UNDERSTANDING CROSSOVER CONTROL USING A. THALIANA AND S. CEREVISIAE

Luke E. Berchowitz

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology

Chapel Hill
2009

Approved by:

Gregory P. Copenhaver, Ph.D.

Corbin D. Jones, Ph.D.

Joseph Kieber, Ph.D.

Patricia J. Pukkila, Ph.D.

Jeff Sekelsky, Ph.D.

ABSTRACT

LUKE BERCHOWITZ: Understanding crossover control using *A. thaliana* and *S. cerevisiae*
(Under the direction of Gregory P. Copenhaver)

During meiosis, homologous chromosomes pair, synapse, and recombine to facilitate accurate chromosome segregation in meiosis I. Meiotic recombination is facilitated by programmed double-strand breaks that can be repaired either as crossovers or non-crossovers. In most organisms, crossover distribution along chromosomes is non-random in that crossovers are more evenly spaced than null expectations. The inhibition of closely spaced events is known as interference. Despite the fact that interference was originally observed almost a century ago, fundamental questions regarding its underlying mechanisms still exist. I discuss key unanswered questions regarding interference as well as the most commonly referenced models that have been proposed to explain the interference mechanism.

We have developed a visual assay (the FTL system) for the detection of crossovers, gene conversions and interference in *A. thaliana*. This assay involves monitoring the segregation of fluorescent proteins in the pollen grains of *qrt1* mutants. *qrt1* mutants exhibit pollen tetrads i.e. the fusion of the four meiotic products, which allows for advanced statistical analyses previously only available in yeasts. The development and applications of this system are discussed.

Humans, *S. cerevisiae* and *A. thaliana* have at least two pathways for producing crossovers, which include a primary pathway that is subject to interference and a secondary pathway that is interference-insensitive. Using the FTL system, we demonstrate that *AtMUS81* is a mediator of the interference-insensitive pathway in *A. thaliana*. *Atmus81* mutants are sensitive to a wide range of DNA damaging agents and exhibit decreased pollen viability and crossover frequency. The remaining crossovers in the *Atmus81* mutant are subject to interference.

Meiotic recombination occurs in the context of chromatin and chromatin context is often invoked to explain why recombination occurs preferentially in some genomic regions. Using a technique called FAIRE, we demonstrate that double-strand break hotspots and regions of open chromatin have a positive but complex association in *S. cerevisiae*. We also show that subtelomeric border regions and regions surrounding tRNA genes are enriched for meiosis-specific open chromatin. Centromeres exhibit constitutive enrichment of open chromatin.

ACKNOWLEDGEMENTS

This work is dedicated to my parents David and Gillian, who did a fine parenting job if I do say so myself, as well as my loving and supportive wife Andrea.

TABLE OF CONTENTS

LIST OF TABLES.....ix

LIST OF FIGURES.....x-xi

LIST OF ABBREVIATIONS AND SYMBOLS.....xii-xv

Chapter

I. GENETIC INTERFERENCE: DON'T STAND SO CLOSE TO ME.....1-42

Abstract.....2

Background.....3-7

Unanswered Questions.....8-18

Models.....19-25

Conclusions.....25-28

Acknowledgments.....29

Table Legends.....29

Figure Legends.....29-30

Tables.....31

Figures.....32-33

References.....34-42

II. A POLLEN-TETRAD BASED ASSAY FOR MEIOTIC RECOMBINATION IN ARABIDOPSIS.....43-69

Abstract.....44

Introduction.....45-47

Results and Discussion.....47-54

Materials and Methods.....54-57

Acknowledgements.....57

Figure Legends.....57-59

Tables.....60-61

Figures.....62-66

References.....67-68

Chapter 2 Addendum.....69

III. FLUORESCENT ARABIDOPSIS TETRADS: A VISUAL ASSAY FOR QUICKLY DEVELOPING LARGE CROSSOVER AND CROSSOVER INTERFERENCE DATASETS.....70-110

Abstract.....71

Introduction.....72-81

Materials.....81-82

Procedure.....82-92

Anticipated Results.....92-95

Acknowledgements.....95

Figure Legends.....95-97

Tables.....98-101

Figures.....102-105

References.....106-109

Chapter 3 Addendum.....110

IV. THE ROLE OF ATMUS81 IN INTERFERENCE-INSENSITIVE CROSSOVERS IN A. THALIANA.....111-146

Abstract.....112

Introduction.....113-117

Results.....117-123

Discussion.....124-126

Materials and Methods.....126-132

Acknowledgements.....132

Figure Legends.....132-134

Tables.....135-136

Figures.....137-141

References.....142-145

Chapter 4 Addendum.....146

V. A POSITIVE BUT COMPLEX ASSOCIATION BETWEEN MEIOTIC DOUBLE-STRAND BREAK HOTSPOTS AND OPEN CHROMATIN IN SACCHAROMYCES CEREVISIAE.....147-206

Abstract.....148

Introduction.....149-152

Results.....152-165

Discussion.....165-171

Methods.....171-176

Acknowledgements.....177

Figure and Table Legends.....177-184

Figures.....185-191

Tables.....192

Supplementary Figures.....193-194

Supplementary Tables.....195-198

References.....199-205

Chapter 5 Addendum.....206

VI. CONCLUSIONS AND FUTURE STUDY.....207-228

Using the FTL system to study recombination in *A. thaliana*.....208-212

A. thaliana has at least two pathways for producing COs,
one of which is interference-insensitive.....212-215

How is interference imposed?.....216-217

Using FAIRE to study the relationship between open chromatin
and recombination in *S. cerevisiae*.....218-222

Figure Legends.....223

Figures.....224

References.....225-228

LIST OF TABLES

Table

1.1. CO interference comparisons across model genetic organisms.....31

2.1. Physical position of FTL insertions.....60-61

3.1. Letter codes for classification of tetrads.....98

3.2. Troubleshooting guide.....99

3.3. Example of data produced by one individual plant.....100

3.4. Example of aggregate data comparing a mutant to wild type.....101

4.1. Sequence information for oligonucleotides used.....135

4.2. Fluorescent marker inserts used in this study.....136

5.1. Sites of open chromatin are preferentially associated with DSB hotspots.....192

5.S1. The association between sites of open chromatin and DSB hotspots is significant even when permutations are restricted to intergenic regions.....195

5.S2. Crossover and non-crossover hotspots are associated with open chromatin to the same degree.....196

5.S3. Enrichment of open chromatin at subtelomeric border regions and pericentromeric regions.....197

5.S4. Coordinates of proximal subtelomeric borders.....198

5.S5. DSB hotspots are not enriched at regions associated with with tRNA genes.....198

LIST OF FIGURES

Figure

1.1. The DSBR and SDSA meiotic recombination models.....32

1.2. Interference models.....33

2.1. Fluorescent markers in pollen tetrads.....62

2.2. Map of fluorescent transgenes.....63

2.3. Developmental and environmental effects on recombination.....64

2.4. Measuring interference.....65

2.5. Detecting gene conversion.....66

3.1. Map of fluorescent transgenes.....102

3.2. Single locus segregation patterns in pollen tetrads.....103

3.3. Classification of tetrad fluorescent patterns.....104

3.4. Examples of multi-color fluorescent tetrads.....105

4.1. The *AtMUS81* gene structure, T-DNA insertion mutant, and expression.....137

4.2. Hypersensitivity of *Atmus81* mutants to MMS, cisplatin, and gamma radiation.....138

4.3. Real-time qPCR analysis of *AtMUS81* transcription.....139

4.4. Pollen viability in wt and *Atmus81* mutants.....140

4.5. Meiotic recombination in the *Atmus81* mutant.....141

5.1. Monitoring meiotic entry and progression in the time course used for FAIRE.....185

5.2. The distribution of open chromatin and DSB hotspots across a representative 110 kb

genomic region containing 53 genes.....186

5.3. Chromatin at known hotspots is dynamic.....187

5.4. The degree of chromatin openness
is predictive of DSB hotspots.....188

5.5. Regions proximally adjacent to the subtelomere borders
exhibit meiosis-specific open chromatin.....189

5.6. Centromeres and pericentromeric regions exhibit
constitutively open chromatin.....190

5.7. A 15 bp motif is overrepresented in meiosis-specific
open chromatin surrounding tRNA genes.....191

5.S1. DSB hotspots can be divided into those that occur
in sites of open chromatin and those that do not.....193

5.S2. Sites identified as mitosis-specific open chromatin
are not just under the significance threshold in meiosis.....194

6.1. Using the FTL system to study different properties of GCs.....224

LIST OF ABBREVIATIONS AND SYMBOLS

AUC: Area under the curve

bp: Base pairs

cDNA: complimentary DNA

CEN: centromere

CFP: Cyan fluorescent protein

chIP-Chip: chromatin immunoprecipitation on microarray chip (technique)

cM: Centimorgan (1 cM = 1% recombination)

CO: Crossover

Col: Columbia ecotype of *Arabidopsis thaliana*

COI: Coefficient of interference

DCO: Double crossover

dHJ: Double Holliday junction

DNA: Deoxyribonucleic acid

DSB: Double-strand break

DSBR: Double-strand break repair (model)

dsRED: *Discosoma sp.* red fluorescent protein

ECFP: Enhanced cyan fluorescent protein

EN: Early Nodule

EYFP: Enhanced yellow fluorescent protein

FAIRE: Formaldehyde-assisted isolation of regulatory elements

FTL: Fluorescent tagged line

GC: Gene conversion

GUS: Beta-D-glucuronidase

H3K4me3: H3 histones trimethylated at lysine 4

HJ: Holliday junction

I: interval

JM: Joint molecule

kb: Kilobase (one thousand base pairs of DNA)

Ler: *Landsberg erecta* ecotype of *Arabidopsis thaliana*

LMS-PCR: Ligation-mediated suppression PCR

LN: Late nodule

mb: Megabase (one million base pairs of DNA)

MI: Meiosis I

MII: Meiosis II

MMS: methyle methanesulfonate

n: haploid chromosome number

NCO: Non-crossover

NPD: Non-parental ditype

ORF: Open reading frame

PCR: Polymerase chain reaction

PD: Parental ditype

PGM: Pollen growth media

PMS: Post-meiotic segregation

qPCR: quantitative PCR

qrt: *quartet* mutation

RACE: Rapid amplification of cDNA ends

rDNA: ribosomal DNA

RF: Recombination frequency

ROC: Receiver-operator characteristic

RN: Recombination nodule

RNA: Ribonucleic acid

RT-PCR: Reverse-transcriptase PCR

SC: Synaptonemal complex

SCO: Single crossover

SD: Standard deviation

SDSA: Synthesis-dependant strand annealing (model)

SEI: Single-end Invasion

SGD: Saccharomyces genome database

sHJ: Single Holliday Junction

ssDNA: Single-stranded DNA

T-DNA: Transfer DNA

TEL: telomere

tRNA: Transfer RNA

TT: Tetratype

wt: Wild-type

YFP: Yellow fluorescent protein

YPA: Yeast peptone acetate

YPD: Yeast peptone dextrose

ZMM: ZIP, MSH, MER group of genes

Δ: deletion mutant

X^2 : chi-square

Chapter 1

Genetic interference: don't stand so close to me

Luke E. Berchowitz and Gregory P. Copenhaver

Department of Biology and the Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280 USA

ABSTRACT

Meiosis is a dynamic process during which chromosomes undergo condensation, pairing, crossing-over, disjunction and segregation. Stringent regulation of the distribution and quantity of meiotic crossovers is critical for proper chromosome segregation in many organisms. In most sexually reproducing organisms, crossovers are more evenly spaced than would be expected from a random distribution. This phenomenon, termed interference, was first reported in the early 20th century by Drosophila geneticists and has been subsequently observed in organisms as diverse as fungi, plants and animals including humans. Despite the near ubiquity of interference and its importance in regulating meiotic recombination, the mechanism that mediate it remain poorly understood. In this review we provide historical background to the discovery and development of the concept of interference, explore which meiotic events are subject to interference, examine the leading models that attempt to explain it and pose a series of questions about interference that remain open and ripe for further experimentation.

BACKGROUND

Meiosis, a type of cell division, reduces the chromosomal complement by half to produce gametes that are essential for sexual reproduction. During meiotic prophase, chromosomes pair with their homologs and, in most organisms, undergo a physical exchange of DNA or an exchange of sequence information in a process called recombination [1]. Recombination is initiated by programmed double-strand breaks (DSBs) of chromosomes. During the repair of some DSBs, chromosome arms are exchanged generating crossovers (COs).

In most organisms, COs are not distributed randomly. Closely spaced COs are observed less frequently than would be expected from a random distribution. This phenomenon is known as crossover interference – though the more general term genetic interference may be more useful since there is growing evidence (described below) that other events can also interfere with one another. Alfred H. Sturtevant and Hermann J. Muller are typically given equal billing for the discovery of interference. Sturtevant clearly describes the phenomenon as early as 1913 [2]. Two years later, he coined the term "interference", though in doing so he gives credit to Muller for suggesting the name and also for his influence in discovering the phenomenon [3]. For clarity, in this review 'Interference' will refer to positive interference, which is the spacing of events that departs from random uniformity, as opposed to negative interference, which describes events that are more clustered than the null expectation. Although interference was originally observed almost a century ago and has subsequently been validated in numerous studies, fundamental questions regarding its underlying mechanisms still exist.

The goal of this article is to outline unanswered questions about interference and also to review existing models.

The genomic distribution of COs is regulated in multiple ways. For example, they are distributed such that each chromosome typically receives at least one, which is known as 'CO assurance.' A random distribution of COs among chromosomes predicts a class of chromosomes that have no COs, yet the observed number of chromosomes without a CO is quite small in most organisms [4, 5]. Growing evidence suggests that interference is the result of multiple levels of recombination regulation and CO assurance could be a result of the interference mechanism.

Meiotic DSBs are enzymatically catalyzed by a topoisomerase I-like protein called Spo11 that remains covalently attached to 5' ends of the break (**Figure 1**). Following Spo11 removal and further end processing (resection), the breaks are left with single-stranded 3' tails. One of these tails can then invade a non-sister chromatid, which is known as strand invasion. Stabilized strand invasion intermediates are known as single-end invasion (SEI) intermediates. The free 3'-OH in the SEI structure is used as a substrate by DNA polymerase to extend the 3' tail and the size of the displaced DNA strand (D-loop). The DNA synthesis that occurs at this stage is primed by one chromatid but uses a non-sister chromatid as a template – therefore any polymorphisms that exist at this locus will be copied from the template chromatid to the invading chromatid. This transfer of parental information is called gene conversion (GC) [6, 7]. At this point the invading end can dissociate from the non-sister chromatid and re-associate with the other end of the break in a process called synthesis dependant strand annealing (SDSA) [8]. After additional DNA synthesis and ligation the break is repaired resulting in a non-

crossover (NCO), potentially with associated GC if a polymorphism existed at the locus. It is important to note however that GC would result in heteroduplex DNA – i.e. that Watson and Crick strands of the converted chromatid would have non-complementary bases at the polymorphic site. The heteroduplex DNA can be recognized by the cell's mismatch repair system and either repaired such that the original parental genotype is restored or such that the converted genotype is kept. Alternatively, the mismatch repair system can fail to recognize the mismatch. In this case, both genotypes in the heteroduplex will be propagated during the next mitotic division in a process called post-meiotic segregation (PMS). PMS results in two populations of cells, one that has experienced GC at the polymorphic site and the other that has not.

If SDSA does not occur, then as the D-loop is extended it can hybridize to the single-stranded 3' tail on the other side of the break in a process called second end capture (**Figure 1**). Again, this structure can be acted on by DNA polymerase, which extends the second single-stranded 3' tail. As before, the priming and template DNA are from non-sister chromatids, therefore GC can occur during this stage. Following DNA synthesis and subsequent ligation, an intermediate called a double Holiday junction (dHJ) is formed. In principle, this structure can be resolved to produce either a crossover (CO) or NCO depending on how the individual junctions are cut to release the chromatids. However, evidence in *Saccharomyces cerevisiae* suggests that these intermediates are resolved predominantly as COs [4, 9, 10]. In either case, the resulting COs and NCOs can be associated with GC.

Early Observations

Sturtevant observed that in *Drosophila melanogaster* a three-point cross involving mutations on the X chromosome (yellow, white and miniature wings), a CO in the yellow-white interval occurred at a frequency of 1/69 gametes without the presence of a neighboring CO in white-miniature, but this frequency plummeted to 1/441 when there was a CO in the white-miniature interval (P = ~0.25) [3]. Sturtevant went on to define an "index of interference" as the expected probability of a double CO (the product of the two individual recombination frequencies) divided by the observed frequency of double COs. This marked the beginning of the debate as to the best way to measure interference, which is still being argued to this day [11]. With the index of interference and additional data Sturtevant decreed that "(interference) is less when the intervals are larger and vice versa [3]."

Data generated in Drosophila dominated advancements in interference research for the next 40 years [12-15]. Subsequent papers lacked empirical observations that changed the essential view of interference; instead they confirmed Sturtevant's initial observations and focused on different ways to measure interference. Nonetheless, several key observations were made. Muller first proposed that interference does not act between different chromosomes i.e. a CO on one chromosome does not affect the probability of a simultaneous CO in an interval on another chromosome [13]. Subsequent research by Weinstein established the reach of interference on the X chromosome. In this case, interference does not manifest when the distance between the two intervals being studied is greater than 46 cM [14]. Additionally, COs in intervals on the same chromosomal arm interfere more strongly than intervals on opposite arms that have approximately equivalent genetic distances [14]. This introduced the idea that interference does not

cross the centromere. In 1932, Graubard observed that chromosome 2 carrying an inversion (max size 25 cM) did not affect interference values of intervals on that chromosome [15]. This result was the first to suggest that pre-existing chromosomal features are not the primary determinant of interference, which is widely believed to this day.

Model fungal organisms that form tetrads (fused meiotic products), allow for both the recovery of non-reciprocal recombination products and novel statistical approaches for interference analysis [16-18]. Non-Mendelian 6:2 segregation (the result of post-meiotic mitotic division of a 3:1 tetrad) at a single locus, indicative of a GC, was first observed in spore pigmentation mutants of *Bombardia lunata* [18]. These observations led researchers to question if GCs were born from the same mechanism that produces COs. If true, this would predict that they should exhibit interference. Using *Neurospora crassa* as a model system, Stadler tested this idea by seeing if GC events interfered with the probability of COs in an adjacent interval. He found that GCs did not interfere with COs and (errantly) concluded that GCs and COs arose from different precursors [19]. Subsequent research by Mortimer and Fogel demonstrated that GCs occur with or without flanking marker exchange establishing the idea of NCO and CO repair [20]. These data suggested that all recombination events have a common molecular basis (which we now know to be DSBs and strand invasion [6, 7]) and that initial events are distributed independently of one another, but that the occurrence of a CO at one site will subsequently influence nearby events to be resolved as NCOs, thus resulting in a CO distribution that displays interference [20].

UNANSWERED QUESTIONS

What types of events interfere with one another?

Interference is often referred to as 'CO interference', which does not capture the breadth of the phenomenon, since COs may not be the only recombination-related events to inhibit one another's distribution. Determining what combinations of recombination events are subject to interference is key to understanding both when interference is imposed as well as the mechanism that mediates it.

DSB-DSB interference

Interestingly, in many organisms, not even all COs are subject to interference (discussed below). However, those COs that do interfere could theoretically be a reflection of an inhibition of closely spaced DSBs. Whether or not, and to what degree DSBs interfere with one another remains an open question. Meiotic DSB mapping studies in *S. cerevisiae* have resulted in a detailed understanding of where DSBs are most likely to occur [21-24], but they are not ideal for addressing whether or not DSBs are subject to interference because DSB mapping is an amalgamation of thousands of independent meioses so even if the DSBs in individual meioses were subject to interference, it would likely be obscured by the layering of data. It may be relevant that the hottest DSB hotspots only have a break in ~10% of meioses [21] and periodicity in DSB hotspot distribution has never been reported. However, in *S. cerevisiae*, researchers have clearly shown competitive DSB inactivation whereby insertion of a strong DSB hotspot reduces the frequency of DSB formation and recombination in nearby regions [25, 26]. Ohta *et al.* have reported that insertion of an artificial DSB hotspot results in a parallel decrease of

DSBs in a ~60 kb region around the insertion site [25]. This led to the proposal that strong DSBs sites could outcompete nearby sites for limiting factors essential for DSB formation. Another proposal is that structural features that influence DSB formation create domain boundaries that isolate regions from one another [25].

DSB distribution has also been addressed via mapping the physical position of structures called recombination nodules (RNs). RNs are proteinaceous structures that are associated with the axial elements of the synaptonemal complex (SC) from leptotene to pachytene and are thought to be locations where meiotic recombination reactions are occurring [27]. RNs are divided into two sub-classes: early nodules (ENs) and late nodules (LNs), which differ in respect to timing, size, shape and number. ENs, the smaller of the two, roughly correspond to Rad51/Dmc1 foci [28], while LNs, which most likely arise from a fraction of ENs, are thought to be the molecular machinery that execute COs and thus represent CO sites [27, 29]. In tomato, both ENs and LNs exhibit a distribution indicative of interference, but the strength of interference is much stronger among LNs [30, 31].

Analysis of the distribution of *MSH4* foci in mouse meiocytes strongly supports the idea that DSBs interfere with one another. In zygotene, ~150 *MSH4* foci are initially detected, which reduces to ~50 foci by late pachytene. Early *MSH4* foci co-localize with *RAD51*/*DMC1* and are thought to mark all DSBs in the early repair stages, while late *MSH4* foci co-localize with *MLH1* foci and are thought to mark CO sites only [32-34]. Early *MSH4* foci exhibit a distribution indicative of interference, but do exhibit as strong interference as *MLH1* foci [31, 35, 36].

Taken together the above results argue that interference is the result of multiple layers of control [31, 32, 36]. DSBs may exhibit positive interference over short distances, but DSB-DSB competitive interaction would not be sufficient to explain CO interference over megabase (Mb) distances.

NCO-NCO and NCO-CO interference

Another observation that suggests interference is not merely a reflection of an underrepresentation of closely spaced DSBs is that in *S. cerevisiae*, NCOs do not interfere with other NCOs. Mortimer and Fogel reported that GC events at *ARG4* and *THR1* (~19 kb apart) do not interfere with one another since alleles at the two loci co-converted at a frequency that is indistinguishable from independence predicted by their individual conversion frequencies. More recently, analysis of genome-wide recombination maps based on DNA tiling arrays reaffirmed that the distance between GCs not associated with COs does not differ significantly between experimental samples and randomized control data [37].

Whether or not COs interfere with NCOs (or vice versa) is a more controversial topic. The first results to address this issue came from *S. cerevisiae* when Mortimer and Fogel showed that a GC at *HIS1* or *ARG4* with exchange of flanking markers (CO) results in a decrease of genetic distance in an adjacent genetic interval, whereas a GC without exchange of flanking markers (NCO) actually promoted CO frequency in the adjacent interval (negative interference) [20]. The idea that NCOs do not interfere with COs was supported by Malkova *et al*. who showed that, at the *met13* locus of *S. cerevisiae*, GCs without an accompanying CO did not exert positive interference on

adjacent intervals, while GCs associated with a CO did [38]. However, they did not observe statistically significant negative interference between GCs not associated with CO at *met13* and COs. Recently, Getz *et al*. showed that NCO GCs did not exert interference on adjacent intervals and the map distances in those intervals were actually increased indicative of negative interference, in support of the idea that NCOs do not positively interfere with COs [39]. However, most recently, a contradictory result showing that NCOs and COs interfere with one another was recently presented by Mancera *et al.* who, using the genome-tiling method, showed that inter-event distance between NCO and CO events were on average ~13kb larger than expected from a random distribution, representing statistically significant, albeit weak, positive interference [37]. While locus-by-locus studies have consistently shown negative or no interference between NCOs and COs, this genome-wide approach showed the opposite. It is difficult to reconcile these contradictory results, but because the Mancera data were generated using a genome-wide analysis the advantage appears to be with the idea that COs and NCOs exhibit positive interference. Nonetheless, it is interesting to note that models in which DSBs that are resolved as COs influence nearby DSBs to be resolved as NCOs predict negative interference between COs and NCOs. One reconciliatory possibility is that there exist both interfering NCOs and non-interfering NCOs (as in COs, see below) and the single locus studies mentioned above happened to measure only the latter class.

CO-CO interference and non-interfering COs

In many organisms, there are at least two pathways for producing COs. *Arabidopsis thaliana*, humans, mouse and *S. cerevisiae* have one pathway constituting the majority of

COs that is sensitive to interference and a secondary pathway that produces interference-insensitive (randomly distributed) COs [40]. In these organisms, primary pathway COs are characterized by the Msh4-Msh5 heterodimer while secondary pathway COs are dependent on the Mus81-Eme1 heterodimer [40]. In *S. cerevisiae*, *msh4Δ* or *msh5Δ* deletions have a ~60% reduction in COs and the remaining COs are interference insensitive [41-43], while *mus81Δ* or *mms4Δ* (*eme1*) deletions have a ~25% reduction in COs and the remaining COs are sensitive to interference [40, 41]. In Arabidopsis, an analogous situation exists where interference-sensitive COs mediated by the Msh4-Msh5 pathway make up ~80-85% of the total while interference insensitive COs mediated by the Mus81-Eme1 pathway make up ~15-20% of the total [44-47]. Not all organisms produce both interfering and non-interfering COs. *Schizosaccharomyces pombe* does not have CO interference and ~80-95% of its COs are dependent on Mus81-Eme1 [48, 49]. In contrast, *Caenorhabditis elegans* has 'perfect' interference (exactly one CO per bivalent) and all COs are dependent on Msh4-Msh5 [50]. *D. melanogaster* is not thought to produce interference-insensitive COs and is thought to exhibit absolute interference across distances < ~10 cM [51]. A comparison of interference patterns in different organisms is presented in Table 1.

While it is known that both interfering and non-interfering COs can be produced in the same cell, the mechanistic difference between the two and specifically why one class exhibits interference and the other does not is unknown. One possibility is based on the work in Arabidopsis by Franklin *et al.* in which they found that, during leptotene, ~115 *AtMUS81* foci form on chromosome axes and many of these co-localize with *AtRAD51* and *AtMSH4* foci, each of which form 80-100 foci during leptotene/early

zygotene [47]. By early pachytene, the number of *AtMUS81* foci drops precipitously to ~5. This suggests a 'toolbox hypothesis' where Mus81-Eme1 and Msh4-Msh5 are recruited to all DSBs and most are repaired via the Msh4-Msh5 pathway, while Mus81-Eme1 acts to resolve a subset that may consist primarily of aberrant joint molecules (JMs) as either COs or NCOs that could not be repaired using Msh4-Msh5. This could result in interfering and non-interfering COs (and perhaps non-interfering NCOs) if the Msh4-Msh5 pathway is subject to interference, but the smaller and randomly distributed population of aberrant JMs resolved as COs by Mus81-Eme1 acts later, after interference has already been established. In congruence with this idea, recent biochemical and genetic analysis of *mus81* mutants in *S. cerevisiae* indicates that Mus81-Eme1 acts late in meiotic recombination and likely resolves aberrant JMs that cannot be resolved by the primary Msh4-Msh5 pathway [52, 53] (**Figure 1**). Another related possibility is that primary CO reactions occur first and initiate an 'interference signal' (discussed below), while secondary COs take longer to process, as they must first be bypassed by the primary pathway. In this model, secondary COs do not produce an interference signal, and are resolved after the interference signal has been imposed.

An alternate hypothesis regarding the difference between interfering and non-interfering COs was introduced by Getz *at al*., who proposes two phases of COs. Early 'pairing phase' COs are non-interfering and late 'disjunction phase' COs, that are dependent on *MSH4*, exhibit positive interference [39]. In addition to being independent of *MSH4*, pairing phase COs are hypothesized to be less proficient at repairing mismatches. The hypothesized existence of MSH4-independent 'pairing phase' COs in *S. cerevisiae* is consistent with the lack of non-interfering (Mus81-dependent) COs in *C.*

elegans and Drosophila, because these organisms do not use COs to pair their chromosomes [54-56]. This model is also compatible with the phenotype of *ndj1* mutants, which have decreased interference and increased rates of nondisjunction [57, 58], which can be explained by increased pairing COs at the expense of disjunction phase COs. Since pairing phase COs are proposed to be interference insensitive it follows that they should be *MUS81*-dependent, which predicts that *mus81* and *eme1* mutants will be pairing defective. In yet another layer of complexity the toolbox and two-phase hypotheses are not mutually exclusive as there could be pairing phase non-interfering COs that have nothing to do with *MUS81* as well as non-interfering disjunction phase COs produced by *MUS81* that are reflective of *MSH4*-independent aberrant JM resolution.

What is the timing of events leading to interference?

Determining the timing of events that lead to interference is extremely challenging since diverse cellular processes likely play a role. Chromatin structure (e.g. nucleosome density, protein-DNA complexes, histone modifications etc.) and steric features of the chromosomes could influence recombination complex spacing. These features are not constant along chromosomes and are dynamic in both the mitotic and meiotic cell cycles. Meiotic chromosome condensation, which begins at the start of meiotic prophase and does not end until after recombination is complete, also likely influences interference [1, 29, 59]. However, pre-recombination chromosomal features are not the only important determinants. In many organisms, mutations in the meiosis-specific ZMM (*ZIP*, *MSH*, *MER*) recombination genes, which act after SEI formation, result in the abolition of

interference. This strongly suggests that the assembly and distribution of recombination complexes is critical for the timing of the imposition of interference.

One attractive proposal is that interference is imposed during strand invasion when Msh4-Msh5 complexes stabilize CO-specific (SEI) recombination intermediates [60]. This idea is based on the observations that *msh5 ndt80* mutants result in very low levels of JM accumulation along with absence of interference in *S. cerevisiae* [61], but *spo16 ndt80* mutants, which are defective for synaptonemal complex (SC) extension, result in high JM accumulation but wild-type interference [60]. The *ndt80* mutation was used in this case because it removes the late pachytene checkpoint and results in accumulation of recombination intermediates [60]. The *spo16* mutant also offers important insight as to the latest interference could be acting. Because *spo16* mutants are defective for SC extension and yet have wild-type interference it is likely that interference is fully implemented before late leptotene/early zygotene when the SC is formed [1]. Supporting this idea is the observation that SC initiation complexes exhibit a distribution indicative of interference [62]. Additional support for the idea that interference involves regulation of the strand invasion step comes from analysis of the *tid1* mutant in *S. cerevisiae* [63]. Tid1 is an accessory factor that facilitates strand invasion and the *tid1* mutant displays ~wild-type levels of COs yet interference is significantly weakened [63].

The timing of events leading to interference is different in Drosophila and *C. elegans*, in which pairing and synapsis occurs prior to the initiation of recombination. Thus, interference in these organisms is likely implemented after (though not dependent on) SC formation.

Do COs influence nearby DSBs to be repaired as NCOs?

Analysis of the ZMM mutant phenotypes has led to the early decision model, in which the commitment of a DSB to be repaired as either a NCO or CO is made at, or prior to, stable SEI formation [10, 64]. How this decision is enforced is unknown. ZMM mutants are strongly CO defective, have abolished or greatly reduced interference, yet are not defective for NCO formation [42, 65, 66]. 2-D gel analysis of DNA intermediates has shown that the kinetics of DSB and NCO repair are normal in ZMM mutant backgrounds, but stable SEI production is inhibited. This observation strongly suggests that ZMM genes are not required for designation of DSB repair as either COs or NCOs and that this designation is made prior to SEI formation [10, 64]. NCO/CO designation is a key determinant of CO distribution and thus interference. It is important to determine if designation of a CO at one site subsequently results in NCO designation of nearby DSBs, which is an attractive, but as yet unproven, proposal. The other possibility is that all DSBs are designated as either COs or NCOs independently of one another in a distribution reflective of CO interference.

What is the role of *PCH2* in the mediation of CO interference?

Recent mutant analyses strongly implicate the AAA+ ATPase Pch2 as an important regulator of CO interference, however the pathway by which Pch2 mediates interference is unknown. Two independent studies in *S. cerevisiae* showed that *pch2*Δ mutants exhibit significantly weakened interference at several loci [67, 68]. Interestingly, both groups reported no significant changes in CO frequency on chromosome III (the shortest *S.*

cerevisiae chromosome) indicating that the processes of CO formation and CO interference can be decoupled, at least on short chromosomes. However, the two studies presented incongruent results regarding CO frequency on other chromosomes in that Joshi *et al*. report no significant changes in CO frequency at any loci [67] while Zanders and Alani report significant increases in CO frequency on medium and large chromosomes [68].

*pch2*Δ mutants display elevated Zip3 foci (an early marker of CO designated sites), aberrantly diffuse Hop1 localization and increased SC length [67]. In light of these findings, the researchers propose a short-range interference model (discussed further below) in which Pch2 aids in the establishment of chromosomal domains in which only one CO can occur [67]. In addition to increased CO frequency, *pch2*Δ mutants display no increase in DSBs and exhibit elevated CO:NCO ratios at two GC loci, compared to wild-type [68]. Importantly, the researchers rule out excess COs as the explanation for weakened interference in *pch2*Δ since *pch2*Δ*;* Spo11-hypomorph double mutants display ~74% of the COs of *pch2*Δ while maintaining the interference defect. Additionally, *pch2*Δ; *mms4* double mutants have significantly higher CO frequency than *mms4* mutants, which one would not see if the interference defect and CO frequency increase was solely due to elevated secondary pathway (Mus81-Mms4 dependent) COs in *pch2*Δ [68]. To explain the *pch2*Δ phenotypes of increased CO frequency, increased CO;NCO ratio and weakened interference, the researchers propose that Pch2 acts to repress the CO designation at the CO/NCO bifurcation in the decision pathway [68]. While *PCH2* very likely plays an important role in interference, this gene is a piece in the puzzle since a) the interference defect in *pch2*Δ is not present at all loci that display interference [67] b)

*pch2*Δ mutants have residual positive interference [67, 68] and c) the interference defect in is most prominent in 50-100kb distances and can be mitigated at lower temperatures [67].

Do meiosis-specific cohesins mediate interference?

In *S. cerevisiae*, the meiotic cohesin complex consists of Scc3, Smc1, Smc3 and Rec8, all which are also present in the mitotic cohesin complex except Rec8 [69, 70]. In addition to its role in sister chromatid cohesion, Rec8 has been implicated in a diverse set of meiotic processes including pairing, SC polymerization, recombination, and disjunction [70, 71]. Many of these functions have been shown to be separable from its role in sister chromatid cohesion [71]. In *S. cerevisiae* and *S. pombe*, Rec8 has a role in the binding of Spo11 to DSB sites in a region-specific manner [72, 73]. ChIP-chip reveals an interesting correlation between binding of Spo11 and cohesins. Spo11 has been shown to co-localize with Rec8 in early meiosis, and the frequency of co-localization decreases as meiosis progresses [73]. It has been proposed that Rec8 provides structural landmarks that dictate the proper distribution of Spo11 [73]. Spo11 could transfer from Rec8 binding sites to chromatin loops before initiating DSBs [73]. Since Rec8 has a role in dictating DSB distribution, a role in the establishment of the weak DSB-DSB interference discussed earlier is possible. The Rec8 binding landscape could serve to bias Spo11 distribution toward uniformity. However, a role for Rec8 in the mediation of interference has not yet been demonstrated, so additional work to elucidate this relationship will have to be conducted.

MODELS

The central unanswered question regarding interference is how it is achieved within the cell. The answer to this question has been elusive partly because traditional genetic screens designed to discover interference mutants are labor-intensive and problematic. Isolating the interference machinery is difficult because the relevant players likely have overlapping roles with other critical meiotic processes. Most mutations that affect interference also affect CO frequency and only a handful of mutants has been identified (all in *S. cerevisiae*) that exhibit an interference defect without a concurrent hypo- or hyper-CO phenotype [57, 63, 67]. Further complicating matters is that many of these mutants behave differently with regards to CO frequency in different studies and within studies at different loci [39, 57, 63, 67, 68, 74]. Additional complications arise from the fact that interference is not absolute in most species but instead reflects the reduced probability of an event in a population of events – making the implementation of efficient screens difficult. Models aimed at explaining the interference mechanism have been proposed, but there is no dominant paradigm as no empirical observations vastly favor one particular explanation. Several commonly referenced models are discussed below.

Mechanical stress model

Because interference is maximized at close distances and decreases with increasing distance [3], interference models that rely on a mechanical explanation invoke a signal that spreads from CO sites. Muller's original interference model posited that the stiffness of a chromosome would make it difficult to bend back on itself after a CO to form another in close proximity [13]. The modern mechanical stress model proposed by

Kleckner *et al*., is based on the idea that in many physical systems, any increase or decrease in stress starts at a locus and propagates outward from that point (**Figure 2**) [59]. Stress is generated as meiotic chromosomes compress and expand. COs result in a localized relief of this stress that spreads in both directions down the axis of the SC. This model is attractive because it posits a simple explanation that predicts many properties of interference including CO assurance and CO homeostasis (discussed below). As each chromosome will be under stress, the occurrence of a first event, CO assurance, is easily obtained. Each event defines a domain of inhibition resulting from stress relief spreading from a CO, which acts as an inhibitory signal. Chromosomal features such as centromeres could act as sinks that absorb the signal, which would explain how interference does not cross the centromere. Multiple events could occur on the same chromosome, and could take place in different locations in different nuclei, but would always have the tendency to be evenly spaced. A mathematical simulation of this model has been used to fit CO data from two species [59]. However, the stress model does not make easily testable predictions. Additionally, it is easy to imagine how DSBs would relieve tensile stress, but less so for COs, which would be necessary for this model to explain all aspects of interference. Furthermore, it is also difficult to account for how some COs could mediate stress relief while others don't which would be necessary to explain the data in organisms that appear to have both interfering and non-interfering crossovers.

An additional layer to the mechanical stress model was put forward to explain the role of Pch2 in the mediation of short-range (< 100 kb) CO-interference [67]. Joshi *et al*. propose a one-CO module hypothesis in which Pch2 aids in the establishment of domains

that tile chromosomes and incur one and only one CO per module [67]. Mechanical stress within each module promotes a single CO that that relieves stress within the module. Key features of one-CO modules, presumably mediated by Pch2, are a centrally located Zip3 focus along with Hop1 hyper-abundance that extends to the edges of the domain, establishing the reaches of short-range interference [67]. The hypothesized contents of the modules are based on the observation that *pch2*Δ mutants, along with weakened interference, display diffuse rather than domainal Hop1 staining and aberrant increased Zip3 foci [67]. The one-CO module model is capable of explaining interference across organisms with larger or smaller interference reach by varying the size of modules. For example, in *C. elegans*, which incurs exactly one CO per bivalent, each chromosome could be encompassed in entirety by one module. The one-CO module model does not explain interference over >100 kb distances in *S. cerevisiae* since in the pch2Δ mutant, where module establishment is presumed to be impaired, interference is unaffected in distances over 100 kb.

Polymerization model

The polymerization model describes a situation where early recombination structures are distributed independently of one another and then have an equal chance per unit time of initiating a bi-directional polymerization event (**Figure 2**) [75]. This polymer spreads from the site of initiation and has the ability to block additional early structures from binding to the bivalent. Sites of initiation are hypothesized to mature into LNs (COs) leading to chiasmata. This model is attractive because it explains interference and assurance while predicting a pattern in which interference is strongest nearest to initiation

events with decreasing strength in a distance-dependant manner. A computerized simulation of the parameters described in the polymerization model was fit satisfactorily to CO data from *Drosophila* and *S. cerevisiae* [75].

Part of the rationale behind the polymerization model was that an optimal interference model should be useful in systems that differ by several orders of magnitude in genome size in bp. Physical distances measured in SC lengths rather than bp are much closer among species, thus the polymer is proposed to move down the axis of the SC. The idea that interference mediated over physical distance would be measured in SC length and that the interference signal is propagated along the SC axis is attractive because organisms of vastly different size genomes can be normalized by modifying the size of DNA loops that are associated with the SC axis. The main obstacle to the polymerization model is that the polymer itself has neither been identified nor observed. However, the search for a polymer as the signal may be a red herring as the signal could be a modification such as phosphorylation, methylation, acetylation or ubiquitination of a protein such as a histone or cohesin. SC polymerization itself has been proposed as an attractive mediator of interference [76, 77]. This would have demonstrated a clear role for the SC while being consistent with the observation that *S. pombe* lacks both SC and interference. However, the possibility that the SC is required for interference is extremely unlikely since in *S. cerevisiae* SC extension has been shown to be dispensable for interference [60] and in mouse, SC defective mutants have normal interference [31]. Lastly, in refutation of the idea that the SC is in any way required for interference, is that in Drosophila and *C. elegans* the SC seems to be complete before DSBs are formed.

Counting model

The counting model is a mathematical construct in which COs are separated by a fixed number (*m*) of intervening NCO events (**Figure 2**) [78]. It was proposed in part to reconcile the fact that, in terms of physical distance (bp of DNA or μm of SC), strength of interference varies by several orders of magnitude from organism to organism [78, 79]. The counting model can account for vast differences in genome size, as interference is dictated by genetic distance i.e. the initial density of precursors and the number of intervening events between COs. CO data from *Drosophila* and *N. crassa* fit the counting model predictions extremely well [78], but initially it was less successful at modeling CO distributions in *S. cerevisiae* and humans [79]. Additionally, *m* appeared to vary in the same organism between sexes and chromosomes [80]. Subsequently, a modified version of the counting model that allows a number of non-interfering COs (v) was suggested. This additional parameter allowed the counting model to satisfactorily fit CO data from *S. cerevisiae* [81], *A. thaliana* [45, 82], and humans [80] and in each of these systems, it has been shown that non-interfering COs exist [40, 44, 83].

Proposed biological equivalents for each parameter in the counting model include DSBs as precursors and NCOs as intervening events (*m*), however, the notion that DSBs are what is being 'counted' is not supported by subsequent observations. A prediction of the counting model when DSBs are 'counted' is that when the overall number of DSBs are reduced, COs and NCOs should reduce proportionally. It then follows that larger distances between COs should result. However, recent results failed to meet this expectation [84]. Cells appear to have a mechanism (called CO homeostasis) that ensures a certain number of COs per meiosis even if the pool of DSBs from which they arise is

reduced. A series of *spo11* hypomorphic mutants that produce DSBs in decreasing frequencies compared to wild-type do not show proportional reductions in COs [84]. Instead, these mutants maintain CO levels at the expense of NCOs, which are reduced proportionally to the reduction in DSBs. CO interference is maintained at wild-type levels in all *spo11* hypomorphic backgrounds [84]. CO homeostasis is seen as a major obstacle to the counting model. However, if DSBs were not what is counted, but instead a factor that establishes DSB sites, the counting model could still be possible. Despite its clear predictions and modeling support, the counting model suffers from an absence of both *in vivo* evidence as well as a testable molecular mechanism for its execution.

Other Models

Interference models can be described in terms of point-process in which precursors i.e. 'points' are distributed according to a mathematical function and then COs are 'processed' from these points according to a second mathematical function [85]. The 'hard-core' model is a point-process variant where points are dispersed in a Poisson distribution with a minimum physical distance between any two [85]. At each point, chromatids have a 50% chance of being involved in a CO. The hard-core model is much like a situation where DSB-DSB interference is the driver of CO interference, which current data strongly suggests is not the case (discussed above). This model does not conform particularly well with CO data from *Drosophila* [85].

Another interesting proposal involves a 'chiasma determining mechanism' that moves along the bivalent [78, 86]. This mechanism is hypothesized to move along the bivalent at a constant rate occasionally firing and thus determining CO sites that mature

into chiasmata. After firing, the mechanism requires time to recharge while still moving, resulting in CO interference [86].

CONCLUSIONS

How is interference imposed?

Interference modeling has traditionally been concerned with the synthesis of models capable of explaining CO interference across many organisms with vastly different genome sizes and recombination rates. Beyond this, accurate models must be able to reconcile additional complex properties of interference such as multi-level control, non-interfering COs and early decision. Even what falls under the purview of the term interference is currently in question. In addition to the strong likelihood that recombination events other than COs interfere, it is also possible that both CO homeostasis and CO assurance could be products of an over-arching CO control mechanism that results in the lack of closely spaced COs. While the above models can be altered to account for these complexities, the truth likely lies in combination of the ideas that have been proposed to explain interference.

Since a growing body of evidence supports the idea that interference is imposed at multiple levels, a comprehensive interference model will account for the presence of DSB-DSB interference, CO-CO interference and possibly CO-NCO interference. Although it remains unclear whether or not and to what degree DSBs interfere with one another, it is possible that DSB-DSB interference acts on a smaller scale and is mediated by multiple factors. First, the pattern of Spo11 distribution may be established by the meiosis-specific cohesin subunit Rec8. Rec8 has been shown to colocalize with Spo11 in

S. cerevisiae, and *rec8*Δ mutants exhibit a drastic alteration in Spo11 distribution [73]. It has been speculated that Rec8 not only provides landmarks along the chromosomal axes that guide the distribution of Spo11, but it is also responsible for the transition of Spo11 from the axes to the loops, where breaks are subsequently formed [73] Rec8 could preferentially bind certain locations within the chromatin context of the chromosomal axis with intervening DNA loops, resulting in a minimum physical distance between subunits. This idea requires that only one DSB could be formed at each site of Rec8 localization, which would result in the minimization of closely spaced DSBs thus establishing a pattern of DSB-DSB interference on a small scale. Secondly, DSBs could recruit limiting break forming factors away from nearby sites [25].

In addition to small-scale DSB-DSB interference, we are intrigued by the possibility that CO-CO interference in *S. cerevisiae* and *A. thaliana* is mediated by an 'interference signal' that is initiated by the stabilization of SEIs by ZMM proteins. This hypothetical interference signal propagates either down the SC axial elements or the cohesin axis. A signal propagated down the axis rather than the DNA itself is attractive because it allows interference to act over very large distances of linear DNA by varying the DNA bp loop/axis ratio. One-CO modules can be incorporated as domains established on the chromosomal axes that create favorable conditions for signal propagation. Importantly, the SC itself is not required for interference in this model. Additionally, if we suppose that the amount of DNA traveled by the signal is dependent on the condensation of the chromosome, it could explain the 'lack' of interference in early pairing phase COs and 'presence' of interference in disjunction phase COs. COs could occur at various time points in the meiotic program and thus over a large range of states

of chromosome condensation. The earlier a (interference-sensitive) CO is designated, the less it will interfere (and vice versa) due to the continual condensation of the chromosomes during the meiotic program. This idea predicts that if CO formation can be in some way be delayed, interference will strengthen, a hypothesis that can be most effectively tested in an organism (such as Drosophila), which has a single interference-sensitive CO pathway.

What is the nature of the interference signal? An intriguing and testable explanation is that a protein modification that is propagated from the majority of COs results in the inhibition of primary-pathway crossing over at sites where this modification spreads. The modified protein(s) could include axial elements, cohesins, or histones. This idea predicts that CO-CO interference can be largely eliminated by a mutation that blocks the spread of the inhibitory modification via the receiver or the modifier. Rec8 is a particularly interesting target since phosphorylation of Rec8 is important for many of its meiotic roles. Two *rec8* mutants with mutations at multiple phosphorylation sites exhibit disrupted synapsis and have delayed production of mature recombinants [71]. Since Rec8 phosphorylation is required for recombination in *S. cerevisiae*, the protein modification could be a de-phosphorylation of Rec8 that disables crossing over. A second possibility is de-methylation of H3K4me3, which has been shown to mark sites of meiotic recombination [87]. Thirdly, modification of axial-element protein Hop1, which is localized in discrete hyper-abundant domains on zygotene chromosomes, is an interesting possibility. *pch2*Δ mutants abolish domainal localization of Hop1 while concurrently weakening interference [67].

An alternate model for the interference signal is based on the recent observation that CO hotspots are significantly correlated with DNA methylation [88]. This group proposed that areas undergoing recombination could be secondarily methylated, which could result in the inhibition of further COs in that area [88]. Another possibility is that genomic regions that are methylated are preferential sites of recombination. These possibilities are not mutually exclusive and make testable predictions regarding local DNA methylation states prior to and after recombination has taken place.

Final thoughts

As the initial report of interference reaches the century mark, some questions regarding how it works are in reach. These questions can be posed within the hypothesis that ZMM proteins stabilize CO intermediates at strand invasion (SEI), which are then committed to an interference-sensitive CO pathway. Does the stabilization of SEI complexes activate a spreading interference 'signal' or are CO-designated events evenly spaced to begin with? Do COs occur in two phases, one interfering and the other not? Does Rec8/Spo11 distribution play a role in interference? Does Mus81-Eme1 mediate the resolution of only aberrant recombination intermediates or does it also play a role in mediating a subset of traditional substrates? These questions and many more will have to be answered in order to solve the interference puzzle.

ACKNOWLEDGEMENTS

LEB and GPC thank NSF (MCB-0618691) and DOE (DE-FGO2-05ER15651) for financial support. We would also like to thank Drs. Corbin Jones, Jeff Sekelsky and Frank Stahl for critical comments.

TABLE LEGENDS

1.1. CO interference comparisons across model genetic organisms. Haploid chromosome number (n) and presence or absence of CO interference is noted. Also shown are presence or absence of Msh4-Msh5 (interference-sensitive) and Mus81-Eme1 (interference-insensitive) mediated CO pathways.

FIGURE LEGENDS

1.1. The DSBR and SDSA meiotic recombination models. Single strands of DNA are shown as either green (parent 1) or yellow (parent 2) rods. The Spo11 complex initiates programmed DSBs. DSBs are resected 5' to 3' to produce single ssDNA tails. ssDNA tails invade the homologous template which is aided by the ssDNA filament forming proteins Dmc1 and Rad51. At this stage, intermediates can undergo DSBR (left), which is thought to produce primarily COs or SDSA (right), which only produces NCOs. Also shown is a pathway (center) describing aberrant JMs, that are hypothesized to either be resolved back to the strand invasion stage by Sgs1 or resolved as COs by the Mus81-Eme1 heterodimer [52]. In the DSBR pathway, strand invasion complexes are stabilized (possibly by ZMM proteins) to form SEIs. Prior to stabilization, Sgs1 could wire SEIs back to the strand invasion stage [52, 89]. The displaced strand, called a D-loop is

captured by the resected break of opposite homolog and subsequent DNA synthesis results in a dHJ intermediate. This intermediate is resolved as a CO upon appropriate resolution of the two HJs. in the SDSA pathway, the invading strand dissociates after a patch of DNA synthesis. This strand then re-anneals to the original parent, resulting in repair of the DSB and a patch of heteroduplex DNA. This pathway is always resolved as a NCO, but it can result in GC.

1.2. Interference models. The left panel depicts the beam-film demonstration of the mechanical stress model proposed by Kleckner *et al* [59]. The beam (chromosomal axis; green), film (chromatin fiber; grey), flaws (CO precursors; black dots). Diagrams depicting the stress level are shown under each beam in which the x axis represents beam position and stress level on the y. The center panel depicts the polymerization model proposed by King and Mortimer [75]. Chromatids are shown in green (parent 1) and yellow (parent 2). Small light blue circles represent recombination precursors and CO designates are shown as larger circles marked with 'CO.' The interference polymer is shown as a large arrow emanating from CO sites, and CO precursors removed by the polymer are shown to the right accompanied with a dashed arrow. The right panel depicts the counting model proposed by Foss *et al*. [78]. Chromatids are shown in green (parent 1) and yellow (parent 2). Small light blue circles represent recombination precursors and CO designates are shown as larger circles marked with 'CO.' In this diagram, m=3 and intervening NCOs between COs are outlined in a red box.

TABLES

Table 1. 1.

Organism	N	Interference?	Msh4-Msh5 COs?	Mus81-Eme1 COs?	~COs/meiosis
Saccharomyces cerevisiae [24, 37, 40, 41, 43]	16	Yes	Yes	Yes	90
Schizosaccharomyces pombe [48, 90]	3	No	No	Yes	38
Neurospora crassa [75, 78, 91]	7	Yes	n.d.	n.d.	20
Aspergillus nidulans [92]	8	No	n.d.	n.d.	n.d.
Caenorhabditis elegans [93]	6	Yes	Yes	No	6
Arabidopsis thaliana * [44-47, 82, 94]	5	Yes	Yes	Yes	10
Lycopersicon esculentum [30]	12	Yes	n.d.	n.d.	21
Zea mays [95, 96]	20	Yes	n.d.	n.d.	20
Drosophila melanogaster [27, 51, 78]	4	Yes	No	No	6
Danio rerio * [97, 98]	25	Yes	n.d.	n.d.	25-40
Mus musculus * [31, 32, 35, 99-101]	20	Yes	Yes	Yes	22-28
Homo sapiens * [32, 34, 99, 102-107]	23	Yes	Yes	Likely	50-70

n.d. = no data

* = Reported differences between male and female COs/meiosis

FIGURES

Figure 1.1.

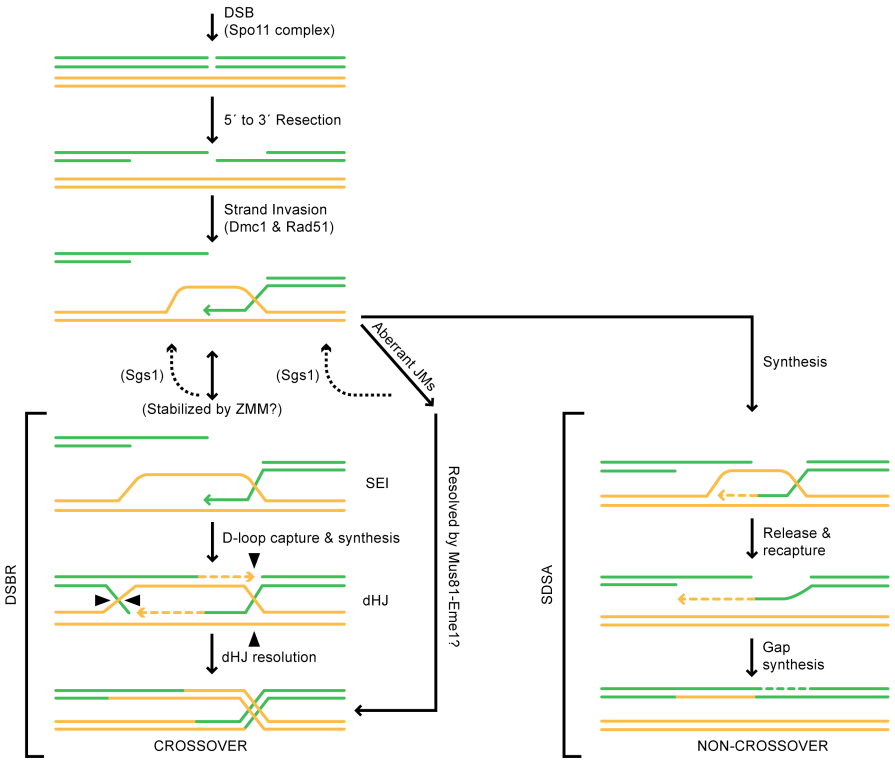
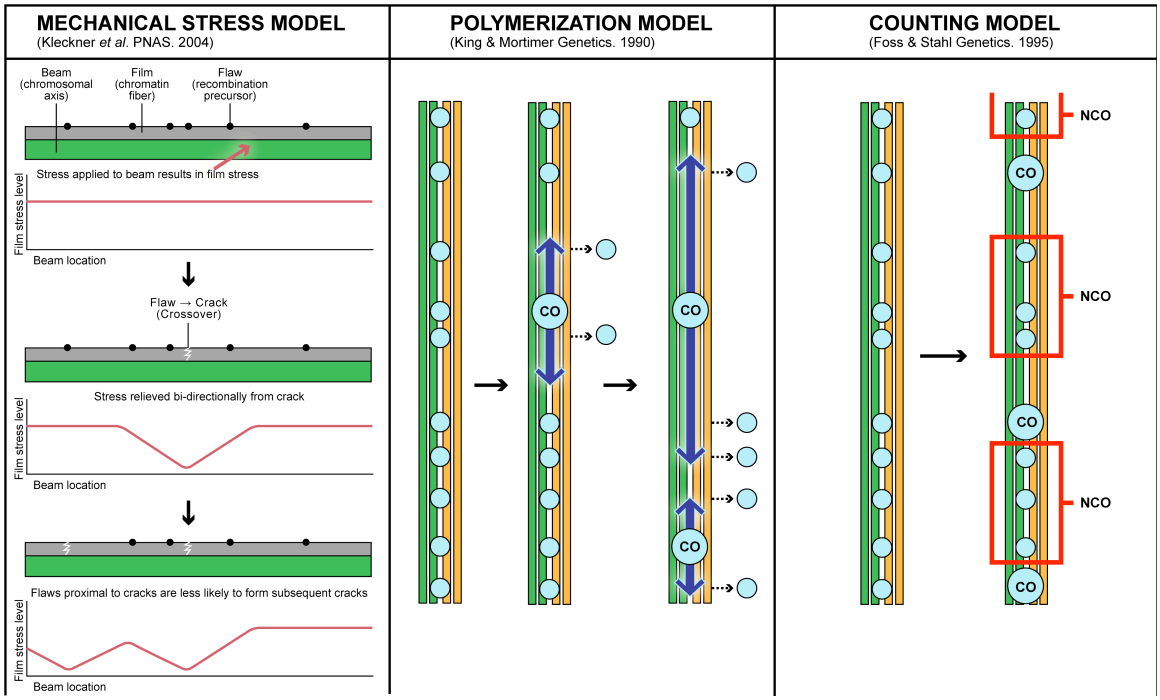


Figure 1.2.


MECHANICAL STRESS MODEL
(Kleckner et al. PNAS. 2004)
POLYMERIZATION MODEL
(King & Mortimer Genetics. 1990)
COUNTING MODEL
(Foss & Stahl Genetics. 1995)
Beam (chromosomal axis)
Film (chromatin fiber)
Flaw (recombination precursor)
Stress applied to beam results in film stress
Film stress level
Beam location
Flaw → Crack (Crossover)
Stress relieved bi-directionally from crack
Flaws proximal to cracks are less likely to form subsequent cracks
CO
NCO

REFERENCES

1. Zickler, D. and N. Kleckner, *The leptotene-zygotene transition of meiosis.* Annu Rev Genet, 1998. **32**: p. 619-97.

2. Sturtevant, A.H., *A Third Group of Linked Genes in Drosophila Ampelophila.* Science, 1913. **37**(965): p. 990-992.

3. Sturtevant, A.H., *The behavior of chromosomes as studied through linkage.* Z. Induct. Abstammungs-Vererbungsl, 1915. **13**: p. 234-287.

4. Borner, G.V., N. Kleckner, and N. Hunter, *Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis.* Cell, 2004. **117**(1): p. 29-45.

5. Jones, G.H., *The control of chiasma distribution.* Symp Soc Exp Biol, 1984. **38**: p. 293-320.

6. Szostak, J.W., et al., *The double-strand-break repair model for recombination.* Cell, 1983. **33**(1): p. 25-35.

7. Sun, H., D. Treco, and J.W. Szostak, *Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site.* Cell, 1991. **64**(6): p. 1155-61.

8. Paques, F. and J.E. Haber, *Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae.* Microbiol Mol Biol Rev, 1999. **63**(2): p. 349-404.

9. Lynn, A., R. Soucek, and G.V. Borner, *ZMM proteins during meiosis: crossover artists at work.* Chromosome Res, 2007. **15**(5): p. 591-605.

10. Bishop, D.K. and D. Zickler, *Early decision; meiotic crossover interference prior to stable strand exchange and synapsis.* Cell, 2004. **117**(1): p. 9-15.

11. Stahl, F.W. and E.A. Housworth, *Methods for Analysis of Crossover Interference in Saccharomyces cerevisiae*, in *Meiosis volume 1*, S. Keeney, Editor. 2009, Springer.

12. Muller, H.J. and J.M. Jacobs-Muller, *The Standard Errors of Chromosome Distances and Coincidence.* Genetics, 1925. **10**(6): p. 509-24.

13. Muller, H.J., *The Mechanism of Crossing Over.* Am Nat, 1916. **50**: p. 193-434.

14. Weinstein, A., *Coincidence of Crossing over in DROSOPHILA MELANOGASTER (AMPELOPHILA).* Genetics, 1918. **3**(2): p. 135-72.

15. Graubard, M.A., *Inversion in DROSOPHILA MELANOGASTER.* Genetics, 1932. **17**(1): p. 81-105.

16. Fogel, S. and D.D. Hurst, *Meiotic gene conversion in yeast tetrads and the theory of recombination.* Genetics, 1967. **57**(2): p. 455-81.

17. Mitchell, M.B., *ABERRANT RECOMBINATION OF PYRIDOXINE MUTANTS OF Neurospora.* Proc Natl Acad Sci U S A, 1955. **41**(4): p. 215-20.

18. Zickler, H., *Genetische untersuchungen an einem heterothallischen askomuzeten (Bombardia lunata nov. spec.).* Planta, 1934. **22**: p. 574-613.

19. Stadler, D.R., *The Relationship of Gene Conversion to Crossing over in Neurospora.* Proc Natl Acad Sci U S A, 1959. **45**(11): p. 1625-9.

20. Mortimer, R.K. and S. Fogel, *Genetical Interference and Gene Conversion*, in *Mechanisms in Recombination*, R.F. Grell, Editor. 1974, Plenum: New York. p. 263-275.

21. Baudat, F. and A. Nicolas, *Clustering of meiotic double-strand breaks on yeast chromosome III.* Proc Natl Acad Sci U S A, 1997. **94**(10): p. 5213-8.

22. Gerton, J.L., et al., *Inaugural article: global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2000. **97**(21): p. 11383-90.

23. Blitzblau, H.G., et al., *Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres.* Curr Biol, 2007. **17**(23): p. 2003-12.

24. Buhler, C., V. Borde, and M. Lichten, *Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in Saccharomyces cerevisiae.* PLoS Biol, 2007. **5**(12): p. e324.

25. Ohta, K., et al., *Competitive inactivation of a double-strand DNA break site involves parallel suppression of meiosis-induced changes in chromatin configuration.* Nucleic Acids Res, 1999. **27**(10): p. 2175-80.

26. Wu, T.C. and M. Lichten, *Factors that affect the location and frequency of meiosis-induced double-strand breaks in Saccharomyces cerevisiae.* Genetics, 1995. **140**(1): p. 55-66.

27. Carpenter, A.T., *Electron microscopy of meiosis in Drosophila melanogaster females: II. The recombination nodule--a recombination-associated structure at pachytene?* Proc Natl Acad Sci U S A, 1975. **72**(8): p. 3186-9.

28. Anderson, L.K., et al., *RecA-like proteins are components of early meiotic nodules in lily.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6868-73.

29. Zickler, D. and N. Kleckner, *Meiotic chromosomes: integrating structure and function.* Annu Rev Genet, 1999. **33**: p. 603-754.

30. Anderson, L.K., K.D. Hooker, and S.M. Stack, *The distribution of early recombination nodules on zygotene bivalents from plants.* Genetics, 2001. **159**(3): p. 1259-69.

31. de Boer, E., et al., *Two levels of interference in mouse meiotic recombination.* Proc Natl Acad Sci U S A, 2006. **103**(25): p. 9607-12.

32. Baudat, F. and B. de Massy, *Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis.* Chromosome Res, 2007. **15**(5): p. 565-77.

33. Santucci-Darmanin, S., et al., *MSH4 acts in conjunction with MLH1 during mammalian meiosis.* FASEB J, 2000. **14**(11): p. 1539-47.

34. Neyton, S., et al., *Association between MSH4 (MutS homologue 4) and the DNA strand-exchange RAD51 and DMC1 proteins during mammalian meiosis.* Mol Hum Reprod, 2004. **10**(12): p. 917-24.

35. de Boer, E., et al., *Meiotic interference among MLH1 foci requires neither an intact axial element structure nor full synapsis.* J Cell Sci, 2007. **120**(Pt 5): p. 731-6.

36. Mezard, C., et al., *The road to crossovers: plants have their say.* Trends Genet, 2007. **23**(2): p. 91-9.

37. Mancera, E., et al., *High-resolution mapping of meiotic crossovers and non-crossovers in yeast.* Nature, 2008. **454**(7203): p. 479-85.

38. Malkova, A., et al., *Gene conversion and crossing over along the 405-kb left arm of Saccharomyces cerevisiae chromosome VII.* Genetics, 2004. **168**(1): p. 49-63.

39. Getz, T.J., et al., *Reduced mismatch repair of heteroduplexes reveals "non"-interfering crossing over in wild-type Saccharomyces cerevisiae.* Genetics, 2008. **178**(3): p. 1251-69.

40. de los Santos, T., et al., *The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast.* Genetics, 2003. **164**(1): p. 81-94.

41. Argueso, J.L., et al., *Competing crossover pathways act during meiosis in Saccharomyces cerevisiae.* Genetics, 2004. **168**(4): p. 1805-16.

42. Chen, S.Y., et al., *Global Analysis of the Meiotic Crossover Landscape.* Dev Cell, 2008.

43. de los Santos, T., et al., *A role for MMS4 in the processing of recombination intermediates during meiosis in Saccharomyces cerevisiae.* Genetics, 2001. **159**(4): p. 1511-25.

44. Higgins, J.D., et al., *The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in Arabidopsis.* Genes Dev, 2004. **18**(20): p. 2557-70.

45. Copenhaver, G.P., E.A. Housworth, and F.W. Stahl, *Crossover interference in Arabidopsis.* Genetics, 2002. **160**(4): p. 1631-9.

46. Berchowitz, L.E., et al., *The Role of AtMUS81 in Interference-Insensitive Crossovers in A. thaliana.* PLoS Genet, 2007. **3**(8): p. 0001-0010.

47. Higgins, J.D., et al., *Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over.* Plant J, 2008. **54**(1): p. 152-62.

48. Cromie, G.A., et al., *Single Holliday junctions are intermediates of meiotic recombination.* Cell, 2006. **127**(6): p. 1167-78.

49. Smith, G.R., et al., *Fission yeast Mus81.Eme1 Holliday junction resolvase is required for meiotic crossing over but not for gene conversion.* Genetics, 2003. **165**(4): p. 2289-93.

50. Meneely, P.M., A.F. Farago, and T.M. Kauffman, *Crossover distribution and high interference for both the X chromosome and an autosome during oogenesis and spermatogenesis in Caenorhabditis elegans.* Genetics, 2002. **162**(3): p. 1169-77.

51. Weinstein, A., *The Geometry and Mechanics of Crossing Over.* Cold Spring Harb Symp Quant Biol, 1958. **23**: p. 177-196.

52. Jessop, L. and M. Lichten, *Mus81/Mms4 endonuclease and Sgs1 helicase collaborate to ensure proper recombination intermediate metabolism during meiosis.* Mol Cell, 2008. **31**(3): p. 313-23.

53. Oh, S.D., et al., *RecQ helicase, Sgs1, and XPF family endonuclease, Mus81-Mms4, resolve aberrant joint molecules during meiotic recombination.* Mol Cell, 2008. **31**(3): p. 324-36.

54. Phillips, C.M., et al., *HIM-8 binds to the X chromosome pairing center and mediates chromosome-specific meiotic synapsis.* Cell, 2005. **123**(6): p. 1051-63.

55. MacQueen, A.J., et al., *Chromosome sites play dual roles to establish homologous synapsis during meiosis in C. elegans.* Cell, 2005. **123**(6): p. 1037-50.

56. McKim, K.S., et al., *Meiotic synapsis in the absence of recombination.* Science, 1998. **279**(5352): p. 876-8.

57. Chua, P.R. and G.S. Roeder, *Tam1, a telomere-associated meiotic protein, functions in chromosome synapsis and crossover interference.* Genes Dev, 1997. **11**(14): p. 1786-800.

58. Conrad, M.N., A.M. Dominguez, and M.E. Dresser, *Ndj1p, a meiotic telomere protein required for normal chromosome synapsis and segregation in yeast.* Science, 1997. **276**(5316): p. 1252-5.

59. Kleckner, N., et al., *A mechanical basis for chromosome function.* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12592-7.

60. Shinohara, M., et al., *Crossover Assurance and Crossover Interference are Distinctly Regulated by the ZMM/SIC Proteins during Yeast Meiosis.* Nat Genet, 2008.

61. Oh, S.D., et al., *BLM ortholog, Sgs1, prevents aberrant crossing-over by suppressing formation of multichromatid joint molecules.* Cell, 2007. **130**(2): p. 259-72.

62. Fung, J.C., et al., *Imposition of crossover interference through the nonrandom distribution of synapsis initiation complexes.* Cell, 2004. **116**(6): p. 795-802.

63. Shinohara, M., et al., *Crossover interference in Saccharomyces cerevisiae requires a TID1/RDH54- and DMC1-dependent pathway.* Genetics, 2003. **163**(4): p. 1273-86.

64. Hunter, N. and N. Kleckner, *The single-end invasion: an asymmetric intermediate at the double-strand break to double-holliday junction transition of meiotic recombination.* Cell, 2001. **106**(1): p. 59-70.

65. Sym, M. and G.S. Roeder, *Crossover interference is abolished in the absence of a synaptonemal complex protein.* Cell, 1994. **79**(2): p. 283-92.

66. Novak, J.E., P.B. Ross-Macdonald, and G.S. Roeder, *The budding yeast Msh4 protein functions in chromosome synapsis and the regulation of crossover distribution.* Genetics, 2001. **158**(3): p. 1013-25.

67. Joshi, N., et al., *Pch2 links chromosome axis remodeling at future crossover sites and crossover distribution during yeast meiosis.* PLoS Genet, 2009. **5**(7): p. e1000557.

68. Zanders, S. and E. Alani, *The pch2Delta mutation in baker's yeast alters meiotic crossover levels and confers a defect in crossover interference.* PLoS Genet, 2009. **5**(7): p. e1000571.

69. Uhlmann, F., *Chromosome cohesion and separation: from men and molecules.* Curr Biol, 2003. **13**(3): p. R104-14.

70. Klein, F., et al., *A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis.* Cell, 1999. **98**(1): p. 91-103.

71. Brar, G.A., et al., *The multiple roles of cohesin in meiotic chromosome morphogenesis and pairing.* Mol Biol Cell, 2009. **20**(3): p. 1030-47.

72. Ellermeier, C. and G.R. Smith, *Cohesins are required for meiotic DNA breakage and recombination in Schizosaccharomyces pombe.* Proc Natl Acad Sci U S A, 2005. **102**(31): p. 10952-7.

73. Kugou, K., et al., *Rec8 Guides Canonical Spo11 Distribution Along Yeast Meiotic Chromosomes.* Mol Biol Cell, 2009.

74. Wu, H.Y. and S.M. Burgess, *Ndj1, a telomere-associated protein, promotes meiotic recombination in budding yeast.* Mol Cell Biol, 2006. **26**(10): p. 3683-94.

75. King, J.S. and R.K. Mortimer, *A polymerization model of chiasma interference and corresponding computer simulation.* Genetics, 1990. **126**(4): p. 1127-38.

76. Egel, R., *Synaptonemal complex and crossing-over: structural support or interference?* Heredity, 1978. **41**(2): p. 233-37.

77. Maguire, M.P., *Crossover site determination and interference.* J Theor Biol, 1988. **134**(4): p. 565-70.

78. Foss, E., et al., *Chiasma interference as a function of genetic distance.* Genetics, 1993. **133**(3): p. 681-91.

79. Foss, E.J. and F.W. Stahl, *A test of a counting model for chiasma interference.* Genetics, 1995. **139**(3): p. 1201-9.

80. Housworth, E.A. and F.W. Stahl, *Crossover interference in humans.* Am J Hum Genet, 2003. **73**(1): p. 188-97.

81. Stahl, F.W., et al., *Does crossover interference count in Saccharomyces cerevisiae?* Genetics, 2004. **168**(1): p. 35-48.

82. Lam, S.Y., et al., *Crossover interference on nucleolus organizing region-bearing chromosomes in Arabidopsis.* Genetics, 2005. **170**(2): p. 807-12.

83. Berchowitz, L.E., et al., *The role of AtMUS81 in interference-insensitive crossovers in A. thaliana.* PLoS Genet, 2007. **3**(8): p. e132.

84. Martini, E., et al., *Crossover homeostasis in yeast meiosis.* Cell, 2006. **126**(2): p. 285-95.

85. McPeek, M.S. and T.P. Speed, *Modeling interference in genetic recombination.* Genetics, 1995. **139**(2): p. 1031-44.

86. Fox, D.P., *The control of chiasma distribution in the locust, Schistocerca gregaria (Forskal).* Chromosoma, 1973. **43**(3): p. 289-328.

87. Borde, V., et al., *Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites.* Embo J, 2009. **28**(2): p. 99-111.

88. Sigurdsson, M.I., et al., *HapMap methylation-associated SNPs, markers of germline DNA methylation, positively correlate with regional levels of human meiotic recombination.* Genome Res, 2009. **19**(4): p. 581-9.

89. Martinez-Perez, E. and M.P. Colaiacovo, *Distribution of meiotic recombination events: talking to your neighbors.* Curr Opin Genet Dev, 2009. **19**(2): p. 105-12.

90. Munz, P., *An analysis of interference in the fission yeast Schizosaccharomyces pombe.* Genetics, 1994. **137**(3): p. 701-7.

91. Perkins, D.D., *Crossing-over and interference in a multiply marked chromosome arm of Neurospora.* Genetics, 1962. **47**: p. 1253-74.

92. Strickland, W.N., *An analysis of interference in Aspergillus nidulans.* Proc R Soc Lond B Biol Sci, 1958. **149**(934): p. 82-101.

93. Tsai, C.J., et al., *Meiotic crossover number and distribution are regulated by a dosage compensation protein that resembles a condensin subunit.* Genes Dev, 2008. **22**(2): p. 194-211.

94. Drouaud, J., et al., *Sex-Specific Crossover Distributions and Variations in Interference Level along Arabidopsis thaliana Chromosome 4.* PLoS Genet, 2007. **3**(6): p. e106.

95. Franklin, A.E., et al., *Three-dimensional microscopy of the Rad51 recombination protein during meiotic prophase.* Plant Cell, 1999. **11**(5): p. 809-24.

96. Li, J., et al., *Functional analysis of maize RAD51 in meiosis and double-strand break repair.* Genetics, 2007. **176**(3): p. 1469-82.

97. Kochakpour, N. and P.B. Moens, *Sex-specific crossover patterns in Zebrafish (Danio rerio).* Heredity, 2008. **100**(5): p. 489-95.

98. Moens, P.B., *Zebrafish: chiasmata and interference.* Genome, 2006. **49**(3): p. 205-8.

99. Holloway, J.K., et al., *MUS81 generates a subset of MLH1-MLH3-independent crossovers in mammalian meiosis.* PLoS Genet, 2008. **4**(9): p. e1000186.

100. Koehler, K.E., et al., *Near-human aneuploidy levels in female mice with homeologous chromosomes.* Curr Biol, 2006. **16**(15): p. R579-80.

101. Koehler, K.E., et al., *Genetic control of mammalian meiotic recombination. I. Variation in exchange frequencies among males from inbred mouse strains.* Genetics, 2002. **162**(1): p. 297-306.

102. Vallente, R.U., E.Y. Cheng, and T.J. Hassold, *The synaptonemal complex and meiotic recombination in humans: new approaches to old questions.* Chromosoma, 2006. **115**(3): p. 241-9.

103. Tease, C., G. Hartshorne, and M. Hulten, *Altered patterns of meiotic recombination in human fetal oocytes with asynapsis and/or synaptonemal complex fragmentation at pachytene.* Reprod Biomed Online, 2006. **13**(1): p. 88-95.

104. Lenzi, M.L., et al., *Extreme heterogeneity in the molecular events leading to the establishment of chiasmata during meiosis i in human oocytes.* Am J Hum Genet, 2005. **76**(1): p. 112-27.

105. Barlow, A.L. and M.A. Hulten, *Crossing over analysis at pachytene in man.* Eur J Hum Genet, 1998. **6**(4): p. 350-8.

106. Lynn, A., et al., *Covariation of synaptonemal complex length and mammalian meiotic exchange rates.* Science, 2002. **296**(5576): p. 2222-5.

107. Anderson, G.W., et al., *The Thrsp null mouse (Thrsp(tm1cnm)) and diet-induced obesity.* Mol Cell Endocrinol, 2009. **302**(1): p. 99-107.

Chapter 2

A Pollen-tetrad Based Visual Assay for Meiotic Recombination in Arabidopsis

Kirk E. Francis, Sandy Y. Lam, Benjamin D Harrison, Alexandra L. Bey, Luke E. Berchowitz and Gregory P. Copenhaver

Department of Biology and The Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, NC 27599

This work was previously published as: Francis, K.E., et al., *A pollen tetrad-based visual assay for meiotic recombination in Arabidopsis.* PNAS, 2007. **104**(10): p. 3913-3918.

ABSTRACT

Recombination, in the form of crossovers (CO) and gene conversions (GC), is a highly conserved feature of meiosis from fungi to mammals. Recombination helps ensure chromosome segregation and promotes allelic diversity. Lesions in the recombination machinery are often catastrophic for meiosis, resulting in sterility. We have developed a visual assay capable of detecting COs, GCs and measuring CO interference in *Arabidopsis thaliana*. This flexible assay utilizes transgene constructs encoding pollen-expressed fluorescent proteins of three different colors in the *qrt1* mutant background. By observing the segregation of the fluorescent alleles in 92,489 pollen tetrads, we demonstrate (i) a correlation between developmental position and CO frequency; (ii) a temperature dependence for CO frequency; (iii) the ability to detect meiotic GC events; and (iv) the ability to rapidly assess CO interference.

INTRODUCTION

Meiosis, the reductive division of the genome in preparation for fertilization, is a critical phase in the life-cycle of sexually reproducing organisms. During meiosis, homologous chromosomes interact, resulting in the heritable rearrangement of DNA through reciprocal exchange between homologous chromosomes (crossing over, CO) or gene conversion (GC). COs serve to stabilize homolog pairing during meiosis and also provide an adaptive mechanism for generating favorable allelic combinations [1, 2]. Conversely, because COs are generated via DNA double-strand breaks, they represent a significant liability to the cell if they are not repaired efficiently and with high fidelity. The cell balances these costs and benefits by tightly regulating meiotic recombination, including the imposition of crossover interference in most organisms. There is surprising diversity in the execution and regulation of meiotic recombination in different organisms, ranging from utilization of different protein components to divergence in the use of entire regulatory pathways, such as interference mechanisms [3, 4]. Because of this diversity, researchers use a variety of model systems to study meiosis, including fungi, plants, insects, nematodes and mammals.

Despite the need to establish a broad-based understanding of meiotic recombination, the preponderance of the current mechanistic data comes from a single organism – *Saccharomyces cerevisiae*. Budding yeast has a number of advantages for studying meiotic processes. It has a convenient collection of easily assayable markers (auxotrophic alleles *etc.*) which makes it reasonable to obtain statistically relevant datasets. Furthermore, the haploid life-cycle of yeast makes it possible to detect recombination events directly in the products of meiosis (more accurately, their clonal

propagants) without having to do subsequent crosses. Perhaps most importantly, tetrad analysis is available in yeast [5]. Each meiosis in yeast produces four spores that remain packaged together, making it possible to track and measure all the products of any recombination or segregation event. For example, tetrad analysis can detect the non-Mendelian 3:1 segregation of markers characteristic of gene conversion unambiguously [6, 7]. Tetrad analysis also allows both products of a reciprocal event (*e.g.* CO) to be observed, providing informational redundancy that reduces the chances of mis-scoring.

Using the advantages offered by yeast as an exemplar, we have developed a novel assay system for detecting meiotic recombination events in *Arabidopsis thaliana*. The system combines the *qrt* mutant background, which yields tetrads of meiotically related pollen grains, with a series of transgenic marker genes that encode pollen-expressed fluorescent proteins excitable by different wavelengths of light [8, 9]. Using this system we are able to visualize CO and GC events directly and reliably in gametes in a way that enables tetrad analysis. Because the system utilizes three different fluorescent proteins, it also facilitates the measurement of CO interference. A different seed-based assay has been described by Melamed-Bessudo et al. [10]. The seed-based assay also has the advantage of being high throughput and using fluorescent proteins as visual reporters, but it lacks the ability to assay recombination directly in the gametes, and it does not incorporate the advantages of tetrad analysis. Our pollen tetrad-based visual assay lends itself to the rapid assessment of COs, GC and interference in virtually any mutant background capable of producing at least some pollen. Here we use this system to show that the developmental position of flowers producing gametes influences CO frequency; we do so by demonstrating that tetrads produced from secondary and tertiary axes have a

higher recombination frequency in comparison to primary bolts. We also show that CO frequency correlates with temperature. Lastly, we demonstrate that the assay sytem can be used to detect meiotic gene conversion and measure crossover interference in Arabidopsis.

RESULTS AND DISCUSSION

Genetic intervals that can be visually assayed for recombination were created by transforming *qrt* plants with marker genes encoding red, yellow or cyan fluorescent proteins under control of the LAT52 promoter (Figure 1A) [11, 12]. Transformed seeds were selected, using either kanamycin or glufosinate to yield 2752 T_1 plants. Each of these lines was screened, using a number of criteria to ensure their usefulness for subsequent meiotic analysis. First, pollen tetrads from each T_1 were observed under the fluorescence microscope. Each T-DNA insertion event in the T_1 will result in a hemizygous locus; therefore, if we observed consistent 2:2 segregation of the fluorescent signal in the pollen tetrads, we assumed a single-locus insertion (Figure 2B). This assumption has two caveats. Multiple tandem copies of the marker gene at a single locus will also segregate in a consistent 2:2 pattern, and silenced or disrupted copies would not be detected. Lines that showed more complicated patterns including 4:0, 3:1, 1:3 and 0:4 were assumed to be the result of multiple insertions, partial silencing, or disruption of the fluorescent marker genes, and were not given further consideration. We also rejected lines that showed very weak expression of the fluorescent protein.

Genomic DNA flanking the T-DNA in single-insert lines with strong expression was obtained using LMS-PCR and was subsequently sequenced to map each insertion

precisely (Figure 2) [13]. Because multiple LMS-PCR fragments can be indicative of multiple or complex integrations, lines resulting in complex LMS-PCR patterns were discarded. This sequencing also allowed us to differentiate insertions in annotated genes from those in intergenic regions. When using these markers to measure recombination frequencies (as described below), it is important that disruption of genes by the marker does not confound the results – particularly when using the markers in mutant backgrounds where synthetic effects would be difficult to predict. Accordingly, we only proceeded with those lines that gave strong 2:2 expression patterns with markers inserted into intergenic regions. These were designated Fluorescent Tagged Lines (FTLs).

Lines selected for further use were also subjected to T_2 segregation analysis. T_1 plants were allowed to self, and their T_2 progeny were visually scored to ensure that plants with 4:0, 2:2 or 0:4 segregation of the transgene in their pollen tetrads were occurring in the expected 1:2:1 Mendelian ratio. This final check serves three purposes: (i) it confirms the previous assumption of a single-locus insertion; (ii) it reveals any phenotypes resulting from carrying the marker in the homozygous state; and (iii) it confirms the stability of the marker in an additional generation. Of the original 2752 positive transformants, 79 lines passed our quality control criteria (Figure 2).

To detect COs, pairs of lines carrying differently colored markers on the same chromosome were crossed to produce F_1 progeny carrying both markers in *trans* and creating a genetic interval bounded by two visible markers. COs can be observed in the pollen tetrads produced by this F_1 plant by visually scoring the segregation of the two fluorescent proteins. Without a CO in the interval the tetrad will have two grains of one color and two grains of the other color. A CO in the interval will result in a tetrad with

one bi-colored grain, one non-colored grain and two mono-colored grains (Figure 1C-D). It should also be noted that two markers in a *cis* configuration would also be useful for assaying COs but would result in a different pattern for recombinant and non-recombinant classes. Several hundred tetrads can be easily scored in an hour using this assay, thereby facilitating the collection of very large data sets.

As specific examples, we crossed two FTL lines carrying insertions on chromosome 1, FTL992 (CFP) and FTL1313 (DsRed) to create "interval 1" (I1) and two insertions on chromosome 3, FTL1500 (CFP) and FTL1371 (DsRed) to create "interval 3" (I3, Figure 2). Assuming an average genomic recombination frequency (RF) of 200kb/cM (119186 Kb/597 cM), these intervals are expected to be 6.4 cM and 19.1 cM respectively (http://www.arabidopsis.org/). We scored 9181 and 8919 tetrads from F1 plants carrying markers that define I1 and I3 respectively, and we observed a map distance of 6.2 cM for I1 and 23.1 cM for I3. Thus, the markers themselves do not appear to inhibit local recombination. I1 and I3 are interstitial regions that do not encompass any special chromosomal domains such as centromeres or rDNA arrays [14]. The collection of FTL lines that we have established provides the basis for establishing test intervals of a variety of sizes on any of the Arabidopsis chromosomes.

Using methods described by Papazian it is possible to calculate the expected number of non-parental (NPD) tetrads resulting from double COs in a single interval using the frequency of single COs which are observed as tetratypes [15]. Applying this method to our data set, we would expect to observe 19 and 304 NPDs for I1 and I3 respectively; we observed only 2 and 51. This paucity of NPDs is consistent with our observation of CO interference using dual interval analysis described below.

Arabidopsis produces multiple flowers, each of which is an independent source of gametes. The production of flowers in a temporal series and in a variety of developmental positions serves as a potential source of natural variability in CO frequencies. To assess this potential, we measured CO frequencies in I1 and I3 (25,249 and 23,830 total tetrads respectively), using flowers from primary bolts (1°), branches on primary bolts (2°) and branches of branches (3°). Surprisingly, we observed that both 2° and 3° flowers produced tetrads with significantly higher CO frequencies compared to 1° flowers (Figure 3 top). This result may indicate a regulatory connection between developmental programs and crossover control in Arabidopsis. It is possible that this difference may result from the fact that 1° flowers are derived from the apical meristem while the 2° and 3° flowers are derived from lateral meristems. It is also possible that on average 2° and 3° flowers are produced when the plant itself is older (an age effect). To explore this possibility, we measured CO frequencies in I1 and I3, using individual flowers along the length of each axis. For each flower we scored at least 100 tetrads. Although there was variation in map distances from flower to flower, the regression from first to last flower revealed no correlation between flower age and map distances on any of the axes (Figure 3, middle). Thus we conclude that the difference in CO frequencies we observed in 1° or 2° and 3° flowers is due to developmental position rather than age. Because plants have multiple independent germlines that arise in different developmental contexts, they provide a unique opportunity to examine the relationship between the regulation of meiotic recombination and development.

Plants, being sessile, often experience dramatic ranges of temperatures in the wild. To measure the sensitivity of CO frequency to temperature, we grew plants in a

20°C growth chamber under long days until they began to bolt, and then we shifted them to 19°C, 22.5°C, 25.5°C or 28°C under constant light. We then assayed RFs in I1 and I3 (10,988 and 20,739 total tetrad respectively). In both intervals we observed significant increases in map distances across the temperature range (Figure 3, bottom). These data are consistent with previous observations in both plants and animals. In *Hordeum vulgare*, *Vicia faba* and *Drosophila melanogaster* there is a general trend toward higher levels of COs with higher temperature [16-18]. This may represent a regulatory connection between temperature sensing and CO control, or the mechanical process of crossing over may be sensitive to temperature. Studies in *Locusta migratoria* and *Allium ursinum* indicate that extreme temperatures can interfere with chromosome mechanics such as synapsis, leading to a decrease in CO frequency [19, 20]. Interestingly, temperature studies in *Neurospora crassa* a temperature optima for COs (in one interval) at moderate temperatures [21]. The fluorescent assay system described here will enable more detailed genetic analysis of the relationship between environmental conditions and meiotic recombination. Whatever their causes, the results of the temperature and the developmental experiments indicate that care must be taken in collecting CO data from Arabidopsis. Growth conditions should be well controlled and data should be collected from equivalent developmental tissues.

To demonstrate that the visual assay system can be used to measure CO interference, we crossed two FTL lines carrying insertions on chromosome 5 - FTL1273 (DsRed) and FTL1659 (YFP) - to create F_1 progeny carrying both markers in *trans.* Self-seeds from these F_1 plants were planted, and pollen from the F_2 plants were visually scored to identify recombinant individuals with the two markers in *cis*. Recombinant F_2

individuals were then crossed to a line carrying a third FTL insertion on chromosome 5, FTL993 (CFP), to produce progeny carrying all three markers (Figure 4). These three markers define two adjacent genetic intervals, I5a and I5b (Figure 2). By following the segregation pattern of all three markers, it is possible to distinguish: (i) meioses lacking COs; (ii) meioses with single COs in either interval; and (iii) all classes (two, three and four-stand) of double crossovers (DCOs, Figure 4). We scored 1247 tetrads from these plants and we observed RFs of .16 in I5a, and .24 in I5b. Multiplying the product of these frequencies by the total number of gametes scored (1247*4 = 4988) we would predict 190 double crossover gametes in the combined interval. But we observed only 60. This data yields an interference value (i.e., 1-[observed DCO/expected DCO]) of 0.68. In the past, measuring interference in Arabidopsis was difficult because the large data sets necessary to accumulate enough DCOs for analysis were time consuming to generate. With our visual assay, interference measurements can routinely be obtained thereby enhancing the value of Arabidopsis as a model system for meiotic processes.

One of the unique advantages of tetrad-based recombination analysis is the ability to identify gene conversion (GC) events unambiguously via their classic non-Mendelian 3:1 segregation pattern. We have created a modified version of the FTL system that makes it possible to detect 3:1 segregation patterns. A plant carrying a non-fluorescent allele of FTL567 (yfp) was generated using EMS mutagenesis (Figure 2). Sequencing this allele revealed a G to A transition at nucleotide 95 in the YFP gene, causing a Gly32 to Asp mutation that resulted in the loss of fluorescence. This yfp/yfp line was crossed to the original parental FTL567 line to generate F1 plants that were heterozygous at the fluorescent transgenic locus (YFP/yfp). Pollen tetrads that do not experience gene

conversion at this locus will have two fluorescent grains and two non-fluorescent grains (Figure 6). A GC event spanning the locus will generate three fluorescent grains and one non-fluorescent grains (or *vice versa*). After scoring 4,033 tetrads, we observed 6 tetrads with a 3:1 segregation ratio. We ignored 1:3 tetrads because of the possibility that occasional non-viable pollen grains might yield false positives. Assuming that the directionality of GC is not biased toward either allele, we conclude that at this locus GC occurred in 1/336 meioses. To control for false 3:1 positives due to reversion of the yfp allele to a fluorescent state, we scored 8783 yfp/yfp tetrads, and we never observed a fluorescent grain ($p<0.0001$ by two-tailed Fishers exact test [22]). As far as we know, this is the first unambiguous observation of meiotic GC in Arabidopsis using tetrad analysis. Additional test loci will need to be developed in order to generate genome-wide estimates of GC frequencies. In contrast to this test locus, which we specifically designed to detect GC, we did not observe any cases of 3:1 segregation in any of the hemizygous lines used to generate the I1 and I3 data described above.

The FTL assay system described here enables rapid and inexpensive detection of COs across the *A. thaliana* genome; measurement of CO interference; and quantification of GC frequencies. These tools will facilitate the analysis of recombination phenotypes in mutant lines. It is particularly advantageous that these analyses can be done in essentially isogenic lines. Moreover, because large numbers of tetrads can be scored easily, experiments which would have been difficult in the past because of laborious crossing schemes, can now be done with relative ease. In demonstrating the use of the assay system, we showed that temperature is correlated with CO frequency in Arabidopsis. More surprisingly, we discovered a correlation between the developmental

position of flowers producing gametes and CO frequency. This finding raises interesting questions about why a connection between developmental position and meiotic recombination should exist and how it might be mediated. This quick and practical assay system will now enable experiments designed to probe how environmental cues like temperature, and developmental signals influence meiotic recombination at a mechanistic level.

MATERIALS AND METHODS

Plant Materials. *A. thaliana qrt1-1* in the Landsberg-0 background (CS8845) and *qrt1-2* in the Columbia-3 background (CS8846) were used to generate the fluorescent-tagged lines in this study. Seeds were sown on Pro-mix (Professional Horticulture, Inc.) and stratified for 3-4 days at 4°C. Plants were germinated and grown under long-day conditions (18 hrs. light) at 20°C unless otherwise noted. Temperatures were monitored with thermometers on the same shelves that the plants were grown on. All parental strains are available from the Arabidopsis Biological Resource Center at Ohio State University (http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/abrchome.htm). All FTL lines are available upon request from the corresponding author.

T-DNA Vector Construction. Plasmids carrying genes encoding the fluorescent proteins EYFP, ECFP and Ds-Red (XFP collectively) were obtained from BD Bioscience Clontech (Palo Alto, CA, USA). The open reading frames for these marker proteins were transferred into the pENTR Gateway vector (Invitrogen, Carlsbad, CA, USA) via PCR-based TOPO cloning using the manufacturer's instructions. The pENTR-based clones were used to transfer the marker genes into the pK2GW7 (kanamycin selection) or

pB7WG2 (glufosinate selection) Gateway vectors that had been modified by removing the 35S promoter and replacing it with the pollen-specific, post-meiotic LAT52 promoter [12]. The transfer was accomplished using the Invitrogen LR Clonase kit following the manufacturer's instructions.

Plant Transformation. *Agrobacterium tumefaciens* strain GV3101 was transformed with the LAT52::XFP T-DNA vectors using electroporation. Plants were transformed using the method of Chung et al. (2000) except that the bacteria were re-suspended in a simplified media consisting of 5% sucrose and 0.05% Silwet L-77 [23]. Positive T_1 transformants harboring an nptII-based construct were selected on 0.5X MS salt medium containing 50mgml^{-1} kanamycin after two weeks of growth and then transferred into soil. Bar-based T_1 transformants were germinated on soil and selected by spraying with 0.0189% glufosinate ammonium from Bayer Crop Science sold under the name Liberty (Research Triangle Park, NC, USA). Expression of the fluorescent marker proteins in the pollen was confirmed using epi-fluorescence microscopy.

Ligation Mediated Suppression PCR. The genomic location of each insertion was established using LMS-PCR as described by Alonso et al. (2003) with the following modifications [13]. DNA from 2-3 cauline leaves was purified as described in Copenhaver *et al.* and digested with either Hind III or EcoR I (NEB, Beverly, MA, USA) [24]. Digested genomic DNA was chloroform extracted, ethanol precipitated and suspended in 20 µl ddH_20. Adapters were then ligated to the library of genomic fragments using T4 ligase (NEB). Adapters for ligation to Hind III ends were made by annealing ADAPS-H3 with a 3' amino terminal end with (acgtcacctgcccgg/3AmMc7/) ADAPL-E1 (ctaatacgactcactatagggctcgagcggccgcccgggcaggtg). Adapters for ligation to

EcoR I ends were made by annealing ADAPS-E1 with a 3' amino terminal end (aattcacctgcccgg/3AmMc7/) with ADAPL-E1. Primary PCR products were generated using primers AP1 (ggatcctaatacgactcactataggc) and PgwLat52LB-WP1 (ctatgttactagatcgaccgg). Secondary PCR products were generated by diluting primary products by 50 fold and amplifying with primers AP2 (tatagggctcgagcggccg) and PgwLat52LB-WP2 (caattcggcgttaattcagtac). Primary and secondary PCR reactions were initiated at 94º for 2 minutes followed by 29 cycles of 30 seconds denaturation (94ºC), 30 seconds annealing (55ºC), 1 minute extension (72ºC) and a final 10 minute 65ºC incubation. All primers were purchased from Integrated DNA Technologies (Skokie, IL, USA). PCR products were sequenced at the UNC Lineberger Comprehensive Cancer Center.

Mutagenesis. Non-fluorescent alleles of transgenes encoding fluorescent proteins were generated following the protocol of Weigel and Glazebrook [25]. 0.5 grams of seed were imbibed in 30 ml of sterile water for 4 hrs., and then mutagenized with 0.2% ethyl methane sulfonate for 16 hours at room temperature with gentle agitation. Mutagenized seeds were rinsed with 30 ml of sterile water eight times and then dried before planting.

Microscopy. Segregation patterns of fluorescent alleles in pollen tetrads were measured using a Nikon E1000 epi-fluorescence microscope (Melville, NY, USA) equipped with filters from Chroma Technology (Rocherster, NY, USA). Pollen was collected by dipping flowers into a drop of 0.1% TritonX-100 on a glass slide. Micrographs were taken with a Nikon Coolpix5000 camera. Multi-color images were obtained by using Adobe Photoshop (San Jose, CA, USA) to merge two or more single-color images taken using the appropriate filter.

Linkage Analysis. To measure the map distance between any two transgenic markers, tetrads were designated parental ditype (PD), nonparental ditype (NPD), or tetratype (T) depending on the segregation pattern of the marker pair. Map distances were then determined using the Perkins formula: 100[(1/2T + 3 NPD)/n] [26]. Expected NPD frequencies for single intervals were calculated using the Papazian foruma: 1/2[(1-T)-(1-(3T/2))$^{2/3}$] (15).

ACKNOWLEDGEMENTS

We would like to thank Daphne Preuss, Frank Stahl, Corbin Jones and Jeff Sekelsky for critical comments and 3 anonymous reviewers for helpful suggestions. Funding for this work was provided by NSF grant number MCB-0618691 and DOE grant number DE-FGO2-05ER15651.

FIGURE LEGENDS

Figure 2.1. Fluorescent markers in pollen tetrads. An *Agrobacterium* T-DNA construct (A) containing either ECFP, EYFP or DsRed driven by the LAT52 promoter and a selectable marker conferring resistance to either kanamycin (nptII) or glufosonate (Bar) was used to transform *qrt1* Arabidopsis seed resulting in T1 plants expressing the fluorescent protein in their pollen (B, shows pollen from three different plants). Crossing lines with differently colored transgenes on the same chromosome (C and D) enables detection of CO events in the interval between the transgenes (E, merge of C & D with arrow indicating recombinant tetrad).

Figure 2.2. Map of fluorescent transgenes. The chromosomal (green bars) insertion site of the transgene carried by each FTL lines is indicated by a red (DsRED), yellow (EYFP), or cyan (ECFP) circle. The genetic intervals (I1, I3, I5a and I5b) used in this study are delineated with brackets and the locus used to detect gene conversion event (GC) is also marked. The TAIR "chromosome map tool" was used to place T-DNA insertion points on the physical map (15).

Figure 2.3. (Top) Map distances in cM were measured using flowers from the primary bolt (1), branches of the primary bolt (2), and branches of branches (3) in both interval 1 (left) and 3 (right). (Middle) As a control for flower age, map distances were also measured in individual flowers, beginning with the first, ending with the last, produced by the primary axis in interval 1 (left) and 3 (right). (Bottom) The influence of temperature during flowering was measured in interval 1 (left) and 3 (right) using flowers from the primary axis. Error bars (top and bottom) are based on standard error.

Figure 2.4. Measuring interference. DsRed (red circle), EYFP (yellow circle) and ECFP transgene (cyan circle) define two adjacent genetic intervals (I5a and I5b, see text) on chromosome 5. The four chromatids present after replication can experience no crossovers (A), single crossovers in either interval (B, C), and double crossovers in the combined interval including two-strand double crossovers (D), both kinds of three-strand double crossovers (E,F) and four-strand double crossovers (G) in the two intervals. Each of these events can be distinguished by observing the segregation of the transgenes.

Figure 2.5. Detecting gene conversion. A heterozygous plant with one fluorescent EYFP allele (yellow circle) and one mutant non-fluorescent eyfp allele (black circle) can be used to detect meiotic gene conversion (GC) events in pollen tetrads.

TABLES

Table 2.1.

Table 1. Physical position of FTL insertions

Chromosome	FTL no.*	Nucleotide position†
1	1285	80614
1	2217	2007279
1	567	3905441
1	2443	5181427
1	1262	5755618
1	1538	5755655
1	1376	6488210
1	1714	9697890
1	992	9850022
1	1313	11130549
1	1405	22606560
1	1230	23371216
1	1321	23599565
1	1377	24508229
1	1134	24645163
1	2385	25956590
1	2376	26746743
1	790	27647234
1	1181	28390370
1	1512	28444753
2	1997	487059
2	1431	1521041
2	1805	5459140
2	2269	8276753
2	1506	12640092
2	1524	13226013
2	1822	14464574
2	965	14675407
2	800	18286716
2	2271	19329549
3	1500	498916
3	1019	1517290
3	2180	2092414
3	1371	4319513
3	1369	6472617
3	1625	9270412
3	1046	9458743
3	2261	9538295
3	2504	10090995
3	2536	16520560
3	2480	16671637
3	975	16895505
3	1268	17242947
3	1632	17860363
3	2201	18319139
3	2456	18743679
3	2066	19560471

3	1413	22743753
3	1056	22882685
4	1323	1305439
4	424	1365848
4	1618	6171524
4	1518	6752928
4	1614	7664346
4	804	8173733
4	1470	11710691
4	1065	12814708
4	1307	13152926
4	1478	16746703
4	2015	18443281
4	2133	18488540
5	1311	2293101
5	1963	2372623
5	1143	3760756
5	1691	5358733
5	2450	5497513
5	2420	5651866
5	1432	6217832
5	826	6622939
5	2375	7798291
5	1922	11750913
5	1442	11752340
5	1273	18164269
5	1659	23080567
5	1066	24083497
5	1248	25071705
5	993	25731311
5	1396	25739005
5	1915	26649857

*The order of FTL lines corresponds to the order shown in Fig. 2 from right to left.

†Nucleotide positions can be referenced at The Arabidopsis Information Resource (www.arabidopsis.org).

FIGURES

Figure 2.1.

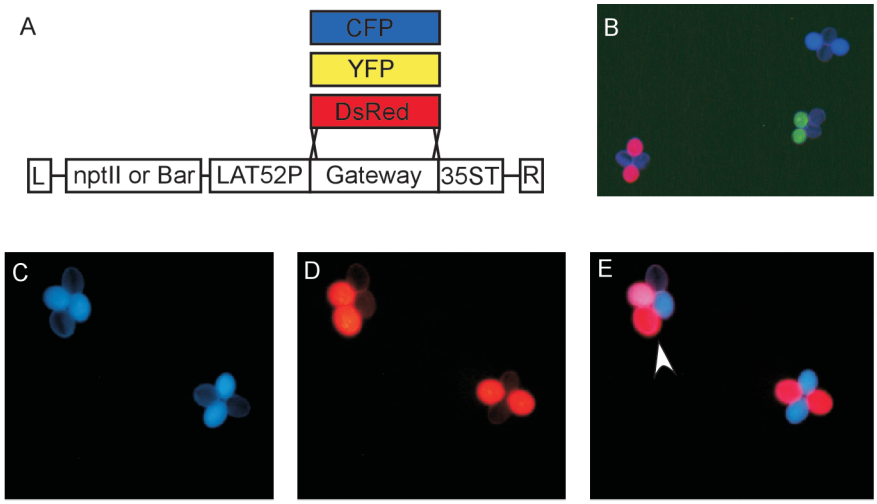


Figure 2.2.

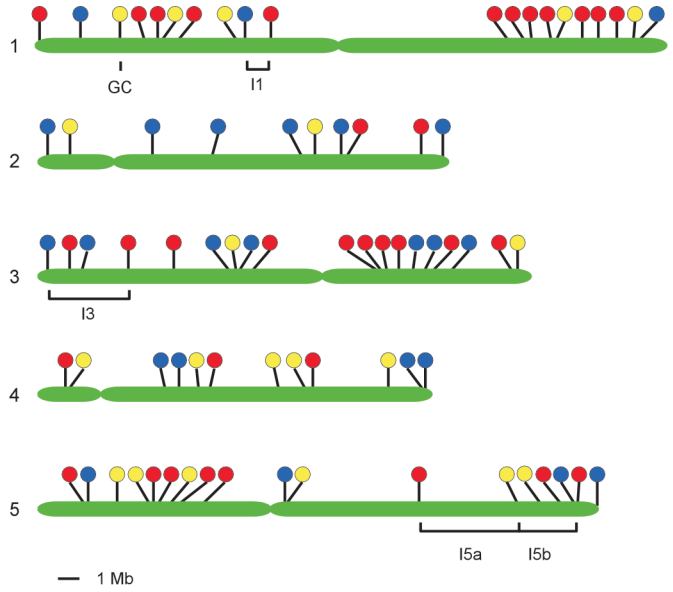


Figure 2.3.

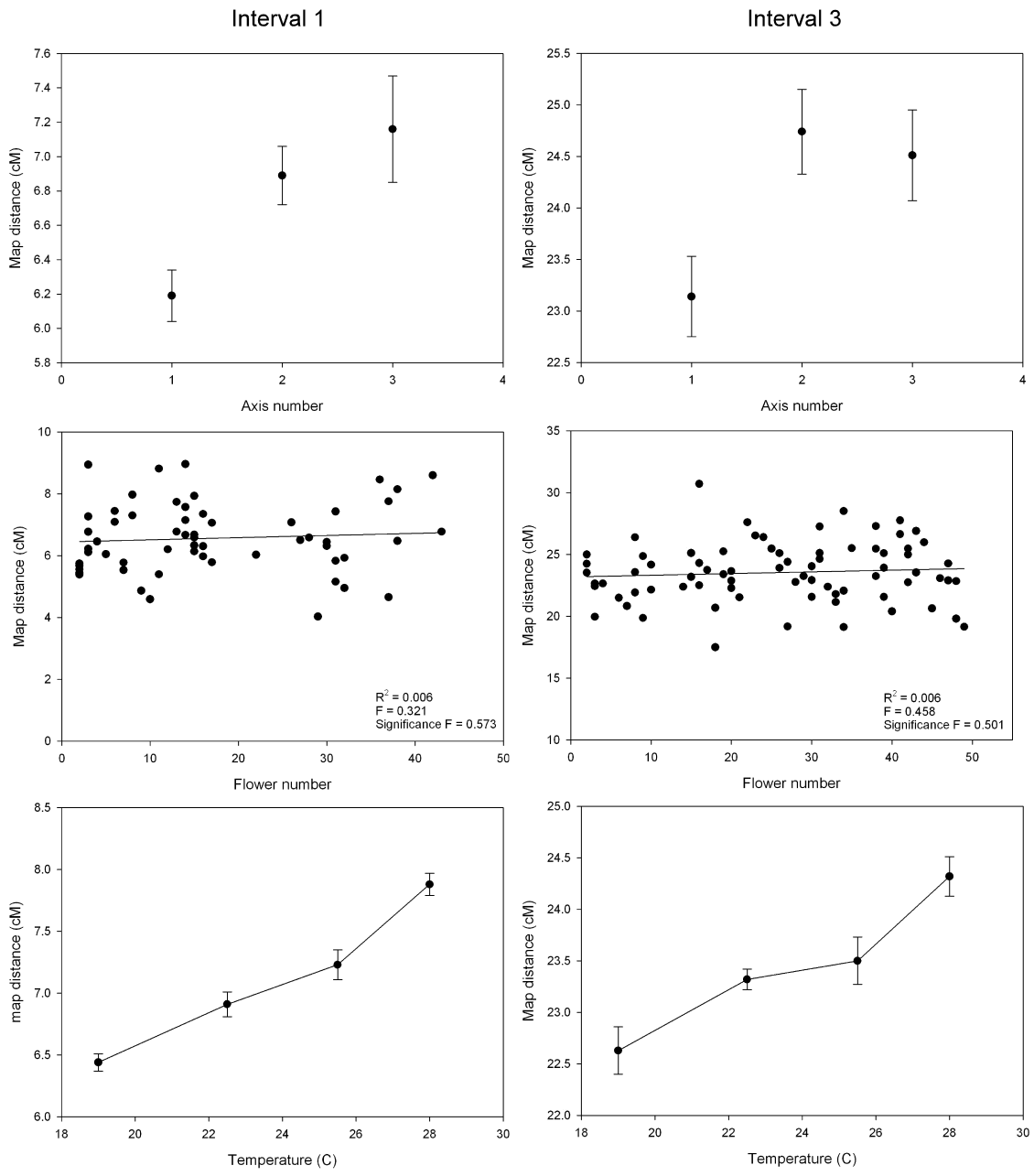

Interval 1
Interval 3
Map distance (cM)
Axis number
Flower number
map distance (cM)
Temperature (C)
R2 = 0.006
F = 0.321
Significance F = 0.573
R2 = 0.006
F = 0.458
Significance F = 0.501

Figure 2.4.

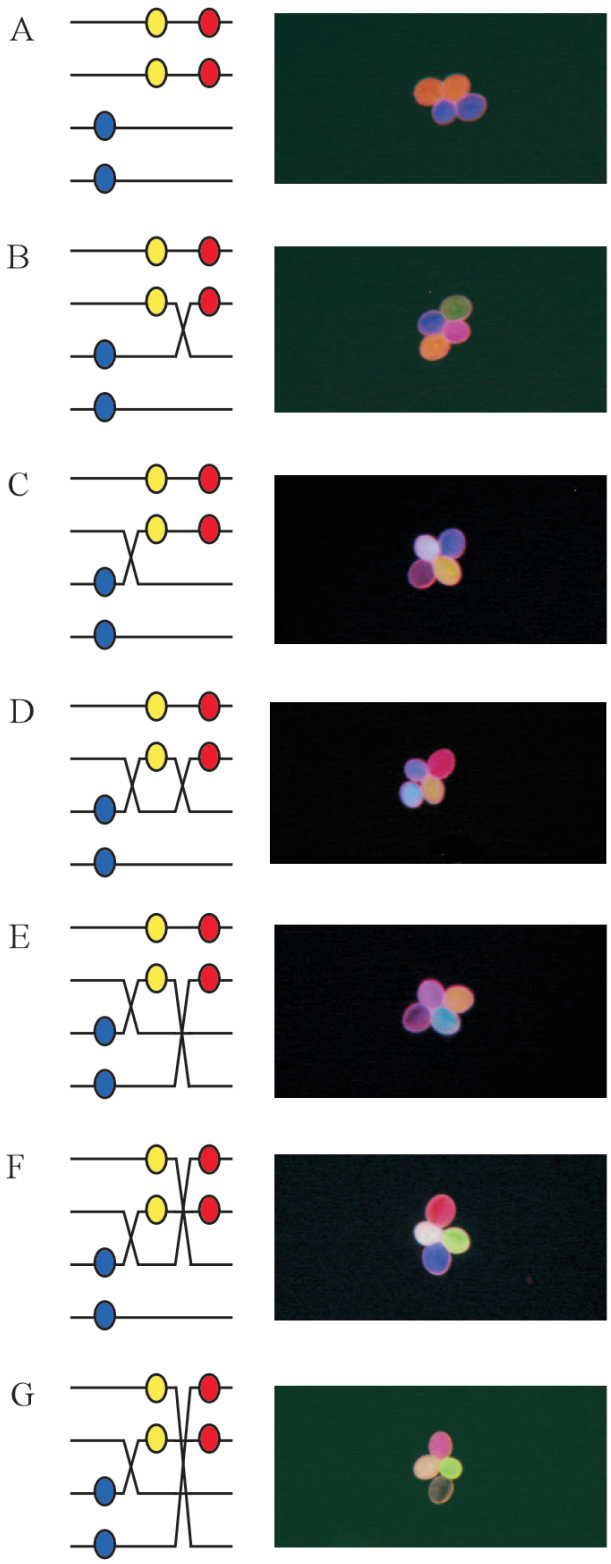
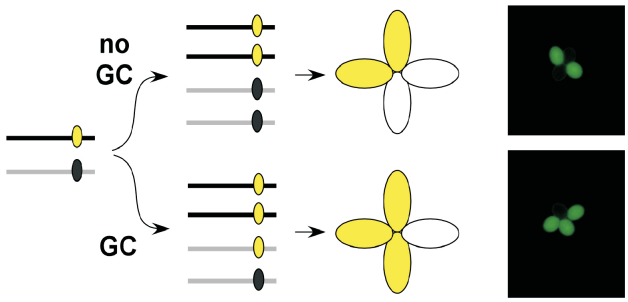


Figure 2.5.



REFERENCES

1. Page, S.L. and R.S. Hawley, *Chromosome choreography: the meiotic ballet.* Science, 2003. **301**(5634): p. 785-9.

2. Felsenstein, J., *The evolutionary advantage of recombination.* Genetics, 1974. **78**(2): p. 737-56.

3. Gerton, J.L. and R.S. Hawley, *Homologous chromosome interactions in meiosis: diversity amidst conservation.* Nat Rev Genet, 2005. **6**(6): p. 477-87.

4. Hillers, K.J., *Crossover interference.* Curr Biol, 2004. **14**(24): p. R1036-7.

5. Copenhaver, G.P., K.C. Keith, and D. Preuss, *Tetrad analysis in higher plants. A budding technology.* Plant Physiol, 2000. **124**(1): p. 7-16.

6. Fogel, S. and D.D. Hurst, *Meiotic gene conversion in yeast tetrads and the theory of recombination.* Genetics, 1967. **57**(2): p. 455-81.

7. Mitchell, M.B., *ABERRANT RECOMBINATION OF PYRIDOXINE MUTANTS OF Neurospora.* Proc Natl Acad Sci U S A, 1955. **41**(4): p. 215-20.

8. Preuss, D., S.Y. Rhee, and R.W. Davis, *Tetrad analysis possible in Arabidopsis with mutation of the QUARTET (QRT) genes.* Science, 1994. **264**(5164): p. 1458-60.

9. Stewart, C.N., Jr., *Go with the glow: fluorescent proteins to light transgenic organisms.* Trends Biotechnol, 2006. **24**(4): p. 155-62.

10. Melamed-Bessudo, C., et al., *A new seed-based assay for meiotic recombination in Arabidopsis thaliana.* Plant J, 2005. **43**(3): p. 458-66.

11. Preuss, D., *Chromatin silencing and Arabidopsis development: A role for polycomb proteins.* Plant Cell, 1999. **11**(5): p. 765-8.

12. Twell, D., J. Yamaguchi, and S. McCormick, *Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis.* Development, 1990. **109**(3): p. 705-13.

13. Alonso, J.M., et al., *Genome-wide insertional mutagenesis of Arabidopsis thaliana.* Science, 2003. **301**(5633): p. 653-7.

14. *Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.* Nature, 2000. **408**(6814): p. 796-815.

15. Papazian, H.P., *The Analysis of Tetrad Data.* Genetics, 1952. **37**(2): p. 175-88.

16. Powell, J. and R. Nilan, *Influence of Temperature on Crossing Over in an Inversion Heterozygote in Barley.* Crop Sci, 1963. **3**: p. 11-13.

17. Berkemeier, J. and G. Linnert, *Effects of High and Low Temperature on Meiotic Recombination in Normal and Oligochiasmatic Lines of Vicia-Faba.* Biol Zentralb, 1987. **106**: p. 219-230.

18. Stern, C., *An Effect of Temperature and Age on Crossing-Over in the First Chromosome of Drosophila Melanogaster.* Proc Natl Acad Sci U S A, 1926. **12**(8): p. 530-2.

19. Loidl, J., Chromosoma, 1989. **97**: p. 449-458.

20. Henderson, S.A. and M.E. Buss, *The superimposition of heat-induced chiasma frequency changes in Locusta migratoria.* Heredity, 1989. **62 (Pt 1)**: p. 77-84.

21. Towe, A.M. and D.R. Stadler, *Effects of Temperature on Crossing over in Neurospora.* Genetics, 1964. **49**(4): p. 577-83.

22. Sokal, R.R. and F.J. Rohlf, *Biometry : the principles and practice of statistics in biological research*. 3rd ed. 1995, New York: Freeman. xix, 887 p.

23. Chung, M.H., M.K. Chen, and S.M. Pan, *Floral spray transformation can efficiently generate Arabidopsis transgenic plants.* Transgenic Res, 2000. **9**(6): p. 471-6.

24. Copenhaver, G.P., W.E. Browne, and D. Preuss, *Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads.* Proc Natl Acad Sci U S A, 1998. **95**(1): p. 247-52.

25. Weigel, D. and J. Glazebrook, *Arabidopsis : a laboratory manual*. 2002, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xii, 354.

26. Perkins, D.D., *Biochemical Mutants in the Smut Fungus Ustilago Maydis.* Genetics, 1949. **34**(5): p. 607-26.

CHAPTER 2 ADDENDUM

Specific contributions by L.E.B.

I helped build the FTL library by selecting T1 candidate lines and mapping insertion sites using LMS-PCR. I also conducted T2 analysis to determine if the FTL lines were segregating in a pattern expected from a single hemizygous insertion. G.P.C, K.E.F, and myself conceived the experimental design. I most notably helped design the experiment that showed that flower age on the primary axis did not significantly affect map distances. Experiments were primarily executed by K.E.F, but I collected recombination data in the flower-age experiment. Data analysis was conducted by G.P.C. and K.E.F. The manuscript was prepared by G.P.C., which I reviewed and edited in conjunction with K.E.F.

Chapter 3

Fluorescent *Arabidopsis* tetrads: A visual assay for quickly developing large crossover and crossover interference datasets

Luke E. Berchowitz and Gregory P. Copenhaver*

The Department of Biology and The Carolina Center for Genome Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

This work was previously published as: Berchowitz, L.E. and Copenhaver, G.P. 2008. Fluorescent Arabidopsis tetrads: a visual assay for quickly developing large crossover and crossover interference data sets. Nat Protoc **3**: 41-50.

ABSTRACT

In most organisms, one crossover (CO) event inhibits the chances of another nearby event. The term used to describe this phenomenon is CO interference. Here, we describe a protocol for quickly generating large datasets that are amenable to CO interference analysis in the flowering plant *Arabidopsis thaliana*. We employ a visual assay that utilizes transgenic marker constructs encoding pollen-expressed fluorescent proteins of three colors in the *quartet* mutant background. In this genetic background, male meiotic products – the pollen grains – remain physically attached thereby facilitating tetrad analysis. We have developed a library of mapped marker insertions that, when crossed together, create adjacent intervals that can be rapidly and simultaneously screened for COs. This assay system is capable of detecting and differentiating single COs as well as 2, 3, and 4-strand double COs. We also describe how to analyze the data that are produced by this method. For a researcher to generate and then score a double interval in a wild-type and mutant background, the entire procedure will take 22-27 weeks.

INTRODUCTION

Meiosis is a specialized type of cell division in which a prior round of DNA replication is followed by two successive rounds of chromosome segregation. The resulting cells have half of the original chromosomal complement. During fertilization, these cells fuse to reconstitute the original complement thereby completing the sexual cycle. In most eukaryotes, a feature of the first meiotic division is crossing over, in which homologous chromosomes physically interact, exchanging material between the paternal and maternal copies [1]. Crossing over provides a connection between homologs that is required in most organisms for the accurate segregation of chromosomes during meiosis I [2]. Chromosomes that fail to have at least one crossover (CO) often segregate aberrantly, resulting in aneuploidy. In addition, crossing over generates genetic diversity by creating new combinations of paternal and maternal alleles [3].

Cells undergoing meiosis regulate recombination at multiple levels. To ensure that each chromosome pair will be physically connected, COs are distributed among chromosomes non-randomly such that each chromosome pair typically undergoes at least one CO even if the total numbers of COs per chromosome are small [4, 5]. The distribution of COs along chromosomes is also tightly regulated. In most organisms, COs are distributed such that one CO event inhibits the chances of another nearby event [6]. The term used to describe this phenomenon, first observed by *Drosophila melanogaster* researchers at the beginning of the 20th century, is CO interference [7, 8].

One of the challenges of studying CO interference is that it is a probabilistic phenomenon of populations. CO interference does not result in a complete lack of closely spaced COs, but instead reduces the proportion of these events in a population

compared to what one would expect if the events were randomly positioned [9]. For this reason, making statistically significant claims often requires large data sets. Patterns of CO interference vary widely from organism to organism with regards to strength vs. distance relationships and in some organisms there appear to be both interfering and non-interfering COs [10-13]. Even within organisms, parameters such as sex and chromosomal location have been shown to have a profound effect on CO interference [14, 15]. Using *A. thaliana*, we present here a method for the quick and relatively easy production of datasets that can be used to analyze CO interference. Using this method, one can assay almost any chromosomal region, customize interval size, and examine the effects of mutants and experimental treatments on CO interference. This method has been used to assay wild type levels of CO interference on a region of chromosome 5 and also to determine a small but detectable difference in CO interference between wild-type plants and *Atmus81* mutant plants [16, 17].

Measuring multiple COs simultaneously

Several methods exist for measuring multiple COs simultaneously, which is necessary to facilitate CO interference analysis. These methods can be sorted into two general categories: genetic, in which the researcher monitors the segregation of markers that define intervals, and; cytological, in which the researcher uses microscopy techniques to visualize structures that mark COs.

Genetic methods: These generally involve a crossing scheme in which an individual heterozygous for multiple markers gives rise to a set of progeny that possess chromosomes reflecting the recombination history of the parent [10, 18, 19]. In some

organisms, researchers can also monitor recombination directly in the products of meiosis by analyzing markers in such cell types as fungal spores [20, 21] or sperm of mammals [22, 23]. Genetic systems in which the four meiotic products are fused in a tetrad (or octad in some cases) are especially powerful because the researcher can account for all of the parental genetic material in each unit of data rather than analyze a pool of random meiotic products [24-27]. This allows, among other things, the identification of non-parental ditype (NPD) tetrads, in which a four strand double crossover (DCO) has occurred in a single interval [24]. Genetic methods allow for the unambiguous characterization of COs in multiple intervals, but the generation of large data sets can be time and labor intensive. The analysis of genetic markers can involve PCR genotyping, DNA blot hybridization, or a combination, which has to be done for every marker in each individual [28, 29].

Cytological methods: These involve looking at structures on chromosomes that are known to be involved in the CO process [30, 31]. The most common method is immuno-staining meiotic prophase I chromosome spreads using antibodies for proteins that are known to be present in the complexes responsible for the formation of COs, such as Mlh1p [32, 33]. Since these proteins are part of the complex present at CO sites, the antibodies that bind these proteins form foci, which can be counted. Often, this is done in conjunction with a differently labeled antibody that recognizes the synaptonemal complex (SC) so that the chromosomes can be visualized and distinguished. This method allows the researcher to monitor all chromosomes at once in each meiotic spread, but only offers a snapshot in time, which may not be indicative of all of the COs that will form. In addition, closely spaced double COs can be potentially difficult to resolve.

Other cytological methods for interference analysis include the monitoring of late recombination nodules (LNs) and chiasmata [34-38]. LNs are electron dense structures that form during late meiotic prophase I and are generally thought to represent CO complexes that will eventually become chiasmata [35, 37]. The correlation between LNs and COs is based on the observation that LN frequency and positional data closely matches that of genetic CO data [38]. Chiasmata are mature COs that have created physical linkages between homologous chromosomes before MI division, which can be visualized in diplotene bivalents [36, 39]. Monitoring chiasmata is a good way to determine total numbers of COs per meiosis [40], and in some systems it is possible to map them to a precise location [34]. Elegant cytological methods allow the researcher to see physical structures involved in crossing over directly. However, the detailed technical nature of these methods as well as the amount of time and labor involved is not amenable to the creation of very large data sets.

Analysis of CO interference

Several methods have been described for analyzing interference [19, 25, 41-43], here, we will use the method of Malkova et al., to measure interference using pollen tetrad data. Briefly, the Malkova et al., method compares the map distance of one interval when an adjacent interval has no CO to the map distance of the first interval when the adjacent interval does have a CO [19]. If the genetic distance in the interval in question is significantly lower with the presence of a CO in the adjacent interval, one can conclude interference extends from one interval to the other. The ratio of distance (without adjacent CO/with adjacent CO) gives a measure of interference. A value of 1

indicates no interference; a value of 0 indicates complete interference and values between 1 and 0 correlate with the strength of interference.

Map distances can be calculated using the Perkins mapping function: $X = 100[(1/2T + 3NPD)/n]$ [44]. This equation calculates an approximation of map distance in a single interval, but its validity diminishes for intervals that sometimes contain more than two crossovers. For this equation tetrads are designated parental ditype (PD), non-parental ditype (NPD), or tetratype (T), depending on the segregation of the marker pair defining each interval.

The Malkova et al., method can be used to compare interference data from two data sets (e. g. wild type and mutant) from the same intervals. One first calculates the ratio of map distance without adjacent CO/map distance with adjacent CO as above for each dataset. These ratios can be statistically compared by obtaining a Z score using the following equation:

$$Z = |R1-R2|/\sqrt{[var\ (R1-R2)]}$$

where R1 is the ratio in treatment 1 (e. g. wild type) and R2 is the ratio in treatment 2 (e. g. mutant).

The significance of the difference between these two ratios can be assessed using a one-tailed test as described on the Stahl Lab Online Tools (http://molbio.uoregon.edu/~fstahl).

When working with tetrad data, it is also possible to calculate interference using a single interval [25]. One can estimate the fraction of NPD tetrads expected in the case of no interference from the fraction of T tetrads observed in an interval using this formula:

Fraction of NPDs expected = $0.5[(1-fT) - (1-(3fT/2))^{(2/3)}]$

Where fT is the fraction of tetratypes observed

As in the methods above, one can then divide the observed fraction of NPDs by the expected to get a measurement of the strength of interference in this interval. The Papazian method, relying as it does on just the T frequency to calculate expected NPDs, uses data inefficiently. A method for detecting interference that uses tetrad data more effieciently is available online at Stahl Lab Online Tools: (http://molbio.uoregon.edu/~fstahl).

Interference can also be measured directly from meiotic chromosomes that have been appropriately labeled such that the chromosomes are distinguishable from one another and CO sites are visible [32, 33]. Inter-CO distances can be measured using computer software such as NIH image (http://rsb.info.nih.gov/nih-image/). The distribution of these distances can be fit to gamma distributions that simulate CO placements with varying degrees of interference. The best fit distribution can be used to estimate the interference parameter *v* [6, 11, 32, 33, 45].

Fluorescent tetrad analysis in A. thaliana

Our lab has developed a unique pollen-based visual assay for meiotic recombination in *A. thaliana* that can be used to quickly generate large data sets that are amenable to interference analysis [16, 17]. *A. thaliana quartet* (*qrt*) mutants produce pollen tetrads in which the four meiotic products are held together, allowing all products of a single meiotic event to be studied relative to one another [46, 47]. This assay system is based on a collection of transgenic lines, in the *qrt* background, each carrying a gene encoding either a red, cyan, or yellow fluorescent protein excitable by different wavelengths of light. Expression of these markers is directed by a post-meiotic pollen-specific promoter (LAT52) [17, 43, 48]. Researchers can construct visually assayable genetic intervals by crossing lines carrying linked markers. Lines carrying two or more marker genes on the same chromosome expressing differently colored proteins produce tetrads that segregate the marker genes (and thus the proteins they encode) in the pollen tetrads in patterns that reflect whether or not a recombination event has happened between them. This system, can detect CO events directly in the pollen grains, and through the construction of double intervals delineated by three colors, can be used to measure CO interference throughout the *A. thaliana* genome.

We have constructed a genome-wide library of single-insertion fluorescent tagged lines (FTLs) by transforming *A. thaliana* (Col) with these marker genes [16]. All of the markers shown in Figure 1 map outside of genes. There are currently 113 FTL lines available including 35 DsRed2, 41 eYFP, and 37 eCFP. Combinations of these three groups can be crossed to create lines with three different markers. Pollen from plants that are heterozygous for these markers is viewed using an epi-fluorescence microscope with three different filters. Such a set of three linked markers on chromosome 5 has been used

by our lab to detect subtle differences in interference between wild-type and *mus81* mutant plants [17].

Advantages and Limitations of FTL interference analysis

The FTL system is capable of generating large data sets rapidly. An experienced student can score ~ 500 three color tetrads in an hour. In addition, the FTL system is flexible in that it is possible to assay interference in almost any genomic location using the extensive collection of marker lines. These characteristics make comparative experiments designed to detect differences in interference levels between sample populations, such as mutants or experimental treatments, routinely feasible.

Although it is possible to score interference in most genomic locations, there are some sections of the genome that have few or no mapped FTL insertions. An insert of each color is needed in any particular region, and some regions do not meet this requirement. Markers in regions near centromeres are particularly sparse, and chromosomes 2 and 4 do not have the marker density of chromosomes 1, 3, and 5 (see Figure 1). Additionally, the FTL system produces data only from male recombination, and in *A. thaliana*, male and female CO levels, as well as interference, are significantly different [14, 28].

Meiotic recombination mutants often exhibit lower levels of pollen viability when compared to wild-type due to inaccurate chromosome segregation or chromosome fragmentation. Chromosomes that experience an aberrant number of exchanges are more likely to segregate improperly and thus produce inviable products, so the viable products of a mutant meiosis can be enriched for COs compared to inviable products. The FTL

system can currently only be used to assay COs in viable pollen grains, so it is possible that when analyzing mutant tetrads, the total change in COs can be underestimated. However, in an example in which a mutant (*Atmsh4*) has been analyzed by both the FTL and cytological methods (which are not thought to artificially enrich for COs) the CO data were very similar and the conclusions also similar [17, 31].

Using this system, it is possible to analyze two intervals simultaneously. Available genetic methods in some organisms allow the researcher to analyze simultaneously as many intervals as they have distinguishable markers, so these other methods may be better for whole-chromosome views of interference. Drouaud et al. recently monitored the segregation of 44 markers on a single *A. thaliana* chromosome and used a sliding window analysis system to analyze sex-specific interference levels on a single chromosome [14]. This type of analysis is not currently possible using the FTL system.

In this protocol, we describe a method for using the FTL system to quickly generate large datasets that are amenable to CO interference analysis in the flowering plant *Arabidopsis thaliana*. We describe a strategy for creating visually scorable adjacent intervals by crossing marker-insertion lines from the FTL library, which can be rapidly and simultaneously screened for COs. This assay system is capable of detecting and differentiating single COs as well as 2, 3, and 4-strand double COs. We also describe multiple ways of analyzing the data that are produced by this method.

Applications

Other than assaying CO interference in adjacent intervals, this system can be used to assay CO rates in single intervals (which is a requisite for interference analysis in this system). The constructs created in this protocol can also be used in a visual assay for pollen viability.

MATERIALS

Reagents

- *A. thaliana qrt1* (Col) seeds (Arabidopsis Biological Resource Center, http://www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm)
- *A. thaliana* FTL seeds (available from G.P.C.)
- (Optional) *A. thaliana* seeds heterozygous for mutant(s) of interest
- Metro-Mix 400 (Sun Gro Horticulture Inc, www.sungro.com)
- Sucrose
- 1M CaCl2
- 0.5 M Boric Acid
- Triton-X detergent (Fisher Scientific, Cat # BP151-100, www.fishersci.com)
- PGM screening solution (see Reagent setup)

Equipment

- Fluorescence microscope (Nikon Eclipse E1000, Nikon, www.nikon.com)
- eCFP, eYFP, DsRed2 filters (Chroma Technologies, www.chroma.com)
- X-Cite 120 Illumination system (EXFO, http://www.exfo-lifesciences.com/x-cite/index.asp)

- 75 mm x 25 mm, 1 mm thick glass microscope slides (Okando via Genesee Scientific, www.geneseesci.com)

- 18 mm x 18 mm cover glass (Fisher Scientific, www.fishersci.com)

- 5-50 ul Finn pipette (Thermo Electron)

- Pipette tips

- Fine point stainless steel tweezers (Biomedical Research Instruments)

- Laboratory counter (Fisher Scientific, www.fishersci.com)

- Soft cotton cloth

Reagent Setup

- PGM screening solution: To 950 ml H20, add 170 g Sucrose (17% wt/vol), 2 ml 1M CaCl2 (2mM), 3.25 ml 0.5M Boric acid (1.625 mM), and 100 ul Triton-X (0.1% vol/vol)

Equipment Setup

A Nikon Eclipse E1000 epi-fluorescence microscope equipped with a motorized filter set including eCFP, DsRed2, and eYFP filters from Chroma Technology was used to develop these methods. No special setup of the microscope is needed. Any epi-fluorescent microscope equipped with these filters will allow the researcher to execute this protocol, but the use of motorized filters is recommended.

PROCEDURE

Selecting appropriate FTLs

1. From the FTL insert library, select the appropriate lines that will define two adjacent intervals. 5 cM per Mb is a reasonable estimate for the average genetic-to-physical ratio in *A. thaliana*.

Δ CRITICAL STEP One line for each of the eCFP, DsRed2, and eYFP markers will need to be selected, but they may occur in any order. The sizes of the two intervals do not have to be similar, but it will slightly complicate analysis if either of the intervals is large enough such that an abundance of NPD tetrads in either single interval is observed. Note that for the remainder of the protocol, the line that carries an insert closest to the North end of the chromosome will be referred to as 'A', the next will be 'B' and the line with an insert closest to the South end will be 'C.'

Construction of first interval (I1)

2. Plant at least 12 seeds each from line A and line B (seeds will usually be heterozygous for the interval). Grow plants on Metro-mix 400 soil in an environmentally-controlled growth chamber with a long photoperiod (16 hrs light at 20º C/ 8hrs dark at 20º C).

CRITICAL STEP: Steps 2-11 can be omitted if the researcher wishes to use interference intervals that our lab has already constructed (see Figure 1). These lines will be supplied as a segregating population of seeds from a ABC/+++ parent and are available by request from G.P.C.

CRITICAL STEP: If the study includes analysis of interference in mutant backgrounds, it will save time to cross the mutants into the *qrt1* background during construction of the three-color interval. We recommend maintaining and crossing meiotic mutants in the

heterozygous state to avoid genomic abnormalities. Useful plants will have the genotype mutant/+; *qrt1*/*qrt1*.

3. As each marker will be segregating 1:2:1 in this generation, it is necessary to determine the marker genotype of each plant by examining the pollen under the epi-fluorescence microscope when plants begin flowering (4-5 weeks). As this step is diagnostic and does not require any counting, use the simplified screening process outlined in Box 1. Plants homozygous (4:0 fluorescent pollen, see Figure 2) for their respective marker will be used for the next crossing step. Collect seeds from 2:2 plants (A/+ and B/+) and 4:0 plants (A/A and B/B) to save as stocks. Store seed at room temperature in labeled 1.5 ml polypropylene tubes with a small hole punctured in the top with a needle to prevent molding.

Box 1: Simplified screening process

i. Rub a microscope slide briskly for ~10 sec with a soft cotton cloth. This is necessary so that the PGM holds its drop formation.
ii. Drop six evenly spaced 10 µl spots of PGM onto the slide.
iii. Using forceps, remove an open flower and place it face down in the first drop of PGM. Let it soak for 20 sec.
iv. Use a gentle tapping motion to release the pollen into the PGM. This is best done on a black bench top that will allow the researcher to see the pollen being released into the solution.
v. Repeat steps iii and iv for 5 more flowers until the slide is full.

CRITICAL STEP: Work quickly so that the drops of PGM do not dry out prior to screening. Drops that dry out can be rehydrated by adding a second drop of PGM.

vi. View and diagnose the pollen under the appropriate filter using the epi-fluorescence microscope. If a plant is homozygous for the marker, all four pollen grains in each tetrad will be fluorescent. If a plant is heterozygous for the marker, each tetrad will fluoresce in a 2:2 fluorescent: non-fluorescent pattern. Likewise, if the plant is homozygous wild-type, no pollen will fluoresce (Figure 2).

END OF BOX 1

4. To create the first interval, cross a line-A plant that is homozygous for one color insert to a line-B plant homozygous for a different colored insert. Make several crosses to be sure of a viable cross. For details and tips on crossing *A. thaliana*, refer to *Arabidopsis: a laboratory manual* [49]. The stigma of the female used for the cross will mature into a silique containing seeds used for the next step.

TROUBLESHOOTING

5. Once the siliques from the crosses (step 4) have dried (2-3 weeks), harvest the seeds and plant them as described in step 2. At the same time, plant at least 12 seeds from line C, which will define the second interval.

6. When these plants begin flowering (4-5 weeks) diagnose the fluorescent genotypes of each plant as described in Box 1. Individuals from the A-B cross should yield pollen tetrads with the same fluorescent pattern (2 color-A: 2 color-B), discard plants that do not. For line C, select plants homozygous for the C marker. Save seed stocks from the A/B plants and the C/+ and C/C plants as described in step 3.

Construction of second interval (I2), and cross into mutant background

7. Cross the A/B plants to the C homozygote as in step 4.

Δ CRITICAL STEP This cross needs to be made several times because the useful plants in the next generation will have all three markers, which can only be achieved when C is fertilized with a recombinant pollen grain harboring both the A and B inserts in a *cis* configuration. The probability (as a percent) of this event can be calculated by dividing the recombinant frequency between A and B (in percent) by 2. This should give a good idea of how many crosses to complete (depending on how many seeds you usually get from your crosses). As a general rule, perform enough crosses to ensure that you have three times as many seeds as would be needed to produce a single three-colored plant. TROUBLESHOOTING

8. After the siliques from these crosses have dried (2-3 weeks), plant all of the seeds as described in step 2. Also plant at least 8 *qrt1*/*qrt1* (non-fluorescent) seeds and/or the mutant/+; *qrt1*/*qrt1* seeds.

9. When these plants begin flowering, diagnose the marker genotype of each plant as described in Box 1. Each of the three fluorescent marker genotypes will have to be determined under the appropriate colored filter in succession. Pollen tetrads from useful plants will fluoresce 2AB:2C, having the genotype AB+/++C. The markers will have had a chance to recombine during the meiosis that produced the tetrads being examined. In useful plants, the tetrads will show many combinations in relation to one another. What is important is that all three colors show a 2:2 pattern.
10. At this stage, if the study includes interference analysis in mutant backgrounds, cross the three color lines to mutant/+; *qrt1*/*qrt1* lines. Otherwise, cross the three color lines to the Col *qrt1* mutants using a similar strategy as in step 7 with the purpose of producing lines that are ABC/+++; *qrt1*/*qrt1* (and optionally mutant/+).

TROUBLESHOOTING

11. When the siliques from these crosses (step 10) have dried (2-3 weeks) plant all of the seeds. Useful plants will have the genotype ABC/+++ and show a majority of pollen tetrads with a 2ABC: 2 non-fluorescent pattern. These plants can be analyzed for interference, but to generate large amounts of data from multiple individuals, it will be necessary to analyze the appropriate progeny from self-crosses. If crosses were made into mutant backgrounds, useful plants will have the genotype ABC/+++; mutant/+; *qrt1*/*qrt1*.

Scoring three-color tetrads for recombination

12. Plant at least 50 seeds collected from ABC/+++; (mutant/+); *qrt1*/*qrt1* individuals. These plants will need to be distinguished from one another throughout the scoring process, so it is useful to sow the seeds in divided flats. We plant in 6 x 4 divided flats (24 cells) where the flats measure 26 cm x 54 cm and each cell measures 6.5 cm x 9 cm. This allows each plant enough room to grow while also being easy to distinguish the plants from one another. At this stage, the plants need to be grown in a temperature-controlled growth chamber.

Δ CRITICAL STEP Columbia FTL lines have been extensively tested in our lab. In general, *A. thaliana* plants are sensitive to environmental changes, so it is important to keep a constant environment as much as possible. Work done in our lab has shown that temperature levels influence recombination, so it is critical to maintain the same levels for all experimental and control populations [16].

13. Once the plants begin flowering (4-5 weeks), it is necessary to determine the fluorescent genotype of each individual plant as described in Box 1. Only plants that are ABC/+++ can be scored for recombination in both intervals simultaneously, discard all others. Note that plants of the genotype ABC/ABC are useful for quantifying pollen viability, see Box 2 for details [17].

Box 2: quantifying pollen viability

i. To begin the scoring process, select an ABC/ABC plant for analysis.

ii. Rub a microscope slide briskly for ~10 sec with a soft cotton cloth.

iii. Place two 10ul drops of PGM onto the slide about 3 cm apart. The drops should form a tight bead on the slide.

iv. Using the forceps, remove a single open flower and place it face down into the first drop of PGM, allow to soak for at least 20 sec. Note the flower number on the plant by counting up the bolt.

v. Use a tapping motion to release the pollen into the PGM. This is best done on a black bench top that will allow the researcher to see the pollen being released into the solution. Remove the flower from the first drop of PGM and repeat the process with the same flower in the second drop. This maximizes the amount of pollen that can be scored from each flower. Gently place a cover slip on top of each (pollen-containing) drop of PGM.

vi. Screen the tetrads on the 10x objective. A pollen grain is classified as dead (inviable) if it is either a) shriveled up/aborted or b) fully developed, but *all* of the fluorescent markers fail to express in the grain (i.e. it appears blank). In any given tetrad, you may observe a combination of live/dead grains. Depending on the pattern of dead:alive grains, each tetrad can be classified as 4:0, 3:1, 2:2, 1:3, 0:4 [17]. Pollen viability can be quantified as a percent of dead grains/total grains, and the percentages of each class of tetrad can be presented in a histogram. Many meiotic mutants show a characteristic elevation in the 2:2 class as compared to wild type, which typically shows a total viability of >95%

END OF BOX 2.

14. To begin the scoring process, select an ABC/+++ plant for analysis.

15. Rub a microscope slide briskly for ~10 sec with a soft cotton cloth.
16. Place two 10ul drops of PGM onto the slide about 3 cm apart. The drops should form a tight bead on the slide.
17. Using the forceps, remove a single open flower and place it face down into the first drop of PGM, allow to soak for at least 20 sec. Note the flower number on the plant by counting up the bolt.

Δ CRITICAL STEP It is best to only score flowers from the primary bolt. Work in our lab has shown that recombination rates are altered in secondary axes [16]. In addition, we recommend scoring only the 5^{th} to the 30^{th} flowers because younger and older flowers tend to have lower pollen counts and elevated lethality. However, we have not observed any relationship between flower number and recombination [16].

18. Use a tapping motion to release the pollen into the PGM. This is best done on a black bench top that will allow the researcher to see the pollen being released into the solution. Remove the flower from the first drop of PGM and repeat the process with the same flower in the second drop. This maximizes the amount of pollen that can be scored from each flower. Gently place a cover slip on top of each (pollen-containing) drop of PGM.

TROUBLESHOOTING

19. The fluorescent markers in the ABC/+++ individuals will be segregating in the pollen. In addition, if a mutant allele has been introduced, the plants will be mixture of wild-type, heterozygotes, and mutant genotypes – these genotypes will be scored in step 20. The researcher can score each four-member pollen tetrad, using the 20X objective on the microscope. The researcher will place each tetrad

in the appropriate recombinant category based on the pattern of fluorescence in the grains (see Table 1 and Figures 3 and 4).

Δ CRITICAL STEP In the examples (Figure 3; Tables 1 and 2), the markers were in the red-yellow-cyan (North to South) configuration. Three-color intervals with a different configuration (such as yellow-cyan-red) are perfectly acceptable, but the North-South order of the different colored markers will determine how the fluorescent tetrads are categorized. To determine how to classify each type of tetrad fluorescence pattern, it may be useful to redraw Figure 3 using the appropriate configuration of markers (see supplemental material for a blank template figure).

Δ CRITICAL STEP Do not score tetrads in clumps, this can lead to erroneously classifying tetrads. The Triton-X detergent is added to the PGM to prevent clumping. Only score 4-member tetrads.

20. If a mutant allele is segregating in this generation, genotype the mutants after scoring is complete. Blind scoring is good experimental practice to avoid bias in the counting process.
21. Repeat for as many plants as the analysis requires – this will differ from experiment to experiment.

TIMING

Steps 1-6: Selection of appropriate FTLs and construction of the first interval, 6-7 weeks

Steps 7-11: Construction of the second interval, and cross into mutant background, 12-14 weeks

Steps 12-20: Scoring three-color tetrads for recombination, 4-6 weeks

TROUBLESHOOTING

Troubleshooting advice is provided in Table 2.

ANTICPATED RESULTS

The type of results for a single plant generated by three-color FTL interference experiments is shown in Table 3. Compiled results from an experiment designed to test the interference phenotype of a mutant is shown in Table 4. If the object of the analysis is to test a mutant phenotype, and the crosses are conducted in the manner described, then the mutant (assuming recessivity) will be segregating 1:3 in the scoring generation. This is ideal because the blind controls for the experiment are built into the design.

To calculate interference levels from the data that are produced using this method, we recommend using the Malkova et al. method for quantifying interference [19]. With this method the researcher determines the map distance of one of the intervals without and with a CO in the adjacent interval, and the ratio between the two represents the level of interference. Using letter codes from Table 1, the following formulas are used:

X_{I1} (w/o adjacent CO) = ((1/2T) + 3NPD)/total = ((1/2 B) + 3(H))/(A+B+H)

X_{I1} (with adjacent CO) = ((1/2T) + 3NPD)/total

= (1/2(D+E+F+G+K) + 3(J+L))/(C+D+E+F+G+I+J+K+L)

Interference ratio = X_{I1} w/o adjacent CO/X_{I1} with adjacent CO

Where, as above, X refers to the map distance generated from the Perkins equation [44].

Ratios between sample populations can be compared using Stahl Lab Online Tools: (http://molbio.uoregon.edu/~fstahl).

Examples of interference analysis using data generated by the FTL system

From an individual plant (table 2):

X I1 (w/o adjacent CO) = ((1/2 B) + 3(H))/(A+B+H) = ((1/2(198)) + 3(4))/(207+198+4) = 27.1 cM

X I1 (with adjacent CO) = (1/2(D+E+F+G+K) + 3(J+L))/(C+D+E+F+G+I+J+K+L) = (1/2(7+5+15+4+0) + 3(0+0))/(119+7+5+15+4+2+0+0+0) = 10.2 cM

Interference ratio = X I1 w/o adjacent CO/X I1 with adjacent CO = 10.2/27.1 = 0.38

Comparing compiled data from two sample populations (table 3):

<u>+/+ and +/- (row 1)</u>

X I1 (w/o adjacent CO) = ((1/2 B) + 3(H))/(A+B+H) = ((1/2 (3881)) + 3(68))/(3729+3881+68) = 27.9 cM

X I1 (with adjacent CO) = (1/2(D+E+F+G+K) + 3(J+L))/(C+D+E+F+G+I+J+K+L) = (1/2(128+161+149+129+6) + 3(9+0))/(2286+128+161+149+129+21+9+6+0) = 10.9 cM

<u>Interference ratio</u>

Interference ratio = X I1 w/o adjacent CO/X I1 with adjacent CO = 0.3885

-/- mutant (row 2)

X I1 (w/o adjacent CO) = ((1/2 B) + 3(H))/(A+B+H)

= ((1/2 (3065)) + 3(62))/(2785+3065+62) = 29.1 cM

X I1 (with adjacent CO) = (1/2(D+E+F+G+K) + 3(J+L))/(C+D+E+F+G+I+J+K+L)

= (1/2(110+112+89+103+1) + 3(1+0))/(1712+110+112+89+103+16+1+1+0) = 9.8 cM

Interference ratio = X I1 w/o adjacent CO/X I1 with adjacent CO = 0.3378

Test of significance between two ratios of map distances (Stahl Lab Online Tools)

The following data were input:

Interval name = wt with adjacent CO

PD = A = 3729; T = B = 3881; NPD = H = 68

Interval name = wt without adjacent CO

PD = C + I = 2307; T = D+E+F+G+K = 573; NPD = J+L = 9

Interval name = mutant with adjacent CO

PD = A = 2785; T = B = 3065; NPD = H = 62

Interval name = mutant without adjacent CO

PD = C + I = 1728; T = D+E+F+G+K = 415; NPD = J+L = 1

The program gives this output:

Ratio Definitions R1 = wt with/wt without R2 = mutant with/mutant without

Ratios R1 = 0.3885 R2 = 0.3378

Variance of Ratios varR1= 0.00032444 varR2= 0.00026894

Standard Error of Ratios S.E. R1= 0.0180122 S.E. R2= 0.01639943

Var[R1 - R2] = 0.00059338

|R1-R2| = 0.05075571

(Two Tailed) Is |R1-R2| > 1.96 x sqrt Var [R1-R2]? 0.0508 > 0.0477 Significant

(One Tailed) Is |R1-R2| > 1.65 x sqrt Var [R1-R2]? 0.0508 > 0.0402 Significant

In this case, the mutant ratio was significantly different from the wild-type ratio and we concluded that the two genotypes had different interference levels. The mutant has significantly stronger interference than the wild-type.

ACKNOWLEDGMENTS

We would like to thank Corbin Jones and Frank Stahl for critical reading. We would also like to thank the NSF (MCB-0618691) and DOE (DE-FGO2-05ER15651) for financial support.

COMPETING INTERESTS STATEMENTS

The authors declare no competing financial interests.

FIGURE LEGENDS

Figure 3.1. Map of fluorescent transgenes. The chromosomal (green bars) insertion site of the transgene carried by each FTL line is indicated by a red (DsRed2), yellow (eYFP), or cyan (eCFP) circle. The identification number of each insertion is given above the circle. The genetic intervals (I1a, I1b, I1c, I2a, I2b, I2c, I3a, I3b, I3c, I3d, I5a, I5b, I5c, and I5d) that are available by request from G.P.C are delineated by brackets. The

Arabidopsis Information Resource (TAIR) "chromosome map tool" (http://www.arabidopsis.org/) was used to place T-DNA insertion points on the map.

Figure 3.2. Single locus segregation patterns in pollen tetrads. A locus with a transgenic marker construct encoding a fluorescent protein (in this case DsRed2) can be homozygous (A/A) or hemizygous (A/+) for the marker or it can be wild-type (+/+). The fluorescence signal in the pollen tetrads will reflect the marker genotype and yield 4:0 (left) 2:2 (middle) or 0:4 (right) pollen tetrads respectively.

Figure 3.3. Classification of tetrad fluorescent patterns. DsRed2 (red oval), eYFP (yellow oval), and eCFP (cyan oval) transgenes define two hypothetical adjacent genetic intervals (I1 and I2). The letters (A-L) correspond to the classes of tetrads described in Table 1. The four chromatids present after DNA replication can experience no crossovers (a), single crossovers (b and c), and DCOs in the combined interval including two strand DCOs (d), both kinds of three-strand DCOs (e and f), and four strand DCOs (g). A four-strand DCO in either of the single intervals can also be observed as an NPD tetrad either in the absence of an adjacent CO (h and i) or the presence of an adjacent CO (j and k). A four strand DCO in both single intervals will result in a tetrad that is NPD for each (l). Each of these events can be distinguished by observing the segregation of the fluorescent markers in the pollen tetrads. In each panel the top, second, third and bottom chromatids are shown segregating into the pollen grains at the top, right, bottom and left positions respectively.

Figure 3.4. Examples of multi-color fluorescent tetrads. To assess three-color intervals each tetrad must be visualized through each of three different fluorescent filters (red, yellow, cyan) on the epi-fluorescence microscope. These individual images can be merged (right column) into a composite image using graphics software such as Adobe Photoshop. A plant that is heterozygous for three markers in *cis* configuration can yield non-recombinant pollen tetrads that have all three colors in the same two pollen grains (top row). A single crossover (in I1 in this case) will yield a pollen tetrad that has one grain that has all three colors, one grain that has two colors, one grain that has one color and one that has no color (middle row). Double crossovers will yield pollen tetrads with a variety of segregation patterns (see Figure 3), in this case a four-strand double crossover results in a tetrad with one red, one yellow/blue, one blue and one red/yellow grain (see Figure 3g).

TABLES

Table 3.1. Letter codes for classification of tetrads

A	B	C	D	E	F	G	H	I	J	K	L
NR	SCO I1	SCO I2	2 st* DCO	3 st* DCOa	3 st* DCOb	4 st* DCO	NPDI1 NCOI2	NCOI1 NPDI2	NPDI1 SCOI2	SCOI1 NPDI2	NPDI1 NPDI2

*st = strand

Table 3.2. Troubleshooting guide

Step	Problem	Possible Reason	Solution
4, 7, 10	Inviable crosses, crosses not producing many seeds	Stigma damaged during emasculation, plants too old	Use new forceps that are sharp. Use a jeweler's magnifying glass. Emasculate young buds that are as large as possible.
18	Few scorable tetrads	Tetrads broken apart, few tetrads released into PGM, mutant causes lethality,	Be gentle when tapping pollen out of flower. Cover slip should be dropped gently, never squashed. It may help to let flower soak in PGM for longer than 20 sec. Some mutants do not produce a lot of viable tetrads, and we find that it helps to score these mutants on lower magnification (10X) so that viable tetrads can be identified more quickly. Most mutants produce some viable tetrads, so low viability can be overcome by scoring a lot of flowers from these individuals.

Table 3.3. Example of data produced by one individual plant

plant	screen date	flower	NR	SCO-I1 (R/Y)	SCO-I2 (Y/C)	2st* DCO	3st* DCOa	3st* DCOb	4st* DCO	NPD-I1 NCO-I2	NCO-I1 NPD-I2	NPD-I1 SCO-I2	SCO-I1 NPD-I2	NPD-I1 NPD-I2	total
2B2	7Mar	4	54	45	25	2	2	3	0	0	0	0	0	0	131
	9Mar	11	49	38	19	0	1	2	1	1	1	0	0	0	112
	9Mar	12	35	29	20	2	0	3	0	0	0	0	0	0	89
	12Mar	21	41	54	23	2	2	4	3	0	1	0	0	0	130
	12Mar	22	28	32	32	1	0	3	0	3	0	0	0	0	99
		totals	207	198	119	7	5	15	4	4	2	0	0	0	561

Table 3.4. Example of aggregate data comparing a mutant to wild type

NR	SCO-I1 (R/Y)	SCO-I2 (Y/C)	2st DCO	3st DCO a	3st DCO b	4st DCO	NPD-I1 NCO-I2	NCO-I1 NPD-I2	NPD-I1 SCO-I2	SCO-I1 NPD-I2	NPD-I1 NPD-I2	total	Geno-type
3729	3881	2286	128	161	149	129	68	21	9	6	0	1056 7	+/+; +/-
2785	3065	1712	110	112	89	103	62	16	1	1	0	8056	-/-

FIGURES

Figure 3.1.

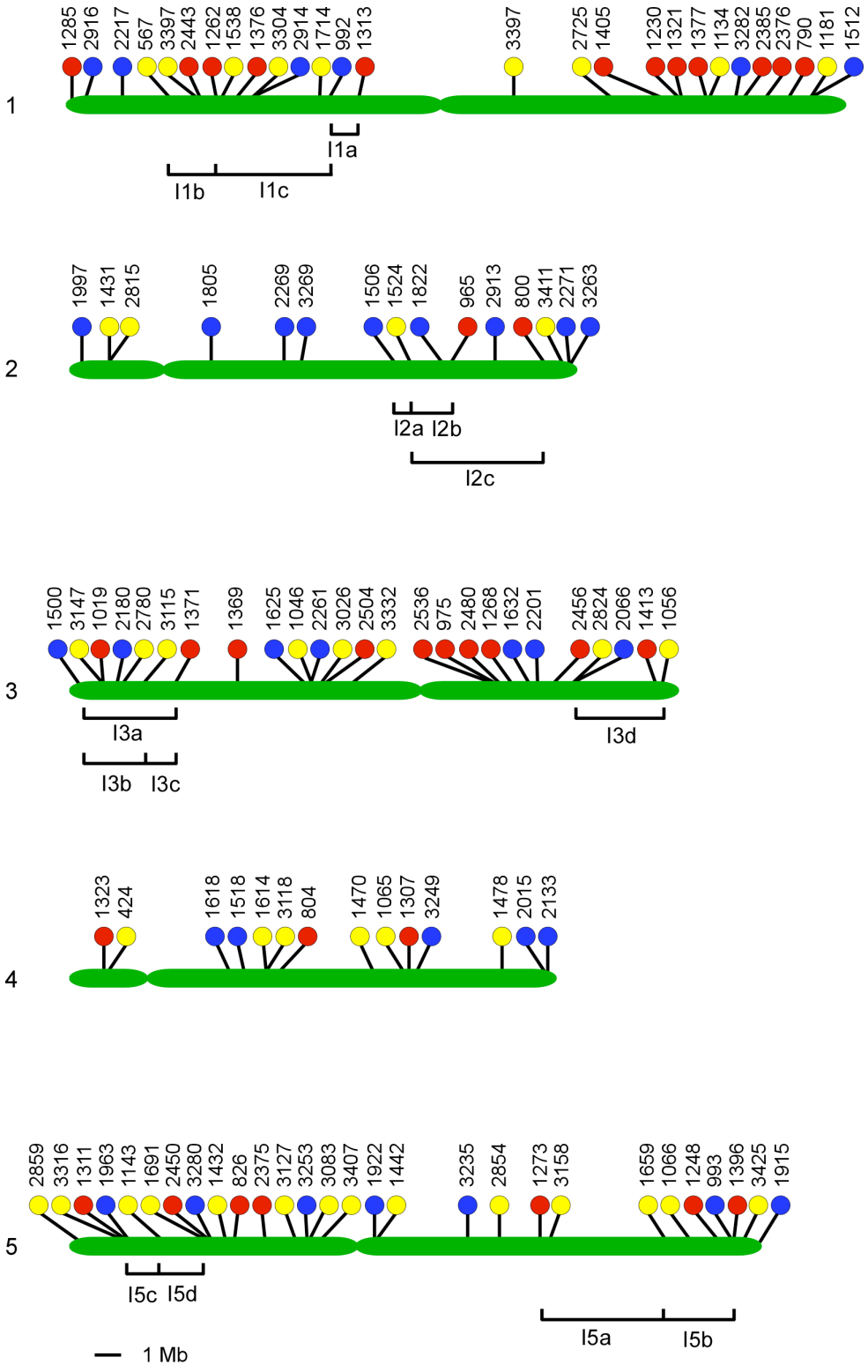


Figure 3.2.

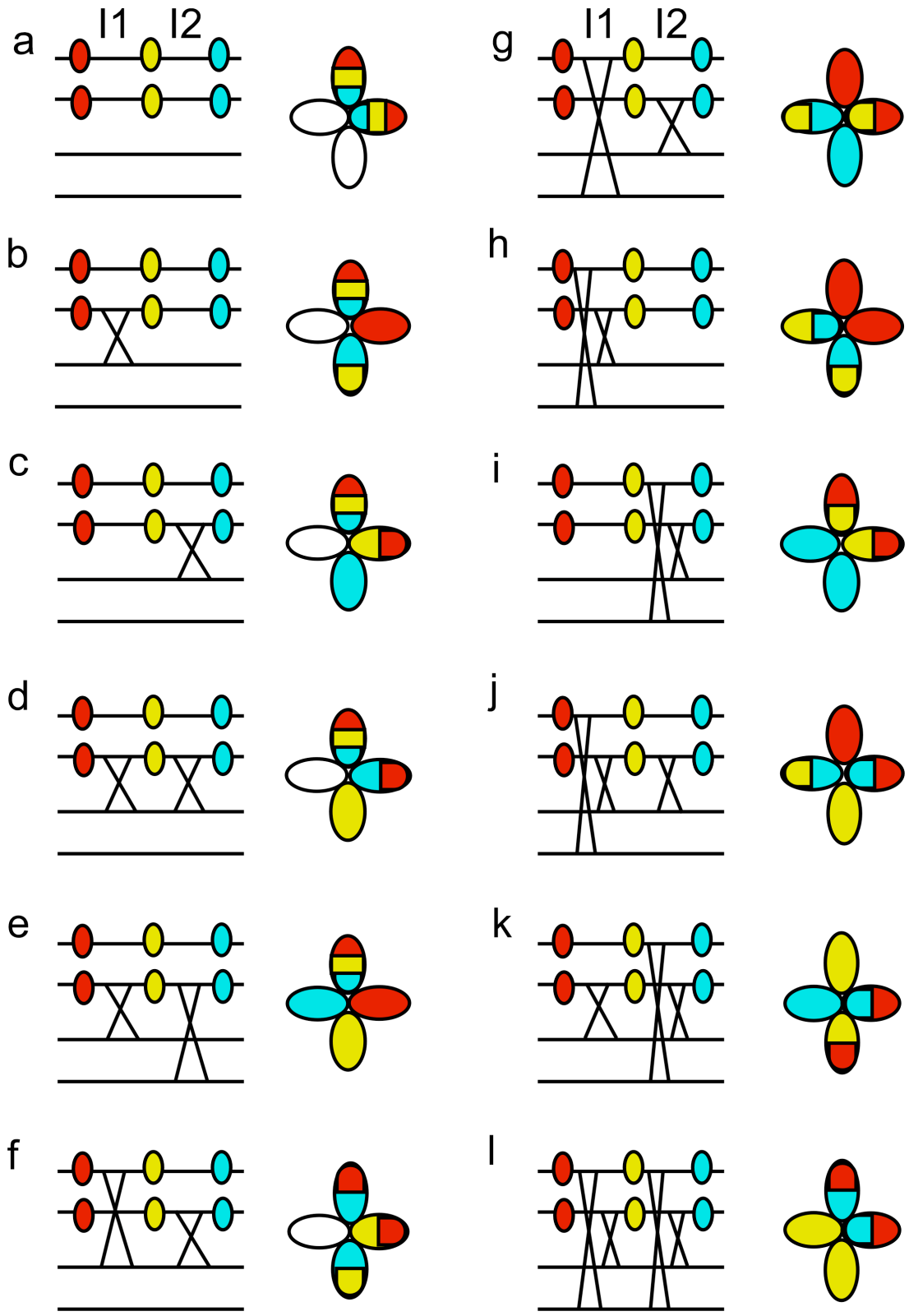


Figure 3.3.

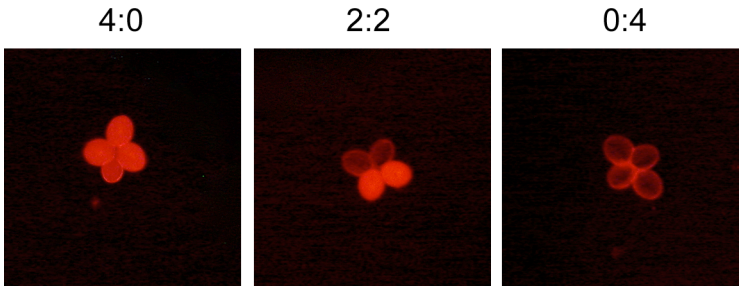
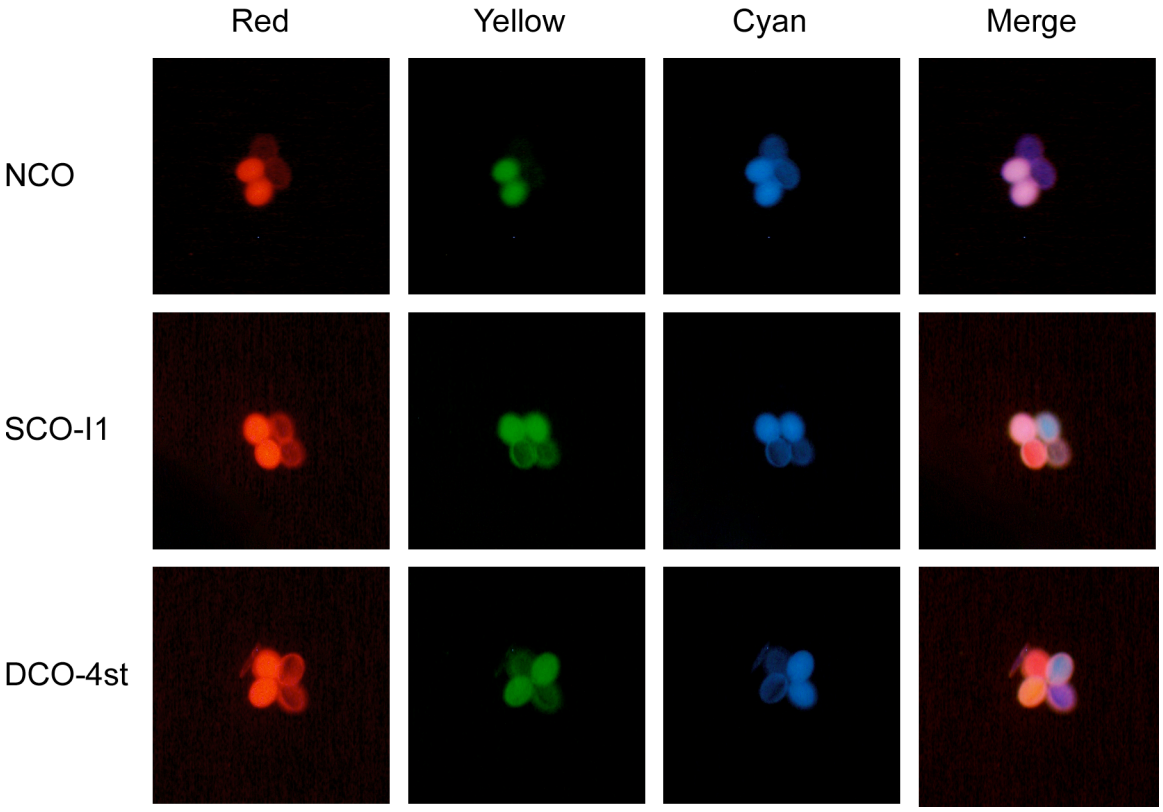


Figure 3.4.



REFERENCES

1. Zickler, D. and N. Kleckner, *The leptotene-zygotene transition of meiosis.* Annu Rev Genet, 1998. **32**: p. 619-97.

2. Page, S.L. and R.S. Hawley, *Chromosome choreography: the meiotic ballet.* Science, 2003. **301**(5634): p. 785-9.

3. Felsenstein, J., *The evolutionary advantage of recombination.* Genetics, 1974. **78**(2): p. 737-56.

4. Jones, G.H. and F.C. Franklin, *Meiotic crossing-over: obligation and interference.* Cell, 2006. **126**(2): p. 246-8.

5. Borner, G.V., N. Kleckner, and N. Hunter, *Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis.* Cell, 2004. **117**(1): p. 29-45.

6. Foss, E., et al., *Chiasma interference as a function of genetic distance.* Genetics, 1993. **133**(3): p. 681-91.

7. Sturtevant, A.H., *The behavior of chromosomes as studied through linkage.* Z. Induct. Abstammungs-Vererbungsl, 1915. **13**: p. 234-287.

8. Muller, H.J., *The Mechanism of Crossing Over.* Am Nat, 1916. **50**: p. 193-434.

9. Hillers, K.J., *Crossover interference.* Curr Biol, 2004. **14**(24): p. R1036-7.

10. Copenhaver, G.P., E.A. Housworth, and F.W. Stahl, *Crossover interference in Arabidopsis.* Genetics, 2002. **160**(4): p. 1631-9.

11. Housworth, E.A. and F.W. Stahl, *Crossover interference in humans.* Am J Hum Genet, 2003. **73**(1): p. 188-97.

12. Foss, E.J. and F.W. Stahl, *A test of a counting model for chiasma interference.* Genetics, 1995. **139**(3): p. 1201-9.

13. McPeek, M.S. and T.P. Speed, *Modeling interference in genetic recombination.* Genetics, 1995. **139**(2): p. 1031-44.

14. Drouaud, J., et al., *Sex-Specific Crossover Distributions and Variations in Interference Level along Arabidopsis thaliana Chromosome 4.* PLoS Genet, 2007. **3**(6): p. e106.

15. Hawley, R.S., K.S. McKim, and T. Arbel, *Meiotic segregation in Drosophila melanogaster females: molecules, mechanisms, and myths.* Annu Rev Genet, 1993. **27**: p. 281-317.

16. Francis, K.E., et al., *A pollen tetrad-based visual assay for meiotic recombination in Arabidopsis.* PNAS, 2007. **104**(10): p. 3913-3918.

17. Berchowitz, L.E., et al., *The Role of AtMUS81 in Interference-Insensitive Crossovers in A. thaliana.* PLoS Genet, 2007. **3**(8): p. 0001-0010.

18. Lam, S.Y., et al., *Crossover interference on nucleolus organizing region-bearing chromosomes in Arabidopsis.* Genetics, 2005. **170**(2): p. 807-12.

19. Malkova, A., et al., *Gene conversion and crossing over along the 405-kb left arm of Saccharomyces cerevisiae chromosome VII.* Genetics, 2004. **168**(1): p. 49-63.

20. Mortimer, R.K. and D.C. Hawthorne, *Genetic Mapping in Saccharomyces IV. Mapping of Temperature-Sensitive Genes and Use of Disomic Strains in Localizing Genes.* Genetics, 1973. **74**(1): p. 33-54.

21. Fogel, S. and R.K. Mortimer, *Recombination in yeast.* Annu Rev Genet, 1971. **5**: p. 219-36.

22. Yauk, C.L., P.R. Bois, and A.J. Jeffreys, *High-resolution sperm typing of meiotic recombination in the mouse MHC Ebeta gene.* Embo J, 2003. **22**(6): p. 1389-97.

23. Tiemann-Boege, I., et al., *High-resolution recombination patterns in a region of human chromosome 21 measured by sperm typing.* PLoS Genet, 2006. **2**(5): p. e70.

24. Perkins, D.D., *The Detection of Linkage in Tetrad Analysis.* Genetics, 1953. **38**(2): p. 187-97.

25. Papazian, H.P., *The Analysis of Tetrad Data.* Genetics, 1952. **37**(2): p. 175-88.

26. Hawley, R.S., *Meiosis in living color: Fluorescence-based tetrad analysis in Arabidopsis.* PNAS, 2007. **104**(10): p. 3673-3674.

27. Copenhaver, G.P., *A compendium of plant species producing pollen tetrads.* J. North Carolina Acad. Sci, 2005. **121**: p. 17-35.

28. Copenhaver, G.P., W.E. Browne, and D. Preuss, *Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads.* Proc Natl Acad Sci U S A, 1998. **95**(1): p. 247-52.

29. Jeffreys, A.J., et al., *Meiotic recombination hot spots and human DNA diversity.* Philos Trans R Soc Lond B Biol Sci, 2004. **359**(1441): p. 141-52.

30. Janssens, F.A., *Spermatogenese dans les batraciens. V. La theorie de la chiasmatypie, nouvelle interpretation des cineses de maturation.* Cellule, 1909. **25**: p. 387-411.

31. Higgins, J.D., et al., *The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in Arabidopsis.* Genes Dev, 2004. **18**(20): p. 2557-70.

32. de Boer, E., et al., *Meiotic interference among MLH1 foci requires neither an intact axial element structure nor full synapsis.* J Cell Sci, 2007. **120**(Pt 5): p. 731-6.

33. de Boer, E., et al., *Two levels of interference in mouse meiotic recombination.* Proc Natl Acad Sci U S A, 2006. **103**(25): p. 9607-12.

34. Jones, G.H., *The control of chiasma distribution.* Symp Soc Exp Biol, 1984. **38**: p. 293-320.

35. Carpenter, A.T., *Recombination nodules and synaptonemal complex in recombination-defective females of Drosophila melanogaster.* Chromosoma, 1979. **75**(3): p. 259-92.

36. Sherman, J.D. and S.M. Stack, *Two-dimensional spreads of synaptonemal complexes from solanaceous plants. VI. High-resolution recombination nodule map for tomato (Lycopersicon esculentum).* Genetics, 1995. **141**(2): p. 683-708.

37. Carpenter, A.T., *Synaptonemal complex and recombination nodules in wild-type Drosophila melanogaster females.* Genetics, 1979. **92**(2): p. 511-41.

38. Carpenter, A.T., *Electron microscopy of meiosis in Drosophila melanogaster females: II. The recombination nodule--a recombination-associated structure at pachytene?* Proc Natl Acad Sci U S A, 1975. **72**(8): p. 3186-9.

39. Sanchez Moran, E., et al., *Chiasma formation in Arabidopsis thaliana accession Wassileskija and in two meiotic mutants.* Chromosome Res, 2001. **9**(2): p. 121-8.

40. Wijeratne, A.J., et al., *The Arabidopsis thaliana PARTING DANCERS gene encoding a novel protein is required for normal meiotic homologous recombination.* Mol Biol Cell, 2006. **17**(3): p. 1331-43.

41. Preuss, D., S.Y. Rhee, and R.W. Davis, *Tetrad analysis possible in Arabidopsis with mutation of the QUARTET (QRT) genes.* Science, 1994. **264**(5164): p. 1458-60.

42. Copenhaver, G.P., K.C. Keith, and D. Preuss, *Tetrad analysis in higher plants. A budding technology.* Plant Physiol, 2000. **124**(1): p. 7-16.

43. Twell, D., J. Yamaguchi, and S. McCormick, *Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis.* Development, 1990. **109**(3): p. 705-13.

44. Perkins, D.D., *Crossing-over and interference in a multiply marked chromosome arm of Neurospora.* Genetics, 1962. **47**: p. 1253-74.

45. Stahl, F.W. and R. Lande, *Estimating interference and linkage map distance from two-factor tetrad data.* Genetics, 1995. **139**(3): p. 1449-54.

46. Mezard, C., et al., *The road to crossovers: plants have their say.* Trends Genet, 2007. **23**(2): p. 91-9.

47. Perkins, D.D., *Biochemical Mutants in the Smut Fungus Ustilago Maydis.* Genetics, 1949. **34**(5): p. 607-26.

48. Stahl, F.W., et al., *Does crossover interference count in Saccharomyces cerevisiae?* Genetics, 2004. **168**(1): p. 35-48.

49. Weigel, D. and J. Glazebrook, *Arabidopsis : a laboratory manual*. 2002, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xii, 354.

CHAPTER 3 ADDENDUM

Specific contributions by L.E.B.

With the help of G.P.C., I developed the protocol for measuring crossover interference using the FTL system including crossing, data collection and analysis. I prepared the manuscript, which was reviewed and edited by G.P.C.

Chapter 4

The Role of *AtMUS81* in Interference-Insensitive Crossovers in *A. thailana*

Luke E. Berchowitz, Kirk E. Francis, Alexandra L. Bey, Gregory P. Copenhaver

Department of Biology and The Carolina Center for Genome Sciences, The University of North Carolina at Chapel Hill, CB 3280, Coker Hall 305, Chapel Hill, NC 27599

This work was previously published as: Berchowitz, L.E., et al., *The Role of AtMUS81 in Interference-Insensitive Crossovers in A. thaliana.* PLoS Genet, 2007. **3**(8): p. 0001-0010.

ABSTRACT

MUS81 is conserved among plants, animals and fungi and is known to be involved in mitotic DNA damage repair and meiotic recombination. Here we present a functional characterization of the *Arabidopsis thaliana* homolog *AtMUS81* which has a role in both mitotic and meiotic cells. The *AtMUS81* transcript is produced in all tissues, but is elevated greater than 9 fold in the anthers and its levels are increased in response to gamma radiation and MMS treatment. An *Atmus81* T-DNA insertion mutant shows increased sensitivity to a wide range of DNA damaging agents confirming its role in cells proliferating mitotically. To examine its role in meiosis, we employed a pollen tetrad-based visual assay. Data from genetic intervals on chromosomes 1 and 3 show that *Atmus81* mutants have a moderate decrease in meiotic recombination. Importantly, measurements of recombination in a pair of adjacent intervals on chromosome 5 demonstrates that the remaining crossovers in *Atmus81* are interference-sensitive, and that interference levels in the *Atmus81* mutant are significantly greater than those in wild type. These data are consistent with the hypothesis that *AtMUS81* is involved in a secondary subset of meiotic crossovers that are interference-insensitive.

INTRODUCTION

During meiotic prophase I, homologous chromosomes pair, synapse, and recombine (via crossing over or gene conversion), all of which are required for proper chromosome segregation during the subsequent stages of meiosis, in which haploid gametes are produced from diploid progenitor cells. Extensive genetic and molecular data from the budding yeast *Saccharomyces cerevisiae* has led to the "double strand break repair" (DSBR) model of meiotic recombination in which chromosomes are subjected to programmed double strand breaks [1-3]. In all sexually reproducing organisms studied to date, these breaks are dependant on the topoisomerase Spo11p [4]. These breaks are then processed leading to either crossovers (COs) or non-crossovers (NCOs). In most organisms, COs are distributed non-randomly such that one CO event inhibits the chances of another nearby event and each chromosome pair usually has at least one crossover. The term used to describe this phenomenon is CO interference [5].

Statistical and experimental evidence suggests that *Arabidopsis thaliana*, like yeast and humans, has two recombination pathways: one that exhibits crossover interference and the other, which does not [6-9]. This is in contrast to organisms such as *Caenorhabditis elegans* and *Drosophila melanogaster*, in which all COs are thought to be subject to interference [10,11]. In the organisms studied to date with both interfering and non-interfering COs, the majority of events are thought to be generated by the primary, interference-sensitive pathway. In both *S. cerevisiae* and *A. thaliana*, several genes active in the interference-sensitive pathway such as the *MSH4/MSH5* heterodimer [12] and the DNA helicase encoding *MER3* [13] have been identified. In *A. thaliana*, disruption of these genes causes a reduction of approximately ~85% of COs [6,14]. Analysis of the

distribution of the residual chiasmata in *msh4/msh4* meiocytes has led to the suggestion that the remaining 15% of COs are processed by a secondary pathway that is not subject to interference. We report here on the mitotic and meiotic characterization of *AtMus81*(At4g30870), which strongly suggests a role for this gene in the interference insensitive crossover pathway in *A. thaliana*.

Our lab has developed a unique pollen-based visual assay for meiotic recombination in *A. thaliana* that has facilitated these investigations [15,16]. This assay system is based on a series of transgenic lines, each carrying a gene encoding either a red, cyan, or yellow fluorescent marker protein excitable by different wavelengths of light. Transcription of these markers is directed by a post-meiotic pollen-specific promoter (LAT52) in the mutant *qrt1* background that produces tetrads of meiotically related pollen grains [17,18]. We constructed visually assayable genetic intervals by crossing lines carrying linked markers. Lines carrying two or more markers of different colors on the same chromosome produce tetrads that segregate the marker genes (and thus the proteins they encode) in patterns that reflect whether or not a recombination event has happened between them. Using this system, we can detect CO events directly in the gametes, and through the construction of double intervals delineated by three colors, we can assay CO interference. We used this system to assay the meiotic recombination phenotypes of the *Atmus81* and the *Atmsh4* mutants. We have also monitored production of fluorescent protein (or lack thereof) in homozygous constructs to quantify pollen viability.

The Methyl Methansulfonate UV Sensitive (*MUS81*) gene was originally identified in *S. cerevisiae* in a two-hybrid screen for protein products that interact with

the recombination factor Rad54p [19]. It was independently isolated in a screen for genes that are essential in *SGS1* and *TOP3* null backgrounds [20]. *Mus81* mutants show sensitivity to DNA damaging agents in yeast [19] and in mammals [21]. In *S. cerevisiae*, mutant alleles of *MUS81* reduce, but do not eliminate meiotic recombination. In these mutants, COs are reduced 1.1 to 1.8 fold and the residual COs are interference sensitive [8]. This is in contrast to other meiotic mutants such as *msh4*, in which the reduction in COs is greater, and the remaining COs are interference insensitive. This has led to the suggestion that in *S. cerevisiae*, there are two distinct classes of COs: those that require *MSH4/MSH5* and are interference sensitive, and those that that require *MUS81* and are interference insensitive [8]. Double mutants in both *msh4* and *mus81* in *S. cerevisiae* result in a severe reduction in COs, reinforcing the two-pathway model [22]. But, even this combination, which reduced COs by 13-15 fold, had some residual CO activity implying a possible third pathway in this organism [22].

Not all organisms have both interfering and non-interfering COs. In *Schizosaccharomyces pombe*, an organism that does not have CO interference, *MMS4/MUS81* mutants show a complete lack of COs [23]. Conversely, in the nematode *C. elegans*, an organism which shows complete interference, such that each homolog pair gets exactly one CO per meiosis, *MSH4* mutants show a complete lack of COs [24]. *D. melanogaster* presents another interesting case in which all COs are interference sensitive and *MUS81* mutants show no reduction in COs, but do show sensitivity to DNA damaging agents, suggesting a role confined to the mitotic cycle (J. Sekelsky personal communication).

In many organisms, Mus81p interacts with another protein to form a heterodimer, which is essential for its function. In *S cerevisiae* it is Mms4p, in *S. pombe* it is Eme1p and in *D. melanogaster*, the interacting partner is Mms4p [20]. *In vitro* studies using the fungal or human heterodimer have shown that this complex can cleave 3' flaps and collapsed replication forks [25]. Recent evidence from *S. pombe* suggests a role for Mus81-Eme1 in the resolution of single Holliday Junctions (HJs), which may be the primary recombination intermediates in this organism [26]. This finding is supported by the previous observation that, in *S. pombe*, expression of the bacterial HJ resolvase RusA can partially suppress the *mus81* mutant phenotype [23]. Interestingly, studies in *S. cerevisiae* show that the role of Mus81p as the essential HJ resolvase is not universal. In budding yeast, double (not single) Holliday Junctions (dHJs) may be the primary recombination intermediate, and expression of RusA failed to suppress the *mus81* meiotic phenotype. Physical analysis of the *S. cerevisiae mus81* deletion mutant is not consistent with a HJ resolution defect as dHJs are processed, however the kinetics are delayed about 2 hours. Notably, dHJ intermediates are reduced in the *mms4* mutant background [8].

S. cerevisiae mus81 mutants show synthetic lethality with *sgs1*, a helicase in the RecQ family. Intriguingly, this lethality is dependent on the *RAD52* DSB repair pathway [25]. A similar synthetic lethality has been demonstrated in *A. thaliana* where *Atmus81* mutants are also synthetically lethal with mutants of the *A. thaliana* homolog of *SGS1: AtRecQ4A*. *AtMus81* mutants were also shown to be sensitive to the DNA damaging agents MMC and MMS confirming a mitotic role for *AtMUS81* as well [27]. In this study we show that this sensitivity phenotype extends to other damaging agents and results in an increase in *AtMUS81* transcript levels.

A previous study found that *Atmus81* mutants do not have a detectable meiotic defect [27]. Using our visual assay for recombination, we have found that the *Atmus81* mutant has a moderate meiotic defect in the form of reduced COs. We also observed an elevation in the expression of the *AtMus81* transcript in tissue types undergoing meiosis and decreased pollen viability in the *Atmus81* mutant. Importantly, interference is stronger in the *Atmus81* mutant, an observation that is consistent with the role of *AtMus81* in the secondary interference insensitive CO pathway in *A. thaliana*.

RESULTS

Isolation of the *AtMus81* gene (*At4g30870*) and identification of a T-DNA insertion mutant

The *A. thaliana* homolog of *MUS81*, *AtMus81*(At4g30870), was identified using a BLAST search using *S. cerevisiae* and *S. pombe* sequences. The cDNA for this gene has been isolated and cloned by Hartung et al. and is 2303 bp [27]. This contains a 5' untranslated region of 146 bp, a 3' untranslated region of 177 bp, and an open reading frame of 1980 bp. The full-length cDNA consists of 15 exons and 14 introns [27]. Using RT-PCR with internal primers as well as 5'/3' RACE to amplify the ends, we confirmed the published sequence with one small difference. The sequence of the molecule obtained from our 5' RACE was 5 bp shorter than what is reported in Genbank (http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi). This could be explained by two alternate transcriptional start sites, a feature of the *MUS81* transcript that was recently observed in the *Oryza sativa* homolog, *OsMUS81* [28]. Apart from this 5 bp difference, a

composite molecule that matched the published sequence was constructed using our amplification products.

Identification of T-DNA insertion mutants in AtMUS81

The SALK laboratories SIGnAL database of T-DNA insertions contains two intronic T-DNA insertions within the open reading frame of *AtMUS81* (http://signal.salk.edu/cgi-bin/tdnaexpress) SALK_107515, and SALK_113F11 (Fig1a). For our analysis, we used the former, as we were able to show that this is a *bona fide* insertion using PCR with primers spanning the insert (Fig1a). We used PCR to genotype individuals as homozygous wild type, heterozygous for the insert, or homozygous for the insert. Under normal growth conditions, *Atmus81* mutants are visually indistinguishable from wild type. Sequence analysis of PCR products from the homozygous insertion line confirmed the T-DNA insertion junction reported in the SALK database. RT-PCR using primers spanning the insertion in lines homozygous for this insertion showed no product (data not shown). RT-PCR using primers downstream of the insertion showed that transcript levels were greatly reduced in lines homozygous for the insertion (Fig 1b).

***Atmus81* mutants are sensitive to a range of DNA damaging agents**

A feature of *mus81* mutants in other organisms is that they exhibit increased sensitivity to many DNA damaging agents [19,21,29]. To determine if the *Atmus81* insertion mutant has elevated sensitivity to DNA damage, we exposed seedlings to the radiomimetic methyl-methanesulfonate (MMS). We assayed the growth of both mutant and wild type individuals on a gradient of MMS concentrations from 0 to 75 ppm (fig

2a). Visual analysis showed that although both wild type and homozygous mutants became more sickly with increased concentration of MMS, the *Atmus81* mutants consistently died at 40 ppm while the wild type lines were much healthier at this concentration and survived at even the highest dosages we tested. These results are consistent with those reported by Hartung et al [27].

To confirm the DNA damage sensitivity phenotype and establish that it was not MMS specific, we conducted a similar assay using cisplatin, which is thought to form inner strand crosslinks and bulky adducts by binding to nitrogen atoms in DNA bases [30]. These adducts subsequently interfere with DNA replication, transcription, and repair [31]. Growth of wild type and *Atmus81* mutants on different concentrations of cisplatin showed that the mutants have an increased sensitivity (fig 2b). We also used gamma radiation as a non-chemical source of DNA damage. Consistent with the previous experiments, *Atmus81* mutants showed increased sensitivity to gamma radiation. Exposure of two-week old seedlings to approximately 100-125 grays disrupted the growth of the *Atmus81* mutant to an extent similar to the known radiation sensitive mutant *atm-2* [32] which we used as a positive control (fig 2c). These results demonstrate that the *Atmus81* mutant has increased sensitivity to a range of DNA damaging agents each with a different mode of action suggesting that *AtMUS81* is active during the somatic cycle and that it has a function in DNA repair.

AtMUS81 transcript levels increase in response to DNA damage

To examine whether the *AtMUS81* transcript was upregulated in response to MMS or gamma radiation treatment we used real-time qPCR. Levels of *AtMUS81*

transcript were not significantly different between untreated wild-type plants and plants harvested immediately after gamma radiation treatment (125 Gys), but were significantly different after letting the plants recover for 6 hours (fig 3a) indicating transcriptional upregulation in response to DNA damage induced by gamma radiation. To assay the levels of *AtMUS81* transcript in response to MMS, two-week old seedlings growing on normal growth media (MS plates) were transferred to liquid MS media containing 50 ppm MMS and allowed to incubate for twelve hours post transfer. Compared to seedlings transferred to liquid media containing no MMS for twelve hours, MMS treated individuals showed a significant increase in *AtMUS81* transcript levels. Fold increase over no treatment control was calculated using the $2^{-\Delta\Delta Ct}$ method [33]. The ubiquitously expressed elongation factor *EF1* was used as a reference gene for this analysis [34,35].

***AtMUS81* transcript levels are upregulated during meiosis**

We have hypothesized that *AtMUS81* is active during meiosis. To test this hypothesis, we used real-time quantitative PCR (qPCR) to assess the levels of *AtMUS81* transcript in various tissue types (fig3b). *AtMUS81* transcript levels were measured relative to transcript levels of the elongation factor *EF1*, which was used as a reference gene. *EF1* is expressed equally in all tissues and has been used for this purpose in other quantitative analyses of tissue specific transcript levels [34,36]. Roots, stems, leaves and inflorescences showed consistent moderate levels of expression. In contrast, the expression levels in anthers were considerably higher. Using the levels in the silique, which showed the lowest levels of *AtMUS81* transcript as a reference point set at 1.0, the transcript levels in the anthers were increased 9.68 +/- 0.33 fold.

Pollen viability is decreased in the *Atmus81* mutant

Other *A. thaliana* meiotic genes such as *RAD51*, *MND1*, *MEI1* and *SPO11* exhibit decreased pollen viability when disrupted, either as a result of chromosome fragmentation or segregation defects [37-39]. This is also true in many other organisms, as a failure to recombine at wild type levels often leads to gametic abnormalities. To measure pollen viability in the *Atmus81* mutant, we monitored production of dsRED, YFP, and CFP fluorescent proteins in the pollen tetrads that were homozygous for all three markers (Fig4a). Pollen grains were scored as non-viable when they were morphologically aberrant (small and misshaped) or when all three color markers failed to express (Fig4a). In our hands, this fluorescent marker based assay for pollen viability is more robust than staining procedures (e.g. flourescein diacetate, propidium idodide or Alexander's stain), producing more consistent results [40]. In contrast to other published work [27], we found that the *Atmus81* mutant shows a statistically significant decrease in overall pollen viability, 87% (788 tetrads) vs. 96% in wild type (853 tetrads; fig 4b). We also found that the *Atmsh4* mutant, which has been previously reported to have a pollen viability defect [12] does indeed exhibit high levels of pollen lethality with a preponderance of 2:2 viable:non-viable tetrads (data not shown), common for chromosome segregation defects and a hallmark of many meiotic mutants in *S. cerevisiae*.

AtMUS81 is essential for a subset of meiotic COs

The elevated *AtMUS81* transcript levels in meiotic tissues and the decreased pollen viability in the *Atmus81* mutant are both consistent with a role in meiotic recombination. To assess this role, we characterized the meiotic defects in both *Atmus81* and *Atmsh4* by measuring meiotic CO levels. To facilitate the use of tetrad analysis, we crossed mutant *Atmus81* and *Atmsh4* individuals into the *qrt* tetrad-producing background. *Atmus81*/*Atmus81*; *qrt1-2*/*qrt1-2* or *Atmsh4*/+; *qrt1-2*/*qrt1-2* plants were then crossed to lines that were homozygous for fluorescent marker genes flanking a genetic interval. Initially, intervals on chromosome 1 (6.7 cM) and chromosome 3 (23.7 cM) were used (fig 5). These intervals were previously described in Francis et al [15]. F1 progeny from these crosses that were heterozygous for *Atmus81* or *Atmsh4* and heterozygous for the fluorescent interval were allowed to self and marker/+ individuals were scored in the F2 generation for recombination. Recombination in these individuals was scored in the pollen tetrads before the genotypes of the individuals were determined (blind scoring). After all recombination data had been collected, the plants were PCR genotyped for the *Atmus81* and *Atmsh4* T-DNA insertions. The *Atmus81* and *Atmsh4* mutant alleles segregated in this generation such that the homozygous mutant: heterozygote: wild-type ratio was approximately 1:2:1, as expected.

The resulting collection of tetrad recombination data was analyzed using the Perkins mapping function of $X = 100[(1/2T + 3NPD)/n]$ [41]. The *Atmus81* mutants showed a 12% reduction in genetic distance in the interval on chromosome 1 (12,323 tetrads) and a 9% reduction in the interval on chromosome 3 (6,710 tetrads; fig 5a). The *Atmsh4* mutant showed a 71% reduction in genetic distance in the chromosome 3 interval. Interval 1 analysis was not conducted for the *Atmsh4* mutant.

The remaining COs of the *Atmus81* mutant are interference sensitive

The observation that CO levels are reduced in *Atmus81* plants is consistent with *AtMUS81* playing a role in an interference insensitive CO pathway [6]. This view would predict that interference would still be intact in the remaining COs in these mutants. To test this prediction we crossed *Atmus81*/*Atmus81*; *qrt1-2*/*qrt1-2* lines to a line with three linked inserts on chromosome 5 (also in the *qrt1-2* background), each encoding a different color fluorescent protein (fig5b). These inserts define two adjacent intervals, which could be simultaneously assayed for recombination. This enabled us to do a type of interference analysis in which we measured genetic distances with and without the presence of a simultaneous event in the neighboring interval, a type of analysis that has been applied to *S. cerevisiae* tetrad data [7,42]. The ratio of genetic distance with the presence of an adjacent CO relative to the distance when an adjacent CO is absent gives a value that is similar to but distinct from the coefficient of interference.

Within the smaller of the two adjacent intervals, we observed a genetic distance of 20.0 cM and 6.8 cM with and without an adjacent CO, respectively, in data pooled from wild type and *Atmus81*/+ heterozygotes (12,718 tetrads). In the *Atmus81* mutant, we observed distances of 20.0 cM and 5.9 cM, with and without an adjacent crossover, respectively (5,905 tetrads). A statistical comparison of wild type and *Atmus81* mutants was significantly different with a one-tailed P value of 0.032 when we compared ratios of genetic distances in the interval with and without an adjacent CO (Fig 6b). This is consistent with the hypothesis that *AtMUS81* is involved in a subset of COs that are interference insensitive.

DISCUSSION

In order to understand the role of *AtMUS81* in DNA damage repair, we examined the specificity of the defect using agents with different mechanisms of action for producing genotoxic damage. In a previous study by Hartung et al.l, *Atmus81* mutants were shown to be sensitive to MMC but were not sensitive to the double strand break inducing agent bleomycin [27]. Our results demonstrate sensitivity of *Atmus81* mutants to cisplatin, which creates bulky adducts as well as MMS and gamma radiation, each of which produce a wide range of DNA damage including double strand breaks (gamma radiation). This phenotype can be explained by a role for *AtMUS81* in repairing collapsed replication forks, a somatic repair role postulated for *MUS81* in *S. cerevisiae* [25], but it does not rule out a role in other kinds of damage repair.

Our result does demonstrate the non-redundancy of the *AtMUS81* gene (At4g30870) in *A. thaliana*. Build 6.0 of the *A. thaliana* genome in the NCBI database (http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi) includes another annotated *MUS81-like* gene present in the *A. thaliana* genome (*At5g39770*). This sequence is most likely a non-functional pseudogene as the putative protein has 3 in-frame stop codons in two different exons [27]. A likely promoter insertion line with a T-DNA insertion 554 bps (SALK_051926) upstream of the first predicted exon did not show sensitivity to MMS at any concentration compared to wild type (data not shown). Hartung et al. report no success in trying to amplify even partial cDNA from this predicted gene [27].

Looking at two different genetic intervals on chromosomes 1 and 3, we observed an approximately 10% decrease in meiotic homologous recombination in the *Atmus81*

mutant background. This result implies that AtMus81 has a meiotic role and is involved in the processing of a subset of meiotic COs. This meiotic role is consistent with the observed decrease in pollen viability. Our qPCR result showing that *AtMUS81* transcript levels are increased by over 9 fold in the anthers is also consistent with this view. Previous statistical modeling of interference in *A. thaliana* suggested that 15% of COs in *A. thaliana* should be free of interference and thus randomly distributed [6,14]. The modest reduction that we observed in these two intervals is consistent with our hypothesis that *AtMUS81* is a mediator of the secondary interference-free pathway. Interference measurements in *Atmus81* plants confirmed that not only were the remaining crossovers interference sensitive, but the overall levels of interference were significantly increased in the mutant background. This result suggests that like *S. cerevisiae*, *A. thaliana* has interfering COs coexisting with non-interfering COs. In *S. cerevisiae*, non-interfering *MUS81* mediated COs are thought to be ~30% of the total [8], while our result shows ~10% of *AtMUS81* mediated non-interfering COs in the two intervals we examined in *A. thaliana*.

Other methods of assaying recombination in *A. thaliana* have been developed, and one in particular involving a tandem disrupted GUS gene [43] was recently applied to study recombination in *Atmus81* mutants [27]. Hartung et al. used a GUS reporter construct to assay inter and intrachromosomal recombination and found no significant difference between wild type and *Atmus81* mutants. Interestingly, it was observed that mitotic recombination was increased in the *Atmus81* mutant after bleomycin treatment. These results are compatible with the meiotic defect observed in this study. Events producing a functional GUS construct arise either by unequal exchange or

intrachromosomal recombination and thus do not measure simple meiotic COs. Hartung et al. state that the COs postulated to exist (~15%) in the secondary recombination pathway [6,12] would be undetectable using their GUS reporter assay [27]. Our fluorescent system allows us to visualize true homologous meiotic COs and provides the opportunity to generate large data sets with relative ease [15]. These large data sets also make possible the ability to detect small differences between mutant and wild type lines.

A recent result from *S. pombe* suggests that in this organism, in which all COs are mediated by Mus81-Eme1, the predominant meiotic recombination intermediate is a single Holliday Junction (sHJ) [26]. This is in contrast to data from *S. cerevisiae* showing that the predominant meiotic recombination intermediate is a double Holliday Junction (dHJ) [44]. Intriguingly, these dHJs have been reported to coexist with a significant number of sHJs [45]. This leads to the possibility that the two pathways are biochemically distinct in that *MSH4/MSH5* mediated COs go through a dHJ intermediate while *MUS81* mediated COs go through a sHJ intermediate, an idea recently proposed by Cromie et al [26]. Direct detection of physical recombination intermediates using 2D electrophoresis or electron microscopy in *A. thaliana* may help determine the structure of recombination intermediates in this two pathway organism.

MATERIALS AND METHODS

Plant and Growth Conditions

The qrt1-2 (Columbia-3 background) line was kindly provided by Dr. Daphne Preuss (The University of Chicago). The SALK_107515 (*Atmus81*) and SALK_136296

(*Atmsh4*) lines were obtained from the Arabidopsis Biological Resource Center at The Ohio State University (http://www.arabidopsis.org/). Seeds were sown on either Pro-mix (Professional Horticulture) or Metromix-400 (Sun-Gro, Bellevue, WA) and stratified for 3-4 days at 4 C. In each experiment, the same soil type was used. Plants were germinated and grown under long day conditions (18h day 6h night) at 20 C. Temperatures were monitored with thermometers on the same shelves where the plants were grown. Plants used in the DNA sensitivity assays were surface sterilized in 10% bleach with 0.1% Triton-X and grown on either 24 well tissue culture plates (Fisher Scientific, Pittsburgh, PA) or on 60mm Petri dishes (Applied Scientific) in MS media with 5.0g Phyto agar/L (Research Products International, Mt. Prospect, IL). All parental strains are available from The Arabidopsis Biological Resource Center at Ohio State University (Columbus, OH). All fluorescent tag lines are available on request from G.P.C.

DNA extraction and PCR analyses

Genomic DNA was extracted from 2-3 cauline leaves as described by Copenhaver et al [46]. 30-cycle PCR was used to identify plants that were homozygous and heterozygous for the two T-DNA insertions, SALK_107515 (*Atmus81*) and SALK_136296 (*Atmsh4*). For the SALK_107515 line the wild type allele was amplified using SALK_107515F (see supplemental Table 1 for oligonucleotides used in this study) and SALK_107515R while the mutant allele was amplified using LBB1, a primer specific to the left border, and SALK_107515R. For the SALK_136296 line, the wild type allele was amplified using SALK_136296F and SALK_136296R while the mutant allele was amplified using LBB1 and SALK_136296R. The mutant allele products were

purified and sequenced to confirm their identity. The *atm-2* mutant was the same as characterized by Garcia et al. and homozyogous mutant plants were identified using the primers LBa1 and ATM126 [32].

Mutagen assays

MMS: wild type and homozygous *Atmus81* mutant were surface sterilized and plated directly in 24 well tissue culture plates (1 ml/well) containing solid MS media with the respective concentration of MMS (0-75 ppm). Photos were taken at 14 days after plating.

Cisplatin: wild type and homozygous Atmus81 mutant were surface sterilized and plated directly on 60 mm Petri dishes containing solid MS media (15ml/plate) with the respective concentration of cisplatin (0-25 ppm). Photos were taken at 12 days after plating.

Gamma Radiation: wild type, homozygous *atm-2* (positive control) [32], and homozygous *Atmus81* mutant were surface sterilized and plated directly on 60 mm Petri dishes containing solid MS media. At 7 days, these plates were placed in a Gammator Cesium-135 irradiator (Radiation Machinery Co., Piscataway, NJ) for times corresponding to the appropriate dosages. Photos were taken 13 days after removal (20 days total).

RNA isolation, AtMUS81 cDNA analysis, and real time qPCR

All RNA used in this study was isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. RT-PCR was performed using the Thermoscript cDNA synthesis kit (Invitrogen, Carlsbad, CA)

according to manufacturer's protocol using ~1 ug RNA and maximum incubation times. RNA used for confirmation of published cDNA sequence was isolated from young inflorescence tissue. 5' and 3' rapid amplification of cDNA ends (RACE) was performed using the Takara 3' and 5' RACE core kits using manufacturer's protocol (Takara, Madison, WI) using ~1 ug RNA and maximum incubation times. Primer sets used to confirm the published sequence of the cDNA were M81_ALB1F, M81_ALB2R, PS3 and PR3 [27]. The 3' RACE forward primer was M81S1_F and the 5' RACE primers used were M81RT/Phos, S1, S2, A1, and A2. Analysis of transcription levels in the *Atmus81* mutant was conducted using RT-PCR as above with the following primer sets: M81S1_F and M81S1_R, M81S2_F and M81S2_R, oMC571 and oMC572 [47] (APT1 control).

Real-time qPCR: to test *AtMUS81* transcript levels after gamma radiation, two-week old seedlings growing on 60mm solid MS media plates were placed in a Gammator as above and were allowed to receive a dosage of 125 Gys. RNA was isolated from these plants directly after removal from the radiation source as well as 6 hours after removal. Treated samples were compared to plants grown in the exact same manner that were not subject to radiation. To test *AtMUS81* transcript levels after MMS treatment, two-week old seedlings were transferred either to liquid MS media containing 50 ppm MMS (treated) or 0 ppm MMS (untreated control) and allowed to incubate for 12 hours post transfer. qPCR was conducted on these samples as above. Tissue specific transcription analysis was conducted using ~100 mg each of root, stem, leaf, young inflorescence, anther, and silique for RNA preparation.

First strand cDNA synthesis was conducted using 2 µl of RNA sample with Superscript III system according to the manufacturer's protocol (Invitrogen). A 40-cycle

real-time PCR reaction with optical reads after each cycle was performed on the Opticon Real-time thermal cycler (MJ Research, Reno, NV) with the SYBR Perfect real time qPCR premix (Takara, Madison, WI) according the manufacturer's protocol. The specificity of the reaction was determined using a melting curve from 60ºC to 95ºC with reads every 0.2ºC. Opticon 3 software (MJ Research) was used as the interface for execution and initial analysis. Reaction sizes were 25µl containing 2µl of undiluted cDNA. Specific primer sets for detection of *AtMUS81* transcript were M81RTRT3_F and M81RTRT3_R which produces a 189 bp fragment. The results obtained were standardized to the constitutive EF1A4α gene expression level [34], amplified with EF1F and EF1R which produces a 73 bp fragment. Efficiencies of the different primer sets were determined by dilution sets to be equal. The $2^{-\Delta\Delta Ct}$ method [33] was used to quantify fold increase of *AtMUS81* transcript. DNA damage assays were compared relative to no treatment controls and the tissue specific samples were compared to silique, which had the lowest levels of *AtMUS81* transcript and was set at 1.0 for the purpose of relative comparison. For all assays, samples were run in quadruplicate and standard error of the mean increase was calculated.

Microscopy

Segregation patterns of fluorescent alleles in pollen tetrads were measured by using a Nikon (Melville, NY) E1000 epifluorecence microscope equipped with filters from Chroma Technology (Rochester, NY). Pollen was collected by dipping flowers into a 10 µl drop of PGM media (34% sucrose, 4mM $CaCl_2$, 3.25mM boric acid, 0.1% Triton-X, pH 7.5) on a glass slide with cover slip. Pollen viability was assayed by monitoring

the production of dsRED, YFP, and CFP fluorescent proteins in pollen tetrads in the homozygous interval 5 background (FTL_1273; FTL_1659; FTL_993). All photographs were taken using Nikon coolpix5000 color digital camera. Figures were prepared using either Photoshop (Adobe, San Jose, CA) or Canvas (Deneba, Miami, FL).

Linkage and Statistical Analysis

See Supplemental Table 2 for a list of the color and chromosomal position of all fluorescent transgenic markers used in this study. To measure the genetic distance between any two transgenic markers, tetrads were designated parental ditype (PD), non-parental ditype (NPD), or tetratype (T), depending on the segregation of the marker pair. Map Distances were then calculated by using the Perkins formula: $X = 100[(1/2T + 3NPD)/n]$ [41]. Interference was measured by measuring CO frequencies in adjacent intercvals and dividing the tetrad data for one interval into groups based on the presence or absence of a CO in an adjacent reference interval [7]. If the genetic distance in the interval in question is significantly lower with the presence of a CO in the adjacent interval, we conclude interference extending from one interval to the other [7,42]. For a pair of adjacent intervals, we compared interference between data sets (wild type and mutant in this case). We calculated the ratio of genetic distances by dividing distance in an interval with an adjacent CO by the distance without a presence of an adjacent CO. These ratios were statistically compared by obtaining a Z score using the following equation:

$Z = |R1-R2|/\sqrt{[var (R1-R2)]}$ where R1 is the ratio in wt and R2 is the ratio in the mutant.

The significance of the difference between these two ratios was assessed using a one-tailed test as described on the Stahl Lab Online Tools (http://molbio.uoregon.edu/~fstahl).

ACKNOWLEDGEMENTS

We would like to thank Drs. Jeff Sekelsky and Frank Stahl for critical reading. We would also like to thank the Department of Energy (DE-FG02-05ER15651) and The National Science Foundation (MCB-0618691) for financial support. The authors declare no other competing interests.

FIGURE LEGENDS

Figure 4.1. The *AtMUS81* gene structure, T-DNA insertion mutant, and expression. (A) An illustration of the *AtMUS81* (At4g30870) locus showing the exon/intron organization of *AtMUS81*. Solid boxes represent transcribed regions including protein coding (black) and untranslated regions (gray). The T-DNA insertion site for the mutant used in this study is shown. Conserved domains are shown above. Below are products for the following primers which were used for genotyping (M81_F/M81_R), RT-PCR (M81S1_F/M81S1R, M81S2_F/M81S2_R) and real-time qPCR (M81RTRT3_F/M81RTRT_R). (B) Whole seedling (10-day) RT-PCR of wild type and the *Atmus81* mutant. Primers (S1 and S2) downstream of the T-DNA insertion site were used in the RT-PCR reaction with and without reverse transcriptase (RT) using RNA

from wild type (WT) and mutant (mus) plants. The APT1 transcript was used as a control.

Figure 4.2. Hypersensitivity of *Atmus81* mutants to MMS, cisplatin, and gamma radiation. (A) Wild type Col-0 (rows one and three) and mutant *Atmus81*/*Atmus81* (rows two and four) were subjected to a gradient (0-75 ppm) of MMS. The photograph was taken after 20 days. The wild type plants can grow at each concentration tested while the mutants cannot grow at >30 ppm MMS. (B) Wild type Col-0 (left of plate) and mutant *Atmus81*/*Atmus81* (right of plate) were subjected to various concentrations of cisplatin (0-15 ppm). The photograph was taken after 12 days. Wild type plants consistently outperformed the mutants at all concentrations tested. (C) Wild type Col-0 (upper left third), mutant *Atmus81*/*Atmus81* (upper right third), and gamma-hypersensitive mutant *atm-2*/*atm-2* (bottom third) were exposed to various levels of gamma radiation (0-150 Grays). At 75-100 Grays, the *Atmus81* mutants resembled the *atm-2* mutants rather than the wild type plants.

Figure 4.3. Real-time qPCR analysis of *AtMUS81* transcription. (A) RNA from untreated (NT) plants, plants harvested immediately after gamma radiation treatment (R0), plants harvested 6 hours after gamma radiation treatment (R6), and plants treated with 50 ppm MMS (MS) was used to measure the induction of the *AtMUS81* transcript. (B) Real time qPCR analysis of RNA from wild type root (RT), stem (ST), leaf (LF), inflorescence (IN), anthers (AN), and silique (SL) tissue as a measure of tissue-specific

AtMUS81 expression. The *EF1* gene was used as a control. Error bars are +/- the standard error of the mean calculated from four replicates.

Figure 4.4. Pollen viability in wt and *Atmus81* mutants. (A) Pollen tetrads from plants homozygous for three different fluorescent markers were examined. Pollen was classified as non-viable if grains were aborted (top row) or if all of the fluorescent proteins were not expressed (bottom row). (B) The *Atmus81* mutant has lower levels of pollen viability. Viability of wild type (open bars) is compared to Atmus81/Atmus81 plants (gray bars) and is also broken into tetrad categories (4:0, 3:1, 2:2, 1:3, 0:4; viable:non-viable). Error bars are +/- the standard error of the mean.

Figure 4.5. Meiotic recombination in the *Atmus81* mutant. (A) Characterization of the meiotic recombination phenotype using intervals on chromosomes 1 (left) and 3 (right). Green bars represent pooled data from wt and heterozygotes, orange bars represent *Atmus81* mutants, and blue bars represent *Atmsh4* mutants. Error bars are +/- the standard error of the mean. (B) Interference analysis of the *Atmus81* mutant using three linked markers on chromosome 5. Each pair of graphs (wt and heterozygotes top; *Atmus81* bottom) shows the genetic distances of an interval without and with an adjacent CO. The ratios of these genetic distances with adjacent CO: without a CO were significantly different between the pooled wt/heterozygotes and Atmus81 mutants with a one-tailed p value of 0.032 (see materials and methods for calculation of the p value).

TABLES

Table 4.1. Sequence information for oligonucleotides used

Primer Name	Sequence (5'-3')
SALK_107515_LP (M81_F)	gacagttgaaggtcgggaag
SALK_107515_RP (M81_R)	aattttccacaaaccctttgg
LBb1	gcgtggaccgcttgctgcaact
SALK_136296_LP	tggaatggatcaatgagttcc
SALK_136296_RP	aatcggtgagtcaggtttcag
oMC571 (APT1_F)	tcccagaatcgctaagattgcc
oMC572 (APT1_R)	cctttcccttaagctctg
LBa1	gcgtggaccgcttgctgcaact
ATM126	tctctccttgtttcaagctctgc
M81S1_F	acaaatattccacgcctccctcct
M81S1_R	tcctcagcacatcaaatccctcca
M81S2_F	tgctcggcataagtaccttgagac
M81S2_R	tgccggtttgctctggtcattatc
PS3	gacttggacactctaagag
PR3	aagcaagagacaaaagcgttg
M81_ALB1F	tcccaatcagcaaggaattt
M81_ALB2R	cgaacaggtgggattttgat
M81RT-Phos	-/5Phos/atcaataagctcttg
S1	ccgcaaagaaattcagtggg
S2	acagaaggactacaaatggg
A1	ttcccagcacgattatcagg
A2	cttgctcagaactttcaaca
M81RTRT3_F	tctctgtaatgtgccatctt
M81RTRT3_R	ttgaacctgatgaacatgga
EF1_F	ctggaggttttgaggctggtat
EF1_R	ccaagggtgaaagcaagaaga

Table 4.2. Fluorescent marker inserts used in this study

Name	Type	Chromosome position
FTL_992	CFP	I_9,850,022
FTL_1313	dsRED	I_11,130,549
FTL_1500	CFP	III_498,916
FTL_1371	dsRED	III_4,319513
FTL_1273	dsRED	V_18164269
FTL_1659	YFP	V_23080567
FTL_993	CFP	V_25731311

FIGURES

Figure 4.1.

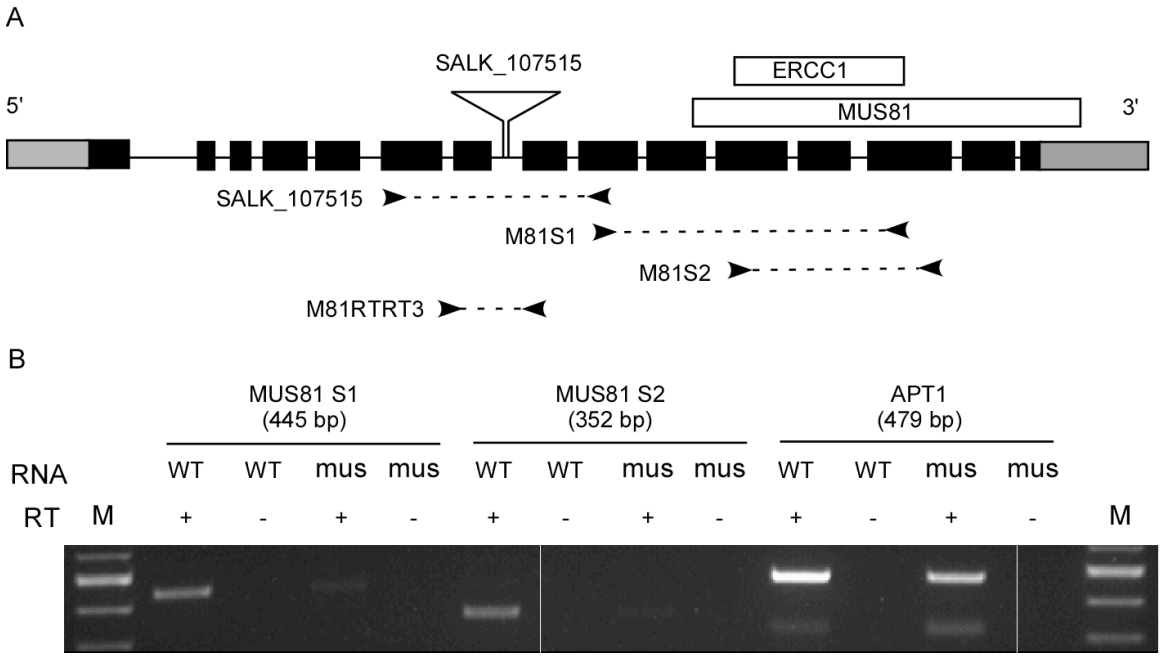


Figure 4.2.

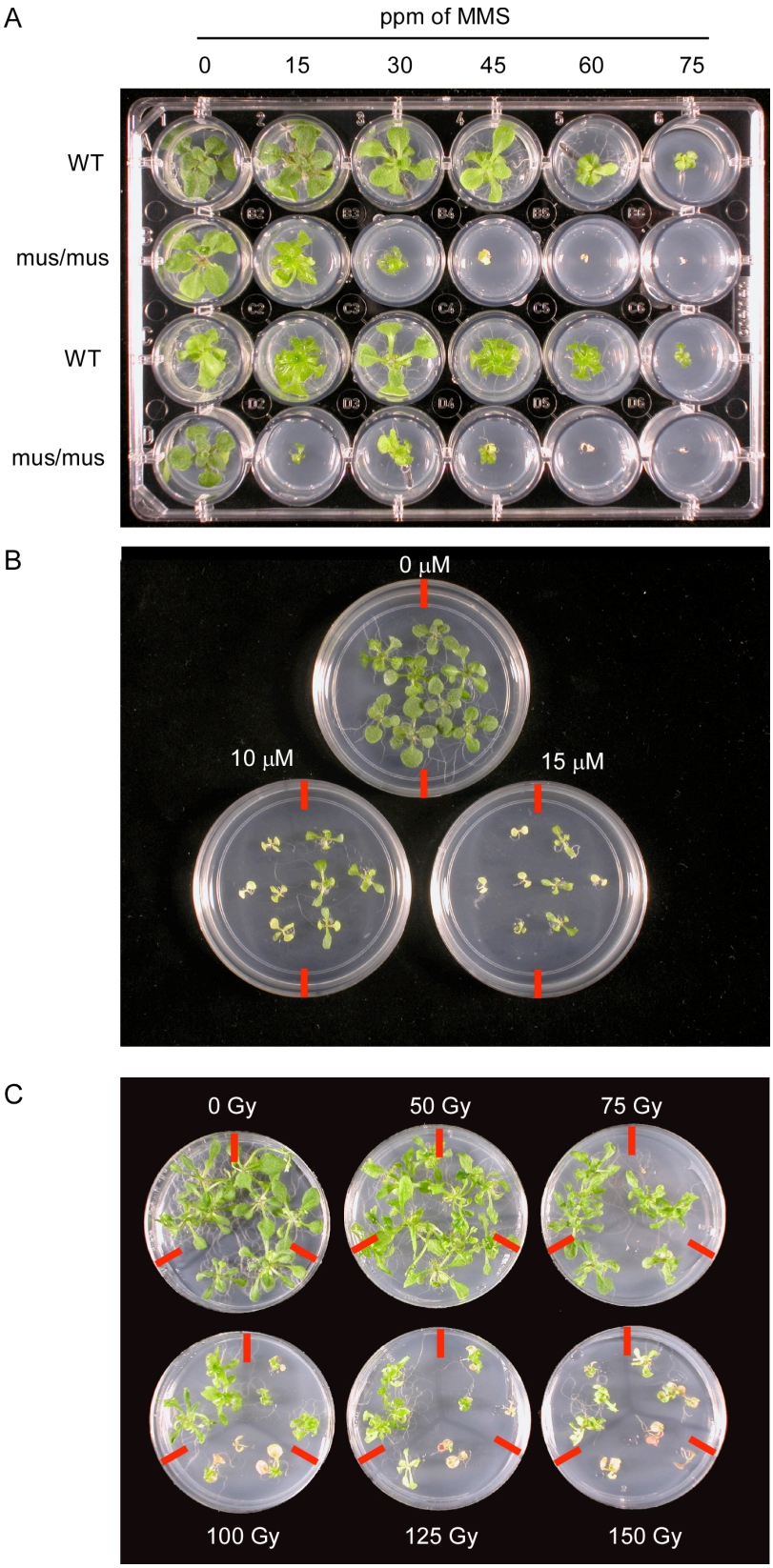

A
ppm of MMS
0
15
30
45
60
75
WT
mus/mus
WT
mus/mus
B
0 μM
10 μM
15 μM
C
0 Gy
50 Gy
75 Gy
100 Gy
125 Gy
150 Gy

Figure 4.3.

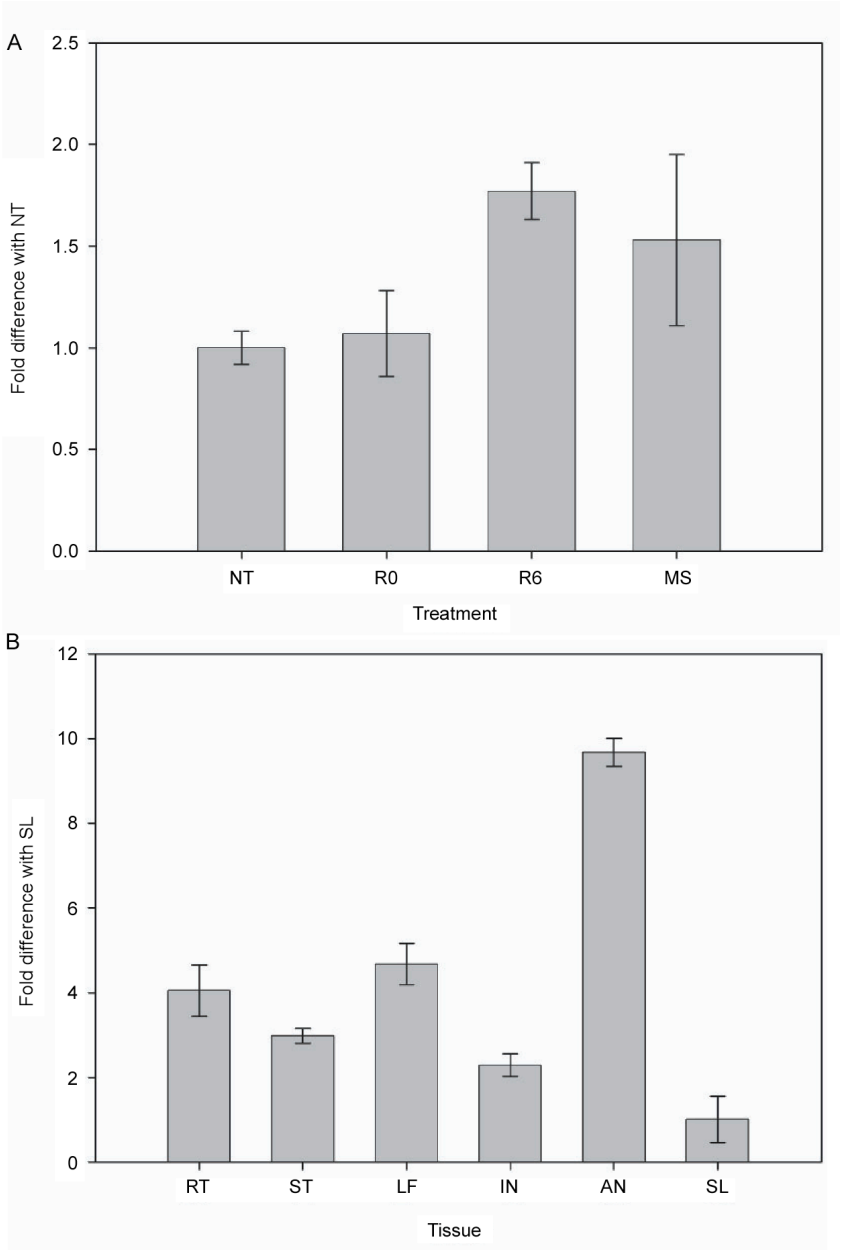

A
2.5
2.0
1.5
1.0
0.5
0.0
Fold difference with NT
NT
R0
R6
MS
Treatment
B
12
10
8
6
4
2
0
Fold difference with SL
RT
ST
LF
IN
AN
SL
Tissue

Figure 4.4.

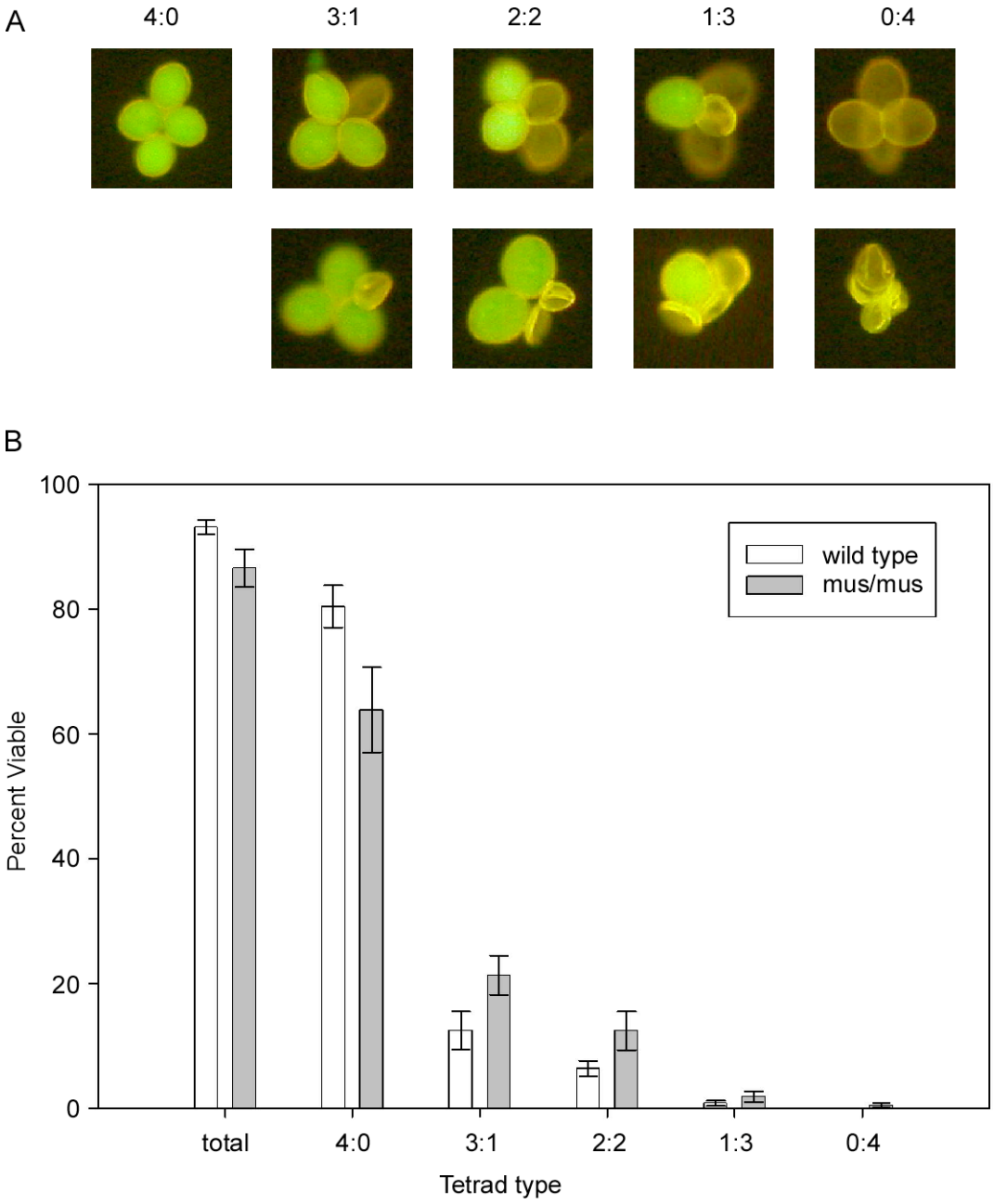
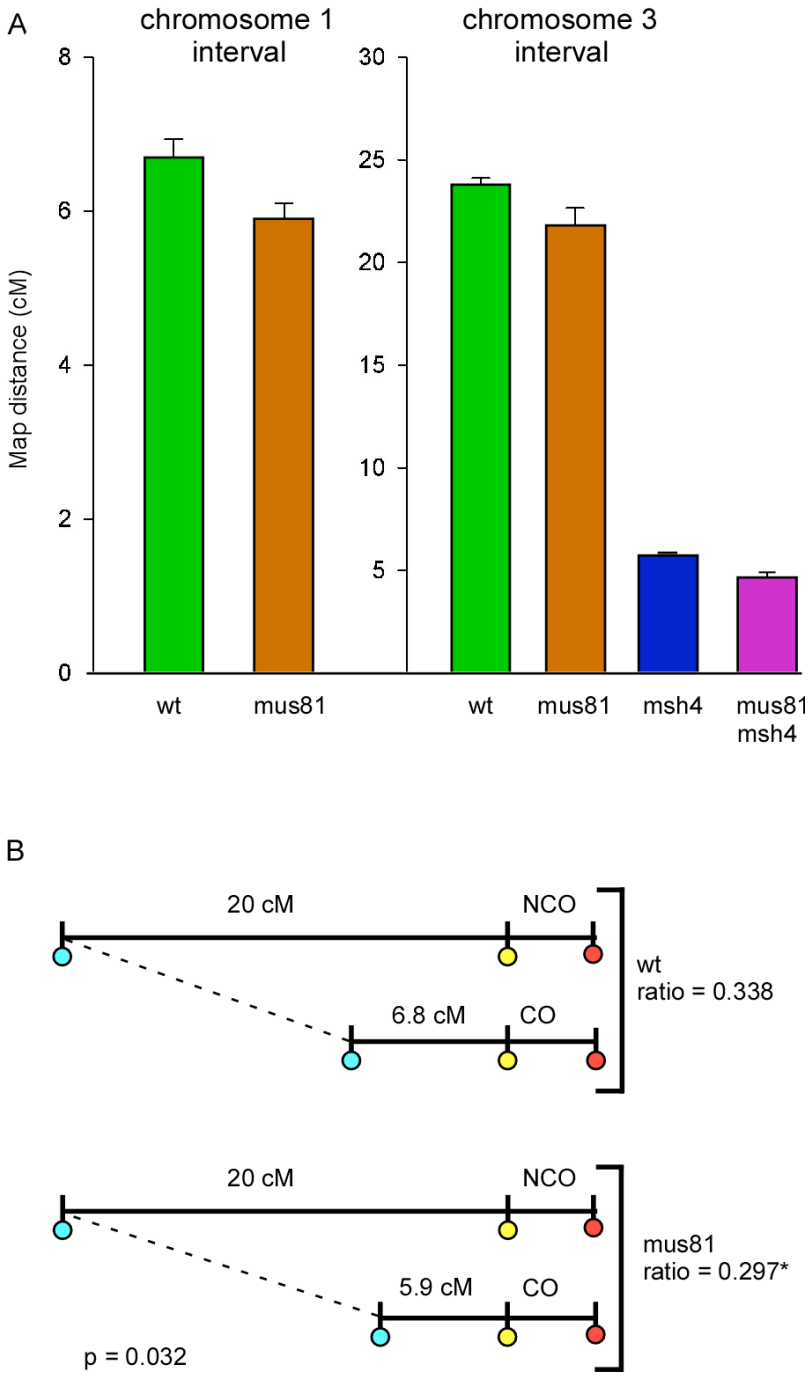

A
4:0
3:1
2:2
1:3
0:4
B
100
80
60
40
20
0
Percent Viable
wild type
mus/mus
total
4:0
3:1
2:2
1:3
0:4
Tetrad type

Figure 4.5.


A
chromosome 1 interval
chromosome 3 interval
Map distance (cM)
0
2
4
6
8
5
10
15
20
25
30
wt
mus81
wt
mus81
msh4
mus81 msh4
B
20 cM
NCO
6.8 cM
CO
wt
ratio = 0.338
20 cM
NCO
5.9 cM
CO
mus81
ratio = 0.297*
p = 0.032

REFERENCES

1. Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., Stahl, F.W. *The double-strand-break repair model for recombination*. Cell, 1983. **33**(1): p. 25-35.

2. Zickler, D., Kleckner, N. *The leptotene-zygotene transition of meiosis*. Annu Rev Genet, 1998. **32**: p. 619-697.

3. Sun, H., Treco, D., Szostak, J.W. *Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site*. Cell, 1991. **64**(6): p. 1155-1161.

4. Keeney, S., Giroux, C.N., Kleckner, N. *Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family*. Cell, 1997. **88** (3): p. 375-384.

5. Sturtevant, A.H. *The behavior of chromosomes as studied through linkage*. Z Induct Abstammungs-Vererbungs, 1915. **13**: 234-287.

6. Copenhaver, G.P., Housworth, E.A., Stahl, F.W. *Crossover interference in Arabidopsis*. Genetics, 2002. **160**(4): p. 1631-1639.

7. Malkova, A., Swanson, J., German, M., McCusker, J.H., Housworth, E.A., et al. *Gene conversion and crossing over along the 405-kb left arm of Saccharomyces cerevisiae chromosome VII*. Genetics, 2004. **168**(1): p. 49-63.

8. de los Santos, T., Hunter, N., Lee, C., Larkin, B., Loidl, J., et al. *The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast*. Genetics, 2003. **164**(1): p. 81-94.

9. Housworth, E.A., Stahl, F.W. *Crossover interference in humans*. Am J Hum Genet, 2003. **73**(1): p. 188-197.

10. Zhao, H., McPeek, M.S., Speed, T.P. *Statistical analysis of chromatid interference*. Genetics, 1995. **139**(2): p. 1057-1065.

11. Meneely, P.M., Farago, A.F., Kauffman, T.M. *Crossover distribution and high interference for both the X chromosome and an autosome during oogenesis and spermatogenesis in Caenorhabditis elegans*. Genetics, 2002. **162**(3): p. 1169-1177.

12. Higgins, J.D., Armstrong, S.J., Franklin, F.C., Jones, G.H. *The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in Arabidopsis*. Genes Dev, 2004. **18**(20): p. 2557-2570.

13. Chen, C., Zhang, W., Timofejeva, L., Gerardin, Y., Ma, H. *The Arabidopsis ROCK-N-ROLLERS gene encodes a homolog of the yeast ATP-dependent DNA helicase MER3 and is required for normal meiotic crossover formation*. Plant J, 2005. **43**(3): p. 321-334.

14. Lam, S.Y., Horn, S.R., Radford, S.J., Housworth, E.A., Stahl, F.W., et al. *Crossover interference on nucleolus organizing region-bearing chromosomes in Arabidopsis*. Genetics, 2005. **170**(2): p. 807-812.

15. Francis, K.E., Lam, S.Y., Harrison, B.D., Bey, A.L., Berchowitz, L.E., et al. *A pollen tetrad-based visual assay for meiotic recombination in Arabidopsis*. PNAS, 2007. **104**(10): p. 3913-3918.

16. Hawley, R.S. Meiosis in living color: *Fluorescence-based tetrad analysis in Arabidopsis*. PNAS, 2007. **104**(10): p. 3673-3674.

17. Preuss, D., Rhee, S.Y., Davis, R.W. *Tetrad analysis possible in Arabidopsis with mutation of the QUARTET (QRT) genes*. Science, 1994. **264**(5164): p. 1458-1460.

18. Twell, D., Yamaguchi, J., McCormick, S. *Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis*. Development, 1990. **109**(3): p. 705-713.

19. Interthal, H., Heyer, W.D. *MUS81 encodes a novel helix-hairpin-helix protein involved in the response to UV- and methylation-induced DNA damage in Saccharomyces cerevisiae*. Mol Gen Genet, 2000. **263**(5): p. 812-827.

20. Kaliraman, V., Mullen, J.R., Fricke, W.M., Bastin-Shanower, S.A., Brill, S.J. *Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease*. Genes Dev, 2001. **15**(20): p. 2730-2740.

21. Dendouga, N., Gao, H., Moechars, D., Janicot, M., Vialard, J., et al. *Disruption of murine Mus81 increases genomic instability and DNA damage sensitivity but does not promote tumorigenesis*. Mol Cell Biol, 2005. **25**(17): p. 7569-7579.

22. Argueso, J.L., Wanat, J., Gemici, Z., Alani, E. *Competing crossover pathways act during meiosis in Saccharomyces cerevisiae*. Genetics, 2004. **168**(4): p. 1805-1816.

23. Boddy, M.N., Gaillard, P.H., McDonald, W.H., Shanahan, P., Yates, J.R. 3rd, et al. *Mus81-Eme1 are essential components of a Holliday junction resolvase*. Cell, 2001. **107**(4): p. 537-548.

24. Zalevsky, J., MacQueen, A.J., Duffy, J.B., Kemphues, K.J., Villeneuve, A.M. *Crossing over during Caenorhabditis elegans meiosis requires a conserved MutS-based pathway that is partially dispensable in budding yeast*. Genetics, 1999. **153**(3): p. 1271-1283.

25. Bastin-Shanower, S.A., Fricke, W.M., Mullen, J.R., Brill, S.J. *The mechanism of Mus81-Mms4 cleavage site selection distinguishes it from the homologous endonuclease Rad1-Rad10*. Mol Cell Biol, 2003. **23**(10): p. 3487-3496.

26. Cromie, G.A., Hyppa, R.W., Taylor, A.F., Zakharyevich, K., Hunter, N., et al. *Single Holliday junctions are intermediates of meiotic recombination*. Cell, 2006. **127**(6): p. 1167-1178.

27. Hartung, F., Suer, S., Bergmann, T., Puchta, H. *The role of AtMUS81 in DNA repair and its genetic interaction with the helicase AtRecQ4A.* Nucleic Acids Res, 2006. **34**(16): p. 4438-4448.

28. Mimida, N., Kitamoto, H., Osakabe, K., Nakashima, M., Ito, Y., et al. *Two Alternatively Spliced Transcripts Generated from OsMUS81, a Rice Homologue of Yeast MUS81, Are Upregulated by DNA-Damaging Treatments*. Plant Cell Physiol, 2007. **48**(4): p. 648-654.

29. Boddy, M.N., Lopez-Girona, A., Shanahan, P., Interthal, H., Heyer, W.D., et al. *Damage tolerance protein Mus81 associates with the FHA1 domain of checkpoint kinase Cds1*. Mol Cell Biol, 2000. **20**(23): p. 8758-8766.

30. Heinen, E., Bassleer, R. *Mode of action of cis-dichloro-diammine platinum(II) on mouse Ehrlich ascites tumour cells*. Biochem Pharmacol, 1976. **25**(16): p. 1871-1875.

31. Nguyen, H.N., Sevin, B.U., Averette, H.E., Perras, J., Ramos, R., et al. *Cell cycle perturbations of platinum derivatives on two ovarian cancer cell lines*. Cancer Invest, 1993. **11**(3): p. 264-275.

32. Garcia, V., Bruchet, H., Camescasse, D., Granier, F., Bouchez, D., et al. *AtATM is essential for meiosis and the somatic response to DNA damage in plants*. Plant Cell, 2003. **15**(1): p. 119-132.

33. Livak, K.J., Schmittgen, T.D. *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-408.

34. Baud, S., Guyon, V., Kronenberger, J., Wuilleme, S., Miquel, M., et al. *Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in Arabidopsis*. Plant J, 2003. **33**(1): p. 75-86.

35. Liboz, T., Bardet, C., Le Van Thai, A., Axelos, M., Lescure, B. *The four members of the gene family encoding the Arabidopsis thaliana translation elongation factor EF-1 alpha are actively transcribed*. Plant Mol Biol, 1990. **14**(1): p. 107-110.

36. Osakabe, K., Abe, K., Yoshioka, T., Osakabe, Y., Todoriki, S., et al. *Isolation and characterization of the RAD54 gene from Arabidopsis thaliana*. Plant J, 2006. **48**(6): p. 827-842.

37. Li, W., Chen, C., Markmann-Mulisch, U., Timofejeva, L., Schmelzer, E., et al. *The Arabidopsis AtRAD51 gene is dispensable for vegetative development but required for meiosis*. Proc Natl Acad Sci, 2004. U S A **101**(29): p. 10596-10601.

38. Kerzendorfer, C., Vignard, J., Pedrosa-Harand, A., Siwiec, T., Akimcheva, S., et al. *The Arabidopsis thaliana MND1 homologue plays a key role in meiotic homologous pairing, synapsis and recombination*. J Cell Sci, 2006. **119**(Pt 12): p. 2486-2496.

39. Grelon, M., Gendrot, G., Vezon, D., Pelletier, G. *The Arabidopsis MEI1 gene encodes a protein with five BRCT domains that is involved in meiosis-specific DNA repair events independent of SPO11-induced DSBs*. Plant J. 2003. **35**(4): p. 465-475.

40. Shivanna, K.R., Rangaswamy, N.S. *Pollen biology: a laboratory manual*. 1992. Berlin; New York: Springer-Verlag. xi, 119 p.

41. Perkins, D.D. *Biochemical Mutants in the Smut Fungus Ustilago Maydis*. Genetics, 1949. **34**(5): p. 607-626.

42. Martini, E., Diaz, R.L., Hunter, N., Keeney, S. *Crossover homeostasis in yeast meiosis*. Cell, 2006. **126**(2): p. 285-295.

43. Molinier, J., Ramos, C., Fritsch, O., Hohn, B. *CENTRIN2 modulates homologous recombination and nucleotide excision repair in Arabidopsis*. Plant Cell, 2004. **16**(6): p. 1633-1643.

44. Schwacha, A., Kleckner, N. *Identification of double Holliday junctions as intermediates in meiotic recombination*. Cell, 1995. **83**(5): p. 783-791.

45. Bell, L.R., Byers, B. *Homologous association of chromosomal DNA during yeast meiosis*. Cold Spring Harb Symp Quant Biol, 1983. **47** (Pt 2): p. 829-840.

46. Copenhaver, G.P., Browne, W.E., Preuss, D. *Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads*. Proc Natl Acad Sci, 1998. U S A **95**(1): p. 247-252.

47. Wijeratne, A.J., Chen, C., Zhang, W., Timofejeva, L., Ma, H. *The Arabidopsis thaliana PARTING DANCERS gene encoding a novel protein is required for normal meiotic homologous recombination*. Mol Biol Cell, 2006. **17**(3): p. 1331-1343.

48. Pfaffl, M.W. *A new mathematical model for relative quantification in real-time RT-PCR*. Nucleic Acids Res, 2001. **29**(9): p. e45.

CHAPTER 4 ADDENDUM

Specific contributions by L.E.B.

I conceived, executed and analyzed the results of all experiments in this manuscript. Experimental design and data analysis were aided by G.P.C. Data collection was done by myself and A.L.B. I prepared the manuscript, which was reviewed and edited by G.P.C. and Jeff Sekelsky. K.E.F. helped develop the experimental systems used.

Chapter 5

A positive but complex association between meiotic double-strand break hotspots and open chromatin in *Saccharomyces cerevisiae*

Luke E. Berchowitz[1], Sean E. Hanlon[1,3], Jason D. Lieb[2,3], and Gregory P. Copenhaver[2,3]

Department of Biology and the Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280 USA

1: These authors contributed equally

2: Co-corresponding authors

3: Lineberger Comprehensive Cancer Center

ABSTRACT

During meiosis, chromatin undergoes extensive changes to facilitate recombination, homolog pairing and chromosome segregation. To investigate the relationship between chromatin organization and meiotic processes, we used Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) to map open chromatin during the transition from mitosis to meiosis in the budding yeast *Saccharomyces cerevisiae.* We found that meiosis-induced opening of chromatin is associated with meiotic DSB hotpots. The positive association between open chromatin and DSB hotspots is most prominent 3 hours into meiosis when the early meiotic genes *DMC1* and *HOP1* exhibit maximum transcription, and the early recombination genes *SPO11* and *RAD51* are strongly up-regulated. While the degree of chromatin openness is positively associated with the occurrence of recombination hotspots, many hotspots occur outside of open chromatin. Of particular interest, many DSB hotspots that fell outside of meiotic open chromatin nonetheless occurred in chromatin that had recently been open during mitotic growth. Finally, we find evidence for meiosis-specific opening of chromatin at the regions adjacent to boundaries of subtelomeric sequences, which exhibit specific crossover control patterns hypothesized to be regulated by chromatin.

Microarray data generated for the purpose of this manuscript is available via Gene Expression Omnibus (GEO) accession #GSE16163. Coordinates of all regions of open chromatin referred to in this manuscript are available in the supplement. There are two supplementary figures and five supplementary tables.

INTRODUCTION

During meiosis, chromosomes undergo a highly orchestrated series of movements and reorganizations including pairing, synapsis, recombination, and two successive rounds of segregation [1]. The end result is the basis for sexual reproduction: haploid gametes. Meiotic recombination contributes to haplotype diversity by generating new combinations of alleles not present in the parental chromosomes [2]. Crossing over results in physical linkages between homologous chromosomes, which are critical for proper chromosome segregation in many organisms [3]. In this manuscript, we explore how the location of recombination events, and their precursors double-strand breaks (DSBs), are associated with how DNA is packaged into chromatin.

In *Saccharomyces cerevisiae*, recombination occurs during prophase I and is initiated by programmed DSBs catalyzed by the topoisomerase-related protein Spo11 [4]. Single-stranded DNA (ssDNA) filaments coated with Rad51 and Dmc1 catalyze strand invasion of homologous duplex DNA, resulting in four-strand intermediates known as joint molecules [5-7]. After strand invasion, DSB repair proceeds and culminates in either crossover or non-crossover products [8-10]. Regions of gene conversion may also be produced. It has been estimated from ssDNA mapping and high-resolution recombination maps that 90 of the 140-170 DSBs that occur during each *S. cerevisiae* meiosis are repaired as crossovers [11, 12].

To a large degree, crossover distribution is governed by the initial placement of meiotic DSBs, which has been examined in *S. cerevisiae* using three different techniques. First, meiotic DSBs that accumulate in a *rad50S* background have been precisely mapped on chromosome III by Southern blot analysis [13]. The *rad50S* strain is useful for

studying DSB distribution because Spo11 remains covalently linked to the 5′ end of the DSBs, causing DSBs to accumulate instead of being converted rapidly into crossovers and non-crossovers [14]. However, *rad50S* mutants may have reduced DSB frequency and altered DSB placement relative to wild-type [12]. More recently, methods that do not require the *rad50S* background were used to map DSBs [12, 15]. As a normal part of DSB processing following cleavage by Spo11, the 5' ends of DSBs are resected, leaving 3' single-stranded tails that participate in the subsequent strand-invasion process. These 3' ssDNA tails can be used as a molecular tag to map DSBs. ssDNA was isolated from a *dmc1Δ* mutant, in which DSBs accumulate prior to arrest in late meiotic prophase [7]. Together, these studies provide an initial map of the DSB landscape and clearly demonstrate that DSBs occur preferentially in intergenic regions (particularly in transcriptionally active promoters) while also showing that chromosome ends and centromeres exhibit substantial DSB activity [12-15].

In most organisms, neither DSBs nor crossovers occur with uniform frequency throughout the genome. Some regions, called hotspots, have very high rates of recombination [16, 17]. Conversely, coldspots are regions of the genome with lower than average rates of recombination [17]. The kinetics and location of meiotic recombination are likely to be influenced by chromatin organization [18-20]. The fundamental unit of chromatin, the nucleosome, is defined as approximately 146 base pairs (bp) of DNA wrapped around a histone octamer [21]. Modulation of nucleosome occupancy provides the cell a mechanism to regulate access to DNA. Regions of DNA that are depleted of nucleosomes are known as “open” chromatin and are more accessible to DNA-binding proteins [22, 23]

During *S. cerevisiae* meiosis, hotspots at *ARG4* and *CYS3* show a meiosis-specific increase in sensitivity to micrococcal nuclease (MNase), a common assay for open chromatin [24]. Additionally, recombination rates can be artificially elevated at specific loci by inserting sequences that exclude nucleosomes, such as bacterial DNA or *S. cerevisiae* telomere fragments [20, 25, 26]. Recently, the distribution of a histone modification associated with transcriptionally active chromatin, H3 histones trimethylated at lysine 4 (H3K4me3), has been shown to be enriched at DSB sites [27].

Further evidence for a chromatin role comes from *S. pombe*. At the *S. pombe ade6-M26* locus, MNase sensitivity increases prior to DSB initiation, and nuclease sensitivity patterns at *ade6-M26* change during meiosis [28, 29]. A study that examined three distinct *S. pombe* DSB hotspots showed one that was constitutively open (*mbs1*), while two others become open only as meiosis proceeds ($tdh1^{+}$, *ade6-M26*) [29]. Mutations in *S. pombe* chromatin remodeling proteins such as Snf22, Gcn5, and Ada2 disrupt meiotic MNase cleavage patterns at *ade6-M26* and also result in decreased recombination at this site [28]. *snf22Δ* mutants fail to exhibit the ~2.4 fold meiotic induction of MNase sensitivity at the *ade6-M26* break point seen in wild-type cells and have a concurrent ~6.2 fold site-specific reduction in recombination at *ade6-M26* [28].

While studies at selected individual loci suggest that DSBs and crossovers occur preferentially at nucleosome-depleted regions [14, 18, 24, 30-32], whether this relationship would hold in an unbiased survey of nucleosome-depleted regions across the genome was unknown. To address the relationship between nucleosome occupancy and meiotic recombination directly, we used Formaldehyde Assisted Isolation of Regulatory Elements (FAIRE) to assess genome-wide associations between DSB hotspots and open

chromatin in yeast. We also assessed whether the distinct regulation of recombination at subtelomeric borders and pericentromeric regions is accompanied by particular chromatin configurations in meiosis.

RESULTS

Monitoring of meiotic entry and progression

To validate meiotic entry, synchrony, and progression in our time course, we analyzed the transcriptional profile of several key meiotic genes by expression microarray (**Figure 1**). We analyzed RNA isolated in a meiotic time course from each FAIRE biological replicate. This included the timepoints used in the FAIRE assay (YPD, 0 hr, 1.5 hr and 3 hr), in addition to RNA isolated from the same cultures at other timepoints. These include a timepoint prior to our "time 0", consisting of pre-meiotic cells that had been starved of a fermentable carbon source for 8 hours (-8hr acetate) and RNA collections at 4.5, 6, 9 and 12 hours, which were taken subsequent to those analyzed by FAIRE. To assess meiotic entry, we monitored the master regulators of the meiotic program *IME1* and *IME2*, which were upregulated at the 0 hr sample relative to YPD and the -8hr acetate sample. Both *IME1* and *IME2* exhibited maximum expression 1.5 hours into meiosis (**Figure 1A**). We also monitored the early meiotic genes *HOP1*, *SPO11*, *RAD51* and *DMC1*. All early genes exhibited strong up-regulation 1.5 hours into meiosis and generally were maximally expressed by 3 hours.

Based on these profiles, we infer that DSB formation and strand invasion had begun in our 3-hour sample. Although transcription analysis is an indirect means of determining the kinetics of DSB and joint molecule formation, our timing inferences

match closely with those determined from 2D gel analysis [6]. Additionally, our averaged expression profiles are strikingly similar to previous meiotic expression array analysis [33] (**Figure 1B**). Finally, by counting spores at each time point and 24 hours past sporulation, we show that over three-quarters of cells in each replicate successfully completed meiosis (**Figure 1C**).

Identification of meiotic open chromatin using FAIRE

We employed FAIRE to study changes in chromatin during the transition from mitotic growth through the early stages of meiosis. FAIRE is a simple and inexpensive method for identifying sites of open chromatin that relies on phenol-chloroform extraction of sonicated, formaldehyde-crosslinked chromatin to isolate nucleosome-depleted regions of DNA [34-36]. FAIRE has been used to positively select for nucleosome-depleted genomic sites throughout the mitotic cycle of *S. cerevisiae* [34] and to isolate nucleosome-depleted active regulatory sites in human cells [36, 37]. FAIRE has a strong anti-correlation with histone H3 and H4 ChIP-chip in mitotic yeast cells [34].

We targeted the first 3 hours of meiosis, which corresponds to the time DSBs are made and joint molecules are beginning to form [6]. Four sample treatments were subjected to FAIRE: cells undergoing normal mitotic growth (YPD), pre-meiotic cells that had been starved of a fermentable carbon source for 16 hours (YPA, 0hr), synchronized cells that had been incubated in sporulation media for 90 minutes (1.5hr) and cells that had been incubated in sporulation media for 180 minutes (3hr). Hereafter, we define a site of "open chromatin" as a FAIRE peak called at $P < 0.01$ (see **Methods**). We identified 1204 sites that were open in at least one timepoint (**Table 1 and Figure 2**)

and 510 sites that were open in all samples (e.g. *CDC19,* **Figure 3**). To further establish a firm relationship between FAIRE and nucleosome depletion in our samples, we compared our FAIRE enrichment profiles to H3 ChIP-chip enrichment profiles generated in an independent study from 0hr, 1hr and 2hr meiotic samples [27]. The H3 ChIP-chip data independently corroborated FAIRE enrichment profiles at individual loci (**Figure 3)** and globally (overall Pearson correlation -0.381).

Genomic regions that undergo high rates of meiotic recombination are generally associated with open chromatin

Evidence from locus-specific studies strongly suggests that meiotic recombination occurs preferentially in regions of open chromatin [18, 24, 30, 31]. To examine this relationship genome-wide, we compared our chromatin profiles to sites of meiotic ssDNA enrichment (DSB hotspots) from two published studies [12, 15] (**Figure 2, Figure 3**). Of the 1157 [12] and 258 [15] defined DSB hotspots, 591 (51%) and 157 (61%) respectively fall within a site of open chromatin from any of the samples (27% and 28% were expected by chance, respectively). To provide a more quantitative measure of overlap, we calculated the percentage of base pairs (bp) in sites of open chromatin that intersected with DSB hotspots at each point in the time course (YPD, 0hr, 1.5hr, and 3hr). Strength of association is expressed as the ratio of observed to expected bp overlap, where the "expected" value is the average overlap measured in 10 permutations in which the genomic position of the open chromatin domains was randomized (**Table 1, Methods**). By this measure, open chromatin was strongly associated with both DSB hotspots ($P << 0.0001$) (**Methods**) and in no case were permuted open chromatin sites as strongly

associated with hotspots as an experimental sample (**Methods**). This was true whether we examined sites that were open in at least one timepoint, or sites that were open in all timepoints.

In the 3 hr sample, when we infer through upregulation of *SPO11*, *DMC1* and *RAD51* and previous studies [6] that cells were processing DSBs into joint molecules, open chromatin had the strongest association with DSB hotspots: 2.5 [12] and 3.3 [15] times more highly associated than random. Sites that were open in all samples had an equally strong association with DSB hotspots, being 2.5 [12] and 3.5 [15] times more highly associated than random, implying that many hotspots exist in sites of constitutively open chromatin. Supporting this interpretation, meiosis-specific sites of open chromatin did not show a stronger association with DSB hotspots than sites that were constitutively open.

The association of open chromatin with DSB hotspots cannot be explained solely by the prevalence of open chromatin at gene promoters

Intergenic regions in yeast are typically depleted of nucleosomes [38, 39], and DSB hotspots are known to be associated with intergenic regions [12, 15-17]. To address the possibility that the association between open chromatin sites and DSB hotspots might be accounted for entirely by their co-occurrence in intergenic regions, we permuted the genomic location of sites of open chromatin such that the permuted location was restricted to intergenic regions. Even under these conditions, open chromatin identified by FAIRE was more highly associated with DSB hotspots than any of the permutations (**Methods and Table S1**). Thus, the open chromatin sites identified by FAIRE provide

information regarding hotspot distribution that is independent of, and in addition to, the information provided by stratifying the genome into coding and non-coding regions.

Among open chromatin sites, the degree of chromatin openness is predictive of DSB hotspots

We asked whether FAIRE sites with higher signal, or "openness" were more likely to be DSB hotspots than FAIRE sites with lower signal. To this end, we conducted ROC (Receiver-Operator Characteristic) curve analysis [40] (**Methods**) on each timepoint, using either regional *P*-values or probe *z* scores to quantify openness (**Figure 4**). "True positives" were defined as DSB hotspots identified in the literature [12, 15]. Five categories of open chromatin sites were analyzed: (I) open in any of our samples; (II) open in all of our samples; (III) open in at least one meiotic sample (1.5 hr or 3 hr) but not YPD; (IV) open in all meiotic samples (1.5 hr and 3 hr) but not YPD; and (V) open in mitotic growth but closed in all meiotic samples.

The degree of openness (as measured by *P*-value) in meiosis-specific open sites (categories III and IV) was strongly predictive of DSB hotspots, with ROC AUC (area under the curve) values of 0.712 (category III) and 0.734 (category IV) (**Figure 4A**). Degree of openness was also predictive of DSB hotspots in sites that were open in any sample (category I; AUC = 0.700) and sites that were open in all samples (category II; AUC = 0.687). Of the individual timepoints, openness was most predictive of DSB hotspots in the 1.5 hr timepoint (AUC = 0.718; **Figure 4B**). Conversely, the degree of openness in sites that are mitosis-specific (category V) was not predictive of DSB hotspots (AUC = 0.466).

We also quantified the ability of FAIRE to predict DSB hotspots independently of peak calling. Using the FAIRE enrichment *z* score of each probe as an ordering statistic, we plotted the True Positive Fraction versus the False Positive Fraction, and determined the likelihood ratio (TPF/FPF) at decreasing *z* score cutoffs at each timepoint (**Figure 4C**) [40]. The degree of FAIRE enrichment as measured by probe values is indeed predictive of DSB hotspots, with the 1.5 hr timepoint again being the most predictive. A given probe measuring FAIRE enrichment at the 1.5 hr timepoint with a *z* score of over 3 has a 5:1 true positive to false positive ratio with respect to DSB hotspot prediction.

A class of DSB hotspots occurs within closed chromatin as defined by FAIRE

Over one third of DSB hotspots, including some well-studied recombination hotspots such as *ARG4* [18, 24], *HIS2* [41], *BUD23* (also known as YCR047C) [18, 42] and *CYS3* [24] occur within chromatin that is not open, as measured by FAIRE, in any sample (*ARG4*, **Figure 3)**. This suggests that neither meiosis-specific opening nor a level of open chromatin beyond what is found intrinsically at promoter regions is strictly required for the establishment of a DSB hotspot. We tested the possibility that the subset of DSB hotspots that do not occur in open chromatin actually occur in relatively open chromatin that falls just below our classification threshold. We divided the 296 consensus DSB hotspots that were held in common between the two datasets into 167 that occurred in a site of open chromatin and 129 that did not. In each grouping, we aligned the hotspots at their start coordinates and calculated the moving average of FAIRE enrichment at each position (**Figure S1, Methods**). Hotspots associated with a site of open chromatin exhibited maximum openness at the center of the defined hotspot region. In contrast,

hotspots not associated with open chromatin exhibit a chromatin profile indistinguishable from the genomic average. We interpret this finding to mean that the DSBs that occur in sites we classify as closed, are in most cases truly closed according to FAIRE.

FAIRE failed to detect open chromatin at *ARG4*, which has previously been shown to be nuclease hypersensitive by independent groups [18, 24, 43]. This raises the possibility that FAIRE may be unable to identify a subset of open chromatin sites. It is possible that large non-histone molecular complexes prevent otherwise open sites to be isolated by FAIRE. Recent reports have shown that Spo11 binding sites do not necessarily incur DSBs [44] and even the strongest DSB hotspots in *S. cerevisiae* undergo DSBs in fewer than 10% of DNA molecules [13]. Therefore, some hotspots may be highly occupied by recombination complexes regardless of whether they actually incur a DSB. Additionally, DSBs may be required for a transition to open chromatin at some loci. Lastly, DSBs that occur in a small percentage of cells may be detected by nuclease sensitivity assays, which are likely more sensitive in this context, but not FAIRE. However, the assertion that some hotspots occur in closed chromatin sites is independently supported by H3 ChIP-chip data, because hotspots not associated with FAIRE enrichment also exhibited positive enrichment by H3 ChIP-chip [27] (**Figure 2B, Figure S1**).

Many DSBs occur in chromatin that had been recently open during mitotic growth

We next asked if some of the DSB hotspots that occur in closed meiotic chromatin were open during mitotic growth. Indeed, 62 of the 156 sites open during YPD and closed in all of the meiotic samples corresponded to published DSB hotspots [12, 15] (expected =

34.69; $P < 0.0001$). This is unlikely to be due to a threshold effect, whereby the mitosis-specific sites of open chromatin barely miss the meiotic cutoff. The mean FAIRE enrichment z-score of these sites in YPD (2.13) is significantly higher than any of the meiotic timepoints (average of all meiotic samples = 0.96; two tailed t-test $P < 0.0001$; **Figure S2**). While the fact that a locus was open during mitosis was predictive of a hotspot, the degree of mitotic openness was not, in contrast to meiotic openess (**Figure 4A**).

The prevalence of DSB occurrence in recently open chromatin could be due to hotspots specified by pre-meiotic events, such as transcription factor binding or histone modification patterns. For example, the presence of histone modification H3K4me3 is strongly elevated 0-2 kb from DSB sites and this pattern is established in vegetative cells [27]. Additionally, some hotspots are known to require transcription factor binding to stimulate meiotic recombination. A particularly well-studied example is the *HIS4* locus, which requires the binding of transcription factors Rap1, Bas1, or Bas2 for hotspot activity although it is unknown if binding is required during meiosis or mitosis [45-47]. We find the *HIS4* hotspot to be open in mitosis (max *z* score 5) and closed by the 1.5hr timepoint of meiosis. Independent H3 ChIP-chip experiments show the *HIS4* region to be closed from 0 hr to 2 hr in meiosis [27] (*HIS4*, **Figure 3**).

Chromatin adjacent to subtelomeric sequences is open specifically during meiosis

Telomeres play a critical role in the meiotic process. Telomere clustering and attachment to the nuclear membrane has been postulated to be involved in homologous chromosome pairing [48, 49]. However, it has been proposed that since telomeres and subtelomeres are

very high in repetitive and non-unique DNA [50], a homology-based pairing search using these regions would be difficult. Concordantly, in *S. cerevisiae*, crossovers are suppressed at telomeres and heterochromatic subtelomeres, presumably to suppress non-homologous exchanges between the repetitive and non-unique sequences that occur at chromosome ends [50-52]. In contrast, euchromatic sequences directly proximal to the subtelomeres, which could play a role in the homology search (see **Discussion**), have crossover rates that are greater than twice the genome average [51]. Based on this and the proposed role of telomeres for homologous chromosome pairing, we hypothesized that the euchromatic DNA sequences adjacent to the subtelomeres would exhibit meiosis-specific changes in chromatin architecture.

We found significantly elevated levels of meiosis-specific open chromatin proximal to the previously defined [51] subtelomeric-euchromatic borders (defined in **Figure 5A** and legend**, Methods**, **Table S3, Table S4**). Sites open in all meiotic samples, but not open in YPD (Category IV) were enriched most dramatically, occurring 2.8 times more frequently than expected 0-10 kb proximal to the subtelomeric border ($P < 0.001$), 2.1 times greater 10-20 kb from the border ($P < 0.05$), and 2.1 times greater 30-40 kb from the border ($P < 0.05$) (**Figure 5B, Table S3**). Expected values were derived from 10,000 permutations of the open chromatin chromosomal positions, and *P* values were calculated using a X^2 test (**Methods; Table S3**). In DNA located 20-30 kb proximal from the subtelomeric border, the distribution of open chromatin at any of the timepoints was indistinguishable from the average of permuted distributions. DSB hotspots are under-represented from the telomere to 10 kb proximal to the subtelomeric border, and are enriched 10-40 kb proximal to the subtelomeric border (**Figure 5C**) [12, 15].

Interestingly, recombination rate peaks 0-10 kb proximal to the subtelomeric border, declines, and then climbs again at 30-40 kb from the border (**Figure 5D**). Since category IV FAIRE data displays a similar pattern of enrichment, we speculate that meiosis-specific open chromatin may play a role in the elevated rates of recombination in these regions (**Discussion**).

Constitutive open chromatin is found at the centromeres and pericentromeric regions

Crossovers are repressed at *S. cerevisiae* centromeres and pericentromeric regions [53, 54], and exchanges at these regions are implicated in elevated frequencies of nondisjunction at meiosis II [55, 56]. However, in some studies, regions around centromeres have normal levels of meiotic DSBs [15]. This has led to the suggestion that DSBs near centromeres might be repaired by the sister chromatid [54], and that centromere-associated proteins such as cohesins and Sgo1 could create a chromatin environment that favors DSB repair from the sister. We speculated that repression of crossovers and non-crossovers near the centromeres is aided by a dense chromatin configuration present during meiosis, which predicts that these regions will not be enriched by FAIRE.

We analyzed the centromeres and associated regions 30 kb to the left and right of each centromere in 10 kb windows, using the same categorical groupings as above (I-IV, **Figure 6, Table S3**). Contrary to our hypothesis, regions spanning centromeres were often open (12/16 centromeres, 1.29 expected $P < 0.0001$). Ten of the sixteen centromeres were constitutively open ($P < 0.0001$), and none of the open sites that

occurred within a centromere were meiosis-specific. The pericentromeric sequences (defined as the DNA extending 10 kb on either side of the centromeres) were also constitutively open, containing more open chromatin than expected ($P < 0.05$ to the left; $P < 0.01$ to the right). Again, these sites of open chromatin were not specific to meiosis. This result does not support the hypothesis that dense chromatin near the centromeres plays a role in the suppression of crossovers and non-crossovers in the pericentromeric regions. Indeed, it raises the possibility that open chromatin somehow plays a role in this suppression.

A 15 bp DNA sequence motif is overrepresented in meiosis-specific open chromatin

We examined whether sites of open chromatin might be specified by one or more DNA sequence motifs. When the open chromatin sites isolated either in any one of the conditions or all conditions were used as an input for the motif-finding algorithm BioProspector, no strong motif was identified. However, when sites of open chromatin specific for meiosis (category III or IV) were used as an input, we identified a 15 bp DNA sequence motif: SSGGTTCGANYCCSS (**Figure 7**). This motif is similar to the B-Box motif involved in regulation of tRNA expression [57]. Sites of open chromatin containing the B-Box related motif tended to be centered downstream rather than upstream of tRNA genes. Of the 29 meiotic open sites adjacent to a tRNA gene and containing a B-Box related motif, 19 are downstream, 8 straddle, and only 2 are upstream of the tRNA gene. Of the 55 sites that meet the same criteria but are open in any one meiosis timepoint, the numbers are 34, 19, and 2 respectively. Therefore, there is a strong

bias for the B-Box related motif to occur downstream of tRNA genes that are located in meiotic open chromatin.

Based on these results, we asked if tRNA genes were associated with meiosis-specific open chromatin. We examined the chromatin profiles at all 275 cytoplasmic tRNA genes in the *S. cerevisiae* genome (**Figure 7**). 98 tRNA genes are associated with a site of open chromatin and 177 are not. Of these 98 tRNA genes, 65 were not open in YPD. Meiosis-specific open chromatin sites (Category III) were associated with tRNA genes far more than any other gene class, and over 25% of category III sites were associated with a tRNA gene, significantly greater than the expected proportion. However, meiosis-specific open chromatin at tRNA genes is unlikely to have a connection to recombination because DSB hotspots are not strongly associated with either tRNA coding sequences themselves or sequences 500 bp upstream and downstream of tRNAs (**Table S5**). The enrichment of open chromatin downstream of tRNA genes suggests a mechanism other than transcriptional clearing of a tRNA gene promoter may be establishing these meiosis-specific chromatin changes.

SUPPLEMENTARY RESULTS

Open chromatin within coding regions is associated with DSB hotspots to the same degree as open chromatin located elsewhere

Although hotspots are traditionally associated with promoters, they also occur within coding regions [11]. Even among hotspots that occur in intergenic DNA, most, as defined by ssDNA, extend into the coding regions. However, it is important to note that since > 1 kb of ssDNA can be generated on either side of a DSB, the occurrence of a DSB hotspot

that overlaps a coding region does not mean that initiating DSBs primarily occur within the coding region. Open chromatin identified by FAIRE in ORFs is associated with hotspots to the same degree as open chromatin is associated with DSBs in chromatin overall. Of the bases within open chromatin, 21% are in ORFs, and of these bases, 42% [12] and 16% [15] are held in common with DSB hotspot regions. This degree of overlap is indistinguishable from the overlap between all open chromatin and DSB hotspots, which is 45% [12] and 16% [15]. We note that in addition to the low resolution afforded by the ssDNA assays, determination of overlap between open chromatin and coding regions could be further affected by the resolution of our microarrays (~270 bp).

Recombination hotspots that are preferentially resolved as crossovers are associated with open chromatin to the same degree as those resolved as non-crossovers

Recombination hotspots have been identified as sites with more recombination events than would be expected under a homogenous genomic rate ($P < 0.001$) [11]. The location of recombination hotspots may differ slightly from DSB hotspots. While all meiotic recombination is initiated by DSBs [8,9], it is possible that some DSBs produce non-crossovers without associated gene conversion or are repaired from the sister. These cases would not be detected by genome tiling recombination maps but would be detected in DSB mapping.

We investigated whether recombination hotspots in open chromatin produce crossovers at a frequency different from those that occurred in closed chromatin. This analysis included 136 'total' recombination hotspots that are classified as hot for all recombination activity, 92 that are classified as crossover hotspots, and 74 that are non-

crossover hotspots [11]. Overall, the degree of association between open chromatin and recombination hotspots is similar to the degree of association between open chromatin and DSB hotspots [12,15] (**Table 1,Table S2**). In this analysis, the strongest association between open chromatin and recombination hotspots occurred at the 1.5 hr timepoint rather than the 3 hr timepoint, which suggests that variation in association of open chromatin and hotspots between timepoints may not be very significant. Open chromatin was significantly associated (z test; $P < 0.01$) with both crossover hotspots and non-crossover hotspots, with no significant difference in the level of association between the two (**Table S2)**. We conclude that the occurrence of a recombination hotspot in open chromatin does not influence whether DSBs in that region are resolved as crossovers or non-crossovers.

No significant difference in the open chromatin status of DSB hotspots resolved as crossovers vs. non-crossovers was observed. We interpret this to mean that chromatin state is important for the establishment of DSBs, but does not influence the repair pathway. However, we were able to analyze only hotspots for which there was data about the frequency of crossover and non-crossover resolution. If and when new data emerges for additional hotspots, this issue can be investigated more thoroughly.

DISCUSSION

We employed FAIRE on a synchronized meiotic time course in *S. cerevisiae* to investigate patterns of open chromatin during the first 3 hours of meiosis. Upregulation of *SPO11*, *DMC1* and *RAD51* in our samples, and previous studies [6, 58] indicate that cells are processing DSBs into joint molecules during this time. An association between open

chromatin and recombination hotspots had been established through many locus-specific studies, but whether and to what degree this association held on a genome-wide scale had yet to be determined.

Challenges

Several challenges were associated with our analysis of chromatin organization in meiosis. First, the biological function of open chromatin is likely dependent on chromosomal and physiological context. However, we used a strict *P*-value cutoff to define "open chromatin" uniformly across the genome. This cutoff likely sacrifices some biologically relevant true positives for the sake of minimizing false negatives. ROC curve analysis indicates that sites identified at the low end of our P-value threshold are less likely to intersect with a DSB hotspot, and this trend intensifies as the cutoff is made more lenient. A second related challenge was defining a significant change in the openness of chromatin between samples or timepoints. Again, determining the degree of change likely to be linked to a biological outcome is difficult. Furthermore, we categorized sites as "open" or "not open" at individual timepoints, without using quantitative information from adjacent timepoints to aid our calls.

Third, FAIRE may miss some sites such as the *ARG4* locus that have been classified as open by other assays (**Results**). We speculate that sites that are highly occupied by recombination complexes could be identified as closed by FAIRE, but recognized as open by nuclease sensitivity assays. Chromatin profiles generated by FAIRE may differ from those created by other assays such as DNase or MNase sensitivity or histone ChIP-chip. FAIRE probes chromatin with a reactive small

molecule, while DNase and MNase probe chromatin by accessibility to a much larger enzyme. ChIP is dependent on the affinity of an antibody for its substrate. Each assay has its own advantages and disadvantages, and each is likely to be more sensitive than another in measuring some specific aspect of chromatin organization. Therefore, each assay reveals independent, non-redundant information about the chromatin profile at each locus, and carries a different definition as to what "open chromatin" means. DNase and MNase sensitivity assays may indeed be most sensitive in detecting open chromatin that occurs in sites that are highly occupied by recombination complexes but not nucleosomal proteins. Conversely, openness as defined by FAIRE may in some cases fail to recognize sites that are highly occupied by non-histone proteins.

Fourth, we used tiling microarrays with a resolution of ~270 bp, which is slightly less than the size of two nucleosomes. This resolution, together with the 250 bp minimum call size of our peak-finding algorithm, makes it possible that our analysis missed small sites of open chromatin that are relevant to our assertions. For example, DSB hotspot mapping on chromosome III by Southern blot analysis showed that some DSB hotspots can be resolved to ~150 bp [13, 59, 60]. A fifth caveat stems from the low resolution of DSB hotspot mapping provided by the detection of meiosis-specific ssDNA. Although every DSB occurs at a single base, the ssDNA tails are often greater than 1 kb in length, and the corresponding hotspot calls can be 2 kb or greater in length [12, 15].

Conclusions

Despite the challenges, our results provide clear evidence for the following:

(1) DSB hotspots are generally associated with open chromatin.

(2) A class of DSB hotspots occur in regions that are no more open than the genomic average, illustrating that open chromatin is not a requirement for DSB formation. The close spacing of our timepoints in relation to DSB formation, and the incomplete synchrony of meiotic samples makes it unlikely that we failed to capture opening of chromatin at these DSB hotspots. DSB hotspots that occur in closed chromatin could be specified by other properties such as histone modifications.

(3) Among sites of meiotic open chromatin, the degree of chromatin openness as defined by *P*-value is predictive of DSB hotspots. However, among sites of mitotic open chromatin, the degree of openness is not predictive of meiotic DSB hotspots.

(4) Regions 0-20 and 30-40 kb proximal to the subtelomeric borders are enriched for meiosis-specific sites of open chromatin (see next section below).

(5) The chromatin surrounding centromeres is open in mitosis and the first 3 hours of meiosis.

(6) Many tRNA genes are open specifically in meiosis, with the open site usually occurring downstream of the annotated gene.

tRNAs and Chromatin

The biological significance of meiosis-specific open chromatin around tRNA genes is not obvious to us. If this enrichment were connected in some way to recombination, we would expect an association between DSB hotspots and tRNA loci, which is not apparent. However, in *S. pombe* and humans, B-box motifs and tRNA-associated sequences do function in modulating in chromatin modification [61, 62]. They also serve as condensin and cohesin loading sites in mitotic cells in both *S. cerevisiae* and *S. pombe*

[57, 63]. In *S. pombe*, deletion of a ~489 bp sequence containing two tRNA genes that form a cohesin binding site resulted in a complete loss of cohesin loading at that site [63]. Additionally, in *S. pombe*, the tRNA alanine ($tRNA^{Ala}$) gene specifies a heterochromatin boundary, which is important for meiotic chromosome disjunction [62]. Deletion of the centromere-proximal $tRNA^{Ala}$ gene causes H3 histones dimethylated at lysine 9 (H3K9me2) to spread beyond the wild-type boundary at $tRNA^{Ala}$. This mutant also exhibits defective meiotic chromosome segregation [62], suggesting that in *S. pombe* chromatin changes directed by tRNAs are important for accurate meiotic disjunction. A recent study in *S. pombe* showed that long, polyadenylated, non-coding RNA genes often colocalize with meiotic DSB hotspots [64], but did not address tRNA genes.

In *S. cerevisiae*, tRNA genes have been shown to play a role in the specification of heterochromatin boundaries. For example, the tRNA threonine ($tRNA^{Thr}$) gene is necessary and sufficient to stop the spread of heterochromatin from the HMR locus on chromosome III [65]. Similarly, the tRNA glutamine ($tRNA^{Gln}$) is required for halting the spread of heterochromatin extending from rDNA arrays on chromosome XII [66]. The barrier activity of both $tRNA^{Thr}$ and $tRNA^{Gln}$ is abrogated in histone acetyltransferase mutants *sas2* and *gcn5*, reinforcing an important function for tRNA in chromatin organization [66].

A pairing function for meiosis-specific open chromatin at the subtelomeric borders?

We observed a strong enrichment of meiosis-specific open chromatin at the subtelomeric border regions. Approximately 8% of the sites we identified to be open during all meiotic timepoints but closed in YPD exist 0-10 kb proximal to the subtelomeric border, which is

significantly above the expected proportion of 3%. DSB mapping studies based on *rad50S* have reported an under-representation of meiotic DSBs 40-100 kb from the telomeres [14, 67, 68]. However, studies using ssDNA mapping in a *dmc1* mutant strain report that the depletion of DSBs extends only 0-20 kb from telomeres, and that regions 20-120 kb from telomeres exhibit an over-representation of DSBs [12, 15]. It is attractive to propose that meiosis-specific opening of the chromatin in the regions proximal to the subtelomeric border (which are generally 0-40 kb from the telomere) provides access for recombination protein complexes, and a unique substrate for homology searching to facilitate homologous pairing. It will be interesting to see if mutants defective for pairing also show an absence of meiosis-specific open chromatin in these regions. Some factors that promote meiosis-specific open chromatin have been identified in *S. pombe* [28, 29, 69], and the idea that meiosis-specific open chromatin aids chromosome pairing predicts that mutants for this process will be pairing-defective.

Future applications in other systems

FAIRE is a low-cost, reproducible method that can be used to identify sites of open chromatin. Our results demonstrate that chromatin that is constitutively open in mitosis or meiosis is correlated with DSBs, and therefore FAIRE holds promise for the identification of putative hotspots in organisms in which meiotic synchrony cannot be established. It is also useful in organisms that lack convenient mutations that arrest meiosis after ssDNA has formed. FAIRE may be especially useful in organisms in which meiotic prophase staging is possible, such as *Caenorhabditis elegans* and *Arabidopsis thaliana*. In *C. elegans*, oocytes at all stages of meiotic prophase exist within each gonad

in an unambiguous temporal/spatial arrangement [70], although it may be challenging to obtain the estimated 500,000 germ cells required for FAIRE. In *A. thaliana* meiotic staging is more difficult, but cells undergoing meiosis can be quickly and cheaply isolated by capillary collection [71].

METHODS

Strains and growth conditions

Vegetative cells were grown in YPD (1% yeast extract, 2% peptone, 2% dextrose) to an OD_{600} of 0.6-0.8. Synchronous sporulation was carried out for each biological replicate (3 total) by using YPD overnight cultures of wild-type cells (SK1 strain SHy02 (*MAT***a**/*MATα ho::LYS2/ho::LYS2 leu2::hisG/leu2::hisG lys2/lys2 ura3/ura3*) to inoculate YPA (1% yeast extract, 2% peptone, 2% potassium acetate) cultures at an OD_{600} of 0.1. We grew the cultures at 30° C to an OD_{600} of 0.8-1.2 (about 16 hours), collected cells by centrifugation, and washed with SM (2% potassium acetate). We then resuspended cells at an OD_{600} of 2.0 in SM and incubated 600ml cultures in 2L flasks at 30° C shaking at 265 rpm.

Sample collection and FAIRE

Cells used for FAIRE were collected during vegetative growth (100 ml of culture per sample) and after 0, 1.5, and 3 hours after placement in SM (50 ml of culture per sample). Three biological replicates of each timepoint were taken. To monitor meiotic entry, progression and synchrony in our biological replicates, we collected RNA samples for expression microarray analysis at each FAIRE time point. Additionally, we collected

RNA from pre-meiotic cells that had been starved of a fermentable carbon source for 8 hours (Acetate) and from synchronized cells that had been incubated in sporulation media for 4.5, 6, 9 and 12 hours to assay the transcriptional profile of key early, middle and late meiotic genes. Total RNA was isolated from frozen samples using hot acid-phenol as previously described [72]. FAIRE samples were fixed with 1% formaldehyde for 20 minutes at room temperature and washed twice with 1X cold PBS. The samples were then snap frozen in a dry ice/ethanol bath and stored at -80° C. FAIRE was carried out as described [34]. Briefly, cells were thawed and lysed with glass beads and the chromatin was sheared to an average size of ~800 bp by sonication as measured by agarose gel electrophoresis. We conducted a phenol-chloroform extraction on the crosslinked samples to remove proteins and protein-associated DNA. The resulting DNA samples were further purified using Zymo Clean and Concentrator kit per manufacturers protocol (Zymo Research). Genomic reference DNA samples were generated using the YPD samples without formaldehyde-crosslinking.

FAIRE amplification and labeling

FAIRE and reference DNA was amplified from ~100 ng starting material using the Whole-Genome Amplification kit (Sigma cat# 127K1745) [73]. Two micrograms of amplified DNA was labeled with either Cy3 or Cy5 conjugated dUTP using the BioPrime Array CGH kit (Invitrogen). For replicates 1 and 3 the FAIRE samples were labeled with Cy5 and the reference Cy3. A dye swap was performed on replicate 2. DNA concentration and dye incorporation were measured using a Nanodrop spectrometer. Four

micrograms of each fluorescently labeled sample and reference were used in the hybridization.

Expression sample labeling

For expression experiments, total RNA was reverse-transcribed using SuperScript II Reverse Transcriptase (Invitrogen) incorporating amino-allyl dUTP (Sigma) at a ratio of 3:2 with dTTP. Genomic DNA (used as a reference) was amplified with Klenow (NEB) incorporating amino-allyl dUTP (Sigma) at a ratio of 3:2 with dTTP. Reactive Cy3 or Cy5 (Amersham) was coupled to the amino-allyl of the resulting DNA fragments in the presence of sodium bicarbonate.

Microarray hybridization and image acquisition

FAIRE samples were hybridized simultaneously with the reference at 52° C for 16 hours in a Maui hybridization chamber to yeast whole genome 4 x 44k tiling microarrays (Agilent; ~270 bp resolution). Expression samples were hybridized simultaneously with the reference sample to yeast whole genome PCR product spotted microarrays containing coding and non-coding regions at approximately 1 kb resolution. The arrays were scanned with an Axon 4000 scanner, and data was extracted using GenePix 6.0 software. Only spots of high quality by visual inspection, with fewer than 10% saturated pixels in either channel, with a background corrected sum of medians for both channels greater than 500, were used for the analysis.

Data processing and statistical analysis

The ratio of FAIRE intensity (sample) to genomic DNA intensity (reference) for each spot on the microarray was converted to a $\log_2$ ratio, which is referred to as a "FAIRE log ratio". Mitochondrial probes were removed from the analysis prior to data processing. FAIRE log ratios were converted to *z* scores by centering at a mean of zero and scaling the standard deviation to one. The median *z* score was determined for each sample across all 3 biological replicates and the values were then re-centered and scaled. These re-scaled median *z* scores were used as input for the peak-finding algorithm ChIPoTle (V1.02) [74] with the following parameters: Gaussian background distribution, step-size 125 bp, and window size 500 bp [74]. Peaks representing sites of open chromatin were collected from all samples using a Bonferroni corrected threshold of $P < 0.01$. Sites of open chromatin were initially organized as an output from each timepoint (YPD, 0hr, 1.5hr, and 3hr). Among samples, sites were considered to be coincident if they shared greater than 50% bp overlap or an identical high spot, defined as the probe with the highest *z* score. For analysis using these groupings, the site with the most significant P-value was used.

Overlaps with published hotspot coordinates as defined by ssDNA enrichment [12, 15] or by genome tiling recombination maps [11] were calculated using the Galaxy web application [75]. Overlaps with ORFs were calculated by determining the intersecting regions with samples and the Saccharomyces Genome Database (SGD) ORF annotations using Galaxy [75]. Regions were considered to overlap between two datasets if at least 1 bp overlapped. As a control, 10 permutations of each sample set of regions were generated using a PERL algorithm, which centers the start and end coordinates of each region on a random probe from the array [36]. Percent overlap was determined for

each permutation as with the samples using Galaxy to generate a distribution of random overlaps for each sample. We tested the distributions of random overlaps for normality using a Shapiro-Wilk goodness of fit test [76]. The normal distribution could not be excluded in any case, and we used a *Z* test to determine the *P* values of our experimental overlaps [77]. The means of the distributions of random overlaps were used as the expected values.

We presented FAIRE enrichment over DSB hotspots and tRNA associated sequences as moving averages of probe z scores. This was done by first determining all probes that fell within DSB hotspots or within 500 bp of a tRNA annotated start or stop. We then aligned DSB hotspots [12] (all are 2 kb) and tRNAs at their annotated starts. At each position within the aligned DSB hotspots or tRNA associated sequences, we calculated the moving average of the closest 15 probes.

We assessed the ability of FAIRE to predict DSB hotspots using ROC curve analysis [40]. Briefly, a ROC curve is a plot of the rates at which true positives and false positives are compiled with respect to an ordering statistic. In this case, the ordering statistic is the *P*-value of a site of open chromatin. A true positive occurs when a FAIRE site intersects (250 bp overlap) with the coordinates of a DSB hotspot [12, 15] and a false positive occurs when a site does not intersect with a DSB hotspot. We ordered FAIRE sites by *P*-value and plotted the false positive fraction (FPF), scaled to 1, on the X-axis vs. the true positive fraction (TPF), scaled to 1, on the Y-axis. We measured the area under the curve (AUC), which is a summary of the predictive value of the ordering statistic. An AUC value of 0.5 indicates a random predictor while an AUC value of 1.0 indicates a perfect predictor. We compared each result to 10 control ROC curves

generated by random placement of the analyzed FAIRE sites (as above). Additionally, we performed ROC curve analysis to quantify the ability of raw FAIRE data to predict DSB hotspots. We conducted the analysis as above ordering probes by FAIRE enrichment *z* score and defining a true positive as when the probe intersected (44 bp overlap) a DSB hotspot [12, 15]. We calculated likelihood ratios (TPF/FPF) at various *z* score thresholds to express the predictability of FAIRE enrichment *z* scores in each timepoint.

Chromatin profiles at the regions associated with centromeres and telomeres were analyzed by calculating the total counts of sites of open chromatin that intersect the regions of interest (either centromeric or telomeric associated regions). Our experimental values were compared to the counts of intersecting regions expected from 10,000 random permutations of each sample. The mean counts of the random permutations were used as the expected values for experimental samples and the statistical significance of the expected overlaps to observed overlaps were assessed by a X^2 test. This method of analysis was also used to compare the overlap of sites of open chromatin with DSB hotspots where the centers of the permuted sites were restricted to intergenic space.

Motif determination

Overrepresented sequence motifs were determined for each group of regions of open chromatin using BioProspector [78], set to find 10, 13 and 15 bp motifs. Yeast intergenic sequences were used as the background model.

ACKNOWLEDGEMENTS

We thank Arnold Barton and David Kaback for providing the chromosomal coordinates for the computationally defined boundaries between the subtelomeres and euchromatin. We also thank Paul Giresi for help with data analysis. LEB and GPC thank NSF (MCB-0618691) and DOE (DE-FGO2-05ER15651) for financial support. This work was supported by NIH grant R01-GM072518 to JDL.

FIGURE AND TABLE LEGENDS

Figure 5.1. Monitoring meiotic entry and progression in the time course used for FAIRE. (A) To validate meiotic entry and synchrony in each biological replicate, we analyzed the transcription profiles of key meiotic progression genes by expression microarray. Timepoint is indicated on the x-axis and expression relative to YPD is shown on the y. Biological replicates are plotted (1: blue, 2: red and 3: green). If only two samples are shown, the YPD probe was flagged for technical reasons and thus no comparisons could be made. Early meiotic genes include *IME1*, *IME2*, *HOP1*, *DMC1*, *RAD51* and *SPO11*. Middle meiotic genes include *NDT80*, *SPS1*, *SPS2*, *CLB1*, *CLB6* and *SMK*1. **(B)** Comparison to meiotic transcription microarray data of Chu *et al*. [33] Purple indicates positive meiotic enrichment and green indicates negative meiotic enrichment. Transcription data is represented as a ratio between timepoint expression and 0 hr expression. **(C)** Completion of meiosis as measured by spore counts. Spore counts were taken at each timepoint. Timepoint is indicated on the x-axis and the percentage of sporulated cells is indicated on the y. Biological replicates are plotted (1: blue, 2: red, 3: green and average: purple).

Figure 5.2. **The distribution of open chromatin and DSB hotspots across a representative 110 kb genomic region containing 53 genes.** Chromosome coordinates are plotted along the x-axis, with FAIRE enrichment (*z* score) plotted on the y. FAIRE data were loaded into the UCSC Genome Browser along with published DSB hotspots (black and gray) [12, 15]. FAIRE data from YPD (red), 0 hr (green), 1.5 hr (blue), and 3 hr (purple) samples are plotted.

Figure 5.3. Chromatin at known hotspots is dynamic. Chromosome coordinates are plotted along the x-axis, with FAIRE enrichment (*z* score) plotted on the y. Positive z scores indicate relatively open chromatin, and negative numbers indicate relatively closed chromatin. FAIRE data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted. Among the 1204 sites, 151 were open in all the meiosis samples (1.5 hr and 3 hr), but not during mitotic growth (e.g. *YAT1*). Conversely, 156 sites were open during mitotic growth, but were not open in any of the meiotic samples. Meiotic histone H3 ChIP-chip data from 0 hr (orange), 1 hr (brown) and 2 hr (dark brown and dotted) are plotted [27]. Hotspot regions identified by Buhler et al. [12] (light green bars), and by Blitzblau et al [15] (dark green bars) are shown. Arrows represent the coding regions of genes, with those above the x-axis coded from the Watson strand and those below coded from the Crick strand. (**A)** Chromatin dynamics at the *CDC19* hotspot on chromosome I, an example of a hotspot with constitutively open chromatin. **(B)** Chromatin dynamics at the *YAT1* hotspot on chromosome I, an example of a meiosis-specific site of open chromatin. **(C)** Chromatin dynamics at the

ARG4_DED81 hotspot on chromosome VIII [18, 30], an example of a DSB hotspot with no observed open chromatin in any sample. **(D)** Chromatin dynamics at the *HIS4* hotspot on chromosome III, an example of a hotspot at which the chromatin is open in YPD and closes as meiosis progresses.

Figure 5.4. The degree of chromatin openness is predictive of DSB hotspots. (A) ROC curve analysis of the five categories of open chromatin. True positives and false positives are defined with respect to DSB hotspots [12] (**Methods**). The FAIRE sites plotted are ordered such that those to the left are most open according to FAIRE (lowest *P*-value) and those to right, while still significantly open, have higher *P*-values. Data from category I (red), category II (dashed maroon), category III (light blue fine dashes), category IV (navy and dotted) and category V sites (green) are plotted. AUC values are listed next to the sample name, and the minimum and maximum AUC values of the random permutations follow in parentheses. **(B)** ROC curve analysis of sites of open chromatin in individual timepoints. Data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted. **(C)** DSB predictive value of individual probes with FAