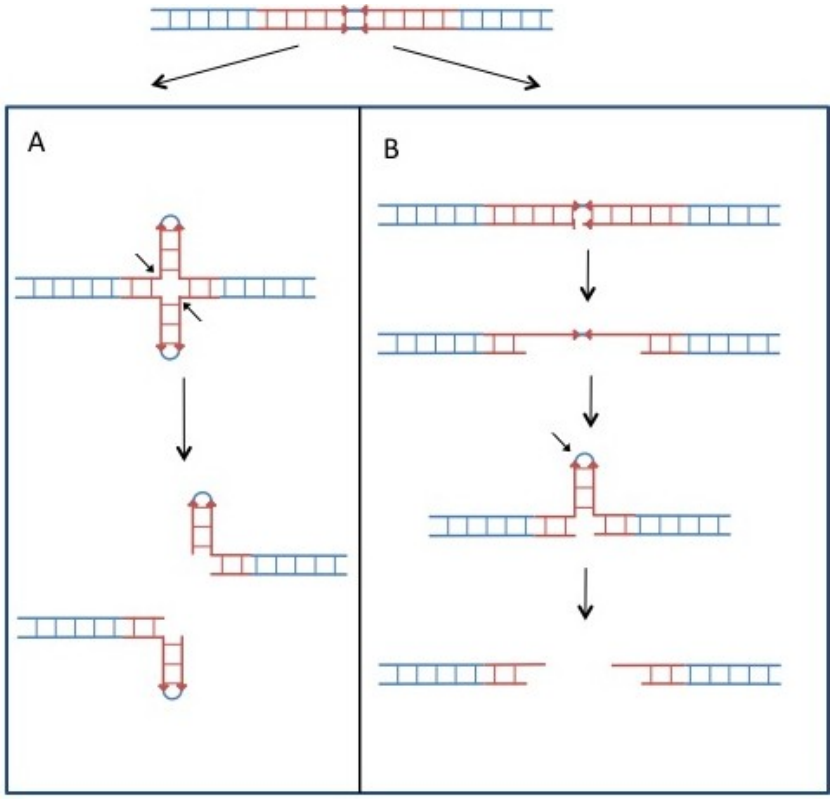


Figure 3.16. Models for a double-strand break occurring at an inverted repeat during G1 of the cell cycle.

The red and blue colors indicate Ty- and non-Ty-related sequences, respectively.

A. Cruciform formation and subsequent cleavage. Arrows show the positions of cleavage, presumably by a Holliday junction-like resolvase. The resulting hairpin-capped ends would have to be processed to allow recombination.

B. Single-stranded nick within the vicinity of the inverted repeat, followed by action of an endonuclease results in a single-stranded gap. Hairpin formation and subsequent cleavage of the spacer on the non-nicked strand would produce a DSB.

An alternative model is that a single-stranded nick near the junction of the inverted repeats is processed into a single-stranded DNA gap, allowing formation of a hairpin on the intact DNA strand (Figure 3.16B). The loop formed by the spacer in the hairpin could then be nicked by a nuclease such as the Mre11p complex (LOBACHEV *et al.* 2002). We do not know the source of the nick that initiates hairpin formation. It is possible that the nick is generated by the repair of a damaged base. Alternatively, since the convergent transcription of Ty elements should result in an increased concentration of positive supercoils, it is possible that the recombinogenic nick reflects that action of topoisomerases that relieves such supercoils (WANG 2002). Since removal of the delta element (the Ty promoter; (MORILLON *et al.* 2002)) located in one of the Ty elements of HS4 reduces, but does not eliminate hotspot activity (Figure 3.13C), transcription of <u>both</u> Ty elements is not required for the recombinogenic effects of the inverted Ty elements.

Although the HS3 and HS4 hotspots appear G1-specific, it is likely that these regions may also be associated with DSBs during the S-period. Replication forks move slowly in genomic regions with inverted repeats in strains lacking Rrm3p (AZVOLINSKY *et al.* 2009), a helicase that aids the replication fork in proceeding through natural impediments. As discussed below in more detail, DSBs occurring during S are

preferentially repaired by sister-chromatid recombination (KADYK and HARTWELL 1992); such events do not result in sectored colonies in our system. In addition, since the inverted Ty repeats at HS3 and HS4 are only present in the W303a-derived homolog, a DSB formed near the center of the inverted repeat would require extensive resection to reveal homologous sequence required for an interaction with the YJM789-derived homolog. This factor would increase the probability for an interaction with the sister chromatid if the DSB occurs in S or G2.

It is also possible that some of the recombination events initiated at HS3 or HS4 could be repaired by ectopic recombination with Ty elements located elsewhere in the genome, although such events would not result in a sectored red/white colonies. Translocations between the Ty elements at HS3 and HS4, and Ty elements located elsewhere in the genome have been observed previously (ARGUESO *et al.* 2008; DUNHAM *et al.* 2002; MCCULLEY and PETES 2010).

3.4.2.3 Mitotic recombination events that are initiated independently of inverted repeats

In addition to inverted repeats, several other factors have been shown to stimulate mitotic recombination or DSB formation. In principle, any cellular process that can produce a DNA lesion (DSBs or single-stranded nicks) could be responsible for initiating a spontaneous mitotic recombination event (AGUILERA *et al.* 2000). In addition, as discussed below, it is likely that some regions of the chromosome, for a variety of reasons, are more prone to recombinogenic DNA damage. Using several different

assays, researchers have found that elevated levels of transcription stimulate recombination (reviewed by AGUILERA (2002)). Several explanations have been offered to explain this phenomenon including: 1) DSBs caused by a collision between converging transcription and replication machinery, 2) recombinogenic lesions induced by topoisomerases in local regions of transcription-induced supercoils, and 3) nicking of the displaced strand in genes in which an R-loop has been formed.

Since mutations in many enzymes involved in DNA replication result in a hyper-Rec phenotype, it is likely that some spontaneous recombination events result from DNA lesions generated during DNA replication (AGUILERA *et al.* 2000; BRANZEI and FOIANI 2010). It has been shown that some DNA sequences or motifs are associated with slow or stalled replication forks including: 1) sequences that form secondary DNA structures (cruciform, hairpins, triplex DNA, G4-quadruplex DNA, left-handed DNA, slipped-strand DNA), 2) sequences that bind replication-blocking proteins (for example, Fob1p), 3) replication-termination (TER) regions, and 4) sites of endogenous DNA damage that lead to bulky DNA lesions (BRANZEI and FOIANI 2010). Consequently, we analyzed our conversion tracts to determine if there was an over-representation of these structure/sequence motifs.

We found that G4 quadruplex motifs were significantly associated with G1-initiated DSBs (Table 3.1). In yeast cells, G4 motifs have been shown previously to be enriched in areas of the genome that have high γ-H2AX binding, a signal associated

with DNA damage (CAPRA *et al.* 2010). G4 motifs have also been associated with recombination-prone regions in mammalian cells. They have a role in VDJ recombination in mammalian B cells when they are located in immunoglobulin switch regions, and their recombination-inducing mechanism is dependent on forming G-loops that are recognized by MutSα (DUQUETTE *et al.* 2004; DUQUETTE *et al.* 2005; LARSON *et al.* 2005). Additionally, G4 motifs have been associated with tumorigenic translocations. The major break region of the t(14;18) follicular lymphoma translocation near the *BCL2* gene contains regions of G-rich sequences that are capable of forming G-quadruplex structure and slowing replication (NAMBIAR *et al.* 2011).

In our study, the events initiated during S or G2 of the cell cycle were significantly associated with replication-termination (TER) regions. This association is especially intriguing because sister chromatids are intercalated following replication, and need to be resolved by Top2p (BERMEJO *et al.* 2007; DINARDO *et al.* 1984; HOLM *et al.* 1985). Since Top2p induces DSBs to allow decatenation of sister-chromatids, it is possible that some of these Top2p induced cleavages are misrepaired, leading to homologous recombination.

We also examined HS1, HS2, and HS5-HS7 for motifs that might explain their activities. No obvious elements were shared among these hotspots, although most of them contained genes that were highly transcribed: *SNF11* and *TPS2* in HS1, *KIN1* and *TRM1* in HS2, *HXT3*, *HXT6*, and *HXT7* in HS5, *VPS74* and *YPR1* in HS6, and *UTP5* and

HPT1 in HS7. In addition, the HS5 and HS6 hotspots contain ARS elements, and the HS6 hotspot contains three short palindromes of at least 16 bp. Although the significance of any of these elements is currently uncertain, it is clear that the mechanism of action of these hotspots is different from that of HS3 and HS4.

3.4.3 Properties of mitotic gene conversion tracts

3.4.3.1 Conversion tract lengths

The median conversion tract length for all analyzed crossovers on the right arm of chromosome IV was 10.6 kb (95% confidence limits: 8.2-13.6 kb). In our previous analysis of conversions on the left arm of chromosome V, we observed a similar average of 11.7 kb (+/- 1.6 kb) (LEE *et al.* 2009). We found that conversion tract lengths varied depending on the cell cycle timing of the initiating DSB lesion. Conversion tract lengths associated with G1-initiated breaks were 14.8 kb (11.7 – 17.6 kb), whereas conversion tracts associated with G2-initiated breaks were 5.4 kb (2.6-9.5 kb). Since the crossovers of both G1- and G2-initiated events occur in G2, the broken ends generated by a G1-initiated lesion have a longer time for 5′-to-3′ resection prior to repair of the DSBs. In addition, the G1-initiated events involve the repair of two broken chromatids. If the conversion tracts associated with these two repair events are propagated in different directions, we would expect that the G1-initiated conversions would be longer than those initiated in G2.

As noted previously (LEE *et al.* 2009), the mitotic conversion tracts characterized in our system are considerably longer than the average meiotic conversion tract of about 2 kb (MANCERA *et al.* 2008). Crossover-associated meiotic gene conversion tracts in yeast are generated by a well-characterized process that produces a double Holliday junction with two regions of heteroduplex (HEYER *et al.* 2010); repair of mismatches within the heteroduplex results in the conversion event (PETES *et al.* 1991). Although it is possible that the same mechanism produces mitotic conversion events, an alternative possibility is that the mitotic events are the consequence of repair of a double-stranded DNA gap, as discussed in the original version of the double-strand break repair model (SZOSTAK *et al.* 1983). Although we do not know which of these two mechanisms generates the very long tracts in our study, in an analysis of mitotic recombination events that integrate plasmids, short mitotic gene conversion events occur primarily by heteroduplex formation and mismatch repair, although gap repair was also observed (MITCHEL *et al.* 2010).

3.4.3.2 Conversion tract length associated with the inverted repeat hotspot

One of the unique characteristics of recombination events associated with HS4 is that their conversion tracts are substantially longer than those that occur at other locations along chromosome IV. The median conversion tract length associated with HS4 is 48.4 kb (95% confidence limits: 17.1-118.3 kb), much longer than the median tract length for all conversions (10.6 kb) and longer than the median length of G1-induced

conversions (14 kb). One factor that would be expected to affect tract length is that HS3 has two intact Ty elements on the W303a-derived homolog whereas the YJM789-derived homolog has only a 2 kb fragment of a Ty (Figure 3.11). Thus, a DSB formed in this hemizygous insertion on the W303a-derived chromosome would need to be processed about 10 kb before exposing homology on the YJM789-derived chromosome. In addition, since the homology would be located internally on the resected strand, a recombination intermediate with nonhomologous single-stranded tails would be formed (PÂQUES and HABER 1999). Removal of these single-stranded tails likely requires the nucleotide excision repair proteins Rad1p and Rad10p (FISHMAN-LOBELL and HABER 1992; IVANOV and HABER 1995), as well as the mismatch repair proteins Msh2p and Msh3p (INBAR and KUPIEC 1999; PÂQUES and HABER 1997; SAPARBAEV *et al.* 1996; SUGAWARA *et al.* 1997). Thus, the hotspot activity of HS4 would also be dependent on these proteins. In summary, it is possible that the more extensive processing of the DSBs occurring in HS4 delays completion of the crossover, resulting in a longer crossover-associated gene conversion tract.

3.4.3.3 Complex conversion tracts

Of the 139 crossovers examined on the right arm of chromosome IV, approximately 24% had conversion tracts with more complex patterns that those shown in Figure 3.1. These complex patterns are depicted in Table S 3.3. Similar complex tracts were observed in our previous study of selected crossovers on chromosome V and

unselected recombination events throughout the genome (ST CHARLES *et al.* 2012). Most of these complex events can be explained by the repair of broken sister chromatids in which one or both of the resulting heteroduplexes has "patchy" repair of mismatches (Figure 3.14). In patchy repair, some of the mismatches are repaired using the invading strand as a template and others are repaired using the invaded strand as a template (ST CHARLES *et al.* 2012). In addition, it is possible that Holliday junctions generated by mitotic recombination undergo branch migration (ESPOSITO 1978). Branch migration during repair of two broken chromatids could result in four chromatids with heteroduplexes. Patchy repair within such structures would result in very complex patterns of conversion. It should be emphasized, however, that the most of the conversion tracts observed in our study can be readily explained by the models depicted in Figure 3.1.

3.4.4 DSBs initiated during G1 versus G2 of the cell cycle

As described above, about two-thirds of the observed conversion events involve the repair of two sister chromatids broken at approximately the same position. We infer these events are a consequence of G1-induced DSB on one of the homologs, followed by replication of the broken chromosome. The repair of two broken chromatids subsequently occurs in G2 (Figure 3.1). This model is supported by the observation that irradiation of G1-synchronized yeast cells results in 4:0 and 4:0/3:1 hybrid tracts whereas irradiation of G2-synchronized cells results in 3:1 tracts (LEE and PETES 2010). Based on

the classic experiments of STERN (1936), many geneticists assume that most mitotic recombination events are initiated in G2. Our model, however, is compatible with Stern's results since the crossovers occur in G2, although the recombinogenic lesions are formed in G1. In addition, other yeast researchers have argued that mitotic recombination events can be induced in G1 and that at least some spontaneous events in yeast are initiated in G1 (ESPOSITO 1978; FABRE 1978; WILDENBERG 1970).

Although the crossovers with 3:1 conversion tracts may reflect DSBs generated on one of the two chromatids in S or G2, two other scenarios consistent with an initiation event in G1 are also possible. First, replication of a chromosome that has a G1-initiated single-strand nick would result in one broken G2 chromatid (GALLI and SCHIESTL 1999). Second, for a G1-initiated DSB, if the repair of one of the two broken sister chromatids involved a very short region of heteroduplex, it may not result in a detectable region of conversion. Our estimate that about two-thirds of the recombination events involving homologs are initiated in G1, therefore, is a minimal estimate.

Although most of the recombination events between homologs are initiated in G1, it is likely that more spontaneous DSBs occur in S than in G1. Rad52 foci, indicative of homologous recombination, form in the S- and G2-phases, but not during G1 (LISBY *et al.* 2001). The RPA proteins, which bind single-stranded DNA intermediates during recombination, form larger foci in G2 than in G1 of the cell cycle (BARLOW *et al.* 2008). A

related observation is that the extent of resection of broken ends is substantially greater in G2 than in G1 (AYLON *et al.* 2004; IRA *et al.* 2004).

One simple model that reconciles these observations is shown in Figure 3.17. We suggest that most spontaneous DSBs are generated in S as a consequence of broken replication forks. These DSBs are preferentially repaired by sister chromatid recombination with only a small fraction involving an interaction with the homolog; this preference is ensured by sister-chromatin cohesion (STROM *et al.* 2007; UNAL *et al.* 2007). In yeast, DSBs generated by ionizing radiation during G2 of the cell cycle are usually repaired using the sister chromatid as the template rather than the homolog (KADYK and HARTWELL 1992). Sister-chromatid repair events would be undetectable by our analysis. In contrast, the DSBs generated in G1, even though less frequent than DSBs initiated in S, are preferentially repaired by recombination between the homologs. Since the two sister chromatids resulting from replicating the broken chromosome have DSBs at the same position, one sister cannot be used as a template to repair the other. Thus, the model accounts for the preferential the use of the homolog in a G1-initiated event.

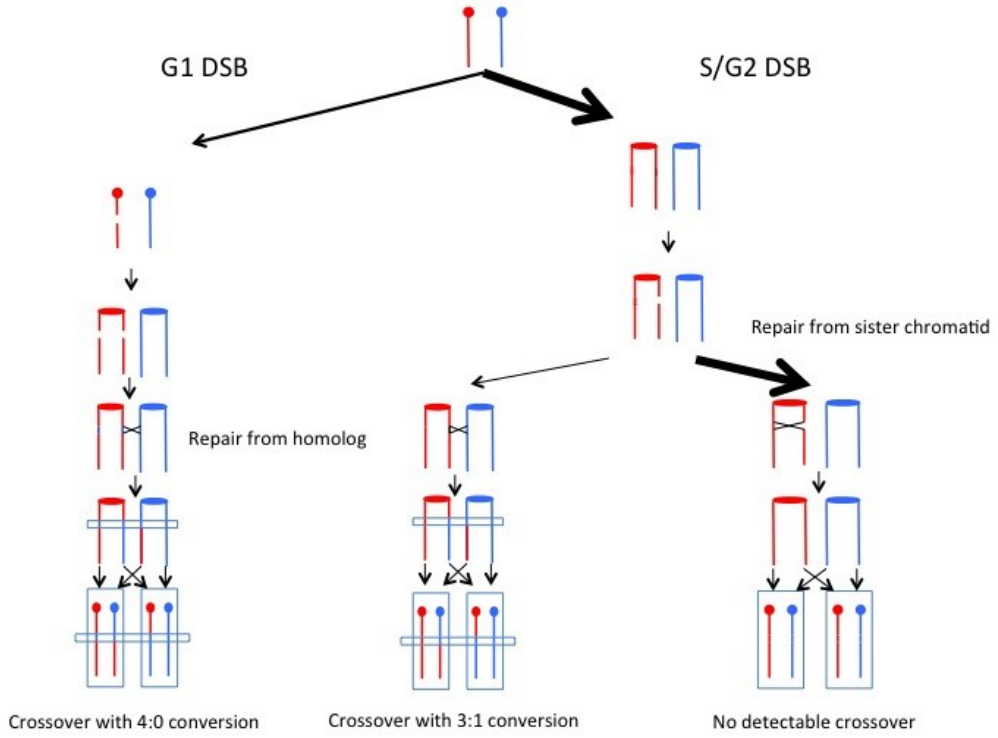


Figure 3.17. Homologous recombination repair choices depending on the timing of the initiating DSB lesion.

We suggest that most spontaneous DSBs occur during S- or G2 of the cell cycle as indicated by the relative widths of the arrows. The DSBs that occur in G1, however, are repaired preferentially by recombination with the homolog, whereas the DSBs that occur in G2 are repaired preferentially by recombination with the sister chromatids. The rationale for these preferences is given in the text.

3.5 Summary

In conclusion, our mapping of mitotic crossovers along the right arm of chromosome IV demonstrates that these events are distributed throughout the chromosome, although there are regional hotspots. The strongest hotspot is associated with an inverted pair of Ty elements separated by only 25 bp. Our analysis of this hotspot indicates that its recombinogenic activity is associated with the ability of this

sequence to form a secondary DNA structure. Other DNA sequence/structure motifs are also associated with the initiation of recombination including G4-quadruplex motifs and replication-termination regions. We demonstrate that most crossovers between homologs are initiated by DSBs in unreplicated chromosomes, and that the distributions of G1- and G2-initiated events along the chromosomes are different. Finally, we show that the two homologs can have different patterns of recombination hotspots.

3.6 Supporting information

3.6.1 Supporting materials and methods

3.6.1.1 Strain construction

In order to examine mitotic reciprocal crossovers on the right arm of chromosome IV, we used the color screening system that is described in the Introduction. This system involved a W303a-YJM789 hybrid strain with an insertion of a *can1-100-KANMX* cassette near the right telomere of chromosome IV (SGD coordinate 1151 kb) on the W303a-derived homolog and insertion of *SUP4-o* on the allelic position on the YJM789-derived homolog. All strains used in this study are listed in Table S 3.1, and all primers used for strain construction are listed in Table S 3.2 unless stated otherwise.

In our previous study of recombination on chromosome V, we selected crossovers in diploids in which the *can1-100* mutation (a nonsense-suppressible allele) was located near the telomere of chromosome V on one homolog and the *SUP4-o* gene

was located at an allelic position on the other homolog, replacing the *CAN1* gene (BARBERA and PETES 2006); the diploid was also homozygous for the *ade2-1* mutation. Crossovers on chromosome V produced canavanine-resistant red/white sectored colonies. In order to generate a similar system for chromosome IV, we deleted the *CAN1* in the W303a-related haploid W1588-4c, replacing it with the *NAT* gene (nourseothricin-resistance). This replacement was done by transforming W1588-4c with a PCR fragment generated by amplifying sequences of the plasmid pAG25 (GOLDSTEIN and MCCUSKER 1999) using the primers V31539::DR F and W CAN1 KO R (listed in Table S 3.2). Transformants were selected on YPD plates containing 0.1 mg/ml of nourseothricin. The resulting strain with this replacement is JSC10-1. The *CAN1* on chromosome V was deleted in the haploid PSL4 (isogenic with YJM789, genotype: *MATa ade2-1 ura3 gal2 ho::hisG*; LEE *et al.* (2009)) by a similar procedure except that the PCR fragment used for transformation was produced by amplifying the *NAT* gene from JSC10-1 genomic DNA with primers CAN1-500 F and W2 CAN1 KO R. The resulting Nat^R transformant was JSC19-1.

A cassette of about 4 kb containing the *can1-100* gene and *KANMX* was made by inserting the geneticin-resistance gene, *KANMX*, near *can1-100* (between bases 31539 and 31540 on chromosome V) in W303a (Table S 3.1). The *KANMX* sequences were amplified from the pFA6-kanMX4 plasmid (WACH *et al.* 1994) using the primers V31539::DR F and V31539::DR R. Transformants were selected on YPD medium

containing 0.2 mg/ml of geneticin. The transformant with this cassette on chromosome V is JSC11-1. The *can1-100-KANMX* cassette was then amplified from JSC11-1 genomic DNA using the primers IV1510386::can1-100 F and IV1510386::can1-100 R, and transformed into JSC10-1. In the resulting transformant (JSC12-1), the *can1-100-KANMX* cassette is inserted onto chromosome IV between bases 1510386 and 1510387.

The *SUP4-o* gene was amplified using the YCPMP2 plasmid (PIERCE *et al.* 1987) as a template, and the primers YJMIV::Sup4-o F and YJMIV::Sup4-o R. The resultant PCR product was used to transform JSC19-1. Since JSC19-1 contains the ochre-suppressible *ade2-1* mutation, we selected Ade^+ transformants. The resulting strain (JSC21-1) had an insertion of *SUP4-o* between bases 1510386 and 1510387 on chromosome IV.

The haploids JSC12-1 (isogenic with W303a except for changes introduced by transformation) and JSC21-1 (isogenic with YJM789 except for changes introduced by transformation) were mated, and the resulting diploid is JSC22-1. To prevent the diploid from undergoing sporulation, we replaced the *MATα* locus with the *HYG* gene. This replacement was performed using a PCR fragment generated by amplifying the pAG32 plasmid (GOLDSTEIN and MCCUSKER 1999) using the primers MATALPHA NATF and MATALPHA NATR (LEE *et al.* 2009). The resulting diploid JSC25-1 was used in our subsequent mapping experiments.

Although JSC25-1 was designed to allow us to select crossovers as Can^R red/white sectored colonies as we did on chromosome V, in this diploid, the suppression of the *can1-100* allele by the *SUP4-o* suppressor was inefficient. Thus, the diploid was not very sensitive to canavanine, resulting in high background growth of the diploid on plates containing canavanine, precluding the selection of canavanine-resistant colonies. Consequently, we screened for crossovers on chromosome IV by identifying red/white sectored colonies non-selectively as described in the main text.

As described in the text, we found that one of the recombination hotspots (HS4) mapped to the location of an inverted pair of Ty elements, near SGD coordinate 980 kb. The HS4 hotspot was specific to the W303a-derived chromosome. In order to test the hypothesis that these Ty elements formed a secondary DNA structure required for the hotspot activity, we created strains that had markers flanking the inverted repeat and various modifications of the inverted repeat structure. Since the conversion tracts associated with HS4 were very long (median length of 48 kb), we inserted the markers flanking the hotspot about 56 kb apart. The control strain with the intact HS4 hotspot was constructed in several steps. JSC52-1 is a W1588-4c derivative with *URA3* inserted centromere-distal to HS4 between SGD coordinates 1013217 and 1013218. The *URA3* gene in JSC52-1 was derived from the strain JAY291 (Table S 3.1) by amplification of genomic DNA using the primers RE HS4 URA3 F and RE HS4 URA3 R. The Hyg^R haploid JSC54-1 was derived from the Ura$^+$ transformant JSC52-1 by inserting the *HYG*

gene at SGD coordinates 957578 to 957579. The fragment used for the transformation was generated by amplifying the plasmid pAG32 with primers RE HS4 HYG F and RE HS4 HYG R. The diploid JSC67-1 was generated by a cross of JSC54-1 to PSL4. The diploid JSC71-1 was derived from JSC67-1 by replacing the *MATα* gene with the *NAT* gene, using the same method described for the construction of JSC25-1. JSC71-1 was the control diploid for analysis of the strains that had alterations of HS4.

Three different diploids with different modifications of HS4 were constructed. The diploid JSC73-2 strain had a deletion of the centromere-proximal Ty element, *YDRWTy2-3*. The first step in the construction of this strain was to replace the Ty element in JSC54-1 with the *KANMX* gene by transforming the strain with a PCR fragment generated by amplifying the pFA6-kanMX4 plasmid with primers HS4 Ty2-3::KAN F and HS4 Ty2-3::KAN R. The resulting Gen^R transformant is JSC57-1. The diploid JSC68-1 is the result of a cross of JSC57-1 with PSL4. JSC73-2 is derived from JSC68-1 by replacement of the *MATα* locus with the *NAT* gene.

In the diploid JSC74-1, the distance between the Ty elements of HS4 was expanded by insertion of the *KANMX* gene between SGD coordinates 987150 and 987225. The first step in the construction was to transform the haploid JSC54-1 with a *KANMX*-containing PCR fragment produced by amplifying the pFA6-kanMX4 plasmid with the primers HS4 Spacer::Kan F and HS4 Spacer::Kan R. The resulting Gen^R transformant was JSC59-2. The diploid JSC70-3 is a cross of JSC59-2 and PSL4, and

JSC74-1 is a diploid with a deletion of *MATα* (performed as described above for strain JSC71-1).

In JSC77-1, the 5′ delta element of *YDRWTy2-3* was replaced with the *KANMX* gene. We first constructed a Gen^R derivative of JSC54-1 (JSC58-2) by transformation with a PCR fragment generated by amplification of the pFA6-kanMX plasmid with the primers HS4 Ty2-3::KAN F and HS4 DELTA19::KAN R. The diploid JSC75-1 is a cross of JSC58-2 and PSL4. The *MATΔ::NAT* derivative of JSC75-1 (JSC77-1) was constructed by transforming JSC75-1 with a PCR fragment generated by amplifying the pAG25 plasmid with the primers MATALPHA NATF and MATALPHA NATR (LEE *et al.* 2009).

3.6.1.2 Design of the SNP microarrays

The SNP microarrays that were used in our study were based on our own design and were generated by Agilent. The arrays contained about 9200 oligonucleotides, allowing us to analyze about 2300 SNPs. The majority of the oligonucleotides used to examine SNP-specific hybridization patterns were designed using the principles described by Gresham et al. (2010) as employed in our previous study (St Charles *et al.* 2012). In addition, we incorporated information from SNP microarray experiments performed by D. R. Georgianna and J. H. McCusker (Duke University; personal communication) and data described by WINZELER *et al.* (2003). The sequences of the oligonucleotides used in construction of our SNP microarrays will be provided on-line

when we submit our paper for publication, but are not provided in the thesis because of page limitations.

3.6.1.3 Conversion tracts associated with DSBs initiated during G1 or S/G2 of the cell cycle

As discussed in the main text, from the patterns of conversions associated with the crossover, we can infer whether the initiating DSB was generated in the chromosome before replication (G1) or during the S-/G2-phases of the cell cycle. For the crossovers associated with simple 3:1, 4:0, or hybrid 3:1/4:0 and 3:1/4:0/3:1 conversion events (approximately two-thirds of the total events), our conclusions about the timing of the DSB are fairly straightforward: simple 3:1 events are a consequence of S- or G2-initiated DSBs and 4:0 or simple hybrid tracts reflect G1-initiated DSBs. For most complex conversion events, however, the interpretation of the events as G1 or S/G2 can be more difficult. In this section, we discuss these more complex events.

The conversion tracts for all of the crossovers are depicted in Table S 3.3. In this table, each sectored colony is represented by two lines, with the upper line showing the patterns of heterozygosity and homozygosity of SNPs in the red sector and the lower line indicating patterns of SNPs in the white sector. Green, red, and black line segments represent heterozygosity for SNPs, homozygosity for W303a-derived SNPs, and homozygosity for the YJM789-derived SNPs, respectively. In Table S 3.3, we divide all of the recombination events into five groups: A (crossovers without detectable conversions), B (simple 3:1 conversions), C (simple 4:0 conversions), D (simple 3:1/4:0 or

3:1/4:0/3:1 hybrid tracts), and E (complex conversion tracts). The Class A events cannot be characterized as G1- or S/G2-initiated, since they lack the diagnostic conversion tract. All other events were classified as G1- or S/G2-initiated as indicated by one or two asterisks, respectively, in Table S 3.3.

One of the simplest indicators of a G1-initiated DSB is the presence of a 4:0 segment within a conversion tract (LEE *et al.* 2009; LEE and PETES 2010). Although, in principle, these events could represent two independent 3:1 conversions, their frequency is much too high to represent two independent events (LEE *et al.* 2009). Based on the criterion of a 4:0 conversion tract segment, all C and D, and most of the Class E events were initiated in G1.

There are a few conversion tracts classified as G1-initiated events that did not contain 4:0 regions. In Classes E7 and E8, there appears to be a crossover within a 3:1 conversion tract (Table S 3.3). The pattern can be explained as a consequence of the conversion-associated repair of two broken chromatids with non-overlapping conversion tracts and a crossover between the converted segments (Fig. S2 in LEE *et al.* (2009)). In 3:1 conversion events, the chromosome with the DSB is the recipient of information. For example, in Figure 3.1A, since the conversion event results in 3 copies of the W303a-derived SNPs ("red" SNPs) and one copy of the YJM789-derived SNPs ("black" SNPs), we infer that the conversion was initiated by a DSB on the YJM789-derived homolog. In addition, the region of the conversion tract that is homozygous for

the W303a-derived SNPs is expected to be in the red sector and the heterozygous region of the tract is expected to be in the white sector (Figure 3.1A). Of the 34 3:1 conversion tracts observed in our study, 29 (Classes B1 and B2) had this pattern. In five tracts (Classes E3 and E4), the homozygous and heterozygous regions were in the "wrong" sectors. These patterns can be explained as a consequence of the repair of two broken chromatids (Fig. S1 in LEE *et al.* (2009)). By similar explanations, Classes E5, E7, E8, E10, E11, and E30 are likely to reflect G1-initiated events.

There are alternative interpretations of the conversion events depicted in Table S 3.3. For example, a 4:0 tract could reflect a DSB that was formed in G2 if the broken chromatid was segregated into the daughter cell without repair. If this broken chromosome was then replicated and repaired in G2 of the subsequent cycle, it would generate a 4:0 conversion. One argument against this scenario is that γ ray-treatment of G2-synchronized yeast cells results primarily in 3:1 events rather than 4:0 events (LEE and PETES 2010). It is also possible that some of the conversions classified as initiated in G2 are initiated in G1, if one of the DSBs was repaired associated with a very short heteroduplex tract. If the heteroduplex does not contain a SNP, it is undetectable as a conversion event. Despite these uncertainties, the most likely interpretation of our data is that both G1- and S-/G-initiated events contribute to the crossovers detected in our study.

3.6.1.4 Associations of gene conversion tracts with the chromosome elements shown in Table S 3.5.

Based on many previous studies (such as TANG *et al.* (2011)), the position of the DSBs that initiate mitotic recombination events are located within the boundaries of the gene conversion events. Therefore, we examined the frequencies of various chromosome elements (replication origins, palindromes, G4 quadruplex motifs, etc.; Table S 3.5) within gene conversion tracts for their over- or under-representation compared to a random distribution. We analyzed the "windows" containing the gene conversion tracts (the sites of potential DSB initiation) using four different methods. For all four methods, the region included for the Class A events (crossover without a gene conversion event; Table S 3.3) was between the heterozygous SNPs and the homozygous SNPs that most closely flanked the crossover.

For Method 1, we summed all conversion tracts of each event to generate the investigated "window." For Method 2, we also summed the conversion tracts for each event, but excluded those events in which the conversion tract was greater than 20 kb. For Method 3, we restricted the "window" of the conversion event to that portion of the conversion event most likely to contain the DSB. More specifically, if the conversion event was a 3:1/4:0 tract, we used a window that was restricted to the 4:0 portion of the tract, since (as shown in Figure 3.1C) the 4:0 region of a hybrid tract is the expected location of the recombinogenic DSB. In complex conversion tracts with multiple 4:0 regions, all 4:0 regions were counted in the window for that sectored colony. In Table S

3.3, the regions that were included in our analysis by Method 3 are outlined in blue rectangles. Method 4 was identical to Method 3, except we excluded those events in which the tract size was greater than 20 kb. The complex recombination events that had conversion events initiated on both homologs (Classes E29-E31) were not included in any of the calculations.

In addition to using different methods of analyzing the data, as described in the main text, we chose five different datasets to examine: 1) all of the data derived from JSC25 ("JSC25 all" in Table 3.1), 2) recombination events initiated in G1 ("JSC25 G1"), 3) events initiated in G2 ("JSC25 G2"), 4) events initiated on the W303a-derived homolog ("W303a all"), and 5) events initiated on the YJM789-derived homolog ("YJM789 all"). All five datasets were examined using all four methods. As an example of the analysis, we will describe the procedure to determine whether there was a significant enrichment of LTR (delta) elements associated with the gene conversion tracts in the dataset "JSC25 all" when examined by Method 1. The Saccharomyces Genome Database shows that there are 33 delta elements in the 1,061,175 bp region between *CEN4* and the *can1-100/KANMX/SUP-o* insertions. The sum of the gene conversions (calculated by Method 1) for the 135 (136 crossovers minus one double crossover) events examined in the "JSC25 all" dataset was 2,660,253 bp. Since the density of delta elements on the right arm of chromosome IV is 1 per 32,157 bp (1,061,175 divided by 33 delta elements), the expected number of delta elements in the conversion window of 2,660,253 bp is 83

(2,660,253 bp conversion window divided by 32,157 bp), and the expected number of delta elements not included in the conversion window is 4,372 (135 events times 33 LTRs per event minus the 83 LTRs expected within the conversion window). The observed number of delta elements in the 2,660,253 bp conversion window was 130 and the observed number of delta elements not included in the conversion window was 4,325 (135 events times 33 LTRs per event minus the 130 LTRs observed within the conversion window). We compared the numbers of expected (83 and 4,372) and observed (130 and 4,325) events by a chi-square goodness-of-fit test using the Microsoft Excel "chisq.test" function. The p value for this comparison was 1.5×10^{-7}. Since we analyzed the dataset for 18 different chromosome elements (listed in Table S 3.5), we determined that this p value was still significant after correcting for multiple comparisons (HOCHBERG and BENJAMINI 1990).

A similar procedure to that described above was used for all of the other comparisons and those comparisons that showed a significant over-representation of elements are shown in Table 3.1; those comparisons that had a the significant under-representation of elements are shown in Table S 3.6. The conversion window sizes (shown in kb in parentheses) used for the different methods (indicated by the number in boldface) and the various datasets were: **1** JSC25 all (2660253); **2** JSC25 all (774888); **3** JSC25 all (1765865); **4** JSC25 all (689319); **1** JSC25 G1 (2129610); **2** JSC25 G1 (489295); **3** JSC25 G1 (1244102); **4** JSC25 G1 (410034); **1** JSC25 G2 (310456); **2** JSC25 G2 (203353); **3**

JSC25 G2 (292886); **4** JSC25 G2 (185983); **1** W303a all (1492068); **2** W303a all (397036); **3** W303a all (932621); **4** W303a all (336974); **1** YJM789 all (947998); **2** YJM789 all (295812); **3** YJM789 all (615629); **4** YJM789 all (270305).

3.6.1.5 Analysis of mitotic recombination hotspot activities.

Although the distribution of recombination events depicted in Figure 3.7A appear strikingly non-random, we did several tests to confirm whether this conclusion. One complication in the depiction of Figure 3.7 (as explained previously in the text) is that the frequency with which a SNPs will be involved in a conversion event is affected by both the frequency of initiating DSBs that occur near the SNP and the size of the conversion tracts extending from the initiating DSB. To eliminate the effect of the size of the conversion event on our calculations, for the bin analysis shown in Figure 3.8, each conversion tract was represented by a single point/coordinate.

The positions of the points used in this analysis for each sectored colony are given in Table S 3.7. For the majority of the conversion tracts, the point that represented the tract was at the center of the DSB region as determined for Methods 3 and 4 described above. For Class A events (crossovers without conversions), the event was represented by the midpoint between the last heterozygous SNP and the first homozygous SNP. For conversion tracts that were discontinuous (regions of conversion separated by regions of heterozygosity such as Classes E2 and E6), the tract was treated

as a single conversion tract with the point located in the middle of the tract. The events in which both homologs were donors (Classes E29-E31) were not used for this analysis.

As described in the main text, we then distributed the "point" conversion events into bins of 200 kb on chromosome IV (Figure 3.8) and tested whether the distribution was random. The analysis was done for all of the data taken together (Figure 3.8A), and for the W303a-associated (G1-initiated and S-/G2-initiated) and YJM789-associated (G1-initiated and S-/G2-initiated) events treated separately (Figure 3.8B). By chi-square goodness-of-fit tests, we found a very significant hotspot for W303a-associated G1-initiated events in the bin with coordinates 845-1045 kb. In addition, by contingency chi-square analysis, we found significant differences in hotspot activity between the homologs and differences between the distributions of G1- and S/G2-initiated events.

3.6.2 Supplementary Tables

Table S 3.1 Strain list and strain constructions.

Standard gene nomenclature is used with a few exceptions. An insertion unaccompanied by a deletion is indicated by a double colon; a replacement is indicated by a Δ, followed by a double colon and the name of the replacing gene. If an insertion is placed within a genomic sequence unassociated with a gene name, we indicate the position of the centromere-proximal base followed by a double colon and the name of the inserted gene. For example, IV10132176::URA3 indicates that the URA3 gene was inserted between bases 10132176 and 10132177 on chromosome IV. For diploid genotypes, the genotype of the W303a-derived haploid is shown above the diagonal and the genotype of the YJM789-derived haploid is shown below the diagonal. The HS4 hotspot (inverted Ty elements) is missing on the YJM789-derived homology; the nomenclature used to describe the missing Ty elements in the genotypes above is: ydrwTy2-3Δ.

Strain	Genotype	Reference/construction
PSL4	*MATα ade2-1 ura3 gal2 ho::hisG*	Isogenic with YJM789 (LEE *et al.* 2009)
W303a	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 rad5*	THOMAS and ROTHSTEIN (1989)
W1588-4c	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5*	Isogenic with W303a (ZHAO *et al.* 1998)
JAY291	*MATa* wild-type haploid	ARGUESO *et al.* (2009)
JSC10-1	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100Δ::NAT RAD5*	W1588-4c with replacement of *can1-100* with *NAT.* Isogenic with W303a.
JSC11-1	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 V31539::KANMX*	W303a with *KANMX* inserted between bases V31539 and V31540.

JSC12-1	*MATa leu2-3,112, his3-11,15 ura3-1 ade2-1 trp1-1 can1-100Δ::NAT RAD5 IV1510386::KANMX-can1-100*	JSC10-1 transformed with *KANMX-can1-100* cassette inserted between bases IV1510386 and 1510387. Isogenic with W303a.
JSC19-1	*MATα ade2-1 ura3 gal2 ho::hisG CAN1Δ::NAT*	PSL4 with *CAN1* replaced with *NAT.* Isogenic with YJM789.
JSC21-1	*MATα ade2-1 ura3 gal2 ho::hisG CAN1Δ::NAT IV1510386::SUP4-o*	JSC19-1 with insertion of *SUP4-o* at IV1510386. Isogenic with YJM789.
JSC22-1	*MATa/MATα leu2-3,112/LEU2 his3-11,15/HIS3 ura3-1 /ura3 GAL2/gal2 ade2-1/ade2-1 trp1-1/TRP1 can1-100Δ::NAT/ CAN1Δ::NAT RAD5/RAD5 IV1510386::KANMX-can1-100/IVI1510386::SUP4-o*	Cross of JSC12-1 and JSC21-1.
JSC25-1	*MATa/MATα::HYG leu2-3,112/LEU2 his3-11,15/HIS3 ura3-1 /ura3 GAL2/gal2 ade2-1/ade2-1 trp1-1/TRP1 can1-100Δ::NAT/ CAN1Δ::NAT RAD5/RAD5 IV1510386::KANMX-can1-100/IVI1510386::SUP4-o*	JSC22-1 with *MATα* replaced with *HYG.*
JSC52-1	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5 IV1013217::URA3*	W1588-4c with *URA3* inserted centromere-proximal to HS4 at position IV1013217. Isogenic with W303a.
JSC54-1	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5 IV957578::HYG IV1013217::URA3*	JSC52-1 with *HYG* inserted centromere-distal to HS4 at position IV957578. Isogenic with W303a.

JSC57-1	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5 IV957578::HYG IV1013217::URA3 YDRWTy2-3Δ::KANMX*	JSC54-1 with a replacement of *YDRWTy2-3* with *KANMX*. Isogenic with W303a.
JSC58-2	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5 IV957578::HYG IV1013217::URA3 YDRWdelta19Δ:: KANMX*	JSC54-1 with a replacement of *YDRWdelta19* with *KANMX*. Isogenic with W303a.
JSC59-2	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5 IV957578::HYG IV1013217::URA3 IV987138:: KANMX*	JSC54-1 with *KANMX* inserted between the two Ty elements of HS4 at position IV987138. Isogenic with W303a.
JSC67-1	*MATa/MATα ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217*	Cross of JSC54-1 and PSL4
JSC68-1	*MATa/MATα ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217 YDRWTy2-3Δ::KANMX/ydrwTy2-3Δ*	Cross of JSC57-1 and PSL4
JSC70-3	*MATa/MATα ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217 IV987138:: KANMX/IV987138*	Cross of JSC59-2 and PSL4

JSC71-1	*MATa/MATαΔ::NAT ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217*	*MATαΔ* derivative of JSC67-1
JSC73-2	*MATa/MATαΔ::NAT ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217 YDRWTy2-3::KANMX/ydrwTy2-3Δ*	*MATαΔ* derivative of JSC68-1
JSC74-1	*MATa/MATαΔ::NAT ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217 IV987138:: KANMX/IV987138*	*MATαΔ* derivative of JSC70-3.
JSC75-1	*MATa/MATα ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217 YDRWdelta19:: KANMX/ydrwdelta19Δ*	Cross of JSC58-2 and PSL4.
JSC77-1	*MATa/MATαΔ::NAT ade2-1/ade2-1 ura3-1/ura3 GAL2/gal2 ho/ho::hisG can1-100/CAN1 leu2-3,112/ LEU2 his3-11,15/ HIS3 trp1-1/TRP1 IV957578::HYG/ IV957578 IV10132176::URA3/IV1013217 YDRWdelta19:: KANMX/ydrwdelta19Δ*	*MATαΔ* derivative of JSC75-1.

Table S 3.2 Names and sequences of primers used in strain constructions and genome analysis.

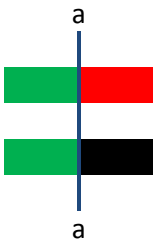
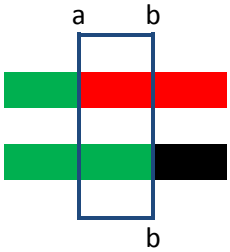
Primer name	Primer sequence[4]
V31539::DR F	AGGTCTGAAGGAGTTTCAAATGCTTCTACTCCGTCTGCTTTC TTTTCGGGACGTACGCTGCAGGTCGAC
W CAN1 KO R	ACTGGCAAGTGCGTATAAATTAAACCTATTTCTTTATCATC ATATTTACTTATCGATGAATTCGAGCTCG
CAN1-500 F	GAGTTCTGCCCTTGGCTTCC
W2 CAN1 KO R	GGATTTGGCGTGGATGAAGG
V31539::DR R	ATAATCTCAGTATATTTATTTTTATCGATGATTCACCACAAC AATCTTGCTATCGATGAATTCGAGCTCG
IV1510386::can1-100 F	CATACGTTATGCACTTCATTCTTCTTGTCGGTTTGATAACAG CAGAATCTACGTACGCTGCAGGTCGAC
IV1510386::can1-100 R	GCGTTTTCGAGGTATGGCTTCTGCCGGGCTAACGTTCAAAT TAAAGGAACTTATATCTTTAACAGATTCCAA
YJMIV::SUP4-o F	CATACGTTATGCACTTCATTCTTTTTGTCGGTTTGATACCAG CAGAATCTAGGATCCGGGACCGGATAAT
YJMIV::SUP4-o R	GCGTTTTCGAGGTATGGCTTCTGTAGGGTTAACGCTCAAAT TAAAGGAACTGGATCCGGAATTCTTGAAAG
IV 980403 F	TTGAACATGGTCGCGACAGC
Ty2 R	GCCGTGCTGTTGGTACTGTCC
Ty1.2 R	TTGAGGAGAGGCATGATGG
IV 993256 F	CTGGTGCTAACCGCTTGTGC
RE HS4 URA3 F	TTTATATAGATAAACAAACTTGCAGGACAGATAGTTAAGC GTCTATATCATAATGTGGCTGTGGTTTCAGG
RE HS4 URA3 R	TCTTTTTTGCCTTTTATCATTTTTGTACTTTTTTCTTCGCTTAA AATACACAGATTCCCGGGTAATAACTG

[4] Primer sequences are written 5′ to 3′.

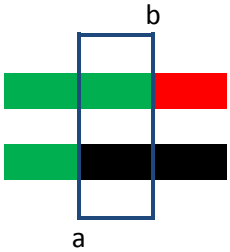
RE HS4 HYG EXT F	CACAAGCACTTCTTGCGAGTCC
RE HS4 HYG EXT R	GGAGTCTCAATTCAGGGATCTGG
HS4 Ty2-3::KAN F	GTTAATAAAACACACATGCAACTTGTTGGAATAAAAATCC ACTATCGTCTTCGTACGCTGCAGGTCGAC
HS4 Ty2-3::KAN R	TTATCTAATTACCCACATATATCTCATAACTATTAGTTGAT AGACGTGTAGATCGATGAATTCGAGCTCG
HS4 DELTA19::KAN R	CTTGATTTGATGGGACTTCCTTAGAAGTAACCGAAGCAGCG GCGCTACCATATCGATGAATTCGAGCTCG
HS4 SPACER:: KAN F	AATGGAATCCCAACAATTACATCAAAATCCACATTCTCTAC ACGTCTATCACGTACGCTGCAGGTCGAC
HS4 SPACER::KAN R	GGAATCCCAACAATTATCTAATTACCCACATATATCTCATA ACTATTAGTTATCGATGAATTCGAGCTCG

Table S 3.3 Depictions of conversion tracts associated with reciprocal crossovers.

All of the different types of conversion tracts observed in this study are shown below. For each conversion tract, the top colored line depicts the red side of the sector and the bottom line depicts the white side of the sector. Green, red, and black indicate heterozygosity for SNPs, homozygosity for W303a-derived SNPs, and homozygosity for YJM789-derived SNPs, respectively. Lowercase letters located above and below the conversion tract mark transitions within the conversion tract. The coordinates of the SNPs on either side of the transitions are listed in Table S 3.4. The blue boxes enclose the regions that are likely to be associated with DSB formation as described in the Supplementary Information section. Our interpretation of subclasses as representing G1- or G2-initiated events are indicated by asterisks to the right of the subclass names. One asterisk and two asterisks indicate a G1- and G2-initiated DSB, respectively. The lengths of the conversion tracts are not drawn to scale.

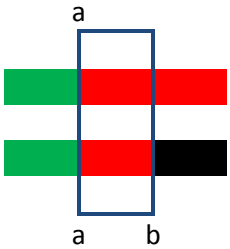
Class	# Class obs.	# Subclass obs.	
A	18		Simple crossover, no conversion tract
			
B	29		3:1 conversion tract
B1**		16	
			

B2** 13

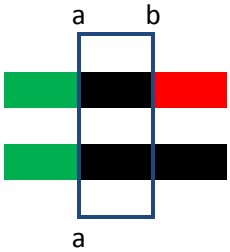


C 7 4:0 conversion tract

C1* 1

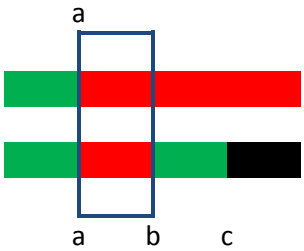


C2* 6

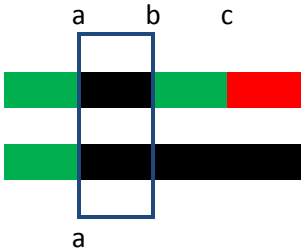


D 46 Simple hybrid conversion tract (4:0/3:1 or 3:1/4:0/3:1)

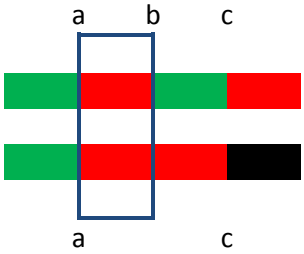
D1* 4



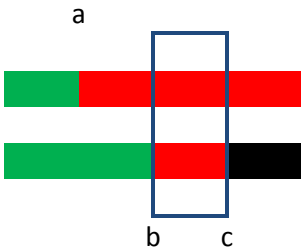
D2* 3



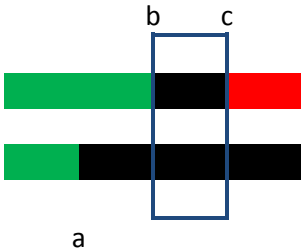
D3* 1



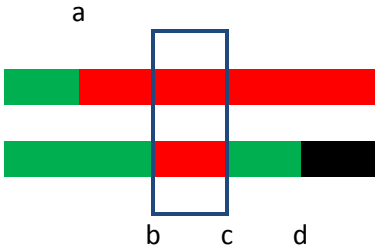
D4* 3



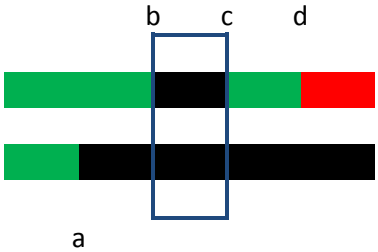
D5* 6



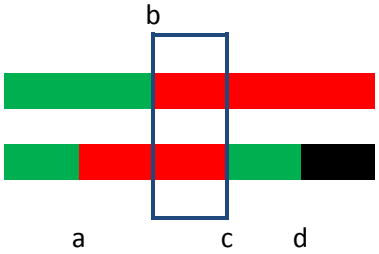
D6* 8



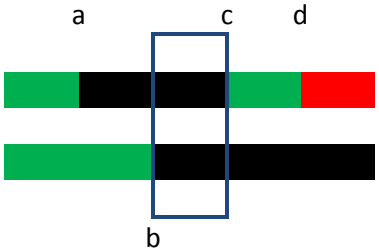
D7* 13



D8* 2

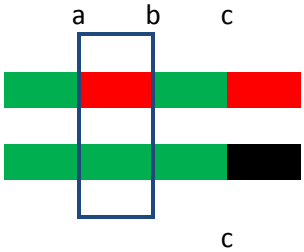


D9* 6

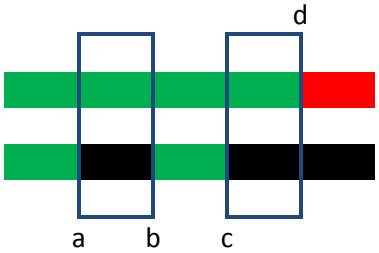


E 39 Complex conversion tracts

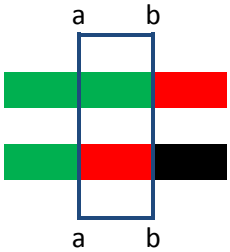
E1** 1



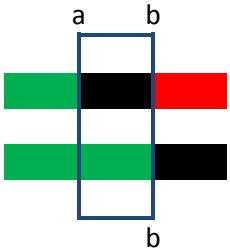
E2** 1



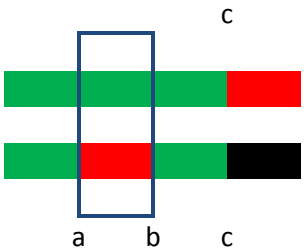
E3* 4



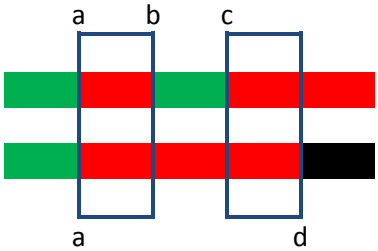
E4* 1



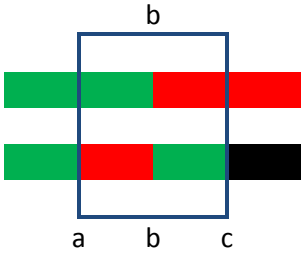
E5* 1



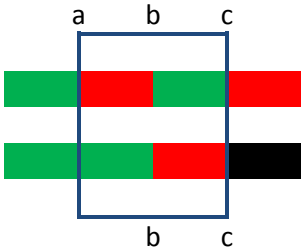
E6* 1



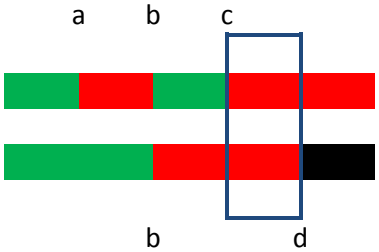
E7* 3



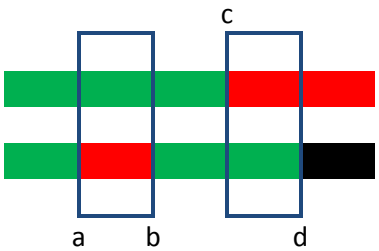
E8* 1



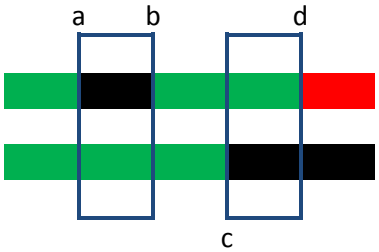
E9* 1



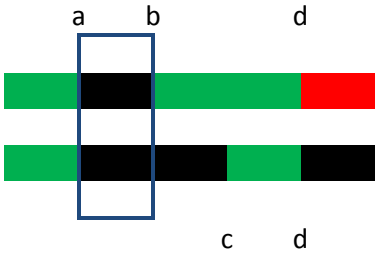
E10* 1



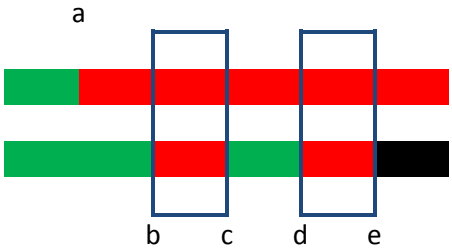
E11* 2



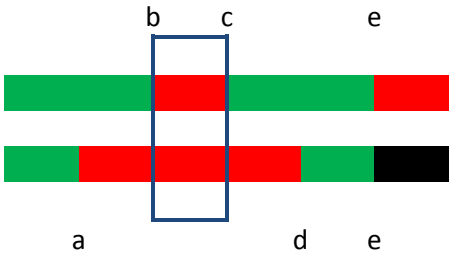
E12* 1



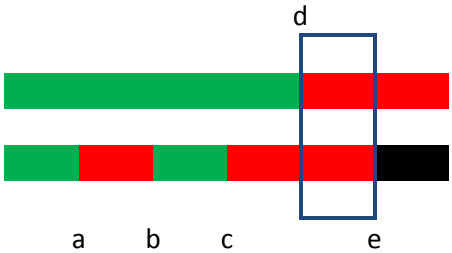
E13* 1



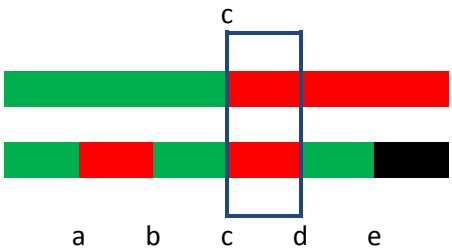
E14* 1



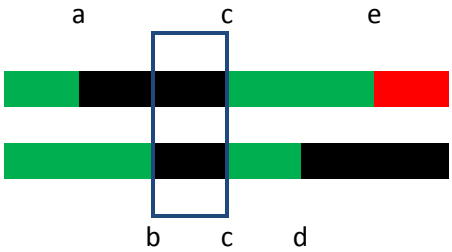
E15* 1



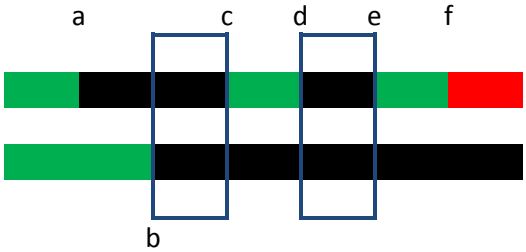
E16* 1



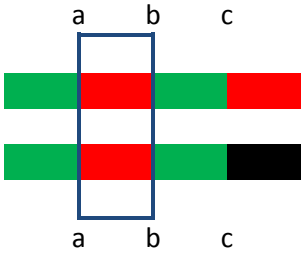
E17* 1



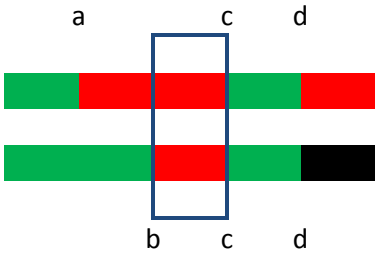
E18* 1



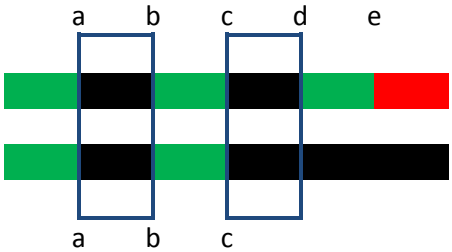
E19* 1



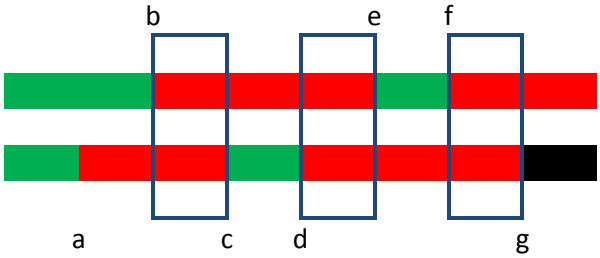
E20* 1



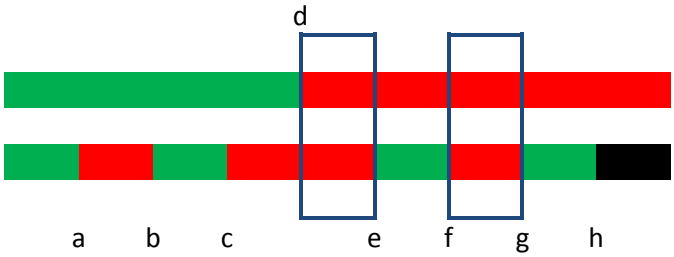
E21* 1



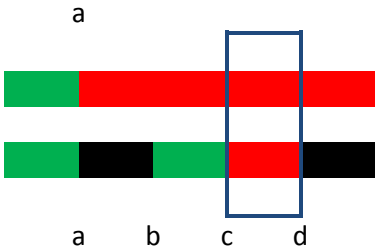
E22* 1



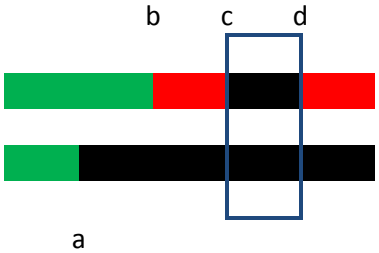
E23* 1



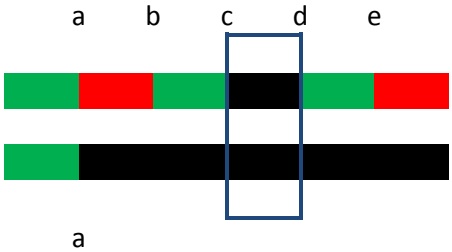
E24* 1



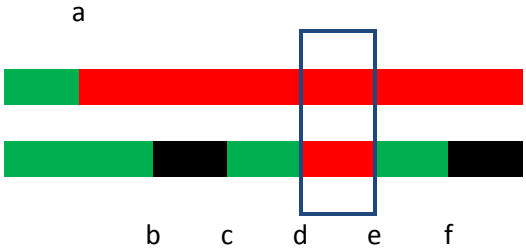
E25* 1



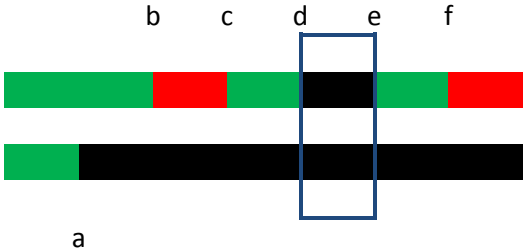
E26* 1



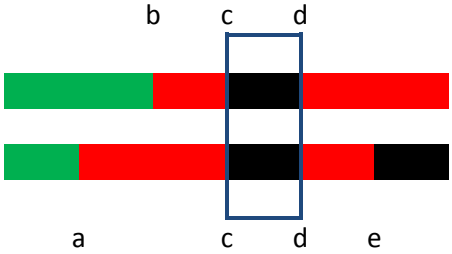
E27* 1



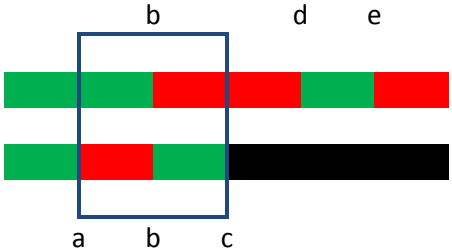
E28* 1



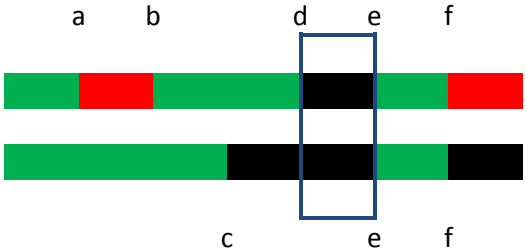
E29* 1



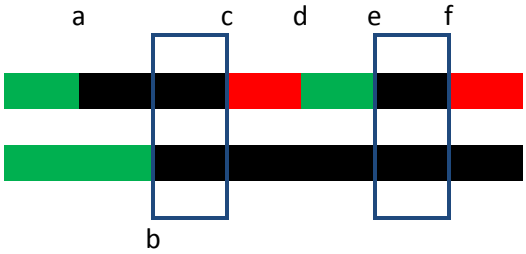
E30* 1



E31* 1



E32* 1



E33* 1

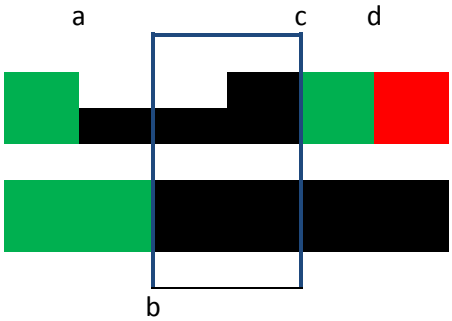


Table S 3.4 SGD coordinates for crossover and conversion transitions on chromosome IV.

Sector	Event class[5]	Transition label[6]	Markers flanking transitions[7]	
			Left	Right
1RW	E7	a	1248863	1248966
		b	1254353	1254489
		c	1255048	1255087
		d	1259633	1259877
		e	1268296	1269558
		f	1273350	1274404
2RW	E31	a	1373752	1374198
		b	1375238	1379772
		c	1380606	1385030
		d	1385030	1386901
		e	1387294	1388991
		f	1388991	1390775
3RW	D1	a	1047033	1049553
		b	1049553	1051286
		c	1055397	1055663
4RW	E4	a	1157254	1161545
		b	1161545	1171413
5RW	E32	a	959463	960043
		b	964391	965740
		c	1022402	1036052
		d	1067771	1069072
		e	1069072	1071135
		f	1077836	1082836
6RW	E3	a	1190838	1216161
		b	1225482	1231783
7RW	E3	a	1429103	1429226

[5] The classes of events are defined in Supp. Table 3.3.
[6] The lower case letters refer to transitions between heterozygous and homozygous regions as shown in Supp. Table 3.
[7] These numbers represent SGD coordinates of SNPs located on each side of the transition. It should be noted that these coordinates are based on SGD coordinates from Feb. 2010, and some of these coordinates may be different from those currently displayed in SGD.

		b	1431061	1431782
9RW	B1	a	1275176	1275710
		b	1275710	1276330
10RW	B2	a	1161545	1171413
		b	1174705	1178332
11RW	D7	a	1308659	1308932
		b	1310877	1311425
		c	1311425	1311652
		d	1315266	1315889
12RW	D5	a	1267082	1267721
		b	1273350	1274404
		c	1275176	1275710
13RW	E13	a	1438183	1441463
		b	1470362	1470765
		c	1476382	1476496
		d	1476496	1477097
		e	1477494	1479384
14RW	E20	a	1465228	1465378
		b	1466345	1466614
		c	1468543	1468701
		d	1470362	1470765
15RW	D7	a	1267082	1267721
		b	1273350	1274404
		c	1278598	1279018
		d	1289357	1289797
16RW	B2	a	1079615	1082836
		b	1089446	1111969
18RW	E5	a	1040924	1044556
		b	1047033	1049553
		c	1049553	1051286
19RW	D1	a	1480252	1483459
		b	1484403	1484580
		c	1492290	1492550
20RW	E3	a	1067771	1069072
		b	1071758	1075942
21RW	B2	a	1401638	1401858
		b	1410413	1414518
23RW	A	a	1161545	1171413
24RW	E11	a	1357190	1357597

		b	1361012	1362046
		c	1364731	1365816
		d	1371117	1371846
25RW	A	a	1201602	1216161
26RW	A	a	1082836	1089446
27RW	D7	a	580548	580807
		b	582114	582453
		c	584836	585304
		d	594021	595015
28RW	D5	a	1357597	1358761
		b	1364731	1365816
		c	1365816	1366726
29RW	E23	a	1292496	1293222
		b	1297117	1297543
		c	1304942	1305105
		d	1310376	1310818
		e	1313530	1313572
		f	1317318	1317884
		g	1320966	1321106
		h	1324027	1324869
30RW	D7	a	1435418	1438183
		b	1441931	1442062
		c	1443817	1443889
		d	1443889	1444237
31RW	A	a	1111969	1142607
32RW	B1	a	1059384	1065071
		b	1089446	1111969
33RW	D5	a	1402542	1402695
		b	1411220	1414470
		c	1415174	1416334
34RW	D2	a	1201602	1216161
		b	1221046	1221462
		c	1225482	1231783
35RW	B2	a	1256206	1256457
		b	1256631	1257078
36RW	A	a	1111969	1142607
37RW	E10	a	980838	992782
		b	993347	993964
		c	999490	999715

		d	1002452	1003257
38RW	E14	a	512801	521609
		b	526963	527305
		c	527510	539130
		d	545386	546290
		e	546605	551364
39RW	A	a	868904	889108
40RW	D7	a	964391	965740
		b	975687	976249
		c	1010105	1011955
		d	1013909	1015206
41RW	A	a	1190182	1216211
42RW	E33	a	868904	887030
		b	976159	977414
		c	980838	992782
		d	996000	996527
43RW	D1	a	1190182	1216161
		b	1217482	1217788
		c	1231783	1235569
44RW	E7	a	1245782	1246274
		b	1249203	1250136
		c	1254154	1254293
45RW	E16	a	1417627	1418072
		b	1420136	1420580
		c	1422118	1422596
		d	1430800	1431782
		e	1433085	1433125
46RW	E1	a	889108	890618
		b	890618	892648
		c	895029	903910
47RW	E29	a	922159	924604
		b	963792	965740
		c	968451	969369
		d	1028494	1036052
		e	1037332	1039370
48RW	B1	a	774835	775273
		b	779147	779952
49RW	D5	a	964391	965936
		b	966485	966599

		c	1010105	1012642
50RW	B1	a	1045316	1046711
		b	1046711	1047033
51RW	B2	a	460205	469765
		b	477827	482169
52RW	D1	a	907253	911910
		b	916127	917134
		c	918488	918591
53RW	A	a	1161545	1171413
54RW	C2	a	855308	855878
		b	911910	912446
55RW	D8	a	629580	630883
		b	633870	633945
		c	643241	643421
		d	644066	651658
56RW	B2	a	1372403	1374333
		b	1379772	1380606
57RW	B1	a	1059444	1065071
		b	1066867	1069072
58RW	A	a	1480252	1483459
59RW	C2	a	1161545	1171413
		b	1201602	1216161
60RW	D6	a	768378	770295
		b	770845	771197
		c	774522	774716
		d	776615	776665
61RW	D4	a	1111969	1142607
		b	1146877	1147303
		c	1161545	1171413
63RW	E7	a	1157254	1161545
		b	1161545	1171413
		c	1174705	1178332
64RW	D9	a	939092	939734
		b	942797	943052
		c	945141	945589
		d	953272	955078
65RW	A	a	1380606	1385030
66RW	D2	a	1111969	1142607
		b	1185793	1186611

		c	1190182	1216161
67RW	C2	a	980838	992782
		b	1003836	1003935
68RW	D7	a	1461572	1461942
		b	1465926	1466198
		c	1468543	1468701
		d	1469252	1469487
69RW	B2	a	1217788	1219699
		b	1221462	1225482
70RW	D6	a	842451	843111
		b	843111	843574
		c	845115	845377
		d	851194	852302
71RW	D5	a	1498068	1498508
		b	1507062	1507520
		c	1508989	1509699
72-1RW	D6	a	788875	789378
		b	789378	790202
		c	795523	795869
		d	802606	802810
72-2RW	A	a	1182080	1183438
73RW	B2	a	541971	542454
		b	544627	544999
74RW	E11	a	739939	741160
		b	747660	748528
		c	748528	749244
		d	752071	753192
75RW	A	a	1346791	1349668
76RW	B2	a	491787	495183
		b	495183	496923
77RW	E6	a	1201602	1216161
		b	1220689	1221462
		c	1221462	1225482
		d	1225482	1231783
78RW	D4	a	729738	730063
		b	730720	731914
		c	733517	735346
79RW	D6	a	671494	672120
		b	684381	684999

		c	690253	690489
		d	698494	699008
80RW	E15	a	1255964	1256206
		b	1261433	1262162
		c	1269558	1269689
		d	1271391	1271682
		e	1279397	1279784
81RW	E17	a	1231783	1235569
		b	1240303	1242166
		c	1251733	1252281
		d	1265332	1266551
		e	1269689	1269955
82RW	D9	a	558167	558882
		b	563201	564349
		c	571796	572478
		d	576294	578154
83RW	B1	a	1391617	1392166
		b	1392166	1394070
84RW	D6	a	1281755	1282224
		b	1282790	1283368
		c	1283368	1283747
		d	1284092	1284918
85RW	B1	a	1071233	1071758
		b	1075942	1077702
86RW	B1	a		449787
		b	454159	454186
87RW	C1	a	713743	714477
		b	715012	715335
88RW	C2	a	1017595	1022402
		b	1022402	1028494
89RW	E21	a	460205	469765
		b	471247	476888
		c	477827	482169
		d	482169	482635
		e	482635	482940
90RW	E30	a	682940	683460
		b	686861	687261
		c	698254	698333
		d	699796	699924

		e	700648	700759
91RW	D7	a	581168	581504
		b	582114	582453
		c	585541	586687
		d	593737	594021
92RW	D7	a	1263246	1264102
		b	1265332	1266551
		c	1267721	1267763
		d	1271793	1271980
93RW	D4	a	1263246	1264102
		b	1266671	1266933
		c	1267082	1267721
94RW	A	a	1111969	1142607
95RW	B1	a	668612	669986
		b	670622	671011
96RW	E22	a	694543	694817
		b	698211	698254
		c	700759	701050
		d	703507	703836
		e	706589	706908
		f	707328	708625
		g	708625	709196
97RW	D9	a	861704	862566
		b	863830	865277
		c	893739	894438
		d	909432	911910
98RW	D3	a	544266	544627
		b	545033	545123
		c	546605	551364
99RW	B2	a	889108	890618
		b	890618	892648
100RW	D6	a	747660	748528
		b	751098	752071
		c	754411	755087
		d	755087	755446
101RW	B2	a	1499309	1499476
		b	1510193	1510396
102RW	D2	a	593628	593737
		b	596803	597003

		c	597003	598170
103RW	E18	a	949439	949782
		b	951654	952255
		c	952316	953272
		d	970231	970419
		e	1000365	1001602
		f	1006070	1006514
104RW	E19	a	868904	887030
		b	895029	901094
		c	901094	907253
105RW	A	a	460205	469765
106RW	E3	a	477827	482169
		b	521609	524793
107RW	D7	a	972912	973113
		b	976249	977414
		c	1007002	1007144
		d	1008785	1008829
108RW	E8	a	777015	777375
		b	794043	794806
		c	796161	802539
109RW	B1	a	491787	495183
		b	495183	496923
110RW	B2	a	460205	469765
		b	470717	471226
111RW	C2	a	863830	865277
		b	868904	887030
112RW	D6	a	663372	663847
		b	674798	675271
		c	696034	696092
		d	700236	700426
113RW	B1	a	666145	666933
		b	667323	667565
114RW	B1	a	564732	565232
		b	567351	567792
115RW	E12	a	853434	855165
		b	863830	865277
		c	868013	868562
		d	868904	887030
116RW	D6	a	796161	802110

		b	808984	809854
		c	811980	812466
		d	816605	817112
117RW	D7	a	966213	966485
		b	980838	992782
		c	1004465	1005641
		d	1013370	1013909
118RW	A	a	796161	802110
119RW	D9	a	495183	496923
		b	509817	512538
		c	524793	525397
		d	527510	539083
120RW	D7	a	754032	754057
		b	755446	755533
		c	761361	761909
		d	765384	766033
121RW	C2	a	706018	706589
		b	712108	712459
122RW	A	a	796161	802110
124RW	E9	a	1004465	1005641
		b	1007819	1008785
		c	1008785	1008829
		d	1010105	1011921
128RW	D7	a	922159	924604
		b	930041	930998
		c	931361	932198
		d	940065	940980
129RW	B1	a	761909	762882
		b	771197	772684
130RW	E25	a	1474547	1474663
		b	1474866	1475232
		c	1477494	1479384
		d	1480252	1483459
131RW	D8	a	1003257	1003678
		b	1003678	1003797
		c	1007819	1008785
		d	1008785	1008829
133RW	E27	a	587919	588369
		b	588547	588694

		c	593197	593449
		d	594021	595514
		e	597003	598170
		f	604773	605544
136RW	E2	a	601671	601816
		b	604427	604773
		c	611081	611611
		d	616610	617827
139RW	D7	a	1007819	1008785
		b	1011921	1012642
		c	1015206	1017595
		d	1017595	1028494
140RW	D9	a	681904	681933
		b	681933	682289
		c	686024	686695
		d	690515	690737
142RW	D9	a	583189	583420
		b	584836	585304
		c	588694	589427
		d	594021	595015
143RW	E7	a	1268296	1268882
		b	1272145	1272729
		c	1278598	1279018
144RW	E24	a	1190182	1216161
		b	1217163	1217192
		c	1217192	1217482
		d	1217788	1219699
145RW	B1	a	640958	641270
		b	657097	658422
146RW	B2	a	922159	924604
		b	934165	934343
147RW	A	a	853434	855165
149RW	E26	a	707328	708625
		b	711298	711895
		c	712684	712843
		d	713491	714477
		e	716627	717363
150RW	D5	a	476888	482169
		b	482635	482940

		c	485707	491787
153RW	B1	a	495183	496923
		b	499309	499633
154RW	B1	a	502953	503234
		b	505869	507475
156RW	A	a	485707	491787

Table S 3.5 Chromosome elements examined for their representation in mitotic gene conversion tracts.

Chromosome element (# of each element present on the right arm of chromosome IV)[8]	Description or reference
Palindromes[9] (46)	LISNIC *et al.* (2005); Z. Zgaga, personal communication
Tandem repeats[10] (97)	Tandem-repeat database (TRDB)[11] (GELFAND *et al.* 2007)
Triplet repeats[12] (10)	Triplet repeats as identified in the TRDB
G4 DNA[13] (34)	Dataset S1 in Capra *et al.* (2010)
tRNA genes (23)	Saccharomyces Genome Database (SGD)
Short (150 bp) intergenic regions (61)	Saccharomyces Genome Database (SGD)
Intron-containing genes (24)	Saccharomyces Genome Database (SGD)
ARS elements (28)	Saccharomyces Genome Database (SGD)
Early- and middle-firing ARS elements (10)	Table S1 in (FACHINETTI *et al.* 2010)

[8] For most of the listed elements, we examined the number of such elements between *CEN4* and the markers near the right telomere used to diagnose crossovers.
[9] Palindromes that were at least 16 bp are counted.
[10] The repeats examined were in the size range of 2 to 213 bp, with a minimum repeat tract of 24 bp.
[11] This database is accessible at the Website: https://tandem.bu.edu/cgi-bin/trdb/trdb.exe
[12] The minimal number of repeats/tract was 8.
[13] The G4 motifs used in this analysis were four tracts of 3 G's separated by spacers of less than 25 bp.

Long terminal repeats (LTRs) (33)	Saccharomyces Genome Database (SGD)
Long terminal repeats (> 1 kb) from a Ty element (11)	Saccharomyces Genome Database (SGD)
Highly-transcribed genes[14] (20)	Table S4 in NAGALAKSHMI *et al.* (2008)
Gamma-H2AX peaks[15] (38)	Supp. Data 1 SZILARD *et al.* (2010)
Rrm3p pause sites[16] (6)	Supp. Table 7 in AZVOLINSKY *et al.* (2009).
Long intergenic regions[17] (17)	Saccharomyces Genome Database (SGD)
Replication-termination regions (TERs) (3)	Table S2 in FACHINETTI *et al.* (2010)
ARS elements flanked by opposing transcripts[18] (12)	Saccharomyces Genome Database (SGD)
Ty elements (8)	Saccharomyces Genome Database (SGD)

[14] Transcription levels were measured for 451 of the 565 ORFs on the right arm of chromosome IV by Nagalakshmi *et al* (2008). Most of these genes were single-copy sequences, since measuring transcription from repeated genes with very similar sequences is difficult. We identified the twenty single-copy genes in this region with the highest level of transcription.

[15] Peaks of accumulation of gamma-H2AX within the genome were mapped; such regions are often associated with DNA damage (Szilard *et al.*, 2010).

[16] Rrm3p is a helicase involved in promoting replication through replication-pause sites (Azvolinsky *et al.*, 2009).

[17] We identified the locations of the 17 longest intergenic regions of a total of 554 on the right arm of chromosome IV.

[18] ARS elements were identified in which the genes flanking the ARS have transcripts that converge on the ARS.

Table S 3.6 Chromosome elements that are under-represented in crossover-associated gene conversion tracts in JSC25.[19]

Element	Data[20]	Method of analysis	p-Value
Long terminal repeats	JSC25 G1	2	1.65E-03
Tandem repeats	JSC25 all	1	7.51E-05
Tandem repeats	JSC25 all	3	1.15E-06
Tandem repeats	JSC25 G1	1	2.39E-04
Tandem repeats	JSC25 G1	3	6.24E-06
Tandem repeats	W303a all	1	6.53E-03
Tandem repeats	W303a all	3	2.95E-03
Tandem repeats	YJM789 all	3	7.10E-04
Short intergenic regions	JSC25 all	1	1.12E-03
Short intergenic regions	JSC25 all	3	5.22E-05
Short intergenic regions	JSC25 G1	1	7.15E-03
Short intergenic regions	JSC25 G1	3	2.80E-04
Short intergenic regions	W303a all	1	4.87E-03
Short intergenic regions	W303a all	3	4.26E-04
Highly-transcribed genes	JSC25 all	3	7.69E-03
Highly-transcribed genes	JSC25 G1	3	5.17E-03

[19] As described in the text, we used four related methods of analysis to determine whether various structure/sequence chromosome motifs (listed in Supp. Table 5) were under-represented in gene conversion tracts associated with crossovers on chromosome IV.

[20] The data examined by Methods 1-4 were: JSC25 all (G1 and G2 events initiated on either W303a- and YJM789-derived homologs), JSC25 G1 (G1 events initiated on either W303a- and YJM789-derived homologs), JSC25 G2 (G2 events initiated on either W303a- and YJM789-derived homologs), W303a all (G1 and G2 events initiated on the W303-derived homologs), and YJM789 all (G1 and G2 events initiated on the YJM789-derived homologs).

Table S 3.7 Coordinates used to represent gene conversion tracts as single points.

For some of the analysis of hotspots, the conversion tracts were represented as to single points. The method for calculating the single point is described in the Supplementary Text. The description of each sectored colony is in Table S 3.3 and Table S 3.4.

Sector	Coordinate of point
1RW	1264086.5
3RW	1049553
4RW	1161545
5RW	1021788
6RW	1220821.5
7RW	1430143.5
9RW	1275710
10RW	1173059
11RW	1311425
12RW	1274790
13RW	1474129.5
14RW	1467578.5
15RW	1276501
16RW	1086141
18RW	1045794.5
19RW	1483931
20RW	1070415
21RW	1406135.5
23RW	1166479
24RW	1364357
25RW	1208881.5
26RW	1086141
27RW	583644.5
28RW	1365816
29RW	1315892
30RW	1442939.5
31RW	1127288
32RW	1077258.5
33RW	1414822

34RW	1218603.5
35RW	1256544
36RW	1127288
37RW	997617
38RW	527407.5
39RW	879006
40RW	993177
41RW	1203196.5
42RW	979126
43RW	1216821.5
44RW	1249669.5
45RW	1426698
46RW	890618
48RW	777210
49RW	988352
50RW	1046711
51RW	473796
52RW	914018.5
53RW	1166479
54RW	883894
55RW	638593
56RW	1377052.5
57RW	1065969
58RW	1481855.5
59RW	1186507.5
60RW	772859.5
61RW	1154424
63RW	1166479
64RW	944096.5
65RW	1382818
66RW	1164200
67RW	998309
68RW	1467370.5
69RW	1220580.5
70RW	844344.5
71RW	1508254.5
73RW	543540.5
74RW	746615.5
75RW	1348229.5
76RW	495183

77RW	1220821.5
78RW	732715.5
79RW	687626
80RW	1275539.5
81RW	1246949.5
82RW	568072.5
83RW	1392166
84RW	1283368
85RW	1073850
86RW	451973
87RW	714744.5
88RW	1022402
89RW	475967
91RW	583997
92RW	1267136
93RW	1267007.5
94RW	1127288
95RW	670304
96RW	703439.5
97RW	879508
98RW	544830
99RW	890618
100RW	753241
101RW	1504834.5
102RW	595270
103RW	976310
104RW	891029.5
105RW	464985
106RW	501889
107RW	992208
108RW	794424.5
109RW	495183
110RW	470241
111RW	867090.5
112RW	685652.5
113RW	667128
114RW	566291.5
115RW	859497.5
116RW	810917
117RW	998623.5

118RW	799135.5
119RW	518665.5
120RW	758447
121RW	709348.5
122RW	799135.5
124RW	1009467
128RW	931179.5
129RW	767039.5
130RW	1479818
131RW	1005808
133RW	596258.5
136RW	609213
139RW	1013924
140RW	684156.5
142RW	586999
143RW	1272437
144RW	1217635
145RW	649183.5
146RW	929384.5
147RW	854299.5
149RW	713167
150RW	484323.5
153RW	498116
154RW	504551.5
156RW	488747
72-1RW	792862.5
72-2RW	1182759

4 Hydroxyurea-induced crossovers

4.1 Introduction

In Chapter 3, I showed that there are at least seven spontaneous recombination hotspots along the right arm of the yeast chromosome IV. I demonstrated that two of these hotspots (HS3 and HS4) were likely to reflect the processing of a hairpin secondary structure formed by inverted Ty elements. The mechanisms of the other hotspots, however, remain unknown. Many proposed mechanisms for producing recombinogenic DNA lesions are related to DNA replication. It has been suggested that DNA damage, secondary DNA structures or DNA-bound proteins might block replication fork progression (AGUILERA and GOMEZ-GONZALEZ 2008), and that the stalled replication forks might be processed to yield single-stranded DNA (ssDNA) gaps or double-stranded DNA breaks (DSBs) (COBB *et al.* 2005; SOGO *et al.* 2002). If some of the hotspots observed for spontaneous mitotic recombination are a consequence of problems related to DNA replication, one would expect that these hotspots might become even more pronounced in conditions of perturbed DNA replication. Consequently, I mapped mitotic crossovers in yeast strains treated with a drug (hydroxyurea) that reduces the rate of DNA replication.

Hydroxyurea (HU) inhibits replication in yeast and mammalian cells by reducing the activity of ribonucleotide reductase (RNR), the enzyme that converts ribonucleotides to deoxyribonucleotides (dNTPs) (KRAKOFF *et al.* 1968; REICHARD 1988; SLATER 1973; TIMSON 1975). The concentration of dNTPs is highest in the G_1/S transition and throughout the S-phase (CHABES *et al.* 2003; KOC *et al.* 2004), and the pools of dNTPs are regulated by ribonucleotide reductase activity (ELLEDGE *et al.* 1992). Consequently, HU treatment results in diminished pools of dNTPs (CHABES *et al.* 2003; KOC *et al.* 2004), and slow-moving replication forks (ALVINO *et al.* 2007; SOGO *et al.* 2002). HU treatment also induces mitotic recombination in yeast (BARBERA and PETES 2006; FERGUSON 1990; GALLI and SCHIESTL 1996), presumably because HU directly or indirectly elevates the frequency of DNA damage.

HU is used for treating several disorders including HIV infection (LORI *et al.* 1994) and sickle cell anemia (MCGANN and WARE 2011). In HIV-positive patients, HU treatment inhibits viral DNA synthesis (LORI *et al.* 1994). In sickle cell anemia patients, HU treatment increases production of fetal hemoglobin; the mechanism for this effect is unknown but is apparently not directly related to the replication-inhibiting activities of HU (MCGANN and WARE 2011; PLATT *et al.* 1984). Although the recombination-inducing activity of HU could result in secondary problems related to treatment, elevated

frequencies of cancers have not been observed in sickle cell anemia patients treated with HU (STEINBERG *et al.* 2010).

In this Chapter, I describe the mapping of HU-induced crossovers on the right arm of chromosome IV. As with the spontaneous crossovers, I found that HU-induced crossovers occur throughout the chromosome arm. Unexpectedly, most of the HU-stimulated crossovers were associated with patterns of gene conversion indicative of G1-induced DSBs, although conversion events representative of S/G2 initiated DSBs were also observed. The crossovers initiated in G2 were strongly correlated with the locations of solo delta elements, the long terminal repeats (LTRs) associated with the Ty retrotransposons.

4.2 Materials and Methods

4.2.1 Strains

The diploid strain JSC25 (*Mata/MATα::HYG leu2-3,112/LEU2 his3-11,15/HIS3 ura3-1/ura3 GAL2/gal2 ade2-1/ade2-1 trp1-1/TRP can1-100::NAT/CAN1Δ::NAT RAD5/RAD5 IV1510386::KANMX-can1-100/IVI1510386::SUP4-o*) was used for all experiments in this chapter. The construction of this strain is described in the Supplementary Materials and Methods of Chapter 3.

4.2.2 Hydroxyurea (HU) treatment

Hydroxyurea (HU) was used to induce crossovers. The diploid JSC25 was inoculated into five ml of liquid rich growth medium and incubated at 30° overnight. The following day, cells were resuspended at a concentration of 3 x 10^8 cells/ml in fresh YPD medium containing 100 mM HU. The culture was grown for six hours at 30° on a rotating drum. Following this incubation, the cells were harvested by centrifugation, and washed twice with water. The cells were then plated on SD-arg medium, and we assayed sectored colonies as described in Chapter 3 Materials and Methods.

4.2.3 Microarray methods and analysis

Microarrays were prepared as previously described in Chapters 2 and 3. Briefly, genomic DNA was isolated (separately) from both red and white sectors and sonicated to fragments of about 200-400 bp; these samples were labeled with Cy5-dUTP. The control DNA was isolated from the JSC25 strain, and labeled with Cy3-dUTP. Labeled experimental and reference samples were then mixed and competitively hybridized on SNP microarrays (whole-genome microarrays previously described in Chapter 2 and chromosome IV microarrays described in Chapter 3) for one to two days at 62°. After this incubation, the microarrays were washed and scanned at 635 nm and 532 nm for the experimental and reference samples, respectively. The ratio of

experimental to reference signal was obtained for each oligonucleotide on the microarray, and used for subsequent analysis as described in Chapters 2 and 3.

4.2.4 Statistical analysis

The chi-square Goodness-of-Fit function of Microsoft Excel (chisq.test) was used to compare expected and observed values. Additionally, the chi-square Test-of-Independence function of Microsoft Excel (chisq.dist.rt) was used to compare different datasets. The Mann-Whitney test from VassarStats website was used to compare the conversion tract lengths of the spontaneous and HU-induced reciprocal crossovers. Table B11 of ALTMAN (1991) was used to calculate the 95% confidence intervals of median conversion tract lengths.

4.3 Results

4.3.1 Crossovers induced by treatment of cells with HU

As described in Chapter 3, we developed a method of detecting and mapping crossovers on the right arm of chromosome IV. This method utilizes the diploid strain JSC25 that is homozygous for the ochre-suppressible *ade2-1* mutation; in the absence of a suppressor, strains with the *ade2-1* mutation form red colonies (Figure 3.2). In JSC25, the ochre suppressor *SUP4-o* is inserted near the telomere of chromosome IV on one homolog and partially suppresses the *ade2-1* mutation, resulting in pink colonies. Following a crossover between *CEN4* and *SUP4-o*, half of the segregation events result in

one daughter cell that is homozygous for the *SUP4-o* insertion (white colony), and one daughter cell that has no copies of the suppressor (red colony). Thus, red-white sectored colonies signal a reciprocal crossover on chromosome IV. Additionally, JSC25 is a hybrid of two sequenced haploid strains, W303a and YJM789, that have approximately 55,000 SNPs relative to each other (WEI *et al.* 2007). These SNPs can be used to map the location of crossovers and conversion tracts (LEE *et al.* 2009; LEE and PETES 2010; ST CHARLES *et al.* 2012).

Asynchronous JSC25 cells were treated for six hours at 30° with 100 mM of HU in a liquid YPD culture, and then plated on medium lacking HU. If the HU-induced crossover occurs in the first cell division on the solid medium following treatment of the cells with HU, a red/white sectored colony will be formed. The average sectoring rate following HU treatment was 8.3×10^{-4}/division, which is about a 26-fold increase over the spontaneous sectoring rate for the same strain (3.1×10^{-5}/division, Chapter 3). Since only 50% of crossovers result in loss of homozygosity (LOH) in the daughter cells, the reciprocal crossover rate induced by the HU treatment is 1.7×10^{-3}/division.

We mapped a total of 52 HU-induced crossovers using the chromosome IV SNP microarrays (Chapter 3). In brief, we examined about 3000 heterozygous SNPs on the 1.1 Mb right arm of chromosome IV using SNP microarrays. This analysis allows us to distinguish regions of heterozygosity from homozygosity, mapping LOH transitions to a

resolution of about 400 bp. This information is used to map crossovers, and to determine that the nature of the conversion tracts associated with each crossover. Schematic depictions of the HU-induced events and the SGD coordinates of the crossovers and conversions are shown in Table S 4.1 and Table S 4.2, respectively. These tables are in the same format as for the spontaneous crossovers described in Chapter 3 (Table S 3.3 and Table S 3.4). Table S 4.1 shows the conversion tracts associated with each crossover using pairs of lines, in which the upper line indicates regions of heterozygosity and homozygosity in the red sector and the lower line shows these patterns in the white sector. Line segments are colored red, black, and green to indicate homozygosity for W303a-derived SNPs, homozygosity for YJM789-derived SNPs, and heterozygosity, respectively. Transitions between heterozygosity and homozygosity are labeled with lower case letters in Table S 4.1, and the SGD coordinates for each transition are presented in Table S 4.2.

The locations of the crossovers and the lengths of the conversion tracts associated with crossovers are summarized in Figure 4.1. In this figure, the conversion tracts that were initiated on the W303a-derived and YJM789-derived chromosomes are shown as vertical lines on the upper and lower parts of the figure, respectively. Our interpretation as to which homolog initiated the conversion event is based on determining the donor of information in the conversion event. Chromosomes with DSBs receive information

during conversion with the donated sequences being derived from the intact chromosome (PÂQUES and HABER 1999). Thus, if one sector has become homozygous for the W303a-derived SNPs and the other sector is still heterozygous, we infer that the YJM789 chromosome had the initiating DNA lesion. By this criterion, we could determine which homolog had the initiating DNA lesion with two exceptions. One sectored colony (not shown in Figure 4.1) had one conversion event that duplicated W303a-derived sequences and one conversion event that duplicated YJM789-derived sequences. In addition, of the 52 sectored colonies examined, 12 had no associated conversion and, therefore, we could not determine which chromosome had the initiating DNA lesion.

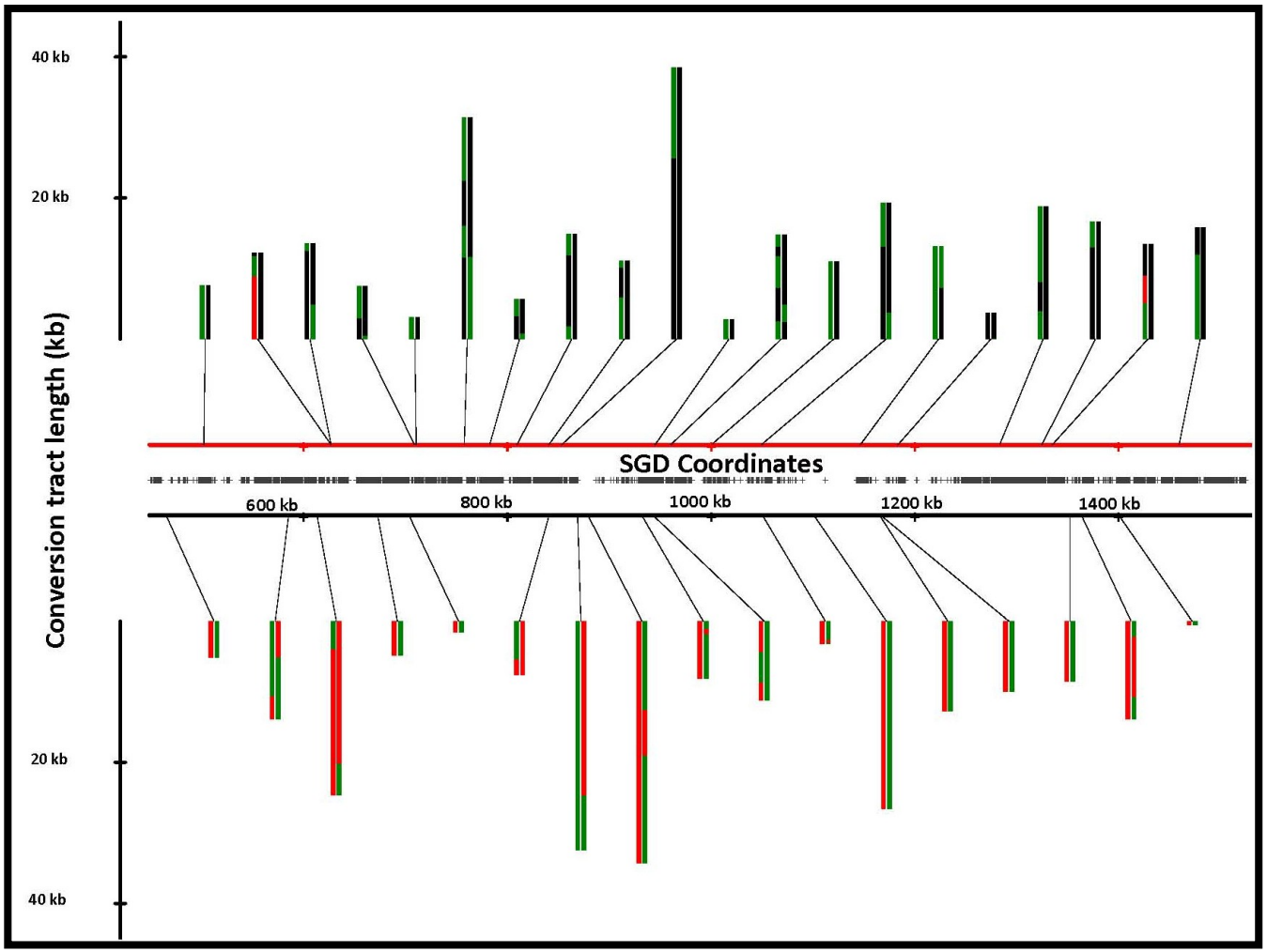


Figure 4.1. Summary of crossovers induced by HU treatment.

This figure shows crossovers with associated conversion tracts on chromosome IV in cells treated with HU. The X-axis indicates the SGD coordinates of the right arm of chromosome IV. The Y-axis indicates the length of each conversion tract. The red and black horizontal lines represent the W303a and the YJM789 homologs, respectively. The line with ticks between the homolog depictions shows the probe density along the chromosome arm, with each tick indicating a set of probes for a single SNP. Crossovers that were initiated on the W303a and YJM789 homologs are depicted above and below the horizontal lines, respectively. The vertical parallel red, green, and black line-segments depict the conversion tracts that were associated with each crossover. The thin black line that connects each conversion tract to the homolog indicates the site of the start of a conversion tract (the site at which the black line meets the homolog line). Red, black, and green in the conversion tract lines depict homozygosity for W303a, homozygosity for YJM789, and heterozygosity, respectively. In each conversion tract line-pair, the left line represents the conversion tract in the red side of the sector and the right line represents the conversion tract in the white side of the sector.

Based on Figure 4.1, it is apparent that the distribution of recombination events between the two homologs is approximately even, demonstrating that neither homolog is more likely to receive a recombination-inducing lesion as a result of HU treatment. Of the 37 sectored colonies for which we could assign a donor in the conversion event, 20 were initiated on the W303a homolog and 17 were initiated on the YJM789 homolog. Additionally, the events are evenly distributed along the chromosome arm, without strong recombination hotspots on either homolog.

Figure 4.2 shows the number of times that individual SNPs on chromosome IV were involved in conversion tracts in the analyzed samples. The average SNP was involved less than one time (0.4 times), and the greatest number of times a single SNP was included in a conversion tract was three. As in Figure 4.1, there appear to be no

strong hotspots for the HU-induced recombination events. Using the criterion described above, we also separately examined crossovers initiated on the W303a-derived homolog and YJM789-derived homolog (Figure 4.2B). Although there appears to be little overlap in the recombination events initiated on the two homologs, the limited dataset do not allow us to determine whether the patterns of crossovers are different on the two homologs. We also examined the distribution of G1- and G2-initiated crossovers (Figure 4.2C). As discussed in Chapter 3 and in previous papers (LEE *et al.* 2009; LEE and PETES 2010; ST CHARLES *et al.* 2012), about two-thirds of spontaneous crossover events are likely to be initiated in G1. As illustrated in Figure 3.1, a conversion event and crossover initiated by a DSB in S or G2 will produce a 3:1 conversion event (one sector homozygous for one form of the SNP and the other sector heterozygous). In contrast, a conversion event initiated by a DSB in G1 (Figure 3.1B and 3.1C) will produce a 4:0 conversion tract (both sectors homozygous for the same SNP) or a hybrid 3:1/4:0 tract. In Figure 4.2C, we show the distribution of G1- and G2-initiated recombination events separately. From this analysis, it is clear the events can be initiated in either G1 or G2.

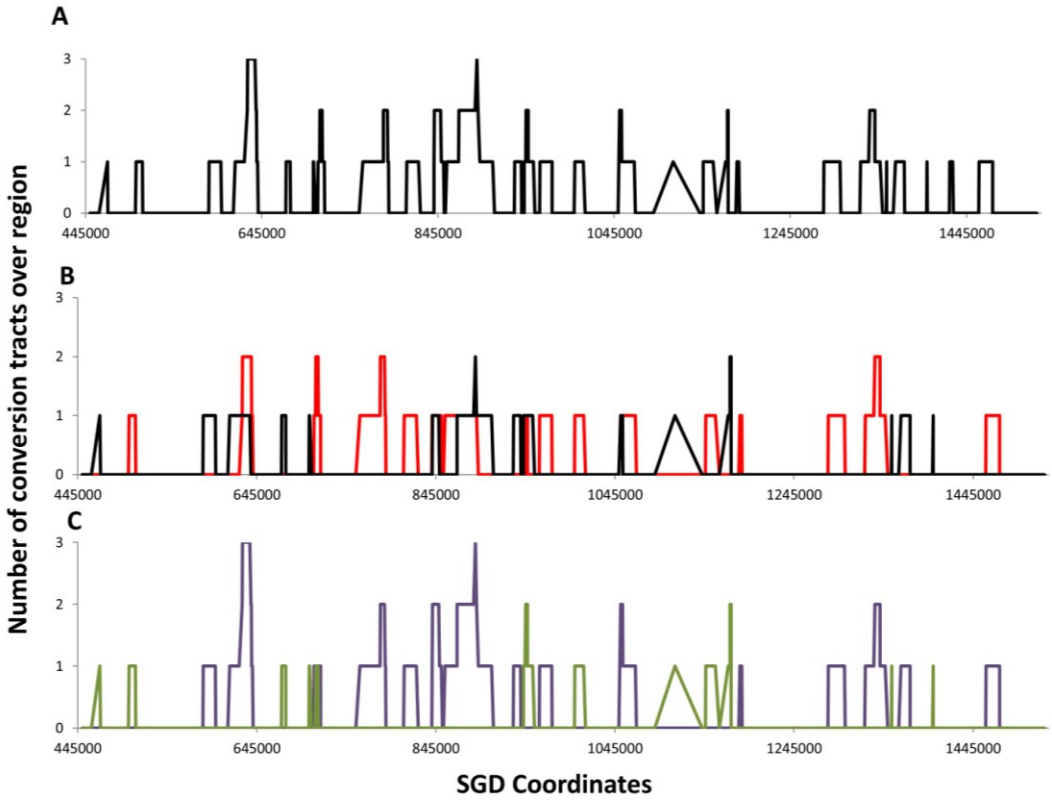


Figure 4.2. Summary of conversion tracts associated with HU-induced crossovers along the right arm of chromosome IV.

As with Figure 4.1, the X-axis represents the SGD coordinates of the right arm of chromosome IV. The Y-axis depicts the number of times each SNP represented on the microarray was involved in a conversion tract.

A. All HU-induced conversions.

B. HU-induced conversions initiated on the W303a-derived (red line) or YJM789-derived (black line) homolog.

C. HU-induced conversions initiated in G1 (purple line) or G2 (green line). Our interpretation of G1 versus G2 initiated events is described in the text.

In order to address the question of whether there were any hotspots for HU-induced reciprocal crossovers by a statistical analysis, we divided the chromosome IV

right arm into six quadrants and assigned each conversion/crossover a single representative point. This representative point was located in the center of the region that was likely to contain the recombination-inducing lesion, as described in the Supp. Materials and Methods of Chapter 3. Table S 4.3 lists each of these points for the HU-induced sectors. The recombination events were further classified by the homolog on which the event was initiated, and the timing of the recombination-initiating lesion. This analysis is shown in Figure 4.3. This analysis, unlike those depicted in Figure 4.1 and 4.2, includes crossovers that were unassociated with conversion tracts in addition to those with conversion tracts.

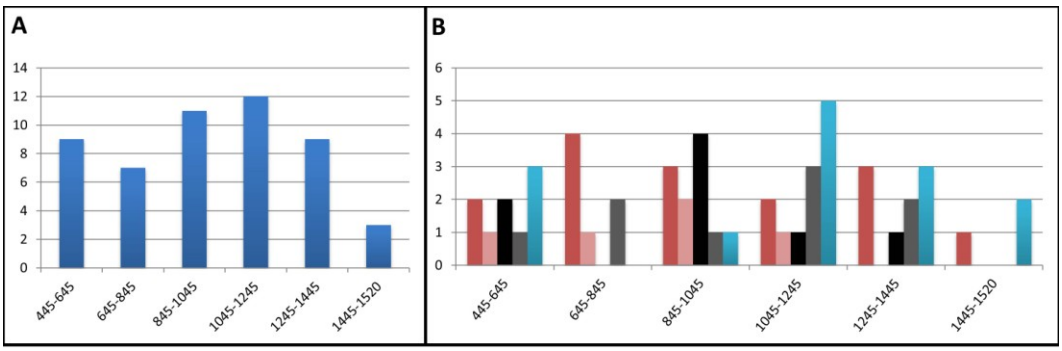


Figure 4.3. Number of HU-induced crossovers located in bins of the right arm of chromosome IV.

In this figure, the Y-axis shows the number of recombination events per bin, and the X-axis indicates the SGD coordinates that define the bin. *CEN4* is located at approximately SGD coordinate 445 kb and the markers used to detect crossovers are located at SGD coordinate 1510 kb.

A. The number of all recombination events within each bin.

B. Numbers of crossovers initiated on the W303a or YJM789 homologs, and initiated in G1 or G2 of the cell cycle within each bin. The colors indicating the various types of events are: red (W303a, G1), pink (W303a, G2), black (YJM789, G1), gray (YJM789, G2), and blue (crossover without an associated conversion).

The distribution of events within the six bins (Figure 4.3A) is not significantly different from that expected by a random distribution (p-value 0.89 by the chi-square goodness-of-fit test). Similarly, the distribution of events initiated on the W303a and YJM789 homologs, and the distribution of G1-initiated and G2-initiated events are not significantly different from a random distribution.

Most of the sectors were examined using SNP arrays that were specific to chromosome IV. In order to check for unselected recombination events induced by HU on other chromosomes, we analyzed two HU-induced sectors using whole-genome SNP microarrays (ST CHARLES *et al.* 2012). These sectored colonies had no additional regions of LOH.

4.3.2 Comparison of HU-induced sectors and spontaneously occurring sectors

We initially expected that the HU treatment would exacerbate a sub-set of the spontaneous recombination hotspots (Chapter 3). Figure 4.4 compares the distribution of HU-induced recombination events with the distribution of spontaneous events. Seven hotspots were detected in the spontaneous data set (Chapter 3, HS1-HS7 in Figure 4.4B). With the possible exception of HS3, none of these hotspots appear to be present in the

HU-induced events. HS3 represents an inverted pair of Ty elements located about 500 bp apart (Chapter 3). While this hotspot appears to be shared between spontaneous and HU-induced events, the HU-induced events are different from the spontaneous events. The spontaneous events are initiated on the W303a-derived homolog, but not the YJM789-derived homolog (Chapter 3); this specificity reflects the presence of the inverted pair Ty elements on the W303a-derived homolog and the absence of these elements on the YJM789-derived homolog. In contrast, the HU-induced events near HS3 appear to initiate on both homologs (Figure 4.2C). A more detailed comparison of the properties of spontaneous and HU-induced events would necessitate the accumulation of more HU-induced events, since we examined about 140 spontaneous events and only 52 HU-induced events.

Both HS3 and HS4 hotspots are associated with inverted repeats separated by short spacer sequences; such structures are likely to form intrastrand "hairpins" (Chapter 3). Since such hairpins would more readily form if the DNA containing inverted repeats was single-stranded (NAG and PETES 1991), and since HU treatment causes the accumulation of single-stranded DNA (Alvino *et al.* 2007; Sogo *et al.* 2002), it is surprising that we did not detect strong hotspots at the positions of HS3 and HS4 in HU-treated cells. Possible rationales for this observation will be outlined in the Discussion.

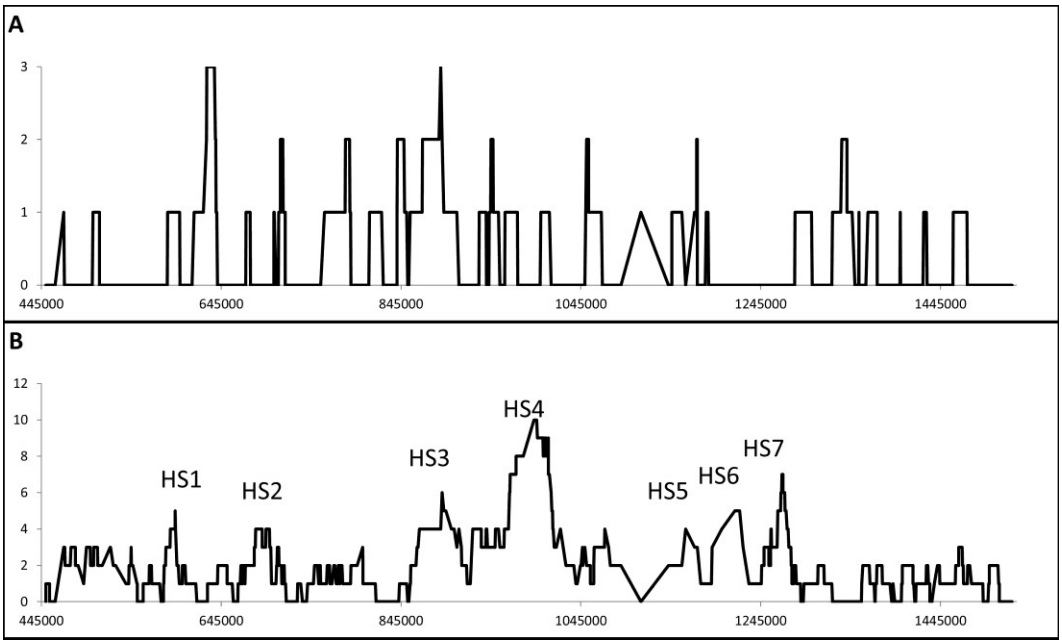


Figure 4.4. Comparison of the location of recombination events for HU-induced and spontaneous crossovers.

The X-axis represents the SGD coordinates and the Y-axis represents the number of times a SNP along the right arm of chromosome IV was involved in a conversion tract.

A. HU-induced crossovers.

B. Spontaneous crossovers. The hotspots that were discussed in Chapter 3 are labeled.

4.3.3 Conversion tracts associated with HU-induced crossovers

Since HU affects the rate of DNA replication, we expected that most of the HU-induced conversion events would have the characteristics of an S/G2-initiated event rather than a G1-initiated event. However, of the crossovers associated with conversion tracts, only 37% (14/37) had conversion tracts consistent with a G2 DSB. In

comparison, 25% (31/121) of the spontaneous crossovers had conversion tracts consistent with a G2 DSB. This difference was not statistically significant (p-value 0.18 by the chi-square Test of Independence). As observed for the spontaneous events, most (23 out of 37) of the HU-induced gene conversions were interpretable as reflecting a G1-initiated event. This observation implies that either the HU treatment induces DSBs in G1 in addition to DSBs in S/G2, or that chromosomes with DSBs initiated by HU treatment in G2 can be segregated into daughter cells without repair. These mechanisms will be discussed further below.

A pie chart depiction of spontaneous and HU-induced crossovers is given in Figure 4.5. The main differences in the two distributions are that the spontaneous events have a smaller proportion of crossovers unassociated with conversion (13% compared to 27%, Figure 4.5) and a greater proportion of complex events (28% versus 17%). The increase of crossovers unassociated with conversion tracts is statistically significant (p-value 0.02 by the chi-square test of independence).

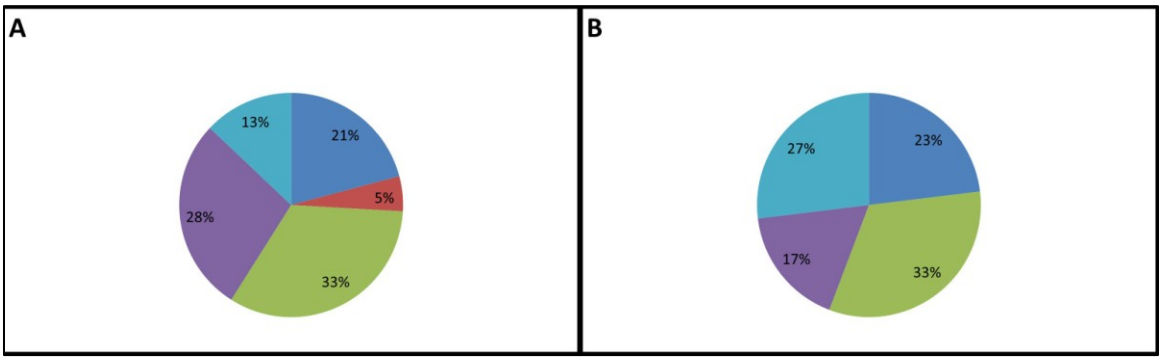


Figure 4.5. Comparison of conversion tract types between the spontaneous crossovers and HU-induced crossovers.

4:0, hybrid tracts, and most complex conversions represent G1-initiated events as discussed in Chapter 3. The colors that represent the various classes of events are: dark blue (3:1 conversion), light blue (crossovers unassociated with conversions), red (4:0 conversions), green (3:1/4:0 or 3:1/4:0/3:1 simple hybrid tracts), and purple (complex conversion events).

A. Spontaneous reciprocal crossovers.

B. HU-induced reciprocal crossovers.

We also compared the conversion tract lengths associated with spontaneous and HU-induced crossovers (Figure 4.6). Conversion tract lengths were calculated as described in Chapter 3. The median conversion tract length (95% confidence intervals in parentheses) associated with the HU-induced crossovers was 11 (7.5-13.5) kb. The median spontaneous conversion tract length was 16.7 (8.2 – 13.6) kb (Chapter 3). These lengths are not statistically different (p-value 0.28 by the Mann-Whitney test). The median conversion tract lengths of the HU-induced G1- and G2-initiated crossovers are 13.9 (11.1 – 18.8) kb and 8.0 (2.8 – 12.8) kb, which are significantly different from each

other (p-value 0.002 by the Mann-Whitney test). We observed similar differences between G1- and G2-initiated spontaneous events, in which the tract lengths were 14.8 (11.7 – 17.5) kb and 4.7 (2.6 – 9.6) kb, for the G1- and G2-induced events, respectively. No significant differences were observed between the G1-initiated conversion tract lengths for HU-induced and spontaneous events (p-value 0.48 by the Mann-Whitney test). Similarly, the G2-initiated conversion tract lengths for HU-induced events were not significantly different from the G2-initiated tract lengths for spontaneous events (p-value 0.18 by the Mann-Whitney test). The very long conversion tracts observed in the bin between coordinates 845-1045 kb (Figure 4.6) for the spontaneous events are those associated with HS4, as discussed in Chapter 3. These very long tracts were not evident for HU-induced events.

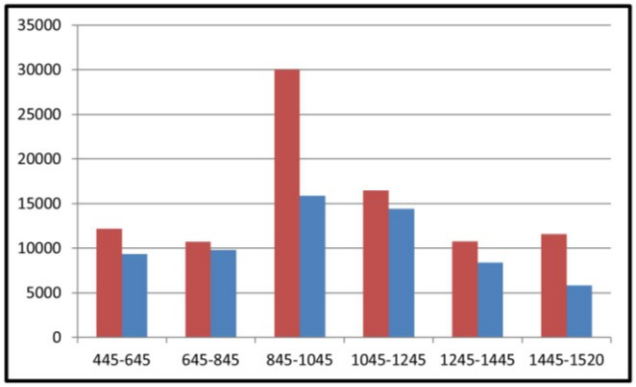


Figure 4.6. Conversion tract lengths associated with spontaneous and HU-induced crossovers.

The X-axis shows the SGD coordinates of different bins on the right arm of chromosome IV; the Y-axis shows the average conversion tract length of events within the bin. The red bars indicate conversion tract lengths associated with spontaneous

crossovers, and the blue bars indicate conversion tract lengths of HU-induced crossovers.

4.3.4 Association of HU-induced crossovers with genomic elements

Previous studies showed that the site of the DSB that initiates mitotic recombination is located near or within the associated conversion tract (PÂQUES and HABER 1999). For example, (TANG *et al.* 2011) found that almost all crossovers stimulated by a tract of 230 GAA/TTC repeats included the repeat within the conversion tract (TANG *et al.* 2011). Therefore, we performed statistical tests to determine whether various chromosome elements (replication origins, tRNA genes, sites of Rrm3p accumulation, etc.) were over-represented within windows that included the HU-induced conversion tracts; as described below, a variety of methods were used to define these windows. We examined a total of 18 chromosome elements (listed in Table S 3.5). This analysis was performed as described in the Results and Supplementary Materials and Methods of Chapter 3.

As with the analysis performed with the spontaneous events (Chapter 3), we used four different methods of calculating the conversion tract windows for the HU-induced events: Method 1 (all conversion tracts were summed), Method 2 (all conversion tracts less than 20 kb were summed), Method 3 (the conversion tracts most likely to containing the initiating DSBs [Supp. Materials and Methods of Chapter 3] were summed), and Method 4 (the conversion tracts most likely to containing the initiating

DSBs that were less than 20 kb were summed). These methods were used to investigate five different datasets: all conversion-associated crossovers, G1-initiated crossovers initiated on both homologs, G2-initiated crossovers initiated on both homologs, G1- and G2-initiated events on the W303a-derived homolog, and G1- and G2-initiated events on the YJM789-derived homolog. All elements with significant enrichments are listed in Table 4.1.

The most striking and unexpected association was with long terminal repeats (LTRs), both associated and unassociated with Ty elements. Early-to-mid firing ARS elements were also significantly enriched in the G2 conversion tracts, which is not unexpected since HU affects replication rates. Additionally, triplet repeats and G4 DNA motifs were significantly enriched in several of the datasets. The significance of these associations will be discussed below.

Table 4.1 Element enrichment in HU-induced crossover-associated conversion tracts.[1]

Element	Data[2]	Method of analysis	p-Value
LTRs (deltas) unassociated with Ty elements	JSC25	4	0.00014
LTRs (deltas) unassociated with Ty elements	JSC25 G2	1	4.38×10^{-11}
LTRs (deltas) unassociated with Ty elements	JSC25 G2	2	6.51×10^{-14}
LTRs (deltas) unassociated with Ty elements	JSC25 G2	3	5.72×10^{-12}
LTRs (deltas) unassociated with Ty elements	JSC25 G2	4	8.99×10^{-16}
LTRs (deltas) unassociated with Ty elements	YJM789	2	1.48×10^{-5}
LTRs (deltas) unassociated with Ty elements	YJM789	3	0.0001
LTRs (deltas) unassociated with Ty elements	YJM789	4	7.44×10^{-6}
Early-to-mid firing ARS elements	JSC25 G2	1	0.0012
Early-to-mid firing ARS elements	JSC25 G2	2	0.0029
Early-to-mid firing ARS elements	JSC25 G2	3	0.0083
LTRs (deltas)[3]	JSC25 G2	1	0.0024
LTRs (deltas)[3]	JSC25 G2	2	0.0019
LTRs (deltas)[3]	JSC25 G2	3	0.001
LTRs (deltas)[3]	JSC25 G2	4	0.00049
Triplet repeats	JSC25 G1	4	0.0014
Triplet repeats	YJM789	1	0.00038
G4 motifs	YJM789	1	0.0024

[1] As described in the text, we used four related methods of analysis to determine whether various structure/sequence chromosome motifs (listed in Supplementary Table 3.5) were over-represented in gene conversion tracts associated with crossovers on chromosome IV.

[2] The data examined by Methods 1-4 were: JSC25 all (G1 and G2 events initiated on either W303a- and YJM789-derived homologs), JSC25 G1 (G1 events initiated on either W303a- and YJM789-derived homologs), JSC25 G2 (G2 events initiated on either W303a- and YJM789-derived homologs), W303a all (G1 and G2 events initiated on the W303-derived homologs), and YJM789 all (G1 and G2 events initiated on the YJM789-derived homologs).

[3] Includes delta elements both associated and unassociated with Ty elements.

4.4 Discussion

As described in Chapter 3, we detected seven hotspots for spontaneous recombination on the right arm of chromosome IV. The mechanism of the HS4 hotspot is related to its ability to form a hairpin secondary structure; it is likely that the hotspot activity of HS3 reflects a similar mechanism. The mechanism(s) of the other hotspots, however, remain unexplained. We noted that two of these hotspots, HS5 and HS6, contain ARS elements that could contribute to their instability. In addition, HS6 (Figure 3.7) co-localizes with several highly-expressed *HXT* genes. One of these genes, *HXT6*, is located on the Crick strand centromere-distal to the nearest early-firing ARS element, *ARS432*. This orientation of the gene and the ARS element could cause a head-on collision between the transcription and replication machineries, one of the suggested sources of recombination-inducing lesions (AGUILERA and GOMEZ-GONZALEZ 2008). Based on the observation that HU treatment reduces the rate of replication, we expected that HU treatment might increase the formation of recombinogenic lesions at one of our previously defined hotspots for spontaneous events. We also expected that HU treatment would produce DSBs in S-phase and, therefore, most of the conversion events associated with HU-induced crossovers would represent 3:1 conversions. Neither of these expectations were fulfilled.

4.4.1 Sectoring rate in the presence of HU

The rate of sectoring in JSC25 cells treated with 100 mM HU was 8.3×10^{-4}, a 26-fold increase over the spontaneous rate of sectoring (3.1×10^{-5}, Chapter 3). In comparison, BARBERA and PETES (2006) detected a 40-fold increase of reciprocal crossovers in cells grown on plates containing 100 mM HU, and GALLI and SCHIESTL (1996) detected a 9-18-fold increase in recombination in response to transient treatment with 150 mM HU. Thus, the degree of stimulation of recombination by HU is similar in all experiments with the minor differences likely reflecting differences in the assays. For example, GALLI and SCHIESTL (1996) measured intrachromosomal deletions and heteroallelic recombination rather than crossovers between homologs.

4.4.2 Comparison of the reciprocal crossover distribution of spontaneous and HU-induced sectors

None of the seven spontaneous recombination hotspots defined in our previous study were strong hotspots following HU treatment. One interpretation of this result is that none of the spontaneous recombination hotspots are generated by replication-dependent mechanisms. This conclusion is consistent with our observation that most of the spontaneous crossovers appear to be induced by a DSB in unreplicated DNA. It is important to emphasize that our study does not directly address the number of DSBs that occur in G1 versus S or G2. We suggest (Figure 3.17) that spontaneous DSBs are likely to be formed more frequently in the S period than in G1. The majority of these

DSBs, however, are repaired using a sister chromatid as a substrate (KADYK and HARTWELL 1992) which would be undetectable by our system. In contrast, if a broken G1 chromosome is replicated, it will generate two chromatids that are broken at the same position; consequently, the only intact substrate for repair is the homologous chromosome.

4.4.3 Unexpected presence of G1-initiated recombination events in HU-treated cells

Based on the perturbed DNA replication in cells incubated in HU (ALVINO *et al.* 2007), we expected an elevated frequency of crossovers associated with 3:1 gene conversion events indicative of DSB formation in S or G2. We observed, however, that most of the crossovers were associated with 4:0 or hybrid gene conversion tracts, suggestive of DSB formation during G1. Below, we will discuss various explanations of this unexpected result.

One possibility is that HU induces DSBs during the S-period, but also prevents the efficient repair of the broken chromatid, possibly because of the low level of dNTPs resulting from inhibition of ribonucleotide reductase (RNR) by HU (CHABES *et al.* 2003; ELLEDGE *et al.* 1992). Inhibition of RNR, either through mutations or HU-treatment, has been shown to impair HR (COVO *et al.* 2012; MOSS *et al.* 2010). Additionally, elevation of RNR activity has been shown to increase HR activity (BURKHALTER *et al.* 2009; MOSS *et al.* 2010). Consequently, if DSBs generated during the S-period fail to get repaired in the

same cell cycle, the broken chromosome could be segregated into one of the daughter cells. Replication of the broken chromatid, followed by its repair during the subsequent cell cycle would produce a crossover associated with a G1-like conversion pattern (4:0).

Another possibility is that HU induces recombinogenic DNA damage by mechanisms that are unrelated to the effect of this drug on DNA replication. KUONG and KUZMINOV (2009) reported that solutions of HU accumulated HCN and NO if they were more than one day old. They also concluded that HU could damage DNA directly. DAVIES *et al.* (2009) found that HU treatment of *E. coli* resulted in elevated levels of hydroxyl radicals that contributed to cell death. Although similar studies have not been done in yeast, it is possible that HU treatment induces recombinogenic DNA damage in yeast in G1.

Finally, it is possible that the HU stimulates a type of replication-associated DNA lesion that results in two broken chromatids. One model is that a stalled replication fork sometimes causes a DSB in the unreplicated portion of the fork; continued replication would generate two broken sister chromatids. An alternative possibility is that a DSB in one chromatid could result in formation of a Holliday junction with the sister chromatid. If this junction is misprocessed by cutting the Holliday junction in two planes (cleaving all four strands), two broken sister chromatids would be formed. In the last chapter of the thesis, I will discuss testing some of these models.

4.4.4 Enrichments of genomic elements within the HU-induced conversion tracts

We examined the conversion tracts associated with HU-induced crossovers to discern whether particular sequence or structural motifs are particularly susceptible to HU-induced DNA damage. Significant associations are listed in Table 4.1. Significant associations were common for the S-/G2-initiated events. For example, early-to-mid-firing ARS elements (FACHINETTI *et al.* 2010) were significantly associated with S/G2-initiated events. This association could be rationalized if HU treatment specifically stalls replication forks that have just begun replication.

One of the most unexpected results in this study was the association of long terminal repeats (LTRs) with crossover-associated conversions (Table 4.1). The most common LTR in yeast is the 330 bp delta elements. There are several hundred delta elements in the genome; about 80 deltas are associated with Ty elements and the remainder are "solo" deltas (KIM *et al.* 1998). We observed a significant enrichment in the conversion tracts for both LTRs that are associated and unassociated with Ty elements. LTRs were specifically enriched in G2 conversion tracts, indicating that the HU-induced recombinogenic lesions were formed during DNA replication or G2. HU has been previously shown to increase the mobility of Ty retrotransposons in yeast (CURCIO *et al.* 2007; O'DONNELL *et al.* 2010). Ty elements have also been shown to be associated with DSBs following HU treatment in *mec1* cells (FENG *et al.* 2011). One possibility is that HU

treatment elevates transcription of Ty elements and solo deltas. By this mechanism, the observed association of LTRs with HU-induced recombination could occur by one of the transcription-associated recombination (TAR) pathways discussed in the Introduction (Chapter 1). Alternatively, HU may (directly or indirectly) stimulate the binding of one or more proteins to the Ty and delta elements. This binding may result in blocked DNA replication forks under conditions in which the replication forks are moving slowly.

The connection between the location of HU-induced recombination events and LTRs is interesting and important for a variety of reasons. First, since it is likely that transient episodes of perturbed DNA replication are common, this connection may be relevant to genome evolution in yeast. It should be noted that DSBs within repeated elements can be repaired by allelic interactions or by interactions with non-allelic elements; the latter type of repair results in deletions, duplications, and translocations. Second, since HU is used in treating various diseases (for example, chronic myelogenous leukemia and sickle cell anemia) and since the human genome contains many LTRs, HU treatment may stimulate chromosome recombination events and chromosome aberrations as a secondary consequence of the HU treatment.

4.5 Summary

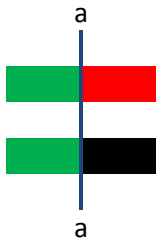
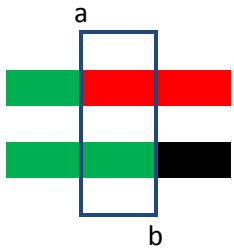
We examined reciprocal crossovers induced by HU treatment. The distribution of HU-induced events showed no strong hotspots. Unexpectedly, most of

the HU-induced crossovers had conversion patterns indicative of initiation by a G1-induced DSB. We saw a very significant association of LTRs with HU-induced conversion tracts.

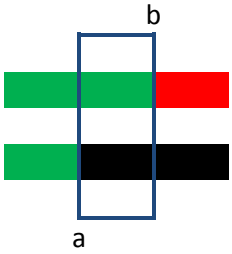
4.6 Supplementary Tables

Table S 4.1 Summary of conversion tracts associated with reciprocal crossovers.

The different types of conversion tracts observed in this study are depicted above. For each conversion tract, the top colored line depicts the red sector and the bottom line depicts the white sector. Green, red, and black depict heterozygosity, homozygosity for W303a-derived SNPs, and homozygosity for YJM789-derived SNPs, respectively. Lowercase letters located above and below the conversion tract depictions represent transitions within the conversion tract. The coordinates of the SNPs on either side of the transitions are listed in Table S 4.2. The blue boxes within the conversion tract depictions represent the DSB regions, as described in the Supplementary Text of Chapter 3. Asterisks to the right of the subclass names indicate whether the event was initiated by a DSB during G1 or G2 of the cell cycle. One asterisk and two asterisks indicate a G1 and G2 DSB, respectively.

Class	# Class obs.	# Subclass obs	
HU A	14		Simple crossover, no conversion tract
			
HU B	12		3:1 conversion tract
HU B1**		8	
			

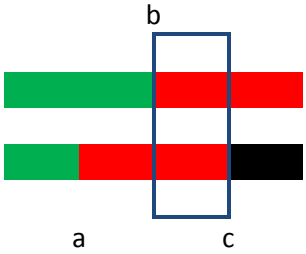
HU B2** 4



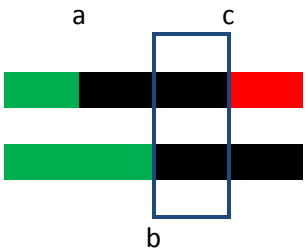
HU C 17

Simple hybrid conversion tracts (3:1/4:0 or 3:1/4:0/3:1)

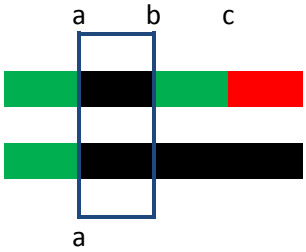
HU C1* 1



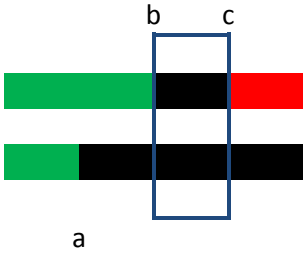
HU C2* 1



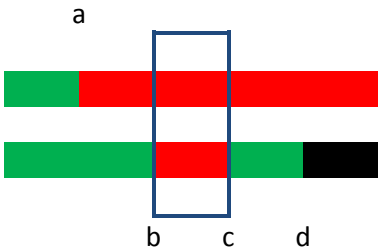
HU C3* 1



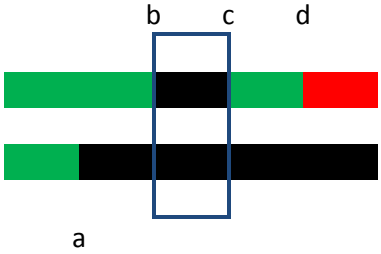
HU C4* 1



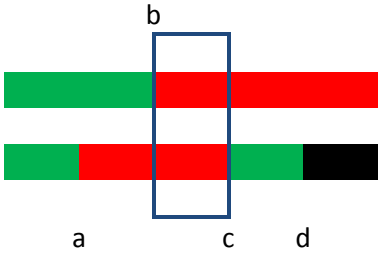
HU C5* 4



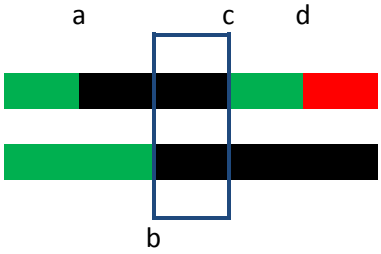
HU C6* 3



HU C7* 1

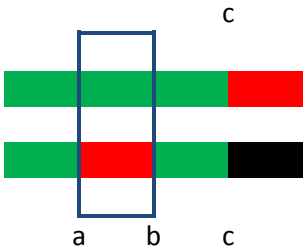


HU C8* 5

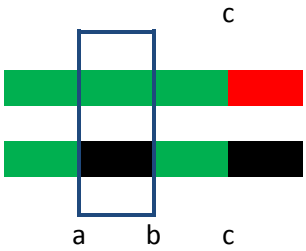


HU D 9 Complex conversion tracts

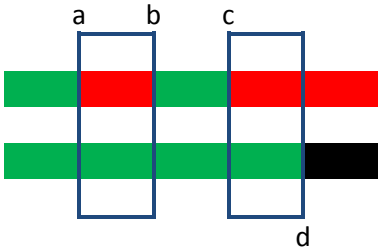
HU D1* 1



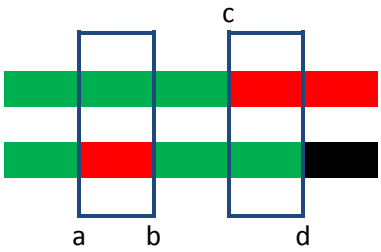
HU D2** 1



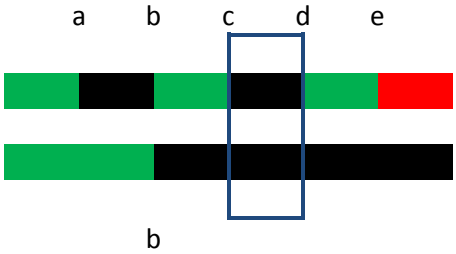
HU D3** 1



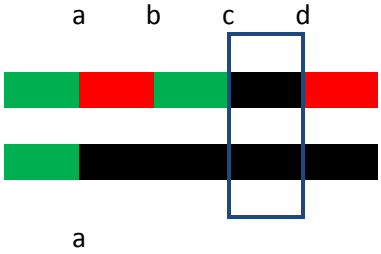
HU D4* 1



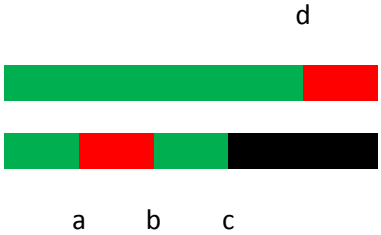
HU D5* 1



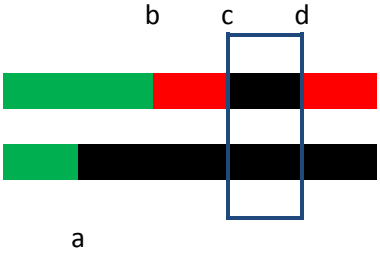
HU D6* 1



HU D7 1



HU D8* 1



HU D9* 1

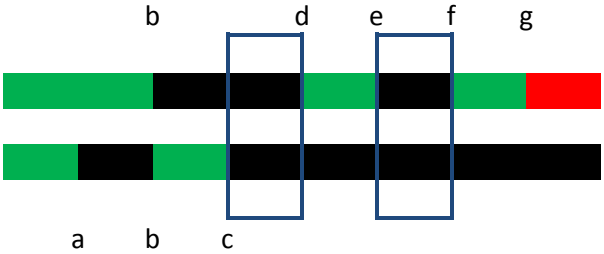


Table S 4.2 SGD coordinates for HU-induced crossover and conversion transitions on chromosome IV.

Sector	Event class[4]	Transition label[5]	Markers flanking transitions[6]	
			Left	Right
HU 1RW	HU A	a	1111969	1142607
HU 2RW	HU A	a	1380606	1385030
HU 3RW	HU D6	a	625236	628829
		b	635286	636074
		c	638445	638827
		d	639029	639208
HU 4RW	HU A	a	1089446	1111969
HU 5RW	HU C5	a	1049553	1051286
		b	1053046	1053271
		c	1053382	1053513
		d	1053513	1053732
HU 6RW	HU B1	a	1399918	1400419
		b	1400419	1400933
HU 7RW	HU D7	a	1425699	1426190
		b	1426190	1427147
		c	1428653	1428998
		d	1429226	1429997
HU 8RW	HU B2	a	710357	710815
		b	713491	713743
HU 9RW	HU A	a	1480252	1483459
HU 10RW	HU D1	a	868562	868904
		b	893025	893739
		c	890529	907253
HU 11RW	HU C6	a	840582	840683
		b	846123	846901

[4] The classes of events are defined in Table S 4.1.
[5] The lower case letters refer to transitions between heterozygous and homozygous regions as shown in Table S 4.1.
[6] These numbers represent SGD coordinates of SNPs located on each side of the transition. It should be noted that these coordinates are based on SGD coordinates from Feb. 2010, and some of these coordinates may be different from those currently displayed in SGD.

		c	850438	851194
		d	851194	852302
HU 12RW	HU B1	a	1349668	1354259
		b	1354259	1366726
HU 13RW	HU A	a	485707	491787
HU 14RW	HU C6	a	808984	809854
		b	809854	812466
		c	820996	821627
		d	823220	825290
HU 15RW	HU B2	a	944549	945141
		b	947397	947901
HU 16RW	HU A	a	1379869	1380606
HU 17RW	HU D3	a	942797	943052
		b	947163	947397
		c	951117	952255
		d	953272	955078
HU 18RW	HU B1	a	1089446	1111969
		b	1111969	1142607
HU 19RW	HU A	a	868904	889108
HU 20RW	HU D5	a	755533	759928
		b	768378	770269
		c	773656	773743
		d	780003	780398
		e	788875	789378
HU 21RW	HU C8	a	1047033	1051286
		b	1052755	1052981
		c	1059444	1065071
		d	1067771	1069072
HU 22RW	HU B2	a	999715	1000365
		b	1010105	1011921
HU 23RW	HU D4	a	584836	585304
		b	589443	590872
		c	595514	596041
		d	598715	599224
HU 24RW	HU B1	a	1161545	1171413
		b	1174678	1178332
HU 25RW	HU C5	a	931521	931973

		b	932502	933384
		c	933384	933740
		d	939734	940054
HU 26RW	HU C8	a	708625	708997
		b	709196	709243
		c	711745	711895
		d	716194	716362
HU 27RW	HU B1	a	672337	672738
		b	677227	677529
HU 28RW	HU A	a	1358761	1361012
HU 29RW	HU B2	a	501652	502459
		b	509473	509716
HU 30RW	HU C8	a	1324027	1324869
		b	1325830	1326832
		c	1337283	1337667
		d	1340963	1341190
HU 31RW	HU C5	a	868904	889108
		b	890618	892648
		c	895029	901094
		d	907253	909432
HU 32RW	HU C5	a	1362046	1364446
		b	1364731	1365816
		c	1373256	1374198
		d	1374571	1379772
HU 33RW	HU A	a	1089446	1111949
HU 34RW	HU C2	a	1181992	1184586
		b	1182080	1184586
		c	1186611	1187255
HU 35RW	HU C7	a	611611	614724
		b	616610	617827
		c	632639	633870
		d	637196	638445
HU 36RW	HU A	a	471247	476888
HU 37RW	HU D8	a	1334123	1335379
		b	1339571	1340268
		c	1343562	1343739
		d	1346791	1349668

HU 38RW	HU B1	a	703372	703507
		b	703891	706018
HU 39RW	HU B1	a	460205	469765
		b	469765	470421
HU 40RW	HU C6	a	1282790	1283368
		b	1286747	1287106
		c	1290703	1291657
		d	1301457	1302344
HU 41RW	HU A	a	1510193	1510396
HU 42RW	HU A	a	1111969	1142607
HU 43RW	HU C4	a	1458992	1459363
		b	1470975	1471237
		c	1474663	1475382
HU 44RW	HU C3	a	851194	855165
		b	868562	889108
		c	890618	892648
HU 45RW	HU A	a	1079615	1142607
HU 46RW	HU D9	a	960043	960710
		b	961833	963726
		c	964391	965936
		d	966937	968316
		e	971858	972225
		f	973142	973779
		g	974614	975606
HU 47RW	HU B1	a	1156998	1173657
		b	1174594	1181603
HU 48RW	HU C8	a	781765	783149
		b	783149	783271
		c	785651	785867
		d	787635	788578
HU 49RW	HU D2	a	1146144	1146359
		b	1153158	1153949
		c	1157254	1161545
HU 50RW	HU C8	a	625236	628588
		b	630883	632639
		c	639208	639635
		d	640034	640958

HU 51RW	HU C1	a	840683	841060
		b	846123	846463
		c	848122	848926
HU 52RW	HU A	a	521609	524793

Table S 4.3 Coordinates used to represent HU-induced crossover-associated gene conversion tracts as single points.

For some of the analysis of hotspots, the conversion tracts were represented as single points. The method for calculating the single point is described in the Supp. Text of Chapter 3. The description of each sectored colony is in Table S 4.1 and Table S 4.2.

Sector	Coordinate of point
HU 1RW	1127288
HU 2RW	1382818
HU 3RW	638826.5
HU 4RW	1100707.5
HU 5RW	1053279.5
HU 6RW	1400425.5
HU 8RW	712050
HU 9RW	1481855.5
HU 10RW	881150.5
HU 11RW	848658.5
HU 12RW	1358197
HU 13RW	488747
HU 14RW	815740.5
HU 15RW	946225
HU 16RW	1380237.5
HU 17RW	948937.5
HU 18RW	1116026.5
HU 19RW	879006
HU 20RW	777027
HU 21RW	1058913
HU 22RW	1005818
HU 23RW	592030
HU 24RW	1169938.5
HU 25RW	933121
HU 26RW	710545.5
HU 27RW	674933
HU 28RW	1359886.5
HU 29RW	505684
HU 30RW	1331748.5
HU 31RW	895856

HU 32RW	1369464.5
HU 33RW	1100697.5
HU 34RW	1184667.5
HU 35RW	625240
HU 36RW	474067.5
HU 37RW	1346615
HU 38RW	704695
HU 39RW	465313
HU 40RW	1289202
HU 41RW	1510294.5
HU 42RW	1127288
HU 43RW	1473178.5
HU 44RW	870151
HU 45RW	1111111
HU 46RW	969085
HU 47RW	1169300.5
HU 48RW	784508
HU 49RW	1150046.5
HU 50RW	635259
HU 51RW	847524.5
HU 52RW	523201

5 Discussion and future directions

In my thesis research, I have examined loss of heterozygosity (LOH) induced by mitotic recombination. The intent of this research was to look for spontaneous recombination hotspots as a method of determining the mechanisms of generating recombinogenic DNA lesions. In addition to mapping mitotic recombination hotspots, I showed that SNP microarrays could be used to map unselected conversion tracts and crossovers induced by ionizing radiation throughout the genome. Furthermore, this study showed that, as the resolution of mapping increased, the complexity of conversion tracts also increased. A number of these complex tracts could not be explained by the canonical SDSA and DSBR recombination models without invoking "patchy" mismatch repair and/or branch migration of Holliday junctions.

Mapping of spontaneous recombination events on the right arm of chromosome IV (Chapter 3) showed that, although recombination events were distributed throughout the chromosome arm, there were several recombination hotspots. Two of these hotspots are associated with inverted repeats of Ty retrotransposons. Building upon previous studies investigating inverted repeats as hotspots for chromosome rearrangements in strains with low levels of DNA polymerase (LEMOINE *et al.* 2008; LEMOINE *et al.* 2005), my results show that naturally-occurring inverted repeats are hotspots for recombination in the absence of exogenous replication perturbation. I also found that

G4-quadruplex motifs are associated with crossovers that result from the repair of G1-initiated DSBs, and replication-termination regions are associated with G2-initiated DSBs. These findings demonstrate that the cause of recombinogenic DNA lesions is different at different phases of the yeast cell cycle. I also showed that the majority of DSBs that cause LOH are initiated during G1 of the cell cycle.

In addition, I mapped mitotic recombination events on chromosome IV in cells treated with hydroxyurea (Chapter 4). Since hydroxyurea (HU) affects DNA replication, I expected that most LOH events would be associated with G2-like gene conversions. Surprisingly, most of the HU-induced LOH events were associated with G1-like conversions, similar to those observed for spontaneous LOH events.

5.1 High-resolution genome-wide mapping of loss-of-heterozygosity events

In Chapter 2, we mapped crossovers that were induced by levels of ionizing radiation (100 Gy) that had little effect on viability. We found an average of 2.4 unselected LOH events per irradiated cell. In these experiments, cells were synchronized in G1. Since repair of DSBs in G1 would not lead to LOH, our observation that cells irradiated in G1 have elevated levels of LOH supports the conclusion that the broken chromosomes are replicated and repaired in G2. Our study was the first to determine the number of genome-wide LOH events induced by γ radiation treatment.

5.1.1 Complex conversion tracts associated with crossovers

As mentioned above, a subset of the conversion tracts analyzed by SNP microarrays were more complex than previous methods implied. The conversion tracts depicted in Table S 2.1 Class J are those that were unexplained by the models proposed in Figure 2.2 and Figure 2.3. The most complicated conversion tract consistent with the simplest version of the DSBR model is the 3:1/4:0/3:1 hybrid tract (Figure 2.3). In contrast, many of the complex conversion events had many transitions between heterozygosity and homozygosity for markers or conversion events in which both homologs were donors.

If a single DSB occurs in one of the two chromatids in G2, and conversion occurs through the SDSA pathway (Figure 1.3A), we expect to observe a single 3:1 conversion tract that reflects the repair of mismatches within a heteroduplex. It has been observed in both meiosis (MANCERA *et al.* 2008; SYMINGTON and PETES 1988) and mitosis (MITCHEL *et al.* 2010) that, in heteroduplexes with multiple mismatches, the mismatches are repaired in a concerted manner. More specifically, all of the mismatches are removed from one of the two strands and the resulting gaps are filled in using the other strand as a template. However, "patchy" repair events have also been observed at a frequency of about 10% in which mismatches are repaired using both the donor and recipient strands as templates within one heteroduplex (KIRKPATRICK *et al.* 1998; MANCERA *et al.* 2008;

MITCHEL *et al.* 2010; SYMINGTON and PETES 1988). The result of this type of repair in a heteroduplex formed during correction of a single DSB in G2 would be a conversion tract in which 3:1 regions are interrupted by regions of heterozygosity. Patchy repair of two heteroduplexes formed during repair of a G1 DSB could generate more complicated patterns. It is likely that patchy repair accounts for at least some of our observed complex tracts.

It should be pointed out that patchy repair could reflect two different mechanisms. One possibility is that the canonical mismatch repair system sometimes switches the strand that is used as a template for mismatch repair. Alternatively, some mismatches may be repaired by a system that is not the canonical repair system. COIC *et al.* (2000) reported evidence for a short patch repair system in yeast that was independent of the canonical MMR and nucleotide excision repair systems. None of the proteins involved in this system have been characterized.

Additionally, some crossovers (for example, Classes J11 and J12 in Table S 2.1) had conversion tracts in which both the W303a- and YJM789-derived homologs were donors in the formation of the conversion tract. This pattern is inconsistent with the model shown in Figure 2.3B. However, if the Holliday junctions of the DSBR model (Figure 2.1B) migrate away from each other prior to cleavage, symmetric heteroduplexes would be generated in where both the donor and recipient chromatids contain

heteroduplex at the same position. Repair of mismatches within a symmetric heteroduplex can result in multiple-donor conversion tracts. Although branch migration of Holliday junctions has been observed in *Escherichia coli*, there is strong genetic evidence that branch migration does not occur during meiotic recombination in *S. cerevisiae* (PETES *et al.* 1991; WEST 1997). However, no strong evidence in favor or against branch migration during mitotic recombination in yeast has been reported.

As discussed in the Introduction, gene conversion events can result from either the repair of mismatches within a heteroduplex or from the repair of a double-stranded DNA gap. It is unclear whether the long conversion tracts observed in my experiments reflect one or both of these mechanisms. One experimental approach to determining whether gene conversion is a consequence of gap repair or repair of mismatches within heteroduplexes is to examine recombination in strains that are unable to repair base-base mismatches. The Msh2p-Msh6p heterodimer recognizes and repairs base pair mismatches (ALANI 1996; IACCARINO *et al.* 1996; JOHNSON *et al.* 1996). Therefore, one could investigate crossovers in a diploid strain that is homozygous for the *msh6* mutation. If gene conversion is a consequence of mismatch repair within a heteroduplex, in an *msh6* strain, within a single sector, there will be two types of cells. For example, assume that that the W303a-derived chromosome has an A/T base pair at a position where the YJM789-derived chromosome has a G/C base pair. During a recombination

event, if the YJM789-derived chromosome donates a strand to the W303a-derived chromosome, an A/G mismatch may be generated. If this mismatch is not repaired, then it would be segregated into one daughter cells that will give rise to a red or white sector. At the next cell division, one cell will have an A/T base pair and the other a G/C base pair. This pattern of segregation pattern has been observed in MMR-deficient strains during mating type switching (HABER *et al.* 1993); I will term this phenomenon as "post-mitotic segregation."

To look for post-mitotic segregation, one could construct an *msh6* diploid in our hybrid genetic background with markers on chromosome IV. In diploid cells that have had a crossover on IV (indicated by a red/white sectored colony), one could examine five colonies derived from the red sector and five derived from the white sector using microarrays in order to detect post-mitotic segregation events. If most crossovers are associated with post-mitotic segregation events, it would argue that the mitotic conversion events observed in my experiments are a consequence of heteroduplex formation followed by mismatch repair. If conversion events, but no post-mitotic segregation events are detected, it would argue that conversion is a consequence of gap repair.

5.1.2 Homologous repair of γ-radiation-induced DSBs

In Chapter 2, I described an analysis of G1-arrested yeast cells that were treated with 100 Gy of ionizing radiation, a dose that should have induced about 35 DSBs per cell. However, by our microarray and high-throughput DNA sequence (HTS) analyses, we detected only an average of 2.4 LOH events per sector. There are several possible explanations for the low detection rate of homologous repair events: 1) most of the DSBs are repaired with conversion tracts that are too small (<50 bp) to detect by SNP microarrays or HTS, 2) radiation-induced DSBs occurs non-randomly in regions of the genome characterized by low SNP density or by repetitive sequences, 3) DSBs are repaired by an SDSA pathway in which mismatches were corrected to restore heterozygosity, and 4) most DSBs are repaired during G1 of the cell cycle by NHEJ. Simple modifications of our current system will allow us to examine the third and fourth possibilities.

As described above, mutations in *msh6* result in defective mismatch repair. Consequently, if the third possibility is correct, we should detect a very substantial increase in post-mitotic segregation in *msh6* cells irradiated in G1 compared to the level of gene conversion events detected in wild-type cells irradiated in G1. One could look for post-mitotic segregation events in the *msh6* diploid as described above.

NHEJ primarily repairs DSBs in yeast during G1 of the cell cycle (FRANK-VAILLANT and MARCAND 2002; SHRIVASTAV *et al.* 2008). In addition, NHEJ is more active in haploid strains or diploid strains that express only one mating type (ASTROM *et al.* 1999; LEE *et al.* 1999) as in my diploid strain JSC25. The possibility that NHEJ is responsible for the repair of a substantial subset of the ionizing radiation-initiated DSBs can be tested by examining the frequency of crossovers in NHEJ-defective derivatives of JSC25. Lig4p is encoded by the gene *DNL4,* which is required for NHEJ (DALEY *et al.* 2005). If NHEJ is responsible for the repair of radiation-induced DSBs in the wild-type strain, we expect that the *lig4* derivative will have an elevated frequency of sectoring and an increased number of unselected LOH events. This expectation is based on the hypothesis that loss of the NHEJ pathway for DSB repair will result in increased repair by homologous recombination. It is also possible that the *lig4* diploid will be more sensitive to killing by γ rays than the wild-type.

5.2 High-resolution mapping of spontaneous chromosome IV crossovers

We mapped spontaneous crossovers on the right arm of chromosome IV to look for hotspots for recombination. This chromosome arm is about 1.1 Mb, approximately 10% of the yeast genome. We showed that mitotic recombination occurs at many sites along the chromosome arm (Figure 3.6 and Figure 3.7). There are two pairs of inverted retrotransposons present on the chromosome arm that were homolog-specific mitotic

recombination hotspots. In addition to the inverted Ty hotspots, we detected five more recombination hotspots that are due to as yet unknown mechanisms. We were able to determine that the majority of the spontaneous recombination events are initiated by DSBs initiated during G1 of the cell cycle. Repair of DSBs initiated during S- or G2-phases represented only a third of the LOH-inducing recombination events.

We were also able to obtain some information about chromosomal regions that are particularly susceptible to recombination-inducing lesions by examining gene conversion events. The chromosome structures/sequences that cause recombination-inducing DNA lesions are usually contained within the associated conversion tracts (PÂQUES and HABER 1999; TANG *et al.* 2011). Thus, I examined the conversion tracts to look for a statistically significant over-representation of 18 chromosome elements (listed in Table S 3.5) that might be associated with DSB formation. In this analysis, we found that different elements were associated with G1- versus G2-initiated recombination events (Table 3.1). Long terminal repeats, Ty elements, and G4 motifs were associated with G1-induced recombination events, while replication termination regions were associated with G2-induced recombination events. These results indicate that G1- and G2-induced DSBs have different causes.

5.2.1 DSBs repaired by HR using the sister chromatid as a template

Our system is only capable of analyzing repair events that use the homolog as a template. For this reason, there may be hotspots for DSBs that we do not detect. We showed that over 70% of recombination events between homologs are initiated by DSBs formed prior to DNA replication. As shown in Figure 3.17, we hypothesize that the broken G1 chromosome is replicated to create two broken sister chromatids during G2 of the cell cycle. Because both sister chromatids are broken at the same position, these DSBs cannot be repaired by sister-strand exchange. However, if DSBs are initiated during S-phase or G2 of the cell cycle, the sister chromatid is highly preferred as the recombination template (KADYK and HARTWELL 1992). At least part of this preference is a consequence of sister chromatid cohesion (COVO *et al.* 2010; STROM *et al.* 2007; UNAL *et al.* 2007). Consequently, if we reduce sister chromatid cohesion, many of the S-/G2-induced recombination events, currently undetectable with our system, might be re-directed into interactions with the homolog, leading to an increase in observable LOH events.

The sister chromatid cohesion complex includes Mcd1p, Smc3p, Smc1p, and Irr1p (ONN *et al.* 2008). Additionally, the Eco1p is necessary to establish both replication-dependent and DSB-induced sister chromatid cohesion (UNAL *et al.* 2007). Mutants in this pathway can affect cohesion between sister chromatids and homologous

recombination (COVO *et al.* 2010; UNAL *et al.* 2007). The *eco1*$^{ack-}$ mutant, which has reduced acetyltransferase activity, is defective in promoting DSB-induced sister chromatid cohesion, but is capable of promoting replication-dependent cohesion (UNAL *et al.* 2008; UNAL *et al.* 2007). COVO *et al.* (2010) demonstrated that reduced sister chromatid cohesion caused by reduced levels of the Mcd1p and Smc3p in tetraploid yeast strains increased ionizing radiation-induced homologous recombination. They also demonstrated that a diploid strain heterozygous for *MCD1* had increased levels of ionizing radiation-induced recombination.

In order to detect a larger fraction of S/G1-induced DSBs, sister chromatid cohesion could be reduced by use of the *eco1*$^{ack-}$ mutant or by lowering levels of Mcd1p or Smc3p. Decreased sister chromatid recombination can be confirmed with an unequal sister chromatid recombination assay (PÂQUES and HABER 1999), and increased homolog recombination can be confirmed by an increase in the rate of sectoring using the *SUP4-o* red/white sectoring assay (Figure 3.2). In summary, in strains with mutations that reduce sister-chromatid cohesion, we should see an elevation in the frequency of S-/G2-type conversion tracts (3:1) relative to G1-induced tracts (4:0).

5.2.2 Inverted repeats can cause mitotic recombination hotspots

The two strongest hotspots on chromosome IV co-localized with two pairs of inverted Ty elements located at SGD coordinates 880 kb and 980 kb. At both of these

hotspots (HS3 and HS4), one of the repeats is a Ty2 element and the other is Ty1; these two classes of elements share extensive sequence similarity. Based on my analysis of conversion tracts associated with crossovers at this hotspot, the majority of the events were initiated on the W303a homolog during G1 of the cell cycle (Figure 3.7). The observation that both Ty repeats are present only on the W303a homolog (Figure 3.11) supports the idea that the inverted repeats were the cause of the recombination hotspot. When one of the repeats was removed, or the spacer between the repeats increased, the recombination activity decreased, further supporting the idea that the secondary structures formed by the inverted repeat was the cause of the hotspot (Figure 3.13). Various characteristics of these hotspots, however, still require elucidation. Although one would expect that secondary structures involving the inverted repeats could occur either in the S-period or in G2, the events detected in my analysis were initiated in G1. Furthermore, the conversion tracts associated with the hotspot were significantly longer than conversion tracts located elsewhere on the chromosome arm (Figure 3.6 and Figure 3.9). Below, I will describe experiments to investigate these issues further.

5.2.2.1 HS4: G1-specificity and conversion tract length

As shown in Figure 3.11, only a fragment of one of the Ty repeats that compose the HS4 hotspot is present in the YJM789-derived chromosome. Thus, the W303a-derived chromosome has about 10 kb of sequence that is absent in YJM789. For a DSB

event initiated at HS4 to be repaired using homologous sequences on the YJM789 homolog, the broken ends would have to be resected about 4-6 kb in each direction in order to expose homologous single-stranded sequences. This extensive resection step may be one of the factors influencing the relative frequencies of G1-initiated events and S-/G2-initiated events.

One could modify this ratio by altering the region containing HS4 on the YJM789-derived homolog. The purpose of this experiment would be to reduce the amount of processing of the W303a-derived broken chromosome necessary to initiate HR with the homolog. This alteration may result in a larger fraction of G2-initiated events. Using the *delitto perfetto* method (STORICI and RESNICK 2006), partial Ty repeats can be inserted into the YJM789-derived chromosome in such a way that the inverted repeat is not reconstructed (Figure 5.1). Construct 1 consists of a Ty2 element on the Watson strand in which the 3′ end has been truncated, and a 3′ end of a Ty1 inserted on the Crick strand. The partial Ty element that is present in the YJM789 homolog will remain unaltered. Construct 2 consists of the same partial Ty2 element on the Watson strand as Construct 1, but the entire Ty1 element will be reconstructed on the Crick strand. Since Construct 1 contains an internal deletion within Ty1 that could affect its use as a template during HR, Construct 2 has an intact Ty1.

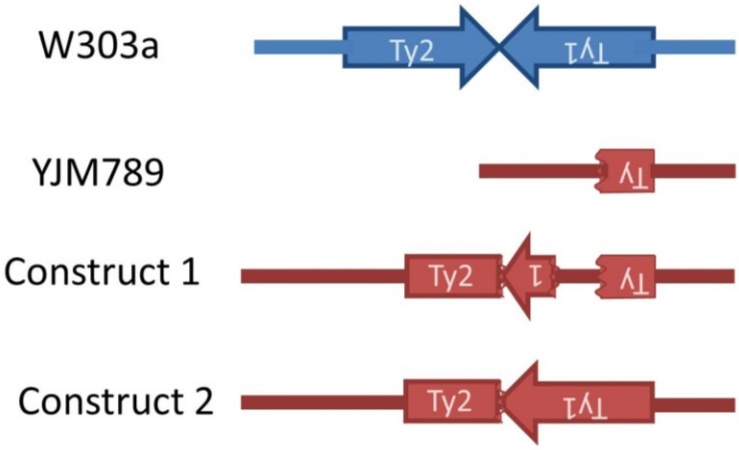


Figure 5.1. Proposed modifications of the YJM789-derived homolog in the region containing HS4.

The inverted repeat at HS4 is only present in the W303a homolog. The YJM789-derived homolog has a 2 kb fragment of a Ty1. In order to determine whether the extent of processing of the broken W303a-derived homolog affects the relative frequencies of G1- and S-/G2-initiated events, additional Ty sequences could be inserted onto the YJM789-derived homolog to increase the shared homology between the two homologs without generating an inverted repeat on the YJM789-derived homolog.

The YJM789-derived strains with the alterations shown in Figure 5.1 can be used to generate a hybrid strain similar to that shown in Figure 3.13, except that one would insert the *SUP4-o* marker distal to the HS4 hotspot, replacing *URA3*. Crossovers between *SUP4-o* and *CEN4* would be identified as red/white sectored colonies, and the drug resistance phenotype could be examined to determine whether the crossover was in the region of HS4. If the efficiency of S-/G2-initiated recombination at HS4 between the two homologs is reduced by the 10 kb heterology in JSC25, there are two expectations: 1) the

frequency of crossovers at HS4 between the homologs should increase in strains with Constructions 1 and 2, and 2) the proportion of G2-initiated events at HS4 should increase. In addition, the absolute frequency of G1-initiated events at HS4 should be unchanged. Furthermore, if the very long conversion tracts at HS4 are, at least in part, a result of the need for extensive resection to reveal homology, conversion tracts at the modified YJM789 HS4 should be shorter than those observed in the unmodified strain.

5.2.2.2 Genetic requirements for HS4 hotspot

As previously discussed, it is likely that the recombinogenic effects of inverted repeats are a consequence of their ability to form secondary DNA structures, either hairpins or cruciforms. The generation of a DNA lesion at these structures should require enzymatic processing (Figure 3.16). Processing of hairpins may occur by nicking of unpaired bases at the tip of the hairpin, whereas processing of cruciforms may require cleavage by a resolvase to create two hairpin-capped ends, followed by nicking of the hairpin-capped ends (COTE and LEWIS 2008; LOBACHEV *et al.* 2002). COTE and LEWIS (2008) showed that Mus81p has a role in the cleavage of *in vivo* cruciform structures, although, in other studies, cruciform processing occurred in the absence of Mus81p (K. Lobachev, personal communication). LOBACHEV *et al.* (2002) demonstrated that Mre11p and Sae2p are necessary for processing of the hairpin-capped ends following cruciform cleavage. The recombinogenic properties of HS4 are likely to require additional

processing steps. Because the inverted repeat is only present on the W303a homolog, removal of non-homologous 3′ ends will be needed to complete a recombination event between the two homologs. Rad1p and Rad10p have been shown to be necessary for the cleavage of 3′ non-homologous ends (FISHMAN-LOBELL and HABER 1992; IVANOV and HABER 1995) in a process that is also dependent on Msh2p and Msh3p (INBAR and KUPIEC 1999; PÂQUES and HABER 1997; SAPARBAEV *et al.* 1996).

In order to get further insights into the mechanism of the HS4 hotspot activity, one could construct diploids homozygous for mutations in *MRE11*, *SAE2*, *RAD1*, *RAD10, MSH2,* or *MSH3*, using a strain in which HS4 is flanked by markers (Figure 3.13). One complication is that some of these mutations might affect recombination in a manner that is independent of HS4 activity. To address this complication, one could compare the recombination activities in a strain with the wild-type HS4 hotspot with a strain that has an inactivated HS4 (Figure 3.13). Mutations that equally affect strains with the active and inactive HS4 are likely due to general effects on recombination. Mutations that affect only the strain with the active HS4 are likely to be involved in the processing of the hairpin structure or in subsequent "trimming" of non-homologous strands.

5.2.2.3 Existence of HS4 in other yeast strains

As was demonstrated in Chapter 3 (Figure 3.11), the inverted pair of intact Ty elements is present only on the W303a-derived homolog. In the allelic position of the YJM789-derived homolog, there was a 2 kb fragment of a Ty element at the position of HS4. The presence of this fragment suggests that a progenitor strain of YJM789 had at least one Ty element. It is possible that the progenitor of YJM789 had the same pair of Ty elements as in W303a, but the inverted repeats were lost. Because of the inherent instability of HS4, its persistence in the W303a-derived homolog is surprising.

The persistence of HS3 and HS4 in the strains W303a and S288c suggests that these pairs of inverted repeats may confer a selective advantage to the cells that contain them. One possibility is that the opportunity to reversibly duplicate the region between HS3 and HS4 is advantageous to cells under certain growth conditions. Duplications of this region have been observed in strains with low levels of DNA polymerase alpha (Wei Song, personal communication). This result suggests that HS4 and HS3 may be fragile sites, regions that have a high frequency of DSBs at a consequence of perturbed DNA replication. Fragile sites are characterized by slow replication (CHA and KLECKNER 2002). Replication near HS4 has been shown to be slow in the absence of Rrm3p (AZVOLINSKY *et al.* 2009), a helicase that helps the replication fork pass through natural impediments. The locations of fragile sites are well conserved in primates and other

mammals, indicating that they may have a functional role (DURKIN and GLOVER 2007). This role, however, is currently unknown.

Another argument for or against the functional importance of HS4 would be provided by the analysis of the sequences of other diverged *S. cerevisiae* strains for the presence or absence of HS4. This analysis could be done using the same primers and PCR reactions previously utilized in Chapter 3 to examine the presence of the repeats in the W303a and YJM789 backgrounds. Alternatively, high-throughput sequence analysis of diverged strains would yield the same information. If HS4 exists in most strains, it is likely that these sequences serve some important function. If both HS3 and HS4 exist in most strains, it would support the hypothesis that duplication of the region between the hotspots may be advantageous under some growth conditions.

5.2.2.4 Recombination activity of other Ty retrotransposon repeats throughout the genome

In the S288c strain, there are three other pairs of inverted Ty elements similar to HS3 and HS4. These pairs are located on chromosome III at 167 kb, chromosome VII from 811 to 823 kb, and chromosome XVI from 844 to 856 kb. The inverted repeat on chromosome III was previously shown to be a hotspot for chromosome rearrangements (termed "FS2") in strains with low levels of DNA polymerase alpha or delta (LEMOINE *et al.* 2008; LEMOINE *et al.* 2005). If the inverted Ty repeat structures are sufficient to cause a recombination hotspot, the inverted repeats on chromosomes VII and XVI should also

be recombination hotspots. This prediction can be tested using a similar system to that depicted in Figure 3.13 (insertion of *HYG* and *URA3* markers flanking the repeats). If these inverted repeats are not mitotic recombination hotspots, it would argue that chromosome context affects the recombinogenic properties of inverted repeats.

5.2.3 Investigation of the spontaneous recombination hotspots, HS5 and HS7

While the likely mechanism of HS3 and HS4 is clear as a result of the studies reported in Chapter 3, the mechanism(s) of the remaining hotspots is undetermined. HS5 and HS7 are two candidate hotspots that would be ideal for further experiments. The HS5 hotspot co-localizes with three closely-spaced *HXT* genes that are highly transcribed and two *ARS* elements. As described in Chapter 1, high levels of transcription are recombinogenic, possibly for several different reasons (formation of R-loops, accumulation of transcription-associated supercoils, collisions with replication machinery). One simple test for determining whether high-level transcription of the *HXT* genes is required for HS5 hotspot activity is to grow the cells in galactose instead of glucose, since the HXT genes are highly transcribed in glucose-containing medium and weakly transcribed in galactose-containing medium (OZCAN and JOHNSTON 1999).

For this experiment, one could use a similar system to that depicted in Figure 3.13. The diploid to be used to examine H5 hotspot activity, however, will be heterozygous for *LYS2* centromere-proximal to HS5 and *URA3* centromere-distal to HS5;

the diploid will be homozygous for mutant *lys2* genes at the normal location of *LYS2*. A crossover at HS5 between the two markers will result in a cell that is 5-FOA^R and Lys$^+$. Thus, this system can then be used to accurately measure the recombination activity of HS5. Cells that are 5-FOA^R and Lys$^-$ are primarily derivatives with a crossover between the heterozygous *LYS2* insertion and *CEN4*. Although chromosome loss events can also result in 5-FOA^R and Lys$^-$ strains, this class of event can be screened out using heterozygous SNPs on the other side of the centromere.

Two types of experiments could be done with this strain. First, as described above, one could monitor HS5 recombination activity in cells grown in glucose and cells grown in galactose. If high levels of transcription are required for the hotspot activity, a significant decrease in the frequency of 5-FOA^R and Lys$^+$ cells is expected. The frequency of 5-FOA^R and Lys$^-$ cells is expected to be unchanged if the effect is HS5-specific. The second type of experiment would involve constructing deletion derivatives of HS5. HS5 co-localizes with *HXT7*, *ARS432*, *HXT6*, *HXT3*, *ARS453* and *SVF1*. *ARS432* and *HXT6*, and *ARS453* and *SVF1*, are positioned such that replication from each origin could potentially lead to a head-on collision with the corresponding gene. One could address this issue by deleting either the *HXT* genes or the ARS elements.

HS7 is also a good hotspot to examine because it is more narrowly defined than the other hotspots (Figure 3.7). The peak of HS7 (SGD coordinates 1268296-1269955) co-

localizes with the *UTP5* gene. Since deletion mutants of *UTP5* are inviable (GIAEVER *et al.* 2002), one could perform deletion analysis of this gene in a strain with a plasmid-borne copy of *UTP5*. The first deletion that would be investigated is one that removes the *UTP5* promoter. If no effect is observed, one could make smaller deletions distributed throughout the gene to determine whether a particular sequence motif is associated with the recombinogenic activity of HS7.

5.3 HU-induced crossovers on chromosome IV

In addition to mapping spontaneous crossovers on the right arm of chromosome IV, we also mapped hydroxyurea (HU)-induced crossovers on the same chromosome arm (Chapter 4). HU inhibits ribonucleotide reductase (RNR)(SLATER 1973), which is necessary for the synthesis of dNTPs in preparation for DNA replication (CHABES *et al.* 2003). In the presence of HU, increased amounts of single-stranded DNA are observed near the replication forks, and recombination rates increase (GALLI and SCHIESTL 1996; SOGO *et al.* 2002). We detected a 26-fold increase in recombination in response to HU, consistent with previous studies of HU-induced recombination (BARBERA and PETES 2006; GALLI and SCHIESTL 1996).

5.3.1 Comparison of HU-induced recombination with spontaneous recombination

I found that HU-induced recombination events were evenly distributed throughout the chromosome arm, and the spontaneous recombination hotspots HS1-

HS7 were not evident in the pattern observed in the HU-treated strains. It has been demonstrated that HU causes DNA replication to proceed at a much slower rate than normal, but replication is eventually completed after an incubation period of about 24 hours (ALVINO *et al.* 2007). In our experiments, cells were treated with HU for 6 hours. It is possible, therefore, that we did not detect the HS1-HS7 hotspots because the replication forks had not reached the hotspot regions. This possibility can be addressed by repeating the experiment with increased HU incubation times. One could also modify the protocol by synchronizing cells in G1 prior to HU-treatment. This synchronization can be done with alpha factor, a pheromone that arrests *MATa* haploids and *MATa/matΔ* diploids in G1 (BUCKING-THROM *et al.* 1973). The cells would be released from the arrest into medium containing 100 mM of HU. Incubation in the presence of HU should proceed for at least 24 hours, in order to allow the cells to complete DNA replication, and then assayed for recombination as described in Chapter 4. If the recombination profile of the extended HU-treatment is indistinguishable from the profile shown in Chapter 4, the possibility of the replication forks not reaching the hotspots in HU-treated cells can be ruled out as a cause for the lack of their detection during HU-treatment.

5.3.2 Timing of recombination initiation lesions in HU-induced crossovers

Many of the HU-induced crossovers were associated with gene conversion tracts indicative of a G1-initiated recombination event. Since HU is thought to induce instability during replication (GALLI and SCHIESTL 1996), we expected to see a substantial elevation of S-/G2-initiated events relative to G1-initiated events. One possible explanation of this unexpected result is that HU treatment causes DSB formation in S and/or G2, but the broken chromatid is segregated into the daughter cell without being repaired. Subsequent replication of the broken chromosome would result in two broken sister chromatids that could be repaired in the next cell cycle resulting in a G1-like conversion. This model is supported by several observations. First, following DNA damage, dNTP levels usually increase 6-8-fold (CHABES *et al.* 2003). However, HU prevents this increase by inhibiting RNR (KRAKOFF *et al.* 1968; REICHARD 1988; SLATER 1973; TIMSON 1975). In addition, homologous recombination is less efficient if RNR is inhibited (COVO *et al.* 2012; MOSS *et al.* 2010). Conversely, elevating the dNTP pools by deletion of the *SML1* gene (an inhibitor of RNR) results in more efficient recombination than observed in wild-type strains (BURKHALTER *et al.* 2009; MOSS *et al.* 2010).

The possibility that the G1-like events actually reflect unrepaired G2 events can be tested by treating G2-arrested cells with HU. Cells arrested in G2 with nocodazole would be treated with γ rays to generate DSBs. Following irradiation, the treated cells

would be incubated in the presence of HU for 3 and 6 hours, before plating on medium lacking HU. One could then examine the resulting sectored colonies to determine whether the crossovers are associated with G1-like or G2-like conversion events. As a control, the same experiment would be done without the HU treatment. Based on previous experiments (LEE and PETES 2010), we expect that control cells will have crossovers with G2-like conversion events. If the G2-irradiated cells incubated in HU have an elevated frequency of G1-like events compared to the control, it would support the conclusion that HU-treatment inhibits repair of G2-induced DSBs, resulting in G1-like conversion tracts.

5.4 Summary

During my thesis research, I have examined both spontaneous and damage-induced mitotic recombination events using SNP microarrays. My research demonstrates that SNP microarrays are a powerful tool for the mapping of homologous recombination events at high resolution throughout the genome. Because of the high resolution of the microarrays, I was able to show that many conversion tracts are more complicated than predicted by the canonical SDSA and DSBR models. In addition to examining conversion tracts, I used SNP microarrays to map the location of crossovers on the right arm of chromosome IV in a W303a-YJM789 hybrid. Strikingly, the recombination profiles of events initiated on the two homologs were quite different from

each other. These differences resulted from hotspots that were chromosome-specific due to inverted repeats that were present on the W303a-derived homolog, but not the YJM789-derived homolog. These inverted repeats were capable of inducing recombination in strains with unperturbed DNA replication; in previous studies, inverted repeats were hotspots for translocations and homologous recombination in strains with mutant DNA polymerases or low levels of the wild-type DNA polymerases (GORDENIN *et al.* 1993; LEMOINE *et al.* 2008; LEMOINE *et al.* 2005). In addition, my results indicate that G1-initiated DSBs cause greater threats to genome stability than G2-initiated DSBs, and confirm that recombination with the sister chromatid is a genome-stabilizing activity relative to recombination with the homolog.

Appendix A- Meiotic chromosome segregation in triploid strains of *Saccharomyces cerevisiae*

Meiosis in triploids results in four highly aneuploid gametes because six copies of each homologue must be segregated into four meiotic products. Using DNA microarrays and other physical approaches, we examined meiotic chromosome segregation in triploid strains of *Saccharomyces cerevisiae*. In most tetrads with four viable spores, two of the spores had two copies of a given homologue and two spores had only one copy. Chromosomes segregated randomly into viable spores without preferences for generating near haploid or near diploid spores. Using single-nucleotide polymorphisms, we showed that, in most tetrads, all three pairs of homologues recombined. Strains derived from some of the aneuploid spore colonies had very high frequencies of mitotic chromosome loss, resulting in genetically diverse populations of cells.

A.1 Introduction

Polyploidy is very common in plants; for example, the frequency of polyploidy in angiosperms is 30-80% (HEGARTY and HISCOCK 2008). Most polyploidy is thought to arise by whole-genome duplication of diploids, producing tetraploids. One mode of producing a triploid is by union of a haploid gamete produced by a "normal" diploid with a diploid gamete produced by a tetraploid. Triploids produced by this mechanism

have been detected in animals as well as plants. In a region with both diploid and tetraploid Palearctic green toads, triploid male toads resulting from crosses of diploid females and tetraploid males were observed (STOCK *et al.* 2010). Triploids can also arise from a cross of diploid individuals in which one individual has tetraploid germinal tissue (BRIDGES and ANDERSON 1925). Although the fertility of triploid Drosophila is reduced, viable offspring between diploids and triploids can be readily obtained. In humans, triploidy is responsible for 15-18% of spontaneous abortions, but only 1 in 1200 fetuses live after birth; all of the live-born individuals die within a few months (ILIOPOULOS *et al.* 2005).

Triploid yeast strains have been observed in the wild (EZOV *et al.* 2006) and have been generated in the lab by forced matings between haploids and diploids (POMPER *et al.* 1954). Since diploids of the *MAT***a**/*MAT*α genotype do not mate, it is likely that the forced matings selected for rare diploids that had become homozygous for *MAT***a** or *MAT*α as a consequence of a mitotic recombination at the mating type locus on chromosome III. In most recent studies, triploids are generated by mating haploids to diploids that have been constructed to be homozygous at the mating type locus (CAMPBELL *et al.* 1981; CAMPBELL and DOOLITTLE 1987; PARRY and COX 1970).

The meiotic products derived from sporulating a triploid would be expected to be highly aneuploid, containing chromosome numbers varying between the haploid

number of 16 and the diploid number of 32. Since even extra copies of single chromosomes often adversely affect cellular growth rates (Torres *et al.* 2007), it is unsurprising that the viability of spores derived from triploids is lower than those derived from diploids. There is, however, a wide range of spore viabilities reported in different labs varying between 15-18% (CAMPBELL *et al.* 1981; CAMPBELL and DOOLITTLE 1987; PARRY and COX 1970) to about 75% (POMPER *et al.* 1954).

The first detailed analysis of chromosome segregation in viable spores derived from triploids was carried out by PARRY and COX (1970). The viable spores were crossed with haploids that had multiple auxotrophic markers and the resulting diploids were sporulated. Departures from 2:2 segregation of the markers were used to diagnose aneuploidy. Markers on 14 of the 16 yeast chromosomes were used. Of 34 segregants examined, three had one copy of all chromosomes tested, and two had two copies of all chromosomes tested. The other 29 segregants had between one and five disomic chromosomes. Parry and Cox noted that disomes of chromosomes V, VIII, IX, XII, and XIII were more frequently observed in the viable spores than disomes of the other chromosomes. In addition, they suggested that the ability of the triploid-derived spores to tolerate aneuploidy was limited, since most of the spores had less than five disomic chromosomes or were diploid.

A more extensive analysis of viable spores derived from triplods was done by CAMPBELL *et al.* (1981). For chromosome III, they observed that about half of the spores had one chromosome and about half had two. They concluded that during meiotic segregation in the triploid, two spores received two copies of each homologue and two received one. For most chromosomes, the frequency of spores with two copies was significantly less than the frequency with one copy. They attributed this difference to loss of disomic chromosomes during mitotic growth of the spore cultures.

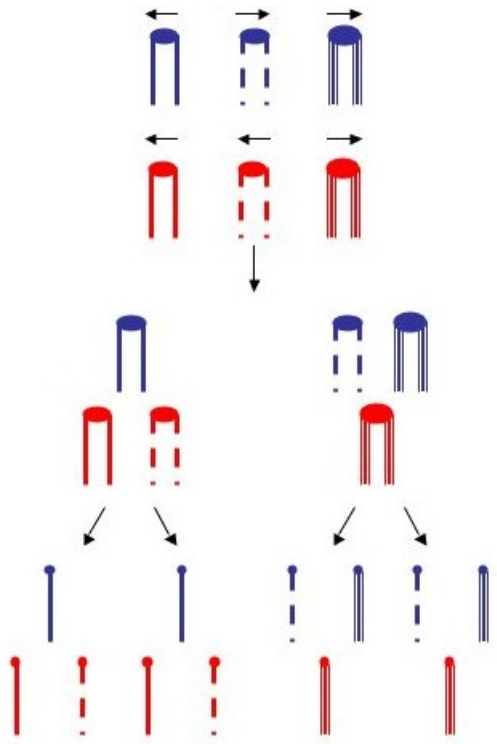


Figure A.1. Expected meiotic segregation pattern of chromosomes of a triploid.

In this figure, we show the meiotic segregation of two homologues shown in blue and in red. The three copies of each homologue are shown as lines with long dashes, short dashes or an uninterrupted straight line. Following the second meiotic division, one would expect to get two disomic spores and two monosomic spores for each chromosome. In this depiction, the sister spores "a" and "b" are disomic for the red homologue and monosomic for the blue homologue, whereas the sister spores "c" and "d" have the reciprocal pattern. With equal frequency, one would observe a pair of doubly disomic sister spores and a pair of spores that are monosomic for both chromosomes. In this figure, we have not shown recombination between the homologues, although such exchanges occur in most tetrads.

In a subsequent study, CAMPBELL and DOOLITTLE (1987) examined chromosome compositions in triploid-derived tetrads that had two viable spores. If these two viable spores are sister spores either both should have two copies of any given chromosome or both spores should have one copy (Figure A.1); below, we will use the term "monosomy" to describe spores that have only one copy of a homologue, although this term is also used to describe diploid cells that lack a single chromosome. These two classes of events should be equally frequent. The observed frequencies of tetrads in which both viable spores were disomic for the same chromosome varied between 0 (chromosome VII) and 0.5 (chromosome XIII) with an average of about 0.19(CAMPBELL and DOOLITTLE 1987). The deviation between the expected (0.5) and observed frequencies of two disomic spores was explained as a consequence of mitotic chromosome loss.

Because of the low spore viability, meiotic recombination in triploid yeast has not been examined previously. However, meiotic recombination in otherwise diploid cells

that are trisomic or tetrasomic for a single chromosome has been studied by a number of researchers. SHAFFER *et al.* (1971) suggested two possible mechanisms by which trisomic chromosomes can pair, recombine, and segregate during meiosis. According to the bivalent/univalent pairing model, two homologues pair, recombine and segregate to opposite poles while the third homologue randomly segregates to either pole. By the second model (trivalent pairing), all three homologues pair, possibly recombine, and randomly segregate to either pole. A distinguishing feature of the trivalent model is that two chromosomes that have recombined with each other are capable of segregating to the same pole during meiosis I. The majority of studies support the trivalent pairing model (CULBERTSON and HENRY 1973; KOLLER *et al.* 1996; LOIDL 1995; RILEY and MANNEY 1978; SHAFFER *et al.* 1971). From the genetic data, it is impossible to determine whether all three homologues pair and recombine at the same time or whether there are multiple cycles of pairing and recombination although, by visualizing the synaptonemal complex in triploid strains, LOIDL (1995) showed that simultaneous pairing of all three homologues was common.

Some of the inferences concerning meiotic segregation and recombination in past studies of triploids were indirect for two reasons. First, not all of the meiotic products from a single tetrad were examined. Second, mitotic chromosome loss occurred during the time-consuming assay of disomic chromosomes during growth of the spore cultures

or in the diploid strains derived from mating the spore cultures. In the current study, we examine the meiotic segregation of chromosomes in triploid yeast in tetrads with four viable spores using physical methods (microarray analysis and a PCR-based examination of single-nucleotide polymorphisms) that are more rapid and sensitive than the methods used previously. We show that meiotic chromosome segregation in triploids is generally accurate, resulting in two spores with two copies and two spores with one copy of each homologue. The mean number of aneuploid chromosomes for each spore is initially close to 8, but we show directly that mitotic chromosome loss in spore cultures can be very rapid. In addition, we show that triploid chromosomes, like trisomic chromosomes, usually undergo trivalent pairing.

A.2 Materials and Methods

A.2.1 Strain construction

Our experiments were conducted using two triploid strains: MH10 and JSC2. The three copies of each homologue in MH10 are identical except for alterations introduced by transformation, whereas the three copies of each homologue in JSC2 are derived from three diverged haploid strains. The genotypes of all strains used in our study are given in Table S A.1. MH10 was generated by crossing MS71 (*MAT*α strain) to JLMy133, an isogenic diploid (*MAT***a**/*mat*Δ*::URA3*). Thus, all three homologues of MH10 have the same DNA sequence except for the mating type locus on chromosome III. The JSC2 triploid was constructed to contain three diverged homologues of each yeast chromosome. JSC2 was generated by mating the haploid PSL4 to the diploid JSC1. The haploid PSL4 (*MAT*α mating type) was derived from the YJM789 genetic background (LEE *et al.* 2009; WEI *et al.* 2007). The JSC1 diploid is a *MAT***a**/*mat*α*Δ::NAT* derivative of JAY306, a diploid generated by crossing FY834 (a haploid isogenic with S288c (WINSTON *et al.* 1995)) with JAY291 (a recently-sequenced haploid derived from an ethanol-producing strain; (ARGUESO *et al.* 2009)).

A.2.2 Genetic methods and media

We used standard genetic and media protocols (GUTHRIE and FINK 1991). Strains were mated by mixing the parental strains on a plate containing rich growth medium

(YPD), and incubating the mixture for six hours at 30°. We then picked zygotes and sporulated the resulting triploids. Following tetrad dissection of sporulated triploids, the dissection plates were incubated at 30° for up to five days. The entire spore colony of tetrads with four viable spores was transferred as a patch to a YPD plate and grown for one to two more days at 30° C. Glycerol stocks were then prepared using the entire patch.

A.2.3 CGH (Comparative Genome Hybridization) microarrays

DNA for CGH microarrays was isolated from 5 ml YPD cell cultures grown for one to two days at 30°. The isolation procedure was the same as that used for preparing DNA samples for Contour-clamped Homogeneous Electric Field (CHEF) gel analysis (NARAYANAN *et al.* 2006). DNA was isolated from the agarose plugs using a QIAgen gel extraction kit. After the plugs were dissolved with buffer QG from the QIAgen kit, the DNA was sonicated. After sonication, DNA was isolated using the standard QIAgen protocol. DNA from spore cultures was labeled with Cy5-dUTP and the control DNA was labeled with Cy3-dUTP (NARAYANAN *et al.* 2006). For the analysis of DNA from MH10-derived spore cultures, MS71 was used as the control DNA; for JSC2-derived spore samples, we used JSC2 DNA for the control. The experimental and control DNAs were competitively hybridized overnight at 65° to microarrays that contained both ORFs and intergenic regions from the yeast genome (ARGUESO *et al.* 2008).

Following hybridization, the slides were scanned using a GenePix 4000B scanner and GenePix Pro 6.0 software. Data files were analyzed using the University of North Carolina Microarray Database (http://genome.unc.edu). The $\log_2$ ratios (hybridization values of experimental and control DNA samples) for each element on the array were first normalized using the $\log_2$ median ratio for the whole data set. If the experimental strain has a mixture of disomic and monosomic chromosomes, following this normalization, the monosomic chromosomes are expected to have an average $\log_2$ ratio of less than zero. Consequently, we determined the average $\log_2$ ratio for all chromosomes with a $\log_2$ ratio less than zero, and added this number to the average $\log_2$ ratio for each individual chromosome. If this sum was greater than 0.17, then the chromosome was considered disomic; if less than 0.17, then the chromosome was considered monosomic. A $\log_2$ ratio of 0.17 represents disomy of a chromosome in 12.5% of the population. The patterns of hybridization were depicted using CGH Miner (http://www-stat.stanford.edu/~wp57/CGH-Miner/).

A.2.4 Diagnostic PCR and restriction digests

The primers listed in Table S A.2 were used to PCR amplify polymorphic regions near the centromere using genomic DNA from JSC2 spores. By treating the resulting fragments with various restriction enzymes (Table S A.2), we could determine the source of the homologues (S288c, YJM789, or JAY291) in each spore. The identity of

chromosomes XIII and XIV were also diagnosed by amplification of centromeric regions that generated DNA fragments differing by 9 to 13 base pairs in size. These PCR products were resolved on a 2% agarose gel. Right and left arm identities were also diagnosed using this method with primers located near the telomeres of each chromosome arm. Primers and restriction enzymes used to diagnose these distal markers are listed in Table S A.3. We also used PCR methods to specifically amplify the centromere-linked JAY291 SNP on chromosome I and the YJM789-specific polymorphism on chromosome XIII. The names and sequences of the primers used for this analysis were: DIST 1.1F (5′ CCACGCCAGGGAATCATCC) and DIST 1.1R (5′ TGCTACAGCATCTCGGCCC) for chromosome I and DIST 13.1F (5′ TTGTCCATAGCAGCAATCATACCAGCACCACC) and DIST 13.2R (5′ AGAAACGCGTCTGGCTTATCTACCGC) for chromosome XIII.

A.2.5 Statistical analysis

The VassarStats Website for Statistical Computation (http://faculty.vassar.edu/lowry/VassarStats.html) was used for the chi-square goodness-of-fit tests and Fisher exact tests. We used the Benjamini and Hochberg test (BENJAMINI and HOCHBERG 1995) to control for the false discovery rate when 12 or more comparisons were made.

A.3 Results

A.3.1 Description of experimental methods

We examined meiotic chromosome segregation in two different types of triploids. In the triploid MH10, all three copies of each homologue were identical in DNA sequence, except for a very small number of changes introduced by transformation. In the triploid JSC2, the three copies of each homologue were derived from three different haploid yeast strains, each of which is diverged in DNA sequence from the other two by about 0.1%. As explained below, this sequence divergence allowed a sensitive assay for disomy in spores derived from JSC2 and also allowed detection of meiotic recombination events. Diploids derived from the diverged haploids had good spore viability, indicating that these strains do not have translocations or inversions (data not shown).

Both triploid strains were sporulated and dissected. Spore viability for both strains was about 50% (Table S A.4). This spore viability is considerably higher than that observed by PARRY and COX (1970) and CAMPBELL *et al.* (1981) (15% and 17.9% respectively) but less than that observed by POMPER *et al.* (1954) (69-83%). In addition, the distribution of the classes of tetrads (4 live: 0 dead, 3 live: 1 dead, 2 live: 2 dead; 1 live: 3 dead, 0 live: 4 dead) deviated significantly from that expected by a binomial expansion (Fisher exact test, $p<0.001$; Table S A.4). We do not know the reasons for the

differences in spore viabilities in different studies. The differences may reflect technical issues (for example, the length of time that asci were treated with glusulase) or variation in the tolerance of aneuploidy in different genetic backgrounds.

About 16% of the tetrads derived from MH10 and JSC2 had four viable spores. All cells in each spore colony were transferred as patches to rich growth solid medium and frozen stocks were prepared. DNA was isolated for microarray analysis by growing large inocula derived from the frozen stocks. This protocol was designed to limit loss or duplication of chromosomes in the mitotic divisions following meiosis. We examined DNA isolated from all four spores for ten tetrads derived from MH10 and six tetrads derived from JSC2. The chromosome compositions of all of these spore cultures were examined by CGH microarrays; we also determined that both of the starting triploids were euploid. In addition, we analyzed the chromosomes in JSC2 spore cultures by examining SNPs as described below.

A.3.2 Analysis of meiotic segregants of the triploid MH10

Figure A.2 shows a representative example of analysis of DNA samples from four spores of an MH10 tetrad by CGH microarrays. DNA from MH10 spores and an isogenic haploid reference strain (MS71) were labeled with different fluorescent nucleotides (Cy3 and Cy5) and competitively annealed to a microarray containing all yeast ORFs and intergenic regions. In the depiction of the microarray shown in Figure

A.2, each of the sixteen chromosomes is represented by a line. Along each line, there are vertical lines whose length and position represents relative hybridization differences between the reference and sample genomic DNA at each element of the microarray. Red and green colors represent duplications or deletions of the experimental strain with respect to the control strain. For example, the spore MH10-32a (Figure A.2) was disomic for chromosomes X and XIII. In many of the microarrays, chromosomes had significantly elevated hybridization signals, although the hybridization level was not two-fold higher than for the other chromosomes. It is likely that this effect is a consequence of loss of the disomic chromosomes in some fraction of the sub-cultured cells during vegetative growth.

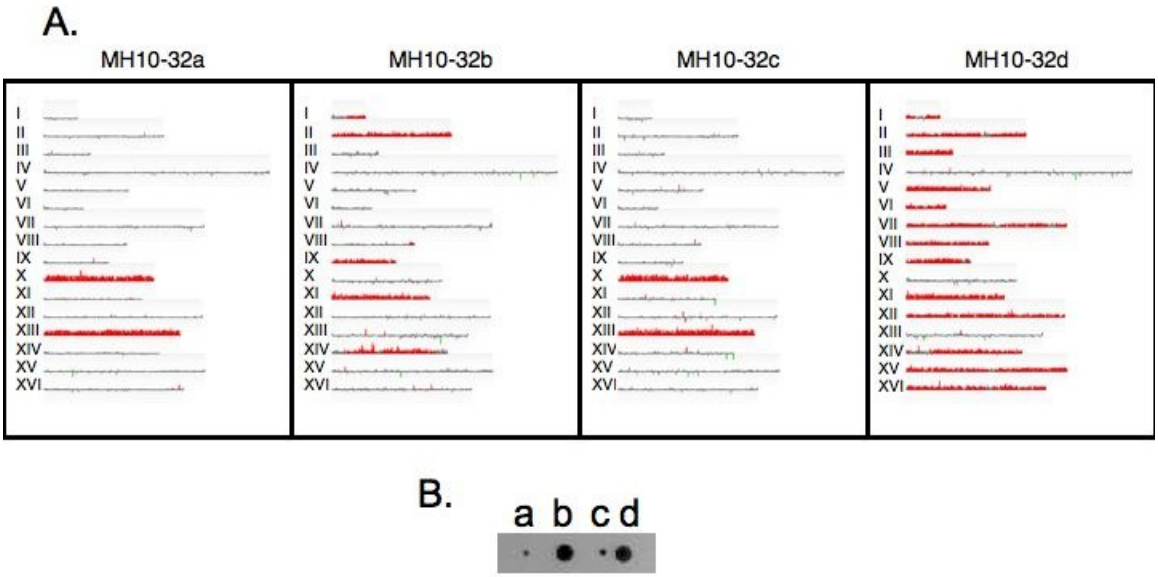


Figure A.2. CGH microarray analysis of DNA derived from four spores of a tetrad derived from the triploid MH10

A. Analysis of the MH10-32 tetrad by CGH microarray. DNA was isolated from individual spore cultures and labeled with Cy5-tagged nucleotides. DNA from the isogenic haploid strain was labeled with Cy3-tagged nucleotides. The individual samples from spore cultures were mixed with the labeled control DNA and the mixture was hybridized to microarrays containing all ORFs and intergenic region. Each chromosome is depicted by a horizontal line whose length is proportional to the length of the chromosome. Red colors indicate gene amplification relative to the reference strain. The meiotic sister pairs of spores are MH10-32a/MH10-32c and MH10-32b/MH10-32d. Extensive loss of disomic chromosomes is evident in spore MH10-32b.

B. Photograph of MH10-32 spore colonies on the tetrad dissection plate.

The spore MH10-32c had the identical chromosome composition as MH10-32a (Figure A.2). The other two spores, MH10-32b and MH10-32d, were both disomic for chromosomes I, II, IX, XI, and XIV. MH10-32d was also disomic for chromosomes III, V-VIII, XII, XV, and XVI. The non-sister pairs of spores had non-overlapping patterns of disomy, and all chromosomes were disomic in at the least one spore except chromosome IV, the largest chromosome. From the pattern of chromosome segregation shown in Figure A.1, one would expect that two pairs of spores (the sister spores) should have reciprocal patterns of disomy, and, for each homologue, one pair of sisters would have two copies and one pair would have one copy. Although chromosomes I, II, IX, X, XI, XIII, and XVI met this expectation, the other chromosomes did not. It should be noted that the pairs of spores that had similar patterns of disomic chromosomes had similar growth properties as spore colonies on the dissection plates (Figure A.2).

One simple explanation of this discrepancy is that the disomic chromosomes are frequently lost during mitotic growth. For example, the sister spores MH10-32a and MH10-32c initially may have had two copies of chromosomes X and XIII and one copy of all of the other chromosomes. The sister spores MH10-32b and MH10-32d had one copy of chromosomes X and XIII, and two copies of all of the other chromosomes. The strain MH10-32b subsequently lost many of the chromosomes originally present in two copies, whereas MH10-32d lost only chromosome IV. It is also possible that sister spores MH10-32a and MH10-32c were originally disomic for chromosome IV, but the extra chromosome was lost in both spore cultures. One alternative explanation of the results is that chromosome loss is very frequent during meiotic segregation in triploids at meiosis I (leading to loss of two copies) or meiosis II (leading to loss of one copy). Another possibility is that there is high rate of chromosome loss in the triploid prior to meiosis. These issues will be discussed in more detail below.

We examined nine other tetrads with four viable spores derived from MH10 by CGH arrays and these data are summarized in Table A.1. For the most part, only meiotic sisters were disomic for the same chromosome in any given tetrad. Most tetrads consisted of colonies that were of two different sizes which corresponded with meiotic sisters as determined by disomy of the same chromosome (Table A.1).

Table A.1 CGH analysis of disomy and monosomy in spores derived from the triploid MH10[1]

		Chromosomes															
Spore	**Colony Size**	I	II	III	VI	V	VI	VII	VII I	IX	X	XI	XI I	XII I	XI V	X V	XV I
4a	small	**1**	1	**1**	1	**1**	1	**1**	1	**1**	**1**	**1**	**1**	**1**	1	**1**	**1**
4b	big	1	2	1	2	1	2	1	2[2]	1	1	1	1	1	2	1	1
4c	small	2	1	2	1	2	2[2]	2	2[2]	2	2	2	2	2	1	2	2
4d	big	1	2	1	2	1	2	1	1	1	1	1	1	1	2	1	1
22a	small	2	1	1	1	2	1	2	2	1	1	1	1	1	1	2	1
22b	big	1	2	1	1	1	1	1	1	1	1	2	1	1	2	1	1
22c	small	2	1	1	1	2	1	2	2	1	1	1	1	1	1	2	1
22d	big	1	**1**	1	1	1	1	1	1	1	1	**1**	1	1	**1**	1	1
24a	big	1	1	2	1	2	1	1	1	2	1	2	1	1	1	1	1
24b	big	1	1	2	1	2	1	1	1	2	1	2	1	1	1	1	1
24c	small	2	2	1	1	1	1	2	2	1	2	1	1	1	**1**	1	1
24d	small	2	2	1	1	1	1	**1**	2	1	2	1	1	1	2	1	1
25a	small	**1**	2	1	**1**	1	**1**	**1**	**1**	1	2	2	1	1	1	**1**	1
25b	big	1	1	2	1	2	1	1	1	2	1	1	2	2	2	1	2
25c	big	1	1	2	1	2	1	1	1	2	1	1	2	2	2	1	2
25d	big	2	2	1	2	2[2]	2	2	2	1	2	2	1	1	1	2	1
28a	small	1	2[2]	2	2	2	1	1	2	2	2	1	1	2	2	1	2
28b	big	2	1	1	1	1	1	2	1	1	1	2	**1**	1	1	1	1
28c	small	1	1	**1**	**1**	2	1	1	2	2	**1**	1	1	**1**	**1**	1	**1**
28d	big	2	2[2]	1	1	1	1	2	1	1	1	2	2	1	1	1	1

[1] Based on CGH microarrays, we determined whether the spore cultures were disomic (indicated by 2 in the table) or monosomic (indicated by 1) for each homologue. Boldface indicates a tetrad in which one sister spore was monosomic, although the other sister was disomic, indicating a chromosome loss event in the monosomic spore.

[2] In these spores, there was a possible chromosome gain. The extra chromosomes were not counted in the averages at the bottom of the table.

29a	med	1	2	1	1	1	2	2	2	1	2	2	2	2	2	2	2
29b	med	2	1	1	**1**	2	1	1	1	2	1	1	1	1	1	1	1
29c	med	2	1	1	2	2	1	1	1	2	1	1	1	1	1	1	1
29d	med	1	2	1	1	1	2	2	2	1	2	2	2	2	2	2	2
30a	big	1	1	1	1	2	1	1	2	2	1	1	1	1	1	2	2
30b	small	2	2	2	2	1	2	2	1	1	2	2	2	2	2	1	1
30c	big	1	1	1	1	2	1	1	2	2	1	1	1	1	1	2	2
30d	small	2	2	2	2	1	2	**1**	1	1	2	2	2	2	2	1	1
32a	small	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1
32b	big	2	2	**1**	1	**1**	**1**	**1**	**1**	2	1	2	**1**	1	2	**1**	**1**
32c	small	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1
32d	big	2	2	2	1	2	2	2	2	2	1	2	2	1	2	2	2
34a	small	**1**	**1**	1	1	1	1	1	1	1	1	1	2	1	**1**	1	1
34b	big	1	1	2	1	2	1	1	2	1	2	2	1	1	1	1	2
34c	small	2	2	1	1	1	1	1	1	1	1	1	**1**	1	2	1	1
34d	big	1	1	2	1	2	1	1	2	1	**1**	**1**	1	1	1	1	**1**
37a	med	2	1	1	1	1	2	2	1	2	1	1	2	2	2	1	2
37b	med	1	2	2	1	2	1	1	2	1	2	2	1	1	1	1	1
37c	med	1	2	2	1	**1**	1	1	2	1	2	2	1	1	1	1	1
37d	med	2	1	1	1	1	**1**	**1**	1	**1**	1	1	**1**	**1**	**1**	1	2
Avg[3]		**5.7**	**5.7**	**5.3**	**4.7**	**5.7**	**4.9**	**5.2**	**5.7**	**5.4**	**5.5**	**5.7**	**5.1**	**5.1**	**5.5**	**4.9**	**5.2**

[3] Average number of individual homologues per tetrad

After we compiled the results of all 10 tetrads analyzed by CGH microarray, it was apparent that spores derived from the triploids are subject to extensive chromosome loss (Table A.1). Assuming that there should be six chromosome copies of each homologue per tetrad (following meiotic DNA synthesis), there should be 96 chromosomes summed over the four spores of the tetrad. On average, however, each tetrad had about 85 chromosomes, representing a loss of about 11 chromosomes distributed over the four spores. Chromosomes IV, VI, and XV were lost at relatively high frequencies and chromosomes I, II, V, VIII, and XI at relatively low frequencies (Table A.1).

Previous studies have already shown that spores derived from triploids have unstable karyotypes (CAMPBELL *et al.* 1981; PARRY and COX 1970), although direct measurements of chromosome stability in the spores were not done. In addition, these studies did not determine whether meiotic chromosome segregation was fundamentally accurate in a triploid (CAMPBELL *et al.* 1981; PARRY and COX 1970). To investigate this issue with a more sensitive assay for aneuploidy than CGH arrays, we analyzed sequence polymorphisms that distinguished the three copies of each homologue in the triploid JSC2.

A.3.3 Analysis of meiotic chromosome segregation in a triploid (JSC2)

We generated the triploid JSC2 with equal contributions of DNA sequences from three different sequenced backgrounds: S288c, YJM789 (MCCUSKER *et al.* 1994), and JAY291 (ARGUESO *et al.* 2009). The details of the construction of the triploid are in Materials and Methods. YJM789 and JAY291 have 6.1 and 5.4 single nucleotide polymorphisms (SNPs) per kb, respectively, relative to S288c (ARGUESO *et al.* 2009; WEI *et al.* 2007). Although the genome alignments of YJM789 to JAY291 have not been completed, in one 49 kb interval aligned in these strains, there were about 10 SNPs/kb(ARGUESO *et al.* 2009). Using a CGH microarray, we confirmed that JSC2 had three copies of every homologue prior to sporulation.

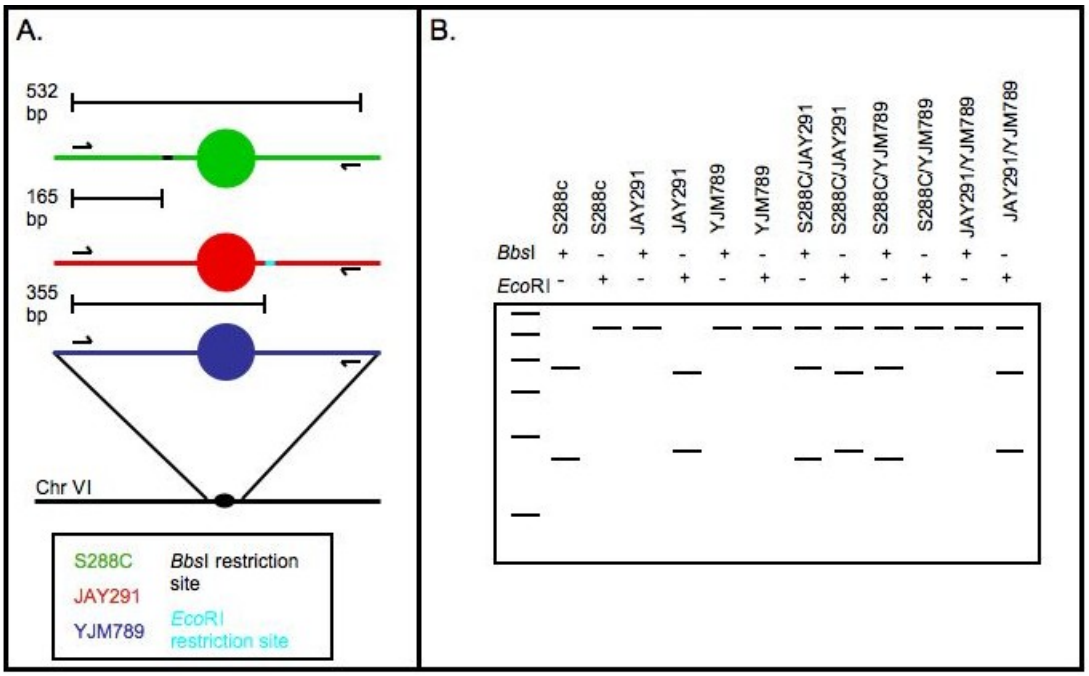


Figure A.3. Use of SNPs to determine the number and identity of homologous chromosome homologues within a JSC2-derived spore

For this analysis, we designed primers that could be used to amplify SNP-containing regions located very near (within 10 kb) the centromere of each of the 16 chromosomes (Table S A.2). By digestion of the resulting PCR products with the relevant restriction enzyme(s), we could diagnose the source of the various copies of each homologue.

A. Depiction of the differential restriction sites present in each strain in the amplified region of chromosome VI. The PCR amplified regions of S288c, JAY291, and YJM789 are depicted in green, red and blue colors respectively. Positions of PCR primers are indicated by horizontal arrows, and circles show the centromere. There is a *Bbs*I restriction site on the S288c PCR fragment (shown as a vertical arrow) that is not present in JAY291 or YJM789, and an *Eco*RI restriction site on the JAY291 PCR fragment that is not present in the other two strains. The *Bbs*I and *Eco*RI restriction sites are located at 165 and 355 bp, respectively, from the left end of the PCR fragment.

B. Demonstration of how differential restriction sites can be used to diagnose the source(s) of homologues in spore colonies. Digestion of each possible monosome and

heterozygous disome with both *Bbs*I and *Eco*RI produces a unique restriction pattern shown as a schematic gel. The leftmost lane is a 100 bp DNA standard ranging from 100 to 600 bp. The horizontal rows above the gel indicate which lane corresponds to which digest with plus signs indicating treatment with that enzyme.

We were able to distinguish each of the three copies of each homologue for all homologues in JSC2 by PCR amplification of sequences within 5 kb of the centromeres, followed by treatment of the PCR fragment with restriction enzymes that distinguish the amplified regions of each strain. For this analysis, we utilized SNPs that altered restriction enzyme recognition sites. In Figure A.3, we illustrate how the three copies of chromosome VI were distinguished by this approach. As shown in Figure A.3, chromosome VI of S288c has a *Bbs*I site near *CEN6* that is not present in JAY291 or YJM789, and JAY291 has a *CEN6*-linked *Eco*RI site that is not present in either S288c or YJM789. Using PCR with primers that flank these heterozygous sites, we generated a DNA fragment of about 530 bp. By examining the sizes of the fragments produced following treatment of the 530 bp fragment with *Eco*RI and *Bbs*I, we can distinguish the three homologues individually (lanes 2-7; Figure A.3) and in pairs (lanes 8-13, Figure A.3). Table S A.2 lists all the primers and restriction enzymes used to diagnose each chromosome. We will refer to this method of analysis as SPA (Single-nucleotide-polymorphism PCR Analysis).

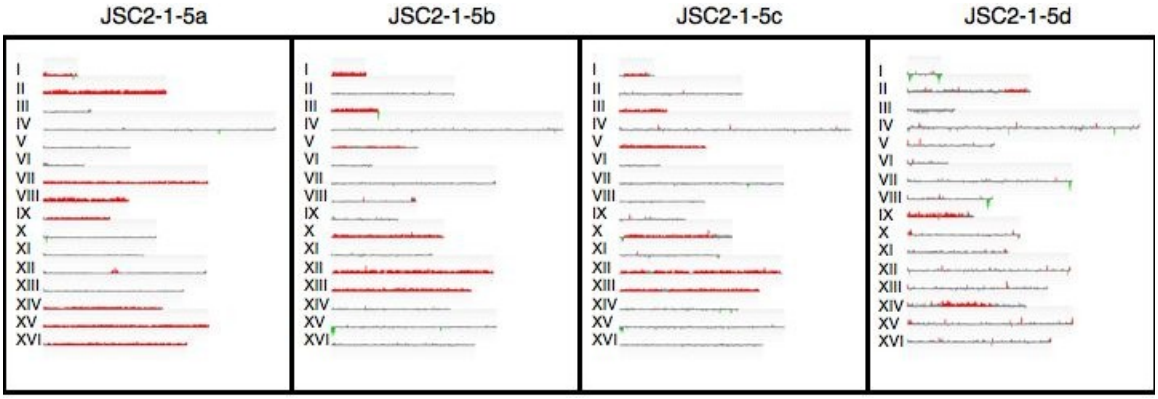


Figure A.4. Analysis of a JSC2 tetrad by CGH microarrays

DNA was isolated from spore colonies of a triploid-derived tetrad of JSC2 and examined by the same method described in Figure A.2. The meiotic sister pairs are JSC2-1-5a/JSC2-1-5d and JSC2-1-5b/JSC2-1-5c. Extensive loss of disomes from spore JSC2-1-5d is evident.

We analyzed 10 JSC2 tetrads that had four viable spores by SPA, and these data are summarized in Table A.2. Additionally, six JSC2 tetrads were also analyzed by CGH microarrays (Table S A.5). Figure A.4 shows a CGH analysis of tetrad 5 of JSC2; the SPA data for this same tetrad are in the first four rows of data of Table A.2

. By both methods, it is clear that the two pairs of meiotic sister spores are JSC2-1-5a and -5d, and JSC2-1-5b and -5c. The two methods of analysis are completely concordant for spores JSC2-1-5b and -5c, both methods showing that these strains are disomic for chromosomes I, III, V, X, XII, and XIII.

Table A.2 Analysis of disomy and monosomy in spores derived from the triploid JSC2-1[4]

	Chromosomes															
Spore	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
5a	J[5]	J Y	Y	Y[6]	Y	S Y	S Y	J Y	S J	J	S J	Y	S	J Y	S Y	S J
5b	S Y	S	S J	S	S J	J	J	S	Y	S Y	Y	S J	J Y	S	J	Y
5c	S Y	S	S J	S	S J	J	J	S	Y	S Y	Y	S J	J Y	S	J	Y
5d	J	J Y	Y	J Y	Y	S Y	S Y	J Y	S J	J	S J	Y	S	J Y	S Y	S J
6a	J	J	Y	S	S	J[5]	J	S	J Y	J	J Y	Y	Y	S Y	S	S
6b	S Y	S Y	S J	J Y	J Y	S Y	S Y	J Y	S	S Y	S	S J	S J	J	J Y	J Y
6c	J	J	Y	S	S	J	J	S	J Y	J	J Y	Y	Y	S Y	S	S
6d	S Y	S Y	S J	J Y	J Y	S Y	S Y	J Y	S	S Y	S	S J	S J	J	J Y	J Y
8a	S	S J	J Y	J Y	J	J	S	S J	Y	S Y	J	Y[5]	S J	S Y	S	S J
8b	J Y	Y	S	S	S Y	S Y	J Y	Y	S J	J	S Y	S J	Y	J	J Y	Y
8c	J Y	Y	S	S	S Y	S Y	J Y	Y	J	J	S Y	S J	Y	J	J Y	Y
8d	S	S J	J Y	J Y	J	J	S	S J	Y	S Y	J	J[7]	S J	S Y	S	S J
9a	Y	S J	J	S	S	S Y	S J	J	S Y	Y	S J	S J	S	S J	S Y	J Y
9b	S J	Y	S	J Y	J Y	J	Y	S Y	J	S J	Y	Y	J Y	Y	J	S

[4] Using SPA (described in text), we determined whether spores were monosomic or disomic, and the source of each homologue (S=S288c; J=JAY291; Y=YJM789).
[5] Extra chromosome homolog detected by CGH microarrays. These extra chromosomes were not counted in the averages shown at the bottom of this table.
[6] Chromosome loss.
[7] Putative gene conversion event.

9c	S J	Y	Y S Y	J Y	J Y	J	Y	S Y	J	S J	Y	Y	J **Y**	Y	J	S
9d	Y	S J	J	S	S	S Y	S[6]	J	S Y	Y	S J	S J	S	S J	S Y	J Y
10a	J Y	J Y	J Y	J	J Y	S	S Y	Y	J	S Y	J	S J	S	S J	J	Y
10b	**J**[8] Y	Y[6]	J[6]	J	J Y	S	S Y	Y	J	S Y	J	S J	S	J[6]	J	Y[6]
10c	S	S	S	S Y	S	J Y	J	S J	S Y	J	S Y	Y[5]	J Y	Y	S Y	S
10d	S	S	S	Y[6]	S	J Y	J	S J	S Y	J	S Y	Y	Y[6]	Y	S Y	S
11a	S Y	Y	Y	J	S Y	S	S J	Y	J	S Y	Y	Y	Y	S Y	S	J Y
11b	J	S J	S J	S Y	J	J Y	Y	S J	S Y	J	S J	S J	S J	J	J Y	S
11c	J	S J	S J	S Y	J	J Y	Y	S J	S Y	J	S J	S J	S J	J	J Y	S
11d	S Y	Y	Y	J	S Y	S	S J	Y	J	S Y	Y	Y	Y	S Y	S	J Y
12a	Y	S J	J	J	J	S Y	Y	S Y	S Y	S	S J	S	J	S Y	S Y	S J
12b	Y	S J	J	J	J	S Y	Y	S[5,6]	S Y	S	S J	S	J	S Y	S Y	S J
12c	S J	Y	S Y	S Y	S Y	J	S[6]	J	J	J Y	Y	J Y	S S[7]	J	J	Y
12d	S **J**	Y	S Y	S Y	S Y	J	S J	J	J	J Y	Y	J Y	S S[7]	J	J	Y
13a	J	S J	S	S J	S Y	Y	J Y	S J	J	S	J	S Y	S	J	S Y	S
13b	S Y	Y	J Y	Y	J	S J	S	Y	S Y	J Y	S Y	J	J Y	S Y	J	J Y
13c	S	Y	J Y	Y	J	S J	S	Y	S	J Y	S	J[5]	J Y	S Y	J	J Y

[8] Boldface indicates that homologues were detected using PCR and primers that exclusively amplified only one of the strain-specific SNPs.

13d	Y J	S J	S	S J[9]	S Y	Y	J Y	S J	Y J	S	Y J	S Y	S	J	Y[6, 9]	S
14a	J	Y	S J	S	Y	J	S	Y	J	J	Y	S J	S	S	S Y	S
14b	J	Y	S J	S	Y	J	S	Y	J	J	Y	S J	S	S	S Y	S
14c	S Y	S J	Y	J Y	S J	S Y	J Y	S J	S Y	S Y	S J	Y	J **Y**	J Y	J	J Y
14d	S Y	S J	Y	J Y	S J	S Y	J Y	S J	S Y	S Y	S J	Y	J Y	J Y	J	J Y
15a	S Y	S J	S J	Y	J	S J	S J	S	J	J	J	J Y	J Y	S Y	Y	J Y
15b	J	Y	Y	S J	S Y	Y	Y	J Y	S Y	S Y	S Y	S	S	J	S J	S
15c	J	Y	Y	S J	S Y	Y	Y	J Y	S Y	S Y	S Y	S	S	J	S J	S
15d	S Y	S J	S J	Y	J	S J	S J	S	J	J	J	J Y	J Y	S Y	Y	J Y
Average[10]	**6**	**5.9**	**5.9**	**5.8**	**6**	**6**	**5.8**	**5.9**	**6**	**6**	**6**	**6**	**5.9**	**5.9**	**5.9**	**5.9**

[9] Translocation between chromosome IV and XV detected by CGH microarray and gel analysis.
[10] Average number of individual homologues per tetrad.

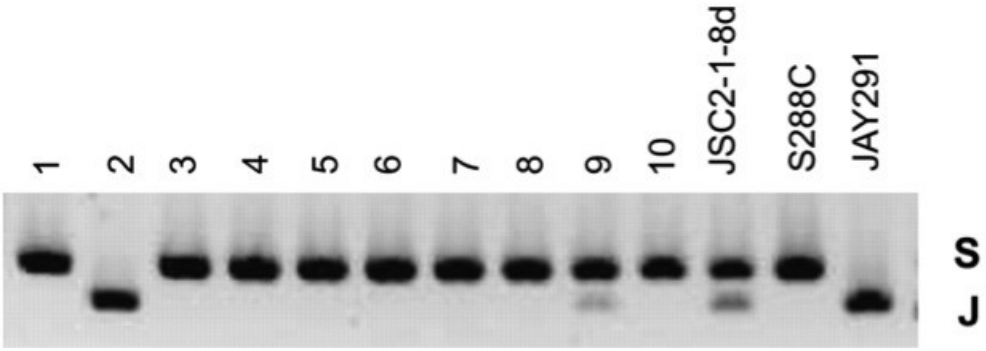


Figure A.5. Analysis of SNPs in colonies derived from a single meiotic spore culture (JSC2-1-8d)

We isolated DNA from single colonies derived from individual cells of the spore culture JSC2-1-8d, and DNA from the original spore culture. By CGH microarrays, this spore had one copy of chromosome II. The region near *CEN2* was amplified in all DNA samples, treated with the restriction enzymes *Dra*I and *Bbs*I, and the resulting fragments separated by gel electrophoresis. By SPA, the spore colony was disomic for chromosome II with copies derived from S288C and JAY291. In this figure, the *Dra*I digests of the PCR fragment are shown. The digest of DNA isolated from the original spore colony (third lane from the right) indicates that more cells in the culture have the S288c chromosome (uncut band) than the JAY291 chromosome (cut band). This conclusion is confirmed by examining single-colony isolates (labeled 1-10).

The spores JSC2-1-5a and -5d are substantially different when examined by CGH and, in addition, show a number of discrepancies when analyzed by CGH versus SPA. By CGH, JSC2-1-5a is disomic for chromosomes I, II, VII-IX, and XIV-XVI, whereas JSC2-1-5d is only disomic for chromosomes II, IX, and XIV. One obvious explanation for this discrepancy is that many of the disomic chromosomes originally present in the JSC2-1-5d spore were rapidly lost during the mitotic divisions required to produce the cultures for the CGH analysis. In support of this explanation, when we examined DNA isolated

from JSC2-1-5d by SPA, we observed disomy for chromosomes II, IV, VI-IX, XI, and XIV-XVI. This increased level of disomy detected by SPA is the result of two factors. First, it is easier to detect disomy present in a small fraction (less than 25%) by SPA than by CGH analysis. Second, we found that some of the spore cultures had two types of cells in approximately equal amounts, one population that had one variant for a particular homologue and one that had a different variant. An example of such a spore culture is shown in Figure A.5. We performed SPA on 10 colonies derived from the JSC2-1-8d spore culture; eight had the SNP characteristic of the S288c-derived chromosome II, one had the SNP characteristic of the JAY291 homologue, and one had both types of homologue, although the S288c-derived homologue predominated. Thus, in this spore colony, although most of cells have only a single copy of chromosome II, the original spore must have had two copies.

As shown in Table A.2, in most of the tetrads, most of the homologues segregated to produce two disomic spores and two monosomic spores. This result demonstrates that meiotic chromosome segregation in triploids is accurate, and most of the chromosome loss events detected by CGH arrays represent mitotic chromosome loss. Analyzing ten MH10 tetrads with CGH arrays, we found 107 chromosome loss events. Analyzing ten JSC2 tetrads using SPA, we found only 15 chromosome loss events, reflecting the greater sensitivity of SPA in detecting disomy. At least some of these 15

loss events also represent mitotic loss (as discussed further below), presumably at a very early stage in the growth of the spore colony. Figure A.6 shows a comparison of the number of recovered disomes for each method. A line depicts the expected binomial distribution of disomes, assuming that the probabilities of a chromosome being disomic or monosomic are both equal to 0.5. The peaks for the JSC2 spores are shifted to the right relative to the MH10 spores, reflecting the more sensitive nature of SPA compared to CGH. Additionally, the line representing JSC2 spores is symmetrical around eight disomes. Since we examined only complete tetrads, this symmetry is expected.

There were two other observations of interest concerning the JSC2 spore analysis. First, in comparing the CGH and SPA data, we found six examples of apparent chromosome duplication. For example, by CGH, the spores JSC2-1-5a, -5b, and -5c are disomic for chromosome I. By SPA, the spores JSC2-1-5b and -5c both have the S288c-derived and the YJM789-derived copies of chromosome I, but the spore JSC2-1-5a has only the JAY291 homologue. We interpret this result as indicating that a single copy of chromosome I originally present in spore JSC2-1-5a underwent a chromosome gain, either by non-disjunction or by re-replication.

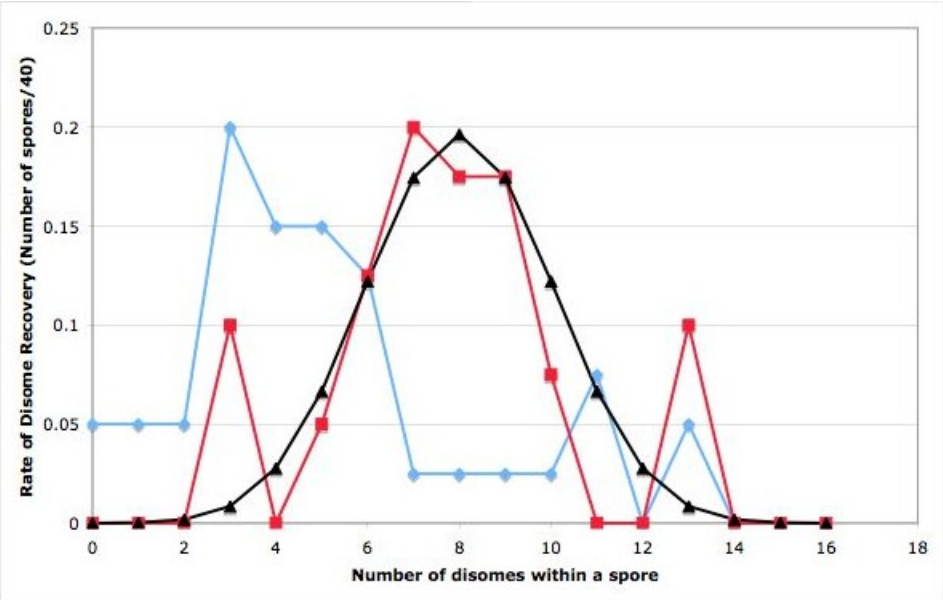


Figure A.6. Comparison of the detection of disomes in spores derived from the triploids MH10 and JSC2

The graph shows the number of disomic chromosomes per spore as a function of the frequency of spores with that chromosome composition. The blue line with diamonds represents the data from the MH10 triploid (based on CGH), and the red line with squares shows the data from the JSC2 triploid (based on SPA). The black line with triangles is the curve predicted by the binomial expansion with the assumption that there is the same probability (0.5) that a spore will be monosomic or disomic.

The last unusual class of tetrad likely reflects meiotic gene conversion, the non-reciprocal transfer of information from one chromosome to the other. One example of conversion is JSC2-1 tetrad 8. In this tetrad, for chromosome XII, there were two chromosomes with the centromere-linked SNP characteristic of S288c, three chromosomes with the SNP characteristic of the JAY291, and only one chromosome with the SNP characteristic of YJM789. This pattern can be explained if there was a gene

conversion that substituted the chromosome XII region on one of the YJM789-derived homologues with information derived from the JAY291-derived homologues. Since the amount of DNA transferred in meiotic gene conversion is usually limited to about 1-3 kb (MANCERA *et al.* 2008; PETES *et al.* 1991), in the spore representing the putative conversion (JSC2-1-8d), we examined (by SPA) a SNP located on chromosome XII on the opposite side of the centromere located about 4 kb from the original SNP (primers described in Table S A.2). As expected, this polymorphism had the SNP characteristic of the YJM789 chromosome. We also observed one tetrad (JSC2-1-12) in which two chromosomes appeared to have undergone gene conversion: four copies of the S288c-associated SNP, no copies of the YJM789-associated SNP, and two copies of the JAY291-associated SNP. Although this tetrad may have undergone a double meiotic event (two conversions of the YJM789 SNP to the S288c SNP), based on the rarity of the single conversion events, it is more likely that the observed pattern reflects a mitotic gene conversion.

A.3.4 Mitotic chromosome loss

Our analysis of JSC2 demonstrates that most tetrads have two spores with two copies and two spores with one copy of each homologue (Table A.2). However, in some tetrads, one or more chromosomes were lost. As discussed above, such losses have three possible sources: 1) loss of one or more chromosomes in the triploid prior to meiosis, 2)

loss of chromosomes during the first or second meiotic divisions, and 3) loss of chromosomes during the mitotic growth of the spores. Several arguments strongly support the third mechanism as the principle factor.

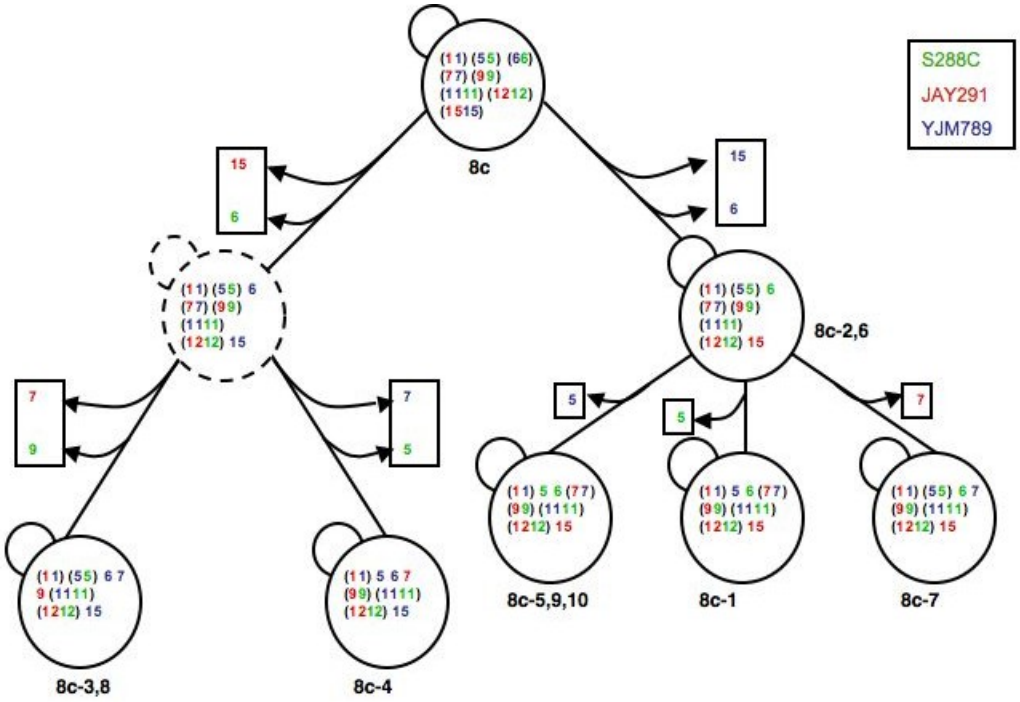


Figure A.7. Mitotic loss of disomic chromosomes from spore JSC2-1-8c

The spore JSC2-1-8c was originally disomic for eight chromosomes (I, V-VII, IX, XI, XII, and XV). We examined 10 single-colony isolates of JSC2-1-8c (labeled 8c-1-10) by SPA, and found six genotypes (shown in solid-line depictions of yeast cells). In this figure, we show a pathway of chromosome loss consistent with these genotypes. The silhouette of a yeast cell with dashed lines is a hypothetical intermediate that was not detected. The numbers within each circle are the chromosome numbers, and the colors indicate the haploid from which the chromosome was derived.

First, we can directly detect chromosome instability in some of the aneuploid spores. We examined the distribution of disomic chromosomes in ten individual

colonies derived from JSC2-1-8c. Our SPA analysis of the chromosome distribution of the original spore colony showed that this spore was disomic for eight chromosomes (I, V-VII, IX, XI, XII, and XV). As shown in Figure A.7, none of the ten colonies (labeled 8C-1 to 8C-10) examined retained disomy for all eight chromosomes. All disomic chromosomes, except I, XI, and XII, were lost in at least one of the ten colonies. In Figure A.7, we show a pathway by which the chromosomes may have been lost; although other pathways could be depicted, the one shown requires the fewest steps.

Although it is difficult to calculate a rate of chromosome loss from this analysis, we estimated the rate of loss of chromosome V in JSC2-8c by another method. The initial spore isolate had two copies of chromosome V, one with the wild-type *URA3* allele and one with the mutant allele. Thus, the spore was initially Ura^+. Since the *URA3* gene is centromere-linked, the rate of sectored Ura^+/Ura^- colonies observed when JSC2-8c colonies cultured on rich growth medium are replica-plated to medium lacking uracil is an estimate of the rate of loss of the chromosome with the wild-type allele. Of 8200 colonies examined, 223 were sectored, a rate of about 0.027. This rate of loss is much higher than the rate of chromosome V loss in a wild-type diploid (10^{-5}) (HARTWELL and SMITH 1985). The frequency of chromosome loss was somewhat spore-specific, and lower rates of loss were observed in JSC2-1-5b (3.5×10^{-3}) and JSC2-1-6d (no sectored colonies observed).

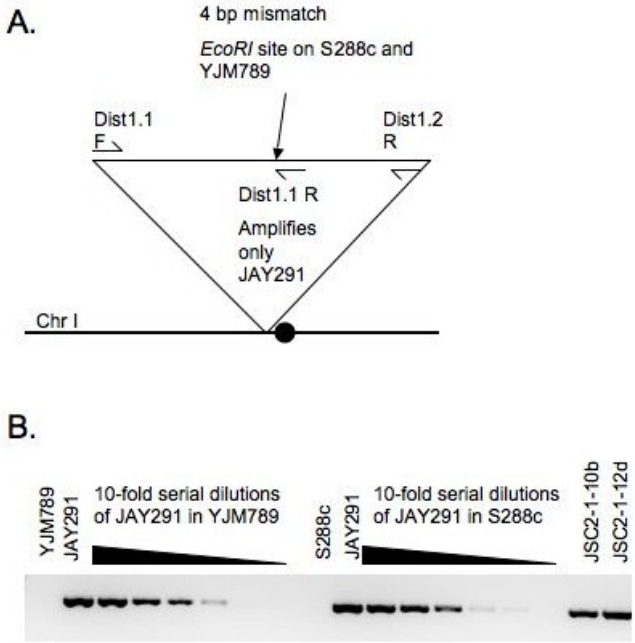


Figure A.8. PCR method of detecting very small amounts of JAY291-derived chromosome I in strains with S288c- or YJM789-derived chromosomes

Although the sister spores of JSC2-1-10b and JSC2-1-12d were disomic for chromosome I and contained the JAY291-derived homologue, we could not detect this homologue using SPA. Consequently, to detect very small amounts of this homologue, we designed primers to specifically amplify DNA with this polymorphism.

A. Primer combinations DIST1.1F/DIST1.1R amplify a region about 12 kb from the centromere of chromosome I. The 3′ end of DIST1.1R anneals to a four bp mismatch region that is only present in the JAY291 strain.

B. 1% agarose gel of a PCR products generated with the DIST1.1F/DIST1.1R primer pair. Genomic DNA from JAY291 was diluted into YJM789 genomic DNA in 10-fold serial dilutions from 10^{-1} to 10^{-6} in lanes 3 through 8. The same analysis is shown with JAY291 genomic DNA diluted into S288c in lanes 11 through 16. Lanes 17 and 18 contain the products of the same PCR reaction with genomic DNA from spores

JSC2-1-10b and JSC2-1-12d, showing that these samples have the JAY291-specific sequence.

Since the triploid JSC2 had one copy of chromosome V with a wild-type *URA3* gene and two copies of chromosome V with mutant *ura3* genes, we also examined sectored colonies in this strain. No sectored Ura^+/Ura^- colonies were observed in 24,000 examined, suggesting a low rate of chromosome loss ($<4 \times 10^{-5}$). This result is consistent with the measurement of chromosome loss rate of 1.4×10^{-5} obtained by MAYER and AGUILERA (1990). Thus, mitotic loss of chromosomes in the JSC2 triploid prior to sporulation is unlikely to be a substantial contributor to the observed loss in the spores derived from JSC2.

A second argument for a high rate of chromosome loss during mitotic growth of spores of the triploid is that we find a lower rate of chromosome loss when we examine the spores with a method of detection that is more sensitive (SPA versus CGH). For two JSC2 spores, we extended our analysis using a method that was even more sensitive than SPA. By our initial SPA analysis, spore cultures of JSC2-1-10b and JSC2-1-12d were monosomic for chromosome I, although their meiotic sisters were disomic. By SPA, JSC2-1-10b had only the "Y" homologue whereas the sister spore had both the Y and J homologues; for JSC2-1-12d, only the S form was detectable by SPA, although the sister spore had both the S and J homologues. To determine whether cells with the J homologue of chromosome I were present in very low amounts in spore cultures of

JSC2-1-10b and JSC2-1-12d, we designed a pair of primers that would specifically amplify a centromere-linked sequence from the J homologue (primer sequence in Materials and Methods). As shown in Figure A.8, by PCR, we found that the J homologue was present in both spore cultures. A similar assay was used to detect missing chromosome XIII in spores JSC2-1-9c and JSC2-1-14c (data not shown). In summary, taking into account these results, there are maximally 11 meiotic chromosome loss events in JSC2 of a total of 960 total chromosomes, establishing an upper limit for the frequency of meiotic chromosome loss events in triploids as 0.01.

A.3.5 Chromosome pairing during meiosis

As discussed in the Introduction, in near-diploid strains with one trisomic chromosome, all three copies of the trisomic chromosome recombine (Shaffer and Fink 1971). To determine whether the three copies of each homologue in triploid also engage in trivalent pairing, we examined meiotic recombination for four chromosomes (IV, V, IX, and XIII) using the same centromere-linked SNPs as in our previous analysis and SNPs located near (within 50 kb) of the left and right telomeres of each chromosome. Table S A.3 lists the primers and restriction enzymes used for the telomere-linked SNPs, and Table S A.6 summarizes the results. The evidence for trivalent pairing is that at least one spore contains a chromosome in which there has been a crossover between the centromere-linked markers for the S, J, and Y SNPs and the telomere-linked SNPs. One

example of this analysis (for chromosome IV of JSC2-1-8) is shown in Table S A.6 and Figure A.9. In spores with a single copy of chromosome IV (spores -8b and -8c), the coupling of the chromosome IV markers is unambiguous. For spores that are disomic for IV, the coupling relationships of the markers are unclear. For some spores of this type, we screened multiple colonies derived from the original spore colony to identify derivatives that had lost one of the two disomic chromosomes. Analysis of the SNPs in these derivatives clarified the coupling relationships. In this tetrad, there is redundant information demonstrating trivalent pairing. For example, the existence of a single copy of chromosome IV in JSC2-1-8a-2 (a colony derived from the -8a spore) with the centromere-linked Y marker, the left telomere J marker, and the right telomere S marker (Table S A.6) is sufficient to argue for trivalent pairing in this meiosis.

Chromosomes IV and V had trivalent pairing in all ten tetrads. Six of ten tetrads had trivalent pairing for chromosome IX. Four tetrads (JSC2-1-5, -6, -12, and -13) could not be completely diagnosed because we were unable to isolate derivatives of disomic spores that had only copy of IX; in such spores, we could not determine the coupling arrangements of the markers. Chromosome XIII underwent trivalent pairing in nine of ten tetrads, but, in one tetrad (JSC2-1-5), the coupling arrangements of the markers on chromosome XIII also could not be determined. In summary, of 35 events in which we

could unambiguously determine the recombination patterns, the chromosomes showed trivalent pairing in all cases.

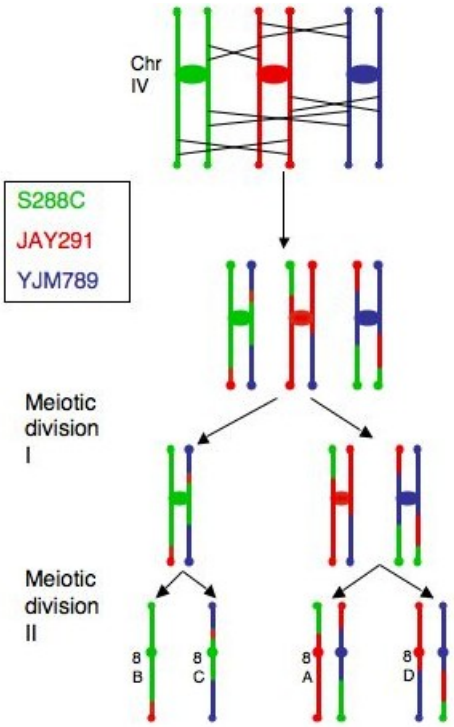


Figure A.9. Chromosome pairing and recombination during meiosis in a triploid

As discussed in the text, using SNPs located near the centromere and telomeres, we mapped crossovers on four chromosomes in ten tetrads (data in Table S A.6). This figure shows patterns of crossovers consistent with the recombination events detected in tetrad JSC2-1-8 on chromosome IV. The centromeres are shown as circles and ovals at the telomeres represent SNPs; the left and right telomeres are at the top and bottom of the chromosome. The individual spore numbers corresponding to those in Table S A.6 are shown at the bottom of the figure. In this tetrad, every chromatid engaged in at least one exchange. The pattern of exchanges is shown in this figure is not the only possible pathway for generating the recombinant chromosomes at the bottom of the picture.

A.4 Discussion

In this study, we showed that meiotic segregation of chromosomes in triploids is usually accurate with two spores receiving one copy and two spores receiving two copies of each homologue. As expected from previous studies, spore viability was low (about 50%) compared to the viability of spores derived from diploids. No particular combinations of two or three disomic chromosomes were favored in the viable spores. The aneuploid strains derived from the spores often had very high rates of mitotic chromosome loss. In most meioses, all three copies of each homologue pair and recombine.

A.4.1 Spore viability

We studied meiotic chromosome segregation in two very different triploids. MH10 has three copies of a haploid set derived from MS71 (STRAND *et al.* 1995) and JSC2 is a triploid derived from three different haploid strains with diverged sequences: S288c, YJM789, and JAY291. Despite this difference, both triploids had similar spore viabilities which were greater than those observed in most past triploid studies (CAMPBELL *et al.* 1981; CAMPBELL and DOOLITTLE 1987; LOIDL 1995; PARRY and COX 1970). The similar spore viability between our strains indicates that the 0.4-1% sequence divergence between homologues in JSC2 does not affect spore viability in triploids.

We examined patterns of chromosome segregation to determine whether certain combination of disomic chromosomes led to spore inviability. All 120 two-chromosome disomy combinations and all 560 three-chromosome disomy combinations within the MH10 and/or JSC2 tetrads were observed in viable spores (Table A.1and Table A.2). No combination of two or three coincident disomies were significantly overrepresented or underrepresented in the viable spores (analysis described in Supporting Information), consistent with previous studies (CAMPBELL *et al.* 1981).

PARRY and COX (1970) observed that most triploid-derived spores were aneuploid for five chromosomes or less, with an average of 2.6 disomes per spore. They suggested that either triploid segregation produced two near-haploid and two near-diploid spores or that aneuploid chromosomes were rapidly eliminated during vegetative growth of the spore cultures. Our results demonstrate that the second explanation is correct, although the strains in our study seem more tolerant of high levels of disomy (Figure A.6) than those examined by Parry and Cox.

A.4.2 Patterns of chromosome loss

Complete or partial loss of disomic chromosomes during the vegetative growth of spore cultures was very common in our study. Two types of observations demonstrate a high rate of chromosome loss. First, the loss events (chromosomes with less than six copies in the tetrad) were much more common when we analyzed the spore

cultures by CGH than when we analyzed them by the more sensitive SPA procedure; for six tetrads of JSC2, both methods were used on the same DNA samples. Second, we directly showed that subcultures of some of the disomic spores had unstable karyotypes (Table S 2.5 and Figure A.7).

In MH10-derived spores, we frequently observed that one sister spore was disomic for a particular chromosome and the other was not, suggesting a chromosome loss event. There were 53 such loss events (indicated in boldface in Table A.1). These loss events were non-randomly distributed among the spores; 70% of the losses were in five of the 40 spores. In addition, the two sister spores often had very different loss rates. One extreme example is the sister spores MH10-4a and MH10-4c. The MH10-4a spore lost 11 chromosomes (becoming a euploid haploid strain), whereas the MH10-4c spore did not lose any chromosomes.

Two different statistical tests were applied to these data (Table S A.7). First, for each of the 40 spores, we determined whether there was a significantly elevated rate of chromosome loss compared to the expected rate of loss per disomic chromosome for all 40 spores (0.2 loss events/disomic chromosome). By the chi-square goodness-of-fit test, spores MH10-4a, -22d, -25a, -28b, -32b, and -37c had significantly elevated rates. Second, we used the Fisher exact test to compare chromosome loss events in pairs of sister spores. By this test, five pairs of sister spores had significantly different numbers of loss

events (Table S A.7). Thus, multiple pairs of sister spores in MH10, which presumably initially had identical karyotypes and genotypes, had different rates of chromosome loss. We suggest two explanations of this interesting observation. One possibility is that certain spores receive a partially defective spindle pole body and this defect persists for a number of generations following spore germination, leading to an elevated rate of chromosome loss. An alternative possibility is that a stochastic loss of one of the disomic chromosomes from a spore results in a much elevated rate of loss of other chromosomes.

Several more points concerning these observations should be made. First, in analyzing the MH10 data, we did not consider chromosome loss events in which all four spores were monosomic, since we could not determine which of the four spores had lost chromosomes. Second, the observation that certain spores have very high rates of chromosome loss in MH10 was considerably less obvious in the JSC2 data (Table S A.8). Of the 24 JSC2 spore cultures examined by microarrays, we detected 33 examples of loss in one sister but not the other (total of 160 events). Five spores had lost ≥ 5 chromosomes, whereas the other 19 spores had lost ≤ 2 chromosomes. Despite this asymmetry, none of the spores had a significantly elevated rate of chromosome loss compared to the average loss rate for the six tetrads (0.21 events/disomic chromosome). Since the sister spores of JSC2 are not isogenic (unlike those of MH10), it is possible that

genotype-dependent mechanisms of chromosome loss prevent the detection of genotype-independent mechanisms of loss.

Another important point to emphasize is that we do not know the mechanism of chromosome loss. If the loss of the disomic chromosome is a consequence of non-disjunction, then we should observe trisomic derivatives at the same frequency as monosomic derivatives. Trisomic strains were not observed, although it is possible that such strains would have a severe growth disadvantage. Although we found a few examples of monosomic chromosomes becoming disomic consistent with chromosome non-disjunction events (Table A.1 and Table A.2), such events were much less frequent than chromosome losses.

An alternative mechanism for chromosome loss is failure to replicate one or both of the disomic chromosomes before cell division. Although we have no direct evidence in favor of this mechanism, the pattern of chromosome loss diagrammed in Figure A.7 has several branch points in which symmetrical chromosome loss (consistent with a defect in chromosome replication) is postulated. For example, we show a symmetric loss of chromosome VI and XV at the first branch point. This pattern is consistent with failure to replicate VI and XV, and segregation of the unreplicated homologues to the two daughter cells. To prove this mechanism, however, one would need to perform pedigree analysis on the aneuploid strains.

TORRES *et al.* (2007) constructed thirteen yeast strains, each disomic for a different yeast chromosome. All disomic strains had proliferative defects, increased glucose uptake, and were sensitive to cycloheximide, a protein synthesis inhibitor. Strains disomic for chromosome IV and XV were among those with the most profound proliferative defect, and these chromosomes were among those most frequently lost from the MH10 tetrads. TORRES *et al.* (2007) were unable to isolate a strain that was disomic only for chromosome VI, and suggested that such strains could not be recovered because of the imbalance between levels of beta tubulin (located on chromosome VI) and alpha tubulin (located on chromosome XIII). In support of this suggestion, they were able to obtain strains disomic for both VI and XIII. Similarly, in our experiments, chromosome VI was frequently lost from MH10 tetrads (Table A.1), although we were able to recover strains that were disomic for VI and monosomic for XIII.

The rate of chromosome loss that we observed in the spore cultures is clearly much greater than that of wild-type diploid cells. The rate of chromosome V loss in a wild-type diploid is about 10^{-5}/cell division (HARTWELL and SMITH 1985). Since many of the disomic chromosomes were lost in less than 50 cell divisions, the lost rate is likely to be considerably greater than 10^{-2}/division. This high loss rate probably reflects defective structures involved in segregation or replication of the chromosomes resulting from

non-stochiometric amounts of the proteins required to form these structures. In addition, since most disomic yeast strains grow more slowly than wild-type (TORRES *et al.* 2007), those strains that lose chromosomes have a selective advantage. In our study, two of the haploid strains derived by sporulating MH10 (MH10-4a and MH10-22d) lost all disomic chromosomes to revert to haploidy (Table A.1).

There is also evidence that chromosome segregation may be inherently less accurate in triploids and tetraploids compared to diploids. GERSTEIN *et al.* (2008) observed that subcultured triploid and tetraploid strains underwent a reduction in chromosome number to diploidy. Since this reduction occurred quickly, they suggested that there was a mechanism of concerted chromosome loss. MAYER and AGUILERA (1990) showed that triploids had chromosome loss rates elevated about 30-fold compared to a diploid (ANDALIS *et al.* 2004; MAYER and AGUILERA 1990). In addition, Mayer and Aguilera found that loss of one chromosome from the triploid or tetraploid increased the likelihood of loss of a second chromosome. STORCHOVA *et al.* (2006) showed that tetraploids have an elevated level of monopolar kinetochore attachments and suggested that tetraploids had partially defective spindle pole bodies, spindles and/or kinetochores as a consequence of defective "scaling" of complex structures in yeast cells with high ploidies.

A.4.3 Chromosome pairing and recombination

We first examined whether there was any evidence of preferential segregation of homologues derived from the three strains used to construct JSC2. More specifically, using the centromere-linked polymorphism data (Table A.2), we tested each of the sixteen homologues for whether the disomic spores had a significant excess of the chromosomes derived from any of the three possible pairs of S288c, YJM789 or JAY291. After correcting for the rate of false discoveries with the Benjamini-Hochberg correction (BENJAMINI and HOCHBERG 1995), no significant strain-specific selection for any pairing combination was observed (detailed discussion of the statistical analysis in Supporting Information).

Although meiotic recombination in triploids has not been examined in any detail, several studies have shown that all three copies of a trisomic chromosome in a near-diploid strain frequently engage in recombination in a single meiosis (CULBERTSON and HENRY 1973; KOLLER *et al.* 1996; RILEY and MANNEY 1978; SHAFFER *et al.* 1971). In our study, we examined meiotic recombination on chromosomes IV, V, IX, and XIII. As discussed above, in all of the 35 events in which we could unambiguously diagnose meiotic exchanges, all three copies of these homologues recombined. Our results are consistent with the cytological studies of LOIDL (1995) in which he observed that all three homologues in triploids were involved in synaptonemal complex formation.

A.5 Acknowledgments

We would like to thank all members of the Petes and Jinks-Robertson labs for their discussions and technical help, and we thank John McCusker, Jennifer McCulley, Hengshan Zhang, Joseph Heitman, and Sue Jinks-Robertson for comments on the manuscript. We thank Lucas Argueso and Piotr Mieczkowski for generating the microarrays, and Einat Hazkani-Covo for help with the statistical analyses. This research was supported by grants from the National Institutes of Health (GM24110, GM52319, and 1RC1ES018091-01).

A.6 Supporting Information

A.6.1 Supporting Materials and Methods

A.6.1.1 Statistical analysis of preferential co-segregation of two different homologues into sister spores

Following a triploid meiosis, there are only two ways to distribute the twelve copies of two different homologous chromosomes within a single tetrad: 1) double disomes (two sister spores disomic for both chromosomes and the remaining two sister spores monosomic for both chromosomes) or 2) mixed disomes (two sister spores disomic for one homologue and monosomic for the other with the other sister pair receiving the reciprocal arrangement of chromosomes). If segregation of the chromosomes into four-viable spore tetrads is random, these two events should occur at

equal frequencies. Among the MH10/JSC2 data set, for all 120 two-chromosome combinations, we found some tetrads with double-disomes (indicating no double-disome combination is lethal). In addition, since we also recovered some tetrads with mixed disomes for all two-chromosome combinations, there is no situation in which disomy for one type of homologue requires disomy for another type of homologue for viability.

We calculated whether the numbers of double disomes and mixed disomes were significantly different for each two-chromosome combination by comparing the number of tetrads in each class for each homologue with the expectation of equality by using the chi-square "Goodness of Fit" test on the Vassar Stat Website (http://faculty.vassar.edu/lowry/VassarStats.html). After correcting for the false-discovery rate (BENJAMINI and HOCHBERG 1995), no significant departures from equality were observed. We point out, however, that because of the relatively small number of tetrads examined and the large number of comparisons examined, this statistical test is not particularly sensitive.

A.6.1.2 Statistical analysis of preferential co-segregation of three different homologues into sister spores

There are 560 combinations of three-chromosome disomes. Among the tetrads of MH10 and JSC2, we found all 560 of these combinations, demonstrating that no combination of three-chromosome disomes is lethal.

There are four possible ways to distribute 18 copies of three homologues within a tetrad as shown below:

	Chromosome 1	Chromosome 2	Chromosome 3
Meiotic pair 1 Meiotic pair 2	monosome disome	monosome disome	monosome disome
Meiotic pair 1 Meiotic pair 2	monosome disome	disome monosome	disome monosome
Meiotic pair 1 Meiotic pair 2	monosome disome	disome monosome	monosome disome
Meiotic pair 1 Meiotic pair 2	monosome disome	monosome disome	disome monosome

As with the two-chromosome combinations, each of these situations should occur at equal frequencies if presence of a disome is independent of all other disomes. Because there were so many missing homologues in the MH10 experiment, we tested this expectation with only the JSC2 data. Using the chi-square "Goodness of Fit" test and the BENJAMINI-HOCHBERG correction as described above, we did not detect any significant deviation from random. As discussed above, because of the relatively small numbers of tetrads examined, this test is not sensitive.

A.6.1.3 Statistical analysis of the non-random segregation of disomic chromosomes in JSC2 based on whether the chromosomes are derived from S288c, JAY291, or YJM789

We also examined whether disomes had a non-random pattern with respect to which genetic background they were derived from. For each spore with a disomic chromosome, there are three possible patterns: 1) one chromosome derived from S288c and one derived from JAY291, 2) one chromosome derived from S288c and one derived

from YJM789, and 3) one chromosome derived from JAY291 and one derived from YJM789. In the absence of preferential segregation, these three patterns should be found with equal frequencies. We examined the JSC2 tetrad data for any deviation from random pairing for all 16 chromosomes with both the chi-square "Goodness of Fit" test and the BENJAMINI-HOCHBERG method as described above. No evidence for preferential segregation was obtained.

A.6.2 Supporting tables

Table S A.1. Strain genotypes and constructions for triploid study

Strain name	Genotype	Strain construction or reference
MS71	*MATα ade-1 his7-2 leu2-3 LEU2 trp1-289 ura3-52*	(STRAND *et al.* 1995)
JLMy100	*MAT***a***/MATα ade5-1/ade5-1 his7-2/his7-2 leu2-3/leu2-3 LEU2/LEU2 trp1-289/trp1-289 ura3-52/ura3-52 + pSAD3-3b/MEC1*	(J. MCCULLEY, Duke University)
JLMy133	*MAT***a***/matα::URA3 ade5-1/ade5-1 his7-2/his7-2 leu2-3/leu2-3 LEU2/LEU2 trp1-289/trp1-289 ura3-52/ura3-52*	Transformation of JLMy100 with PCR fragment with *matα::URA3*. The template was genomic DNA from a *URA3* yeast strain and primers were Malpha2/URA3F and Malpha1/URA3R (BARBERA and PETES 2006)
MH10	*MAT***a***/MATα/matα::URA3 ade5-1/ade5-1/ade5-1 his7-2/his7-2/his7-2 leu2-3/leu2-3/leu2-3 trp1-289/trp1-289/trp1-289 ura3-52/ura3-52/ura3-52 LEU2/LEU2/LEU2 CAN1/CAN1/CAN1*	Cross of JLMy133 and MS71
JAY291	*MAT***a**	(ARGUESO *et al.* 2009)
FY834	*MATα ura3-52 leu-2Δ1 trp1-Δ63 his3-Δ200 lys2-Δ202*	(WINSTON *et al.* 1995)
JAY306	*MATα/MAT***a** *URA3/ura3-52 LEU2/leu2-Δ1 TRP1/trp1-Δ63 HIS3/his3-Δ200 LYS2/lys2-Δ20*	Cross between JAY291 and FY834
PSL4	*MATα ade2-1 ura3 gal2 ho::hisG*	(LEE *et al.* 2009)
JSC1	*matα::NAT/MAT***a** *URA3/ura3-52 LEU2/leu2-Δ1 TRP1/trp1-Δ63 HIS3/his3-Δ200 LYS2/lys2-Δ20*	Transformation of JAY306 with a PCR fragment (*matα::NAT*^{R}). The PCR template was pAG25 (GOLDSTEIN and MCCUSKER 1999). The primers were MatAlpha Nat F and MatAlpha Nat R (LEE *et al.* 2009).

JSC2	*MAT***a**/*MAT*α/*mat*α::*NAT URA3/ura3-52/ura3 LEU2/LEU2/leu2-Δ1 TRP1/TRP1/trp1-Δ63 HIS3/HIS3/his3-Δ200 LYS2/LYS2/lys2-Δ20 ADE2/ADE2/ade2-1 GAL2/GAL2/gal2*	Cross between JSC1 and PSL4

Table S A.2 Primers and restriction enzymes used to distinguish centromeric sequences derived from strains S288c, JAY291, and YJM789[11]

			Diagnostic restriction enzymes and fragment sizes following digestion of PCR product			
SGD coordinates for forward primer[12]	Primer name[13]	Primer sequence (5′ to 3′)	Restriction Enzyme	Size of PCR fragment after treatment with restriction enzyme for each strain (bp)		
				S288C	JAY291	YJM789
149166	CEN1 PM F	TCTTTCAGGATGTA CCTTTGGATAGC	*Xcm*I	549	150, 399	150, 399
	CEN1 PM R	TACAACGGACAGC GATGGAC	*Hpa*I	256, 293	256, 293	549
239655	CEN2 PM F	CCAAAATACTCCTT AATGACGAAAG	*Bsr*I	158, 324	482	158, 324
	CEN2 PM R	AGATCTCAGCCACC ATTCCTAAC	*Dra*I	482	89, 393	89, 393
115864	CEN3 PM F	TCCCAACTCCTTCG GCTAATATG	*Bsl*I	143, 359	502	502
	CEN3 PM R	ATGCGCAGGTGAG CTGATTG	*Pst*I	502	502	131, 371
451010	CEN4 PM F	GGCACATACGCCCT CAAACC	*Bst*XI	543	56, 487	543
	CEN4 PM R	TGGTTGGGACTCCT ATATGATGG	*Bsm*I	119, 424	543	543
151849	CEN5 PM F	TTGGTAAACAAAG GGCCAAGC	*Taq*$^{\alpha}$I	447	100, 347	447
	CEN5 PM R	ATGTGCGGCTTTGT CAGCAG	*Cac*8I	132, 315	132, 315	447
148329	CEN6 PM F	AATTTGGGCGATGG AAGAGG	*Bbs*I	165, 367	532	532

[11] As described in the text, we could diagnose strain-specific SNPs for the S288c-, JAY298-, and YJM789-derived chromosomes by PCR amplification of the chromosomal region containing the SNP, followed by restriction digests of the resulting fragments.
[12] The Saccharomyces Genome Database (SGD) coordinate is for the first base of the forward primer.
[13] As discussed in the text, this primer pair was used to characterize a meiotic gene conversion in tetrad 8 of JSC2.

	CEN6 PM R	TTCTGCTTTCCTTCC CAAACAG	*Eco*RI	532	177, 355	532
499780	CEN7 PM F	CTTTGCATGCCCGT CACTTG	*Mwo*I	341	141, 200	341
	CEN7 PM R	TCAGATTTGAGTTA GCCCGTTTAGAG	*Dra*III	341	121, 220	121, 220
104245	CEN8 PM F	GGAAATAAACGCG CAAATAAACC	*Nco*I	469	178, 291	178, 291
	CEN8 PM R	TTTGGCAATTGAGG AGGTATGG	*Hph*I	469	190, 279	469
355523	CEN9 PM F	CGGATGAATATTGC CAATGACTG	*Dra*I	173, 173	173, 173	346
	CEN9 PM R	CAGCTGAAGCTTGC CTCTGTATG	*Ase*I	154, 192	346	346
434671	CEN10 PM F	TGCGTCTTGTTCCG AGCTAAAG	*Nsi*I	368	161, 207	161, 207
	CEN10 PM R	CGGATTCCGTTGCG AGTG	*Mfe*I	171, 197	368	171, 197
437899	CEN11 PM F	TTTGATGGGTCAAC GACAGG	*Acl*I	274	45, 229	274
	CEN11 PM R	CAAAGTGTTCAACC GAGCTTGC	*Hinf*I	78, 196	274	274
147404	CEN12 PM F	CAGAGGAAGGCGC AATAGAAGC	*Bcg*I	30, 137, 436	604	30, 137, 436
	CEN12 PM R	CCGGAGGATTGTG ACCCAAC	*Pst*I	75, 529	75, 529	604
152181	CEN12.2 PM F[c]	GGATGGATAACTC CCTGGTTGA	*Apo*I	68, 129	68,129	196
	CEN12.2 PM R	TTCGAAGACCTTAT TTATTCCAACA				
267890	CEN13 PM F	AGGGATGTGATGG CCCTTTC	*Mfe*I	363	78, 285	363
	CEN13 PM R	TCAACTCTAGCCAA TTTATCCTGTCG				

267890	CEN13 PM F	AGGGATGTGATGG CCCTTTC	PCR[14]	258	249	249
	CEN13. 2 PM R	TTTAAAGTTCGGAA CGCATACAC				
628382	CEN14. 1 PM F	GGCTGAGTATGGC GCAATTAG	*Bsr*I	299	299	129, 170
	CEN14. 1 PM R	TCCGATTTATTATA GGGTTGCAAAT				
627401	CEN14. 4 PM F	TTCGTTGCTGCATC GTTGTC	PCR	151	164	164
	CEN14. 5 PM R	AACAGTAAACAGA TCTTCCTTGTCCTC				
326813	CEN15. 1 PM F	AAGCACCCAATAA GTGCGTTATG	*Bbv*I	87, 230	317	87, 230
	CEN15. 1 PM R	CAGCTTATCCAGCG AATGGTCTG				
328652	CEN15. 2 PM F	AGTCTGCCTTAGCA AACTTGGAG	*Sau*3AI	137, 173	310	310
	CEN15. 2 PM R	TCGCTAGATGATGG GTCGAATC				
555679	CEN16 PM F	GATGATTATAGGCG CATTATAGAGTCG	*Bsm*AI	46, 149	46, 149	195
	CEN16. 2 PM R	CGACAACATAACG GCTTCAACC	*Dra*I	195	81, 114	81, 114

[14] For these PCR products, we diagnosed the source of the homologue by measuring the size of the PCR product rather than by restriction enzyme digestion of the product.

Table S A.3 Primers and restriction enzymes used to distinguish telomeric sequences of chromosomes IV, V, IX, and XIII derived from strains S288c, JAY291, and YJM789[15]

SGD coordinates for forward primer	Primer Name	Primer sequence (5′ to 3′)	Restriction enzymes used for diagnosis and strains diagnosed			
			Restriction Enzyme	Size of PCR fragment after treatment with restriction enzyme for each strain (bp)		
				S288c	JAY291	YJM789
48202	LA 4.1 F	TTCTGCAAACCTCATTTGAGTAGC	*Hph*I	70, 83, 134	287	287
	LA 4.1 R	AAACCTATGGTTTACGAAATGATCC				
49577	LA 4.2 F	TGATAAAGGAACCATGTGATCTTACG	*Dde*I	169	169	69, 100
	LA 4.2 R	GTGCGTAGAATAACGAACATTCATAGC				
1515307	RA 4 F	AAAGTGGGATTCTGCCTGTGG	*Pvu*I	143, 242	385	385
	RA 4 R	GCGACCTGGGTTTACATTTCC	*Bsa*I	140, 245	140, 245	385
24783	LA 5 F	CACTTGAGGCCACGCATACTG	*Hin*dIII	402	402	120, 282
	LA 5 R	GCAACGTTTGGGAAGAAAACG	*Eco*RV	136, 266	402	136, 266
560490	RA 5 F	TTCTCAGCCGTACAATCATGC	*Taq*$^{\alpha}$I	491	55, 436	491
	RA 5 R	AAACTCCTTCCAAAGGGTCTGG	*Eco*RI	491	491	221, 270
27145	LA 9 F	TTGCCTTTCGTGTCTATTTGC	*Hph*I	380	69, 311	69, 311

[15] As in Table S A.2, we show the primers and restriction enzymes required to detect sequence-specific SNPs from the three haploid strains used to construct JSC2. LA and RA indicate the SNPs located near the left and right telomeres, respectively. The chromosome number is indicated in the primer name.

	LA 9 R	TGCCTCGATATACAAAT CGTACTGC	*Ahd*I	157, 223	157, 223	380
393735	RA 9 F	CGAGGGATAATAACTG ATGAATAGGG	*Taq*$^{\alpha}$I	111, 207	318	318
	RA 9 R	TGGGATTCAAGGCATCA TCG	*Bsm*AI	318	84, 234	318
13879	LA 13.1 F LA 13.1 R	TGTGTGTGTGAACTATG GCTACTGG GCAATGAAGTCACTTTC GTGATGG	*Bam*HI	105, 114	105, 114	219
13088	LA 13.2 F LA 13.2 R	TTTGGTCGGTGGCGATG TAG TTGTTCAATCATCCTTGA AGCTAAG	*Bsr*I	56, 155	211	56, 155
905692	RA 13.1 F RA 13.2 R	GGAAGAACTCGGGTGT CAGAGC TCAACGCAGGAAGGGA ACG	*Mwo*I	77, 148	225	77, 148
904665	RA 13.3 F RA 13.3 R	TGTTGTAGGCTTGGCCA TCG CCATCCTCGCCGTTTAT CTCG	*Hin*dIII	122, 197	319	319

Table S A.4 Patterns of spore viability in the triploid strains JSC2 and MH10[16]

Strain name	Total tetrads dissected	4 live: 0 dead	3 live: 1 dead	2 live: 2 dead	1 live: 3 dead	0 live: 4 dead	Spore viability
JSC2-1	77	13	22	22	12	8	0.56
JSC2-2	79	11	15	28	19	6	0.52
MH10	135	24	17	58	19	17	0.52
Totals	291	48	54	108	50	31	0.53
Observed frequencies		0.16	0.19	0.37	0.17	0.11	
Expected frequencies[17]		0.08	0.28	0.37	0.22	0.05	

[16] Tetrads of JSC2 and MH10 were dissected onto solid rich growth medium and the patterns of spore viability monitored by colony growth. JSC2-1 and JSC2-2 are isogenic and were generated by mating the same two parental strains, and picking two different zygotes.

[17] Expected frequencies were calculated by a binomial expansion based on the observed spore viability frequency of 0.53.

Table S A.5 CGH analysis of disomy and monosomy in spores derived from six tetrads of the triploid JSC2[18]

	Chromosomes															
Spore	I	II	III	IV	V	VI	VII	VIII	XI	X	XI	XII	XIII	XIV	XV	XVI
5a	2[19]	2	1	1	1	1	2	2	2	1	**1**	1	1	2	2	2
5b	2	1	2	1	2	1	1	1	1	2	1	2	2	1	1	1
5c	2	1	2	1	2	1	1	1	1	2	1	2	2	1	1	1
5d	1	2	1	1	1	1	**1**	**1**	2	1	2	1	1	2	**1**	**1**
6a	1	1	1	1	1	2[19]	1	1	2	1	**1**	1	1	**1**	1	1
6b	2	**1**	2	2	2	2	**1**	2	1	**1**	1	**1**	**1**	1	1	1
6c	1	1	1	1	1	1	1	1	2	1	2	1	1	2	1	1
6d	2	2	2	2	2	2	2	2	1	2	1	2	2	1	1	1
8a	1	2	2	2	1	1	1	2	1	1	1	2[19]	2	**1**	1	2
8b	2	1	1	1	2	2	**1**	1	2	1	**1**	2	1	1	1	1
8c	2	1	1	1	2	2	2	1	2	1	2	2	1	1	1	1
8d	1	**1**	**1**	**1**	1	1	1	2	1	1	1	1	**1**	2	1	**1**
10a	2	2	2	1	2	1	2	1	1	2	1	2	1	2	1	1
10b	**1**	**1**	**1**	1	2	1	**1**	1	1	2	1	2	1	**1**	1	1
10c	1	1	1	2	1	2	1	2	2	1	2	2[19]	2	1	2	1
10d	1	1	1	**1**	1	**1**	1	2	2	1	**1**	1	**1**	1	2	1
12a	1	2	1	1	1	2	1	2	2	1	2	1	1	2	2	2
12b	1	2	1	1	1	2	1	2[b]	2	1	2	1	1	2	2	2
12c	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
12d	**1**	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
13a	1	2	1	**1**	2	1	**1**	2	1	1	1	2	1	1	1	1

[18] Based on CGH microarrays, we determined whether the spore cultures were disomic (indicated by 2 in the table) or monosomic (indicated by 1) for each homologue. Boldface indicates a tetrad in which one sister spore was monosomic, although the other sister was disomic, indicating a chromosome loss event in the monosomic spore.

[19] In these spores, there was a possible chromosome gain. The extra chromosomes were not counted in the averages at the bottom of the table.

13b	2	1	2	1	1	1	1	1	2	2	2	1	2	2	1	2
13c	2	1	2	1	1	1	1	1	2	2	**1**	2[19]	2	2	1	2
13d	1	2	1	2	2	1	2	2	1	1	1	1	1	1	1	1
Ave.	**5.7**	**5.5**	**5.7**	**4.8**	**5.7**	**5.0**	**4.8**	**5.8**	**6.0**	**5.2**	**5.2**	**5.5**	**5.2**	**5.5**	**4.8**	**5.0**

Table S A.6 Patterns of marker exchange on chromosomes IV, V, IX, and XIII in spores derived from the triploid JSC2[20]

	Chromosomes											
	IV			V			IX			XIII		
Spore	LT	CEN	RT	LT	CEN	RT	LT	CEN	RT	LT	CEN	RT
5a	S	Y	S	J	Y	J[21]	J Y	S J	S J	S	S	S
5b	J	S	J		S J	J J	S	Y	Y	J Y	J Y	J Y
5c	S	S	S	S Y	S J	S Y	S	Y	Y	J Y	J Y	S Y
5d	J Y	J Y	Y Y	Y	Y	S	J Y	S J	S J	S	S	J
5d-7[c]	J	J	Y									
6a	Y	S	S[21]	S	S	S	S J	J Y	S J	S	Y	S
6b	S J	J Y	S Y	J J	J Y	S J	S	S	S	J Y	S J	S Y
6c	J	S	Y	Y	S	Y	J Y	J Y	J Y	S	Y	J
6d	S Y	Y	S J	S Y	J Y	J Y	Y	S	Y	J Y	S J	J Y
8a	J	J Y	S J	J	J	Y	J	Y	Y	S Y[21]	S J	S Y
8a-2	J	Y	S									
8b	S	S	J	S Y	S Y	J Y	J Y	S J	S J	J	Y	J
8c	Y	S	Y	J Y	S Y	S J	S S	S J	S J	Y	Y	J
8d	J Y	J Y	S Y	S	J	S	Y	Y	Y	J J	S J	S Y
8d-1	J	J	Y									
8d-2	Y	Y	S									
9a	S	S	S	S	S	J	SS	S Y	S Y	J	S	S
9b	S Y	J Y	S J	Y	Y	S Y	J	J	J	S J	J Y	S J
9c	J Y	J Y	Y Y	J J	J Y	S Y	Y	J	J	Y	J	Y
9d	J	S	J	S	S	J	J Y	S Y	S Y	S	S	J
10a	J	J	Y	J J	J Y	J Y	S	J	Y	S[21]	S	Y
10b	S	J	S	S Y	J Y	J Y	Y	J	J	Y	S	J
10c	S Y	S Y	J Y	S	S	S	S Y	S Y	S Y	J Y	J Y	S J
10d	J	Y	S	Y	S	S	J J	S Y	S J	Y	Y	Y

[20] As described in the text, we examined the patterns of marker segregation using SNPs located near the centromere (CEN), the left telomere (LT), and the right telomere (RT). As in other tables, chromosomes that are derived from S288c, JAY291, and YJM789 are indicated by S, J, and Y, respectively. The data from the original spore cultures are indicated by a number and a letter (for example, 15c). Strain names that contain an additional hyphen and a number are single colony isolates derived from the original spore culture (for example, 15c-1).

[21] This superscript indicates a tetrad with a meiotic gene conversion event. For example, in tetrad 5, for the SNP at the right end of chromosome V, the Y-associated SNP was converted to a J-associated SNP.

11a	Y	J	J	S Y	S Y	S Y	J	J	S	J	Y	J
11b	J	S Y	S Y	J	J	Y	S J	S Y	J Y	S Y	S J	S S
11c	S S	S Y	S J	S	J	J	Y Y	S Y	S Y	S J	S J	J Y
11d	Y	J	Y	J Y	S Y	S J	S	J	J	Y	Y	Y
12a	J	J	J	Y	J	S	J Y	S Y	S Y	Y	J[22]	S
12b	S	J	Y	S	J	Y	S Y	S Y	S Y	J	J	J
12c	S J	S Y	S J	J Y	S Y	J Y	J	J	J	J Y	S S	J Y
12d	Y Y	S Y	S Y	S J	S Y	S J	S	J	J	S S	S S	S Y
13a	Y Y	S J	Y Y	S Y	S Y	S J	S	J	J	J	S	S[21]
13a-1				S	Y	J						
13a-2				Y	S	S						
13b	J	Y	J	J	J	Y	J Y	S Y	S Y	S J	J Y	J Y
13c	S	Y	J	J	J	J	J Y	S Y	S Y	Y Y	J Y	Y Y
13d	S J	S J	SS	S Y	S Y	S Y	S	J	J	S	S	S
14a	Y	S	J	J	Y	J[21]	S	J	J	J	S	S
14b	J	S	S	J	Y	S	J	J	S	J	S	S
14c	J Y	J Y	S Y	Y Y	S J	J Y	S J	S Y	S Y	S ?[23]	J Y	Y ?[23]
14d	S S	J Y	J Y	S S	S J	S J	Y Y	S Y	J Y	S Y	J Y	J Y
15a	Y	Y	S	J	J	J	S	J	J	S J	J Y	S Y
15b	S J	S J	J Y	S Y	S Y	S Y	J Y	S Y	S Y	Y	S	S
15c	S J	J	Y	Y	Y	J Y	J Y	S Y	S J	S	S	J
15c-1	J	S	J	Y	S	J						
15c-2	S	J	Y	S	J	Y						
15d	Y	Y	S	J	J	S	S	J	Y	J Y	J Y	J Y

[22] In this tetrad, the CEN-linked SNP for chromosome XIII underwent a double meiotic conversion or a mitotic gene conversion.

[23] The spore culture JSC2-1-14c had much more of the J-derived chromosome than the Y-derived chromosome; the YJM789-derived chromosome could only be detected with primers designed to specifically amplify the YJM789 SNP. Consequently, we detected only one allele for the left and right telomere SNPs JSC2-1-14c.

Table S A.7 Statistical analysis of loss of disomic chromosomes in spores derived from MH10[24]

Spore name[25]	Initial # disomic chrom.[26]	# chrom. Lost[27]	Freq. of loss[28]	Exp. # of lost chrom.[29]	(χ^2)[30]	p value for χ^2	p value for sister spore test[31]
4a	11	11	1.0	2.2	39	<0.0001[32]	**<0.0001[32]**
4c	11	0	0	2.2	1.6	0.2	
4b	4	0	0	0.8	0.1	0.7	1
4d	4	0	0	0.8	0.1	0.7	
22a	5	0	0	1	0.3	0.6	**1**
22c	5	0	0	1	0.3	0.6	
22b	3	0	0	0.6	0.02	0.9	0.1
22d	3	3	1.0	0.6	7.5	0.006[32]	
24a	4	0	0	0.8	0.1	0.7	**1**
24b	4	0	0	0.8	0.1	0.7	
24c	6	1	0.2	1.2	0	1	1
24d	6	1	0.2	1.2	0	1	

[24] This analysis is based on using CGH microarrays to examine disomy and monosomy in spores derived from the MH10 triploid. For this analysis, we count only those chromosomes that were disomic in one spore of the meiotic sister pair and monosomic in the other. This condition allows us to diagnose chromosome loss unambiguously. If a homologue is present in one copy in all spores, we are unable to determine which spores lost the chromosomes.

[25] One pair of sister spores is shown in boldface and the other shown without boldface.

[26] The total number of homologues that are present in two copies in at least one of the two sister spores.

[27] The number of homologues present in one copy in the spore that were present in two copies in the sister spore.

[28] The number of chromosomes lost divided by the initial number of disomic chromosomes.

[29] The expected number of chromosomes lost was calculated by multiplying the initial number of disomic chromosomes by 0.20. The value of 0.20 is the number of chromosomes lost for the whole data set (53) divided by the number of disomic chromosomes for the whole data set (264).

[30] The "goodness of fit" chi-square value was calculated from the observed numbers of chromosomes lost and retained, and the expected numbers of chromosomes lost and retained.

[31] These p values were calculated by comparing the number of lost and retained chromosomes in each member of the pair of meiotic sisters by the Fisher exact test. The boldface numbers represent the calculation for the meiotic sisters shown in boldface in the leftmost column.

[32] Significant difference at a p value of <0.05 after application of the Benjamini-Hochberg correction BENJAMINI, Y., and Y. HOCHBERG, 1995 Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. J. R. Statist. Soc **57:** 289-300..

25a	9	6	0.7	1.8	9.5	0.002[32]	**0.009[32]**
25d	9	0	0	1.8	1.2	0.3	
25b	7	0	0	1.4	0.7	0.4	1
25c	7	0	0	1.4	0.7	0.4	
28a	9	0	0	1.8	1.2	0.3	**0.009[32]**
28c	9	6	0.7	1.8	9.5	0.002[32]	
28b	4	1	0.3	0.8	0	1	1
28d	4	0	0	0.8	0.1	0.8	
29a	11	0	0	2.2	1.6	0.2	**1**
29d	11	0	0	2.2	1.6	0.2	
29b	4	1	0.3	0.8	0	1	1
29c	4	0	0	0.8	0.1	0.8	
30a	5	0	0	1	0.3	0.6	**1**
30c	5	0	0	1	0.3	0.6	
30b	11	0	0	2.2	1.6	0.2	1
30d	11	1	0.1	2.2	0.3	0.6	

Table S A.8 Statistical analysis of loss of disomic chromosomes in spores derived from JSC2[33]

Spore name	Initial # disomic chrom.	# chrom. lost	Freq. of loss	Exp. # of lost chrom.	χ^2	p value for χ^2	p value for sister spore test
5a	8	1	0.1	1.7	0.02	0.9	**0.3**
5d	8	4	0.5	1.7	2.6	0.1	
5b	6	0	0	1.2	0.6	0.5	1
5c	6	0	0	1.2	0.6	0.5	
6a	3	2	0.7	0.6	1.6	0.2	**0.4**
6c	3	0	0	0.6	0.03	0.9	
6b	11	5	0.5	2.3	2.7	0.1	0.04
6d	11	0	0	2.3	1.8	0.2	
8a	7	1	0.1	1.4	0	1	**0.1**
8d	7	5	0.7	1.4	8.1	0.004	
8b	7	2	0.3	1.4	0	1	0.5
8c	7	0	0	1.4	0.8	0.4	
10a	8	0	0	1.7	1	0.3	**0.03**
10b	8	5	0.6	1.7	6.1	0.01	
10c	7	0	0	1.4	0.8	0.4	0.07
10d	7	4	0.6	1.4	3.7	0.06	
12a	8	0	0	1.7	1	0.3	**1**
12b	8	0	0	1.7	1	0.3	
12c	2	0	0	0.4	0	1	1
12d	2	1	0.5	0.4	0.02	0.9	
13a	5	2	0.4	1	0.3	0.6	**0.4**
13d	5	0	0	1	0.3	0.6	
13b	8	0	0	1.7	1	0.3	1
13c	8	1	0.1	1.7	0.01	0.9	
Total	**160**	**33**					

[33] This analysis is based on using CGH microarrays to examine disomy and monosomy in spores derived from the JSC2 triploid; to make the comparison similar to that used for MH10, we did not incorporated our SPA data. The methods and descriptions of each column are identical to those previously described for Table S7.

References

ABBEY, D., M. HICKMAN, D. GRESHAM and J. BERMAN, 2011 High-Resolution SNP/CGH Microarrays Reveal the Accumulation of Loss of Heterozygosity in Commonly Used Candida albicans Strains. G3 (Bethesda) **1:** 523-530.

ABDULOVIC, A., N. KIM and S. JINKS-ROBERTSON, 2006 Mutagenesis and the three R's in yeast. DNA Repair (Amst) **5:** 409-421.

AGUILERA, A., 2002 The connection between transcription and genomic instability. The EMBO Journal **21:** 195-201.

AGUILERA, A., S. CHAVEZ and F. MALAGON, 2000 Mitotic recombination in yeast: elements controlling its incidence. Yeast **16:** 731-754.

AGUILERA, A., and B. GOMEZ-GONZALEZ, 2008 Genome instability: a mechanistic view of its causes and consequences. Nat Rev Genet **9:** 204-217.

AGUILERA, A., and H. L. KLEIN, 1989 Yeast intrachromosomal recombination: long gene conversion tracts are preferentially associated with reciprocal exchange and require the RAD1 and RAD3 gene products. Genetics **123:** 683-694.

AHMAD, M., and H. BUSSEY, 1986 Yeast arginine permease: nucleotide sequence of the CAN1 gene. Curr Genet **10:** 587-592.

ALANI, E., 1996 The Saccharomyces cerevisiae Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. Mol Cell Biol **16:** 5604-5615.

ALANI, E., R. A. REENAN and R. D. KOLODNER, 1994 Interaction between mismatch repair and genetic recombination in Saccharomyces cerevisiae. Genetics **137:** 19-39.

ALLERS, T., and D. R. LEACH, 1995 DNA palindromes adopt a methylation-resistant conformation that is consistent with DNA cruciform or hairpin formation in vivo. J Mol Biol **252:** 70-85.

ALLERS, T., and M. LICHTEN, 2001 Differential timing and control of noncrossover and crossover recombination during meiosis. Cell **106:** 47-57.

ALTMAN, D. G., 1990 *Practical Statistics for Medical Research*. CRC, New York.

ALTMAN, D. G., 1991 *Practical statistics for medical research*. Chapman and Hall, London ; New York.

ALVINO, G. M., D. COLLINGWOOD, J. M. MURPHY, J. DELROW, B. J. BREWER *et al.*, 2007 Replication in hydroxyurea: it's a matter of time. Molecular and cellular biology **27:** 6396-6406.

ANDALIS, A. A., Z. STORCHOVA, C. STYLES, T. GALITSKI, D. PELLMAN *et al.*, 2004 Defects arising from whole-genome duplications in Saccharomyces cerevisiae. Genetics **167:** 1109-1121.

ANDERSEN, S. L., and J. SEKELSKY, 2010 Meiotic versus mitotic recombination: two different routes for double-strand break repair: the different functions of meiotic versus mitotic DSB repair are reflected in different pathway usage and different outcomes. Bioessays **32:** 1058-1066.

ARGUESO, J. L., M. F. CARAZZOLLE, P. A. MIECZKOWSKI, F. M. DUARTE, O. V. NETTO *et al.*, 2009 Genome structure of a Saccharomyces cerevisiae strain widely used in bioethanol production. Genome research **19:** 2258-2270.

ARGUESO, J. L., J. WESTMORELAND, P. A. MIECZKOWSKI, M. GAWEL, T. D. PETES *et al.*, 2008 Double-strand breaks associated with repetitive DNA can reshape the genome. Proceedings of the National Academy of Sciences of the United States of America **105:** 11845-11850.

ARLT, M. F., A. M. CASPER and T. W. GLOVER, 2003 Common fragile sites. Cytogenet Genome Res **100:** 92-100.

ASTROM, S. U., S. M. OKAMURA and J. RINE, 1999 Yeast cell-type regulation of DNA repair. Nature **397:** 310.

AYLON, Y., and M. KUPIEC, 2004 DSB repair: the yeast paradigm. DNA Repair (Amst) **3:** 797-815.

AYLON, Y., B. LIEFSHITZ and G. BITAN-BANIN, 2003 Molecular dissection of mitotic recombination in the yeast Saccharomyces cerevisiae. Molecular and cellular â€¦.

AYLON, Y., B. LIEFSHITZ and M. KUPIEC, 2004 The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. EMBO J **23:** 4868-4875.

AZVOLINSKY, A., P. G. GIRESI, J. D. LIEB and V. A. ZAKIAN, 2009 Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. Molecular cell **34:** 722-734.

BAKER, S. J., A. C. PREISINGER, J. M. JESSUP, C. PARASKEVA, S. MARKOWITZ *et al.*, 1990 p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. Cancer Res **50:** 7717-7722.

BARBERA, M. A., and T. D. PETES, 2006 Selection and analysis of spontaneous reciprocal mitotic cross-overs in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America **103:** 12819-12824.

BARLOW, J. H., M. LISBY and R. ROTHSTEIN, 2008 Differential regulation of the cellular response to DNA double-strand breaks in G1. Mol Cell **30:** 73-85.

BARZEL, A., and M. KUPIEC, 2008 Finding a match: how do homologous sequences get together for recombination? Nat Rev Genet **9:** 27-37.

BENJAMINI, Y., and Y. HOCHBERG, 1995 Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. J. R. Statist. Soc **57:** 289-300.

BENNETT, C. B., A. L. LEWIS, K. K. BALDWIN and M. A. RESNICK, 1993 Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. Proc Natl Acad Sci U S A **90:** 5613-5617.

BENNETT, C. B., J. R. SNIPE and M. A. RESNICK, 1997 A persistent double-strand break destabilizes human DNA in yeast and can lead to G2 arrest and lethality. Cancer Res **57:** 1970-1980.

BENNETT, C. B., T. J. WESTMORELAND, J. R. SNIPE and M. A. RESNICK, 1996 A double-strand break within a yeast artificial chromosome (YAC) containing human DNA can result in YAC loss, deletion or cell lethality. Mol Cell Biol **16:** 4414-4425.

BERMEJO, R., Y. DOKSANI, T. CAPRA, Y. M. KATOU, H. TANAKA *et al.*, 2007 Top1- and Top2-mediated topological transitions at replication forks ensure fork

progression and stability and prevent DNA damage checkpoint activation. Genes & development **21:** 1921-1936.

BLITZBLAU, H. G., G. W. BELL, J. RODRIGUEZ, S. P. BELL and A. HOCHWAGEN, 2007 Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres. Curr Biol **17:** 2003-2012.

BORDE, V., W. LIN, E. NOVIKOV, J. H. PETRINI, M. LICHTEN *et al.*, 2004 Association of Mre11p with double-strand break sites during yeast meiosis. Mol Cell **13:** 389-401.

BORG, A., Q. X. ZHANG, P. ALM, H. OLSSON and G. SELLBERG, 1992 The retinoblastoma gene in breast cancer: allele loss is not correlated with loss of gene protein expression. Cancer Res **52:** 2991-2994.

BRANZEI, D., and M. FOIANI, 2008 Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol **9:** 297-308.

BRANZEI, D., and M. FOIANI, 2010 Maintaining genome stability at the replication fork. Nature reviews Molecular cell biology **11:** 208-219.

BREEN, A. P., and J. A. MURPHY, 1995 Reactions of oxyl radicals with DNA. Free Radic Biol Med **18:** 1033-1077.

BRESSAN, D. A., B. K. BAXTER and J. H. PETRINI, 1999 The Mre11-Rad50-Xrs2 protein complex facilitates homologous recombination-based double-strand break repair in Saccharomyces cerevisiae. Mol Cell Biol **19:** 7681-7687.

BREWER, B. J., and W. L. FANGMAN, 1988 A replication fork barrier at the 3' end of yeast ribosomal RNA genes. Cell **55:** 637-643.

BRIDGES, C. B., and E. G. ANDERSON, 1925 Crossing over in the X Chromosomes of Triploid Females of DROSOPHILA MELANOGASTER. Genetics **10:** 418-441.

BRUNBORG, G., and D. H. WILLIAMSON, 1978 The relevance of the nuclear division cycle to radiosensitivity in yeast. Mol Gen Genet **162:** 277-286.

BUCKING-THROM, E., W. DUNTZE, L. H. HARTWELL and T. R. MANNEY, 1973 Reversible arrest of haploid yeast cells in the initiation of DNA synthesis by a diffusible sex factor. Exp Cell Res **76:** 99-110.

BUHLER, C., V. BORDE and M. LICHTEN, 2007 Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in Saccharomyces cerevisiae. PLoS Biol **5:** e324.

BURCH, L. H., Y. YANG, J. F. STERLING, S. A. ROBERTS, F. G. CHAO *et al.*, 2011 Damage-induced localized hypermutability. Cell Cycle **10:** 1073-1085.

BURKHALTER, M. D., S. A. ROBERTS, J. M. HAVENER and D. A. RAMSDEN, 2009 Activity of ribonucleotide reductase helps determine how cells repair DNA double strand breaks. DNA Repair (Amst) **8:** 1258-1263.

CAMPBELL, D., J. S. DOCTOR, J. H. FEUERSANGER and M. M. DOOLITTLE, 1981 Differential mitotic stability of yeast disomes derived from triploid meiosis. Genetics **98:** 239-255.

CAMPBELL, D. A., and M. M. DOOLITTLE, 1987 Coincident chromosomal disomy in meiotic dyads from triploid yeast. Curr Genet **12:** 569-576.

CAPRA, J. A., K. PAESCHKE, M. SINGH and V. A. ZAKIAN, 2010 G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. PLoS computational biology **6:** e1000861-e1000861.

CASPARI, T., J. M. MURRAY and A. M. CARR, 2002 Cdc2-cyclin B kinase activity links Crb2 and Rqh1-topoisomerase III. Genes Dev **16:** 1195-1208.

CASPER, A. M., P. W. GREENWELL, W. TANG and T. D. PETES, 2009 Chromosome aberrations resulting from double-strand DNA breaks at a naturally occurring yeast fragile site composed of inverted ty elements are independent of Mre11p and Sae2p. Genetics **183:** 423-439, 421SI-426SI.

CASPER, A. M., P. NGHIEM, M. F. ARLT and T. W. GLOVER, 2002 ATR regulates fragile site stability. Cell **111:** 779-789.

CAVENEE, W. K., T. P. DRYJA, R. A. PHILLIPS, W. F. BENEDICT, R. GODBOUT *et al.*, 1983 Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature **305:** 779-784.

CEJKA, P., J. L. PLANK, C. Z. BACHRATI, I. D. HICKSON and S. C. KOWALCZYKOWSKI, 2010 Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3. Nat Struct Mol Biol **17:** 1377-1382.

CHA, R. S., and N. KLECKNER, 2002 ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. Science (New York, NY) **297:** 602-606.

CHABES, A., B. GEORGIEVA, V. DOMKIN, X. ZHAO, R. ROTHSTEIN *et al.*, 2003 Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. Cell **112:** 391-401.

CHAMBERS, S. R., N. HUNTER, E. J. LOUIS and R. H. BORTS, 1996 The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. Mol Cell Biol **16:** 6110-6120.

CHEN, W., and S. JINKS-ROBERTSON, 1998 Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. Mol Cell Biol **18:** 6525-6537.

CHUA, P., and S. JINKS-ROBERTSON, 1991 Segregation of recombinant chromatids following mitotic crossing over in yeast. Genetics **129:** 359-369.

COBB, J. A., T. SCHLEKER, V. ROJAS, L. BJERGBAEK, J. A. TERCERO *et al.*, 2005 Replisome instability, fork collapse, and gross chromosomal rearrangements arise synergistically from Mec1 kinase and RecQ helicase mutations. Genes Dev **19:** 3055-3069.

COIC, E., L. GLUCK and F. FABRE, 2000 Evidence for short-patch mismatch repair in Saccharomyces cerevisiae. EMBO J **19:** 3408-3417.

COLLINS, N., R. MCMANUS, R. WOOSTER, J. MANGION, S. SEAL *et al.*, 1995 Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13. Oncogene **10:** 1673-1675.

CORTES-LEDESMA, F., and A. AGUILERA, 2006 Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange. EMBO Rep **7:** 919-926.

COTE, A. G., and S. M. LEWIS, 2008 Mus81-dependent double-strand DNA breaks at in vivo-generated cruciform structures in S. cerevisiae. Mol Cell **31:** 800-812.

COURCELLE, J., J. R. DONALDSON, K. H. CHOW and C. T. COURCELLE, 2003 DNA damage-induced replication fork regression and processing in Escherichia coli. Science **299:** 1064-1067.

COVO, S., J. W. WESTMORELAND, D. A. GORDENIN and M. A. RESNICK, 2010 Cohesin Is limiting for the suppression of DNA damage-induced recombination between homologous chromosomes. PLoS genetics **6:** e1001006-e1001006.

COVO, S., J. W. WESTMORELAND, A. K. REDDY, D. A. GORDENIN and M. A. RESNICK, 2012 RAD53 is limiting in double-strand break repair and in protection against toxicity associated with ribonucleotide reductase inhibition. DNA Repair (Amst) **11:** 317-323.

CULBERTSON, M. R., and S. A. HENRY, 1973 Genetic analysis of hybrid strains trisomic for the chromosome containing a fatty acid synthetase gene complex (fas1) in yeast. Genetics **75:** 441-458.

CURCIO, M. J., A. E. KENNY, S. MOORE, D. J. GARFINKEL, M. WEINTRAUB *et al.*, 2007 S-phase checkpoint pathways stimulate the mobility of the retrovirus-like transposon Ty1. Mol Cell Biol **27:** 8874-8885.

DAIGAKU, Y., A. A. DAVIES and H. D. ULRICH, 2010 Ubiquitin-dependent DNA damage bypass is separable from genome replication. Nature **465:** 951-955.

DALEY, J. M., P. L. PALMBOS, D. WU and T. E. WILSON, 2005 Nonhomologous end joining in yeast. Annu Rev Genet **39:** 431-451.

DAVIES, B. W., M. A. KOHANSKI, L. A. SIMMONS, J. A. WINKLER, J. J. COLLINS *et al.*, 2009 Hydroxyurea induces hydroxyl radical-mediated cell death in Escherichia coli. Mol Cell **36:** 845-860.

DEEM, A., A. KESZTHELYI, T. BLACKGROVE, A. VAYL, B. COFFEY *et al.*, 2011 Break-induced replication is highly inaccurate. PLoS biology **9:** e1000594-e1000594.

DETLOFF, P., J. SIEBER and T. D. PETES, 1991 Repair of specific base pair mismatches formed during meiotic recombination in the yeast Saccharomyces cerevisiae. Mol Cell Biol **11:** 737-745.

DINARDO, S., K. VOELKEL and R. STERNGLANZ, 1984 DNA topoisomerase II mutant of *Saccharomyces cerevisiae*: topoisomerase II is required for segregation of daughter

molecules at the termination of DNA replication. Proc Natl Acad Sci U S A **81:** 2616-2620.

DUDASOVA, Z., A. DUDAS and M. CHOVANEC, 2004 Non-homologous end-joining factors of Saccharomyces cerevisiae. FEMS Microbiol Rev **28:** 581-601.

DUNHAM, M. J., H. BADRANE, T. FEREA, J. ADAMS, P. O. BROWN *et al.*, 2002 Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A **99:** 16144-16149.

DUQUETTE, M. L., P. HANDA, J. A. VINCENT, A. F. TAYLOR and N. MAIZELS, 2004 Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. Genes Dev **18:** 1618-1629.

DUQUETTE, M. L., P. PHAM, M. F. GOODMAN and N. MAIZELS, 2005 AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. Oncogene **24:** 5791-5798.

DURKIN, S. G., and T. W. GLOVER, 2007 Chromosome fragile sites. Annu Rev Genet **41:** 169-192.

ECKARDT, F., and R. H. HAYNES, 1977 Induction of pure and sectored mutant clones in excision-proficient and deficient strains of yeast. Mutat Res **43:** 327-338.

ELLEDGE, S. J., Z. ZHOU and J. B. ALLEN, 1992 Ribonucleotide reductase: regulation, regulation, regulation. Trends Biochem Sci **17:** 119-123.

ELLWOOD, M., and M. NOMURA, 1982 Chromosomal locations of the genes for rRNA in Escherichia coli K-12. J Bacteriol **149:** 458-468.

ESPOSITO, M. S., 1978 Evidence that spontaneous mitotic recombination occurs at the two-strand stage. Proceedings of the National Academy of Sciences of the United States of America **75:** 4436-4440.

ESPOSITO, R. E., 1968 Genetic recombination in synchronized cultures of Saccharomyces cerevisiae. Genetics **59:** 191-210.

EZOV, T. K., E. BOGER-NADJAR, Z. FRENKEL, I. KATSPEROVSKI, S. KEMENY *et al.*, 2006 Molecular-genetic biodiversity in a natural population of the yeast

Saccharomyces cerevisiae from "Evolution Canyon": microsatellite polymorphism, ploidy and controversial sexual status. Genetics **174:** 1455-1468.

FABRE, F., 1978 Induced intragenic recombination in yeast can occur during the G1 mitotic phase. Nature **272:** 795-798.

FABRE, F., A. BOULET and H. ROMAN, 1984 Gene conversion at different points in the mitotic cycle of Saccharomyces cerevisiae. Mol Gen Genet **195:** 139-143.

FACHINETTI, D., R. BERMEJO, A. COCITO, S. MINARDI, Y. KATOU *et al.*, 2010 Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. Molecular cell **39:** 595-605.

FAIRHEAD, C., and B. DUJON, 1993 Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. Mol Gen Genet **240:** 170-178.

FENG, W., S. C. DI RIENZI, M. K. RAGHURAMAN and B. J. BREWER, 2011 Replication stress-induced chromosome breakage is correlated with replication fork progression and is preceded by single-stranded DNA formation. G3 (Bethesda) **1:** 327-335.

FERGUSON, L. R., 1990 Mutagenic and recombinogenic consequences of DNA-repair inhibition during treatment with 1,3-bis(2-chloroethyl)-1-nitrosourea in Saccharomyces cerevisiae. Mutat Res **241:** 369-377.

FISCHER, J. M., and J. R. STRINGER, 2008 Visualizing loss of heterozygosity in living mouse cells and tissues. Mutation research **645:** 1-8.

FISHMAN-LOBELL, J., and J. E. HABER, 1992 Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1. Science **258:** 480-484.

FRANK-VAILLANT, M., and S. MARCAND, 2002 Transient stability of DNA ends allows nonhomologous end joining to precede homologous recombination. Mol Cell **10:** 1189-1199.

FRANKLIN, W. A., P. W. DOETSCH and W. A. HASELTINE, 1985 Structural determination of the ultraviolet light-induced thymine-cytosine pyrimidine-pyrimidone (6-4) photoproduct. Nucleic Acids Res **13:** 5317-5325.

FREUDENREICH, C. H., S. M. KANTROW and V. A. ZAKIAN, 1998 Expansion and length-dependent fragility of CTG repeats in yeast. Science **279:** 853-856.

FRIEDBERG, E. C., 2006 *DNA Repair and Mutagenesis*. ASM Press.

FRIEDENSON, B., 2007 The BRCA1/2 pathway prevents hematologic cancers in addition to breast and ovarian cancers. BMC cancer **7:** 152.

GALLI, A., and R. H. SCHIESTL, 1996 Hydroxyurea induces recombination in dividing but not in G1 or G2 cell cycle arrested yeast cells. Mutation research **354:** 69-75.

GALLI, A., and R. H. SCHIESTL, 1999 Cell division transforms mutagenic lesions into deletion-recombinagenic lesions in yeast cells. Mutat Res **429:** 13-26.

GELFAND, Y., A. RODRIGUEZ and G. BENSON, 2007 TRDB--the Tandem Repeats Database. Nucleic Acids Res **35:** D80-87.

GERALD, J. N., J. M. BENJAMIN and S. J. KRON, 2002 Robust G1 checkpoint arrest in budding yeast: dependence on DNA damage signaling and repair. J Cell Sci **115:** 1749-1757.

GERSTEIN, A. C., R. M. MCBRIDE and S. P. OTTO, 2008 Ploidy reduction in Saccharomyces cerevisiae. Biol Lett **4:** 91-94.

GERTON, J. L., J. DERISI, R. SHROFF, M. LICHTEN, P. O. BROWN *et al.*, 2000 Global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. Proc Natl Acad Sci U S A **97:** 11383-11390.

GIAEVER, G., A. M. CHU, L. NI, C. CONNELLY, L. RILES *et al.*, 2002 Functional profiling of the Saccharomyces cerevisiae genome. Nature **418:** 387-391.

GLOVER, T. W., C. BERGER, J. COYLE and B. ECHO, 1984 DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet **67:** 136-142.

GOCKE, E., and T. R. MANNEY, 1979 Expression of radiation-induced mutations at the arginine permease (CAN1) locus in Saccharomyces cerevisiae. Genetics **91:** 53-66.

GOLDSTEIN, A. L., and J. H. MCCUSKER, 1999 Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast **15:** 1541-1553.

GONZALEZ-BARRERA, S., F. CORTES-LEDESMA, R. E. WELLINGER and A. AGUILERA, 2003 Equal sister chromatid exchange is a major mechanism of double-strand break repair in yeast. Mol Cell **11:** 1661-1671.

GONZALEZ-BARRERA, S., M. GARCIA-RUBIO and A. AGUILERA, 2002 Transcription and double-strand breaks induce similar mitotic recombination events in Saccharomyces cerevisiae. Genetics **162:** 603-614.

GORDENIN, D. A., K. S. LOBACHEV, N. P. DEGTYAREVA, A. L. MALKOVA, E. PERKINS *et al.*, 1993 Inverted DNA repeats: a source of eukaryotic genomic instability. Molecular and cellular biology **13:** 5315-5322.

GORDENIN, D. A., and M. A. RESNICK, 1998 Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability. Mutat Res **400:** 45-58.

GOTTLIEB, S., J. WAGSTAFF and R. E. ESPOSITO, 1989 Evidence for two pathways of meiotic intrachromosomal recombination in yeast. Proc Natl Acad Sci U S A **86:** 7072-7076.

GOUYER, V., S. GAZZERI, E. BRAMBILLA, I. BOLON, D. MORO *et al.*, 1994 Loss of heterozygosity at the RB locus correlates with loss of RB protein in primary malignant neuro-endocrine lung carcinomas. Int J Cancer **58:** 818-824.

GRAVEL, S., J. R. CHAPMAN, C. MAGILL and S. P. JACKSON, 2008 DNA helicases Sgs1 and BLM promote DNA double-strand break resection. Genes Dev **22:** 2767-2772.

GRESHAM, D., B. CURRY, A. WARD, D. B. GORDON, L. BRIZUELA *et al.*, 2010 Optimized detection of sequence variation in heterozygous genomes using DNA microarrays with isothermal-melting probes. Proceedings of the National Academy of Sciences of the United States of America **107:** 1482-1487.

GUTHRIE, C., and G. R. FINK, 1991 *Guide to Yeast Genetics and Molecular Biology*. Academic, San Diego.

HABER, J. E., B. L. RAY, J. M. KOLB and C. I. WHITE, 1993 Rapid kinetics of mismatch repair of heteroduplex DNA that is formed during recombination in yeast. Proc Natl Acad Sci U S A **90:** 3363-3367.

HARTWELL, L. H., and D. SMITH, 1985 Altered fidelity of mitotic chromosome transmission in cell cycle mutants of S. cerevisiae. Genetics **110:** 381-395.

HEGARTY, M. J., and S. J. HISCOCK, 2008 Genomic clues to the evolutionary success of polyploid plants. Curr Biol **18:** R435-444.

HEYER, W. D., K. T. EHMSEN and J. LIU, 2010 Regulation of homologous recombination in eukaryotes. Annu Rev Genet **44:** 113-139.

HICKS, W. M., M. KIM and J. E. HABER, 2010 Increased mutagenesis and unique mutation signature associated with mitotic gene conversion. Science (New York, NY) **329:** 82-85.

HO, C. K., G. MAZON, A. F. LAM and L. S. SYMINGTON, 2010 Mus81 and Yen1 promote reciprocal exchange during mitotic recombination to maintain genome integrity in budding yeast. Molecular cell **40:** 988-1000.

HOCHBERG, Y., and Y. BENJAMINI, 1990 More powerful procedures for multiple significance testing. Stat Med **9:** 811-818.

HOLLIDAY, R., 1964 Mechanism for Gene Conversion in Fungi. Genet Res **5:** 282-&.

HOLM, C., T. GOTO, J. C. WANG and D. BOTSTEIN, 1985 DNA topoisomerase II is required at the time of mitosis in yeast. Cell **41:** 553-563.

HONIGBERG, S. M., and R. E. ESPOSITO, 1994 Reversal of cell determination in yeast meiosis: postcommitment arrest allows return to mitotic growth. Proc Natl Acad Sci U S A **91:** 6559-6563.

HOPPER, A. K., and B. D. HALL, 1975 Mating type and sporulation in yeast. I. Mutations which alter mating-type control over sporulation. Genetics **80:** 41-59.

HUANG, G. S., and R. L. KEIL, 1995 Requirements for activity of the yeast mitotic recombination hotspot HOT1: RNA polymerase I and multiple cis-acting sequences. Genetics **141:** 845-855.

HUERTAS, P., and A. AGUILERA, 2003 Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol Cell **12:** 711-721.

HUERTAS, P., F. CORTES-LEDESMA, A. A. SARTORI, A. AGUILERA and S. P. JACKSON, 2008 CDK targets Sae2 to control DNA-end resection and homologous recombination. Nature **455:** 689-692.

HUNTER, N., and R. H. BORTS, 1997 Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. Genes Dev **11:** 1573-1582.

IACCARINO, I., F. PALOMBO, J. DRUMMOND, N. F. TOTTY, J. J. HSUAN *et al.*, 1996 MSH6, a Saccharomyces cerevisiae protein that binds to mismatches as a heterodimer with MSH2. Curr Biol **6:** 484-486.

ILIOPOULOS, D., G. VASSILIOU, E. SEKERLI, V. SIDIROPOULOU, A. TSIGA *et al.*, 2005 Long survival in a 69,XXX triploid infant in Greece. Genet Mol Res **4:** 755-759.

INBAR, O., and M. KUPIEC, 1999 Homology search and choice of homologous partner during mitotic recombination. Mol Cell Biol **19:** 4134-4142.

INBAR, O., B. LIEFSHITZ, G. BITAN and M. KUPIEC, 2000 The relationship between homology length and crossing over during the repair of a broken chromosome. J Biol Chem **275:** 30833-30838.

IRA, G., A. PELLICIOLI, A. BALIJJA, X. WANG and S. FIORANI, 2004 DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. Nature **431:** 1011-1017.

IVANOV, E. L., and J. E. HABER, 1995 RAD1 and RAD10, but not other excision repair genes, are required for double-strand break-induced recombination in *Saccharomyces cerevisiae*. Mol Cell Biol **15:** 2245-2251.

IVANOV, E. L., V. G. KOROLEV and F. FABRE, 1992 XRS2, a DNA repair gene of Saccharomyces cerevisiae, is needed for meiotic recombination. Genetics **132:** 651-664.

IVESSA, A. S., B. A. LENZMEIER, J. B. BESSLER, L. K. GOUDSOUZIAN, S. L. SCHNAKENBERG *et al.*, 2003 The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. Mol Cell **12:** 1525-1536.

JACOBS, P. A., T. W. GLOVER, M. MAYER, P. FOX, J. W. GERRARD *et al.*, 1980 X-linked mental retardation: a study of 7 families. Am J Med Genet **7:** 471-489.

JAMES, A. P., and B. J. KILBEY, 1977 The timing of UV mutagenesis in yeast: a pedigree analysis of induced recessive mutation. Genetics **87:** 237-248.

JESSOP, L., T. ALLERS and M. LICHTEN, 2005 Infrequent co-conversion of markers flanking a meiotic recombination initiation site in Saccharomyces cerevisiae. Genetics **169:** 1353-1367.

JOHNSON, R. E., G. K. KOVVALI, L. PRAKASH and S. PRAKASH, 1996 Requirement of the yeast MSH3 and MSH6 genes for MSH2-dependent genomic stability. J Biol Chem **271:** 7285-7288.

KADYK, L. C., and L. H. HARTWELL, 1992 Sister chromatids are preferred over homologs as substrates for recombinational repair in *Saccharomyces cerevisiae*. Genetics **132:** 387-402.

KATOH, K., G. ASIMENOS and H. TOH, 2009 Multiple alignment of DNA sequences with MAFFT. Methods Mol Biol **537:** 39-64.

KEENEY, S., C. N. GIROUX and N. KLECKNER, 1997 Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell **88:** 375-384.

KEIL, R. L., and G. S. ROEDER, 1984 Cis-acting, recombination-stimulating activity in a fragment of the ribosomal DNA of S. cerevisiae. Cell **39:** 377-386.

KIM, H. M., V. NARAYANAN, P. A. MIECZKOWSKI, T. D. PETES, M. M. KRASILNIKOVA *et al.*, 2008 Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. EMBO J **27:** 2896-2906.

KIM, J. M., S. VANGURI, J. D. BOEKE, A. GABRIEL and D. F. VOYTAS, 1998 Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. Genome research **8:** 464-478.

KIM, N., and S. JINKS-ROBERTSON, 2012 Transcription as a source of genome instability. Nat Rev Genet **13:** 204-214.

KIRKPATRICK, D. T., M. DOMINSKA and T. D. PETES, 1998 Conversion-type and restoration-type repair of DNA mismatches formed during meiotic recombination in Saccharomyces cerevisiae. Genetics **149:** 1693-1705.

KLEIN, H. L., 1988 Different types of recombination events are controlled by the RAD1 and RAD52 genes of Saccharomyces cerevisiae. Genetics **120:** 367-377.

KLEIN, H. L., and L. S. SYMINGTON, 2009 Breaking up just got easier to do. Cell **138:** 20-22.

KNUDSON, A. G., 2001 Two genetic hits (more or less) to cancer. Nat Rev Cancer **1:** 157-162.

KNUDSON, A. G., JR., 1971 Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A **68:** 820-823.

KOC, A., L. J. WHEELER, C. K. MATHEWS and G. F. MERRILL, 2004 Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. The Journal of biological chemistry **279:** 223-230.

KOGO, H., H. INAGAKI, T. OHYE, T. KATO, B. S. EMANUEL *et al.*, 2007 Cruciform extrusion propensity of human translocation-mediating palindromic AT-rich repeats. Nucleic Acids Res **35:** 1198-1208.

KOLLER, A., J. HEITMAN and M. N. HALL, 1996 Regional bivalent-univalent pairing versus trivalent pairing of a trisomic chromosome in Saccharomyces cerevisiae. Genetics **144:** 957-966.

KOLODNER, R. D., and G. T. MARSISCHKY, 1999 Eukaryotic DNA mismatch repair. Curr Opin Genet Dev **9:** 89-96.

KRAAKMAN-VAN DER ZWET, M., W. J. OVERKAMP, R. E. VAN LANGE, J. ESSERS, A. VAN DUIJN-GOEDHART *et al.*, 2002 Brca2 (XRCC11) deficiency results in radioresistant DNA synthesis and a higher frequency of spontaneous deletions. Mol Cell Biol **22:** 669-679.

KRAKOFF, I. H., N. C. BROWN and P. REICHARD, 1968 Inhibition of ribonucleoside diphosphate reductase by hydroxyurea. Cancer Res **28:** 1559-1565.

KRAMER, K. M., and J. E. HABER, 1993 New telomeres in yeast are initiated with a highly selected subset of TG1-3 repeats. Genes Dev **7:** 2345-2356.

KREMER, E. J., M. PRITCHARD, M. LYNCH, S. YU, K. HOLMAN *et al.*, 1991 Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science **252:** 1711-1714.

KROGH, B. O., and L. S. SYMINGTON, 2004 Recombination proteins in yeast. Annu Rev Genet **38:** 233-271.

KUNZ, B. A., M. K. PIERCE, J. R. MIS and C. N. GIROUX, 1987 DNA sequence analysis of the mutational specificity of u.v. light in the SUP4-o gene of yeast. Mutagenesis **2:** 445-453.

KUONG, K. J., and A. KUZMINOV, 2009 Cyanide, peroxide and nitric oxide formation in solutions of hydroxyurea causes cellular toxicity and may contribute to its therapeutic potency. J Mol Biol **390:** 845-862.

KUZMINOV, A., 2001 DNA replication meets genetic exchange: chromosomal damage and its repair by homologous recombination. Proc Natl Acad Sci U S A **98:** 8461-8468.

LANG, G. I., and A. W. MURRAY, 2008 Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. Genetics **178:** 67-82.

LARSON, E. D., M. L. DUQUETTE, W. J. CUMMINGS, R. J. STREIFF and N. MAIZELS, 2005 MutSalpha binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. Curr Biol **15:** 470-474.

LAWRENCE, C. W., 2002 Cellular roles of DNA polymerase zeta and Rev1 protein. DNA Repair (Amst) **1:** 425-435.

LEE, P. S., P. W. GREENWELL, M. DOMINSKA, M. GAWEL, M. HAMILTON *et al.*, 2009 A fine-structure map of spontaneous mitotic crossovers in the yeast *Saccharomyces cerevisiae*. PLoS genetics **5:** e1000410-e1000410.

LEE, P. S., and T. D. PETES, 2010 Mitotic gene conversion events induced in G1-synchronized yeast cells by gamma rays are similar to spontaneous conversion events. Proceedings of the National Academy of Sciences of the United States of America.

LEE, S. E., F. PAQUES, J. SYLVAN and J. E. HABER, 1999 Role of yeast SIR genes and mating type in directing DNA double-strand breaks to homologous and non-homologous repair paths. Curr Biol **9:** 767-770.

LEMOINE, F. J., N. P. DEGTYAREVA, R. J. KOKOSKA and T. D. PETES, 2008 Reduced levels of DNA polymerase delta induce chromosome fragile site instability in yeast. Molecular and cellular biology **28:** 5359-5368.

LEMOINE, F. J., N. P. DEGTYAREVA, K. LOBACHEV and T. D. PETES, 2005 Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. Cell **120:** 587-598.

LENGRONNE, A., and E. SCHWOB, 2002 The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). Mol Cell **9:** 1067-1078.

LI, H., B. HANDSAKER, A. WYSOKER, T. FENNELL, J. RUAN *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:** 2078-2079.

LI, X., and J. L. MANLEY, 2005 Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell **122:** 365-378.

LILLEY, D. M., 1980 The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc Natl Acad Sci U S A **77:** 6468-6472.

LIM, D. S., and P. HASTY, 1996 A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. Mol Cell Biol **16:** 7133-7143.

LINDAHL, T., 1993 Instability and decay of the primary structure of DNA. Nature **362:** 709-715.

LINDBLAD-TOH, K., D. M. TANENBAUM, M. J. DALY, E. WINCHESTER, W. O. LUI *et al.*, 2000 Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol **18:** 1001-1005.

LISBY, M., R. ROTHSTEIN and U. H. MORTENSEN, 2001 Rad52 forms DNA repair and recombination centers during S phase. Proc Natl Acad Sci U S A **98:** 8276-8282.

LISNIC, B., I. K. SVETEC, H. SARIC, I. NIKOLIC and Z. ZGAGA, 2005 Palindrome content of the yeast Saccharomyces cerevisiae genome. Curr Genet **47:** 289-297.

LO, K. C., D. BAILEY, T. BURKHARDT, P. GARDINA, Y. TURPAZ *et al.*, 2008 Comprehensive analysis of loss of heterozygosity events in glioblastoma using the 100K SNP mapping arrays and comparison with copy number abnormalities defined by

BAC array comparative genomic hybridization. Genes Chromosomes Cancer **47:** 221-237.

LOBACHEV, K. S., D. A. GORDENIN and M. A. RESNICK, 2002 The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. Cell **108:** 183-193.

LOBACHEV, K. S., A. RATTRAY and V. NARAYANAN, 2007 Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells. Front Biosci **12:** 4208-4220.

LOBACHEV, K. S., B. M. SHOR, H. T. TRAN, W. TAYLOR, J. D. KEEN *et al.*, 1998 Factors affecting inverted repeat stimulation of recombination and deletion in *Saccharomyces cerevisiae*. Genetics **148:** 1507-1524.

LOBACHEV, K. S., J. E. STENGER, O. G. KOZYREVA, J. JURKA, D. A. GORDENIN *et al.*, 2000 Inverted Alu repeats unstable in yeast are excluded from the human genome. The EMBO journal **19:** 3822-3830.

LOIDL, J., 1995 Meiotic chromosome pairing in triploid and tetraploid Saccharomyces cerevisiae. Genetics **139:** 1511-1520.

LONGHESE, M. P., M. FOIANI, M. MUZI-FALCONI, G. LUCCHINI and P. PLEVANI, 1998 DNA damage checkpoint in budding yeast. EMBO J **17:** 5525-5528.

LORI, F., A. MALYKH, A. CARA, D. SUN, J. N. WEINSTEIN *et al.*, 1994 Hydroxyurea as an inhibitor of human immunodeficiency virus-type 1 replication. Science **266:** 801-805.

LYDEARD, J. R., S. JAIN, M. YAMAGUCHI and J. E. HABER, 2007 Break-induced replication and telomerase-independent telomere maintenance require Pol32. Nature **448:** 820-823.

MAGENIS, R. E., F. HECHT and E. W. LOVRIEN, 1970 Heritable fragile site on chromosome 16: probable localization of haptoglobin locus in man. Science **170:** 85-87.

MALKOVA, A., E. L. IVANOV and J. E. HABER, 1996 Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. Proc Natl Acad Sci U S A **93:** 7131-7136.

MALONE, R. E., and R. E. ESPOSITO, 1980 The RAD52 gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast. Proc Natl Acad Sci U S A **77:** 503-507.

MANCERA, E., R. BOURGON, A. BROZZI, W. HUBER and L. M. STEINMETZ, 2008 High-resolution mapping of meiotic crossovers and non-crossovers in yeast. Nature **454:** 479-485.

MARSISCHKY, G. T., S. LEE, J. GRIFFITH and R. D. KOLODNER, 1999 'Saccharomyces cerevisiae MSH2/6 complex interacts with Holliday junctions and facilitates their cleavage by phage resolution enzymes. J Biol Chem **274:** 7200-7206.

MAYER, V. W., and A. AGUILERA, 1990 High levels of chromosome instability in polyploids of Saccharomyces cerevisiae. Mutation research **231:** 177-186.

MCCULLEY, J. L., and T. D. PETES, 2010 Chromosome rearrangements and aneuploidy in yeast strains lacking both Tel1p and Mec1p reflect deficiencies in two different mechanisms. Proceedings of the National Academy of Sciences of the United States of America **107:** 11465-11470.

MCCUSKER, J. H., K. V. CLEMONS, D. A. STEVENS and R. W. DAVIS, 1994 Genetic characterization of pathogenic Saccharomyces cerevisiae isolates. Genetics **136:** 1261-1269.

MCGANN, P. T., and R. E. WARE, 2011 Hydroxyurea for sickle cell anemia: what have we learned and what questions still remain? Curr Opin Hematol **18:** 158-165.

MCMURRAY, M. A., and D. E. GOTTSCHLING, 2003 An age-induced switch to a hyper-recombinational state. Science (New York, NY) **301:** 1908-1911.

MERKER, J. D., M. DOMINSKA and T. D. PETES, 2003 Patterns of heteroduplex formation associated with the initiation of meiotic recombination in the yeast Saccharomyces cerevisiae. Genetics **165:** 47-63.

MICHELSON, R. J., and T. WEINERT, 2000 Closing the gaps among a web of DNA repair disorders. Bioessays **22:** 966-969.

MIMITOU, E. P., and L. S. SYMINGTON, 2008 Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature **455:** 770-774.

MIMITOU, E. P., and L. S. SYMINGTON, 2009 DNA end resection: many nucleases make light work. DNA Repair (Amst) **8:** 983-995.

MIMITOU, E. P., and L. S. SYMINGTON, 2011 DNA end resection--unraveling the tail. DNA Repair (Amst) **10:** 344-348.

MIRKIN, S. M., 2007 Expandable DNA repeats and human disease. Nature **447:** 932-940.

MITCHEL, K., H. ZHANG, C. WELZ-VOEGELE and S. JINKS-ROBERTSON, 2010 Molecular structures of crossover and noncrossover intermediates during gap repair in yeast: implications for recombination. Molecular cell **38:** 211-222.

MIYAZAKI, T., D. A. BRESSAN, M. SHINOHARA, J. E. HABER and A. SHINOHARA, 2004 In vivo assembly and disassembly of Rad51 and Rad52 complexes during double-strand break repair. EMBO J **23:** 939-949.

MODRICH, P., and R. LAHUE, 1996 Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem **65:** 101-133.

MOORE, H., P. W. GREENWELL, C. P. LIU, N. ARNHEIM and T. D. PETES, 1999 Triplet repeats form secondary structures that escape DNA repair in yeast. Proc Natl Acad Sci U S A **96:** 1504-1509.

MORILLON, A., L. BENARD, M. SPRINGER and P. LESAGE, 2002 Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons. Mol Cell Biol **22:** 2078-2088.

MOSS, J., H. TINLINE-PURVIS, C. A. WALKER, L. K. FOLKES, M. R. STRATFORD *et al.*, 2010 Break-induced ATR and Ddb1-Cul4(Cdt)(2) ubiquitin ligase-dependent nucleotide synthesis promotes homologous recombination repair in fission yeast. Genes Dev **24:** 2705-2716.

MOYNAHAN, M. E., and M. JASIN, 2010 Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. Nature reviews Molecular cell biology **11:** 196-207.

NAG, D. K., and T. D. PETES, 1991 Seven-base-pair inverted repeats in DNA form stable hairpins in vivo in Saccharomyces cerevisiae. Genetics **129:** 669-673.

NAGALAKSHMI, U., Z. WANG, K. WAERN, C. SHOU, D. RAHA *et al.*, 2008 The transcriptional landscape of the yeast genome defined by RNA sequencing. Science (New York, NY) **320:** 1344-1349.

NAKADA, D., K. MATSUMOTO and K. SUGIMOTO, 2003 ATM-related Tel1 associates with double-strand breaks through an Xrs2-dependent mechanism. Genes Dev **17:** 1957-1962.

NAKAI, S., and R. K. MORTIMER, 1969 Studies on the genetic mechanism of radiation-induced mitotic segregation in yeast. Mol Gen Genet **103:** 329-338.

NAMBIAR, M., G. GOLDSMITH, B. T. MOORTHY, M. R. LIEBER, M. V. JOSHI *et al.*, 2011 Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma. Nucleic Acids Res **39:** 936-948.

NARAYANAN, V., P. A. MIECZKOWSKI, H. M. KIM, T. D. PETES and K. S. LOBACHEV, 2006 The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. Cell **125:** 1283-1296.

NATHANSON, K. L., R. WOOSTER and B. L. WEBER, 2001 Breast cancer genetics: what we know and what we need. Nat Med **7:** 552-556.

NEGRINI, S., V. G. GORGOULIS and T. D. HALAZONETIS, 2010 Genomic instability--an evolving hallmark of cancer. Nat Rev Mol Cell Biol **11:** 220-228.

NEVO-CASPI, Y., and M. KUPIEC, 1994 Transcriptional induction of Ty recombination in yeast. Proc Natl Acad Sci U S A **91:** 12711-12715.

NICKOLOFF, J. A., D. B. SWEETSER, J. A. CLIKEMAN, G. J. KHALSA and S. L. WHEELER, 1999 Multiple heterologies increase mitotic double-strand break-induced allelic gene conversion tract lengths in yeast. Genetics **153:** 665-679.

O'DONNELL, J. P., M. GEHMAN and J. B. KEENEY, 2010 Regulators of ribonucleotide reductase inhibit Ty1 mobility in saccharomyces cerevisiae. Mob DNA **1:** 23.

ONN, I., J. M. HEIDINGER-PAULI, V. GUACCI, E. UNAL and D. E. KOSHLAND, 2008 Sister chromatid cohesion: a simple concept with a complex reality. Annu Rev Cell Dev Biol **24:** 105-129.

ORR-WEAVER, T. L., and J. W. SZOSTAK, 1983 Yeast recombination: the association between double-strand gap repair and crossing-over. Proc Natl Acad Sci U S A **80:** 4417-4421.

OZCAN, S., and M. JOHNSTON, 1999 Function and regulation of yeast hexose transporters. Microbiol Mol Biol Rev **63:** 554-569.

PAESCHKE, K., J. A. CAPRA and V. A. ZAKIAN, 2011 DNA replication through g-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 dna helicase. Cell **145:** 678-691.

PAGE, S. L., and R. S. HAWLEY, 2003 Chromosome choreography: the meiotic ballet. Science **301:** 785-789.

PALMER, S., E. SCHILDKRAUT, R. LAZARIN, J. NGUYEN and J. A. NICKOLOFF, 2003 Gene conversion tracts in Saccharomyces cerevisiae can be extremely short and highly directional. Nucleic Acids Res **31:** 1164-1173.

PAN, J., M. SASAKI, R. KNIEWEL, H. MURAKAMI, H. G. BLITZBLAU *et al.*, 2011 A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. Cell **144:** 719-731.

PANAYOTATOS, N., and R. D. WELLS, 1981 Cruciform structures in supercoiled DNA. Nature **289:** 466-470.

PÂQUES, F., and J. E. HABER, 1997 Two pathways for removal of nonhomologous DNA ends during double-strand break repair in *Saccharomyces cerevisiae*. Mol Cell Biol **17:** 6765-6771.

PÂQUES, F., and J. E. HABER, 1999 Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. Microbiology and molecular biology reviews : MMBR **63:** 349-404.

PARRY, E. M., and B. S. COX, 1970 The tolerance of aneuploidy in yeast. Genet Res **16:** 333-340.

PAULOVICH, A. G., R. U. MARGULIES, B. M. GARVIK and L. H. HARTWELL, 1997 RAD9, RAD17, and RAD24 are required for S phase regulation in Saccharomyces cerevisiae in response to DNA damage. Genetics **145:** 45-62.

PEAK, M. J., and J. G. PEAK, 1986 DNA-to-protein crosslinks and backbone breaks caused by far- and near-ultraviolet, and visible light radiations in mammalian cells. Basic Life Sci **38:** 193-202.

PETES, T. D., 2001 Meiotic recombination hot spots and cold spots. Nature reviews Genetics **2:** 360-369.

PETES, T. D., and C. W. HILL, 1988 Recombination between repeated genes in microorganisms. Annu Rev Genet **22:** 147-168.

PETES, T. D., R. E. MALONE and L. S. SYMINGTON, 1991 Recombination in yeast., pp. 407-521 in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, edited by J. R. BROACH, E. W. JONES and J. R. PRINGLE. Cold Spring Harbor Press, Cold Spring Harbor.

PETUKHOVA, G., P. SUNG and H. KLEIN, 2000 Promotion of Rad51-dependent D-loop formation by yeast recombination factor Rdh54/Tid1. Genes Dev **14:** 2206-2215.

PIERCE, M. K., C. N. GIROUX and B. A. KUNZ, 1987 Development of a yeast system to assay mutational specificity. Mutat Res **182:** 65-74.

PLATT, O. S., S. H. ORKIN, G. DOVER, G. P. BEARDSLEY, B. MILLER *et al.*, 1984 Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. J Clin Invest **74:** 652-656.

POMPER, S., K. M. DANIELS and D. W. MCKEE, 1954 Genetic Analysis of Polyploid Yeast. Genetics **39:** 343-355.

POSTOW, L., C. ULLSPERGER, R. W. KELLER, C. BUSTAMANTE, A. V. VOLOGODSKII *et al.*, 2001 Positive torsional strain causes the formation of a four-way junction at replication forks. J Biol Chem **276:** 2790-2796.

PRADO, F., and A. AGUILERA, 2005 Impairment of replication fork progression mediates RNA polII transcription-associated recombination. EMBO J **24:** 1267-1276.

PUTNAM, C. D., E. J. JAEHNIG and R. D. KOLODNER, 2009 Perspectives on the DNA damage and replication checkpoint responses in Saccharomyces cerevisiae. DNA Repair (Amst) **8:** 974-982.

REICHARD, P., 1988 Interactions between deoxyribonucleotide and DNA synthesis. Annu Rev Biochem **57:** 349-374.

RELIENE, R., A. J. BISHOP and R. H. SCHIESTL, 2007 Involvement of homologous recombination in carcinogenesis. Adv Genet **58:** 67-87.

REYNOLDS, R. J., 1987 Induction and repair of closely opposed pyrimidine dimers in Saccharomyces cerevisiae. Mutat Res **184:** 197-207.

RILEY, M. I., and T. R. MANNEY, 1978 Tetraploid Strains of SACCHAROMYCES CEREVISIAE That Are Trisomic for Chromosome III. Genetics **89:** 667-684.

ROBINSON, J. T., H. THORVALDSDOTTIR, W. WINCKLER, M. GUTTMAN, E. S. LANDER *et al.*, 2011 Integrative genomics viewer. Nat Biotechnol **29:** 24-26.

ROITGRUND, C., R. STEINLAUF and M. KUPIEC, 1993 Donation of information to the unbroken chromosome in double-strand break repair. Curr Genet **23:** 414-422.

RUDOLPH, C. J., P. DHILLON, T. MOORE and R. G. LLOYD, 2007 Avoiding and resolving conflicts between DNA replication and transcription. DNA Repair (Amst) **6:** 981-993.

SAMADASHWILY, G. M., G. RACA and S. M. MIRKIN, 1997 Trinucleotide repeats affect DNA replication in vivo. Nat Genet **17:** 298-304.

SANDELL, L. L., and V. A. ZAKIAN, 1993 Loss of a yeast telomere: arrest, recovery, and chromosome loss. Cell **75:** 729-739.

SANTOCANALE, C., and J. F. DIFFLEY, 1998 A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. Nature **395:** 615-618.

SAPARBAEV, M., L. PRAKASH and S. PRAKASH, 1996 Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in *Saccharomyces cerevisiae*. Genetics **142:** 727-736.

SAXE, D., A. DATTA and S. JINKS-ROBERTSON, 2000 Stimulation of mitotic recombination events by high levels of RNA polymerase II transcription in yeast. Mol Cell Biol **20:** 5404-5414.

SCHOFIELD, M. J., and P. HSIEH, 2003 DNA mismatch repair: molecular mechanisms and biological function. Annu Rev Microbiol **57:** 579-608.

SCULLY, R., J. CHEN, A. PLUG, Y. XIAO, D. WEAVER *et al.*, 1997 Association of BRCA1 with Rad51 in mitotic and meiotic cells. Cell **88:** 265-275.

SETLOW, R. B., 1966 Cyclobutane-type pyrimidine dimers in polynucleotides. Science **153:** 379-386.

SHAFFER, B., I. BREARLEY, R. LITTLEWOOD and G. R. FINK, 1971 A stable aneuploid of Saccharomyces cerevisiae. Genetics **67:** 483-495.

SHAO, C., L. DENG, O. HENEGARIU, L. LIANG, N. RAIKWAR *et al.*, 1999 Mitotic recombination produces the majority of recessive fibroblast variants in heterozygous mice. Proc Natl Acad Sci U S A **96:** 9230-9235.

SHEN, C. Y., J. C. YU, Y. L. LO, C. H. KUO, C. T. YUE *et al.*, 2000 Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. Cancer Res **60:** 3884-3892.

SHERR, C. J., 1994 G1 phase progression: cycling on cue. Cell **79:** 551-555.

SHIM, E. Y., W. H. CHUNG, M. L. NICOLETTE, Y. ZHANG, M. DAVIS *et al.*, 2010 Saccharomyces cerevisiae Mre11/Rad50/Xrs2 and Ku proteins regulate association of Exo1 and Dna2 with DNA breaks. EMBO J **29:** 3370-3380.

SHKUNDINA, I. S., and M. D. TER-AVANESYAN, 2007 Prions. Biochemistry (Mosc) **72:** 1519-1536.

SHRIVASTAV, M., L. P. DE HARO and J. A. NICKOLOFF, 2008 Regulation of DNA double-strand break repair pathway choice. Cell research **18:** 134-147.

SIEDE, W., A. S. FRIEDBERG, I. DIANOVA and E. C. FRIEDBERG, 1994 Characterization of G1 checkpoint control in the yeast Saccharomyces cerevisiae following exposure to DNA-damaging agents. Genetics **138:** 271-281.

SIMCHEN, G., 2009 Commitment to meiosis: what determines the mode of division in budding yeast? Bioessays **31:** 169-177.

SINDEN, R. R., G. X. ZHENG, R. G. BRANKAMP and K. N. ALLEN, 1991 On the deletion of inverted repeated DNA in Escherichia coli: effects of length, thermal stability, and cruciform formation in vivo. Genetics **129:** 991-1005.

SLATER, M. L., 1973 Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle. J Bacteriol **113:** 263-270.

SMITH, S. A., D. F. EASTON, D. G. EVANS and B. A. PONDER, 1992 Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. Nat Genet **2:** 128-131.

SOGO, J. M., M. LOPES and M. FOIANI, 2002 Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. Science (New York, N.Y.) **297:** 599-602.

SONG, W., and T. D. PETES, 2012 Haploidization in Saccharomyces cerevisiae induced by a deficiency in homologous recombination. Genetics **191:** 279-284.

ST CHARLES, J., E. HAZKANI-COVO, Y. YIN, S. L. ANDERSEN, F. S. DIETRICH *et al.*, 2012 High-resolution Genome-wide Analysis of Irradiated (UV and gamma rays) Diploid Yeast Cells Reveals a High Frequency of Genomic Loss of Heterozygosity (LOH) Events. Genetics.

STEINBERG, M. H., W. F. MCCARTHY, O. CASTRO, S. K. BALLAS, F. D. ARMSTRONG *et al.*, 2010 The risks and benefits of long-term use of hydroxyurea in sickle cell anemia: A 17.5 year follow-up. Am J Hematol **85:** 403-408.

STERN, C., 1936 Somatic Crossing over and Segregation in Drosophila Melanogaster. Genetics **21:** 625-730.

STEWART, S. E., and G. S. ROEDER, 1989 Transcription by RNA polymerase I stimulates mitotic recombination in Saccharomyces cerevisiae. Mol Cell Biol **9:** 3464-3472.

STOCK, M., J. USTINOVA, D. K. LAMATSCH, M. SCHARTL, N. PERRIN *et al.*, 2010 A vertebrate reproductive system involving three ploidy levels: hybrid origin of triploids in a contact zone of diploid and tetraploid palearctic green toads (Bufo viridis subgroup). Evolution **64:** 944-959.

STORCHOVA, Z., A. BRENEMAN, J. CANDE, J. DUNN, K. BURBANK *et al.*, 2006 Genome-wide genetic analysis of polyploidy in yeast. Nature **443:** 541-547.

STORICI, F., and M. A. RESNICK, 2006 The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods in enzymology **409:** 329-345.

STRAND, M., M. C. EARLEY, G. F. CROUSE and T. D. PETES, 1995 Mutations in the MSH3 gene preferentially lead to deletions within tracts of simple repetitive DNA in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A **92:** 10418-10421.

STRATHERN, J. N., B. K. SHAFER and C. B. MCGILL, 1995 DNA synthesis errors associated with double-strand-break repair. Genetics **140:** 965-972.

STRATHERN, J. N., K. G. WEINSTOCK, D. R. HIGGINS and C. B. MCGILL, 1991 A novel recombinator in yeast based on gene II protein from bacteriophage f1. Genetics **127:** 61-73.

STROM, L., C. KARLSSON, H. B. LINDROOS, S. WEDAHL, Y. KATOU *et al.*, 2007 Postreplicative formation of cohesion is required for repair and induced by a single DNA break. Science **317:** 242-245.

SUGAWARA, N., F. PAQUES, M. COLAIACOVO and J. E. HABER, 1997 Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. Proc Natl Acad Sci U S A **94:** 9214-9219.

SUN, H., D. TRECO, N. P. SCHULTES and J. W. SZOSTAK, 1989 Double-strand breaks at an initiation site for meiotic gene conversion. Nature **338:** 87-90.

SUNG, P., 1997 Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. J Biol Chem **272:** 28194-28197.

SUTHERLAND, G. R., 1977 Fragile sites on human chromosomes: demonstration of their dependence on the type of tissue culture medium. Science **197:** 265-266.

SYMINGTON, L. S., 2002 Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol Mol Biol Rev **66:** 630-670, table of contents.

SYMINGTON, L. S., and J. GAUTIER, 2011 Double-strand break end resection and repair pathway choice. Annu Rev Genet **45:** 247-271.

SYMINGTON, L. S., and T. D. PETES, 1988 Meiotic recombination within the centromere of a yeast chromosome. Cell **52:** 237-240.

SZILARD, R. K., P. E. JACQUES, L. LARAMÉE, B. CHENG, S. GALICIA *et al.*, 2010 Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. Nature structural & molecular biology **17:** 299-305.

SZOSTAK, J. W., T. L. ORR-WEAVER, R. J. ROTHSTEIN and F. W. STAHL, 1983 The double-strand-break repair model for recombination. Cell **33:** 25-35.

SZOSTAK, J. W., and R. WU, 1980 Unequal crossing over in the ribosomal DNA of Saccharomyces cerevisiae. Nature **284:** 426-430.

TAKATA, M., M. S. SASAKI, E. SONODA, C. MORRISON, M. HASHIMOTO *et al.*, 1998 Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J **17:** 5497-5508.

TANG, W., M. DOMINSKA, P. W. GREENWELL, Z. HARVANEK, K. S. LOBACHEV *et al.*, 2011 Friedreich's ataxia (GAA)n*(TTC)n repeats strongly stimulate mitotic crossovers in *Saccharomyces cerevisae*. PLoS Genet **7:** e1001270.

THOMAS, B. J., and R. ROTHSTEIN, 1989 Elevated recombination rates in transcriptionally active DNA. Cell **56:** 619-630.

TIMSON, J., 1975 Hydroxyurea. Mutat Res **32:** 115-132.

TORRES, E. M., T. SOKOLSKY, C. M. TUCKER, L. Y. CHAN, M. BOSELLI *et al.*, 2007 Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science (New York, NY) **317:** 916-924.

TSUBOUCHI, H., and H. OGAWA, 1998 A novel mre11 mutation impairs processing of double-strand breaks of DNA during both mitosis and meiosis. Mol Cell Biol **18:** 260-268.

TUNA, M., S. KNUUTILA and G. B. MILLS, 2009 Uniparental disomy in cancer. Trends in molecular medicine **15:** 120-128.

TUTT, A., A. GABRIEL, D. BERTWISTLE, F. CONNOR, H. PATERSON *et al.*, 1999 Absence of Brca2 causes genome instability by chromosome breakage and loss associated with centrosome amplification. Curr Biol **9:** 1107-1110.

TYERS, M., G. TOKIWA and B. FUTCHER, 1993 Comparison of the Saccharomyces cerevisiae G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins. The EMBO journal **12:** 1955-1968.

UHLMANN, F., and K. NASMYTH, 1998 Cohesion between sister chromatids must be established during DNA replication. Curr Biol **8:** 1095-1101.

UNAL, E., J. M. HEIDINGER-PAULI, W. KIM, V. GUACCI, I. ONN *et al.*, 2008 A molecular determinant for the establishment of sister chromatid cohesion. Science **321:** 566-569.

UNAL, E., J. M. HEIDINGER-PAULI and D. KOSHLAND, 2007 DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through Eco1 (Ctf7). Science **317:** 245-248.

VENKITARAMAN, A. R., 2001 Chromosome stability, DNA recombination and the BRCA2 tumour suppressor. Curr Opin Cell Biol **13:** 338-343.

VOINEAGU, I., V. NARAYANAN, K. S. LOBACHEV and S. M. MIRKIN, 2008 Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. Proceedings of the National Academy of Sciences of the United States of America **105:** 9936-9941.

WACH, A., A. BRACHAT, R. POHLMANN and P. PHILIPPSEN, 1994 New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. Yeast **10:** 1793-1808.

WANG, J. C., 2002 Cellular roles of DNA topoisomerases: a molecular perspective. Nature reviews Molecular cell biology **3:** 430-440.

WANG, X., and J. E. HABER, 2004 Role of Saccharomyces single-stranded DNA-binding protein RPA in the strand invasion step of double-strand break repair. PLoS Biol **2:** E21.

WANG, Z. C., M. LIN, L. J. WEI, C. LI, A. MIRON *et al.*, 2004 Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. Cancer Res **64:** 64-71.

WARD, J. F., 1990 The yield of DNA double-strand breaks produced intracellularly by ionizing radiation: a review. Int J Radiat Biol **57:** 1141-1150.

WEI, W., J. MCCUSKER, R. HYMAN, T. JONES, Y. NING *et al.*, 2007 Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. Proceedings of the National Academy of Sciences **104:** 12825-12830.

WEIFFENBACH, B., and J. E. HABER, 1981 Homothallic mating type switching generates lethal chromosome breaks in rad52 strains of Saccharomyces cerevisiae. Mol Cell Biol **1:** 522-534.

WEINERT, T. A., and L. H. HARTWELL, 1988 The RAD9 gene controls the cell cycle response to DNA damage in Saccharomyces cerevisiae. Science **241:** 317-322.

WEISSENBACH, J., G. GYAPAY, C. DIB, A. VIGNAL, J. MORISSETTE *et al.*, 1992 A second-generation linkage map of the human genome. Nature **359:** 794-801.

WELCSH, P. L., and M. C. KING, 2001 BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. Hum Mol Genet **10:** 705-713.

WENG, Y. S., and J. A. NICKOLOFF, 1998 Evidence for independent mismatch repair processing on opposite sides of a double-strand break in Saccharomyces cerevisiae. Genetics **148:** 59-70.

WEST, S. C., 1997 Processing of recombination intermediates by the RuvABC proteins. Annu Rev Genet **31:** 213-244.

WILDENBERG, J., 1970 The relation of mitotic recombination to DNA replication in yeast pedigrees. Genetics **66:** 291-304.

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.

WINZELER, E. A., C. I. CASTILLO-DAVIS, G. OSHIRO, D. LIANG, D. R. RICHARDS *et al.*, 2003 Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. Genetics **163:** 79-89.

WONG, A. K., R. PERO, P. A. ORMONDE, S. V. TAVTIGIAN and P. L. BARTEL, 1997 RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. J Biol Chem **272:** 31941-31944.

XU, H. J., P. CAIRNS, S. X. HU, M. A. KNOWLES and W. F. BENEDICT, 1993 Loss of RB protein expression in primary bladder cancer correlates with loss of heterozygosity at the RB locus and tumor progression. Int J Cancer **53:** 781-784.

YANG, Y., J. STERLING, F. STORICI, M. A. RESNICK and D. A. GORDENIN, 2008 Hypermutability of damaged single-strand DNA formed at double-strand breaks and uncapped telomeres in yeast Saccharomyces cerevisiae. PLoS Genet **4:** e1000264.

YOO, G. H., H. J. XU, J. A. BRENNAN, W. WESTRA, R. H. HRUBAN *et al.*, 1994 Infrequent inactivation of the retinoblastoma gene despite frequent loss of chromosome 13q in head and neck squamous cell carcinoma. Cancer Res **54:** 4603-4606.

ZAKHARYEVICH, K., S. TANG, Y. MA and N. HUNTER, 2012 Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. Cell **149:** 334-347.

ZHAO, X., E. G. MULLER and R. ROTHSTEIN, 1998 A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Mol Cell **2:** 329-340.

ZHENG, H. T., Z. H. PENG, S. LI and L. HE, 2005 Loss of heterozygosity analyzed by single nucleotide polymorphism array in cancer. World J Gastroenterol **11:** 6740-6744.

ZHU, Z., W. H. CHUNG, E. Y. SHIM, S. E. LEE and G. IRA, 2008 Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell **134:** 981-994.

ZIERHUT, C., and J. F. DIFFLEY, 2008 Break dosage, cell cycle stage and DNA replication influence DNA double strand break response. EMBO J **27:** 1875-1885.

ZOU, L., and S. J. ELLEDGE, 2003 Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science **300:** 1542-1548.

Biography

Jordan Anne St. Charles

BORN: August 3, 1983 in Indianapolis, IN

EDUCATION

Duke University Durham, NC

Ph.D. in Pharmacology August 2005 – August 2012

University of Michigan, Ann Arbor, MI

B.S. Biochemistry August 2001 – May 2005

PUBLICATIONS

J. St Charles*, E. Hazkani-Covo*, Y. Yin, S. L. Andersen, F. S. Dietrich, P. W. Greenwell, E. Malc, P. Mieczkowski, and T. D. Petes. (2012) High-resolution Genome- wide Analysis of Irradiated (UV and gamma rays) Diploid Yeast Cells Reveals a High Frequency of Genomic Loss of Heterozygosity (LOH) Events. Genetics 190:1267-84.

*These authors contributed equally to this work.

J. St Charles, M. L. Hamilton, and T. D. Petes. (2010) Meiotic Chromosome Segretation in Triploid Strains of *Saccharomyces cerevisiae.* Genetics 186: 537-550.

Y. Huang, W. Giblin, M. Kubec, G. Westfield, J. St Charles, L. Chadde, S. Kraftson, and J. Sekiguchi (2009) Impact of hypomorphic Artemis disease allele on lymphocyte development, DNA end processing, and genome stability. J Exp Med 206:893-908.

HONORS AND AWARDS

Duke University Graduate School Conference Travel Fellowship 2011

Graduated cum laude, University of Michigan 2005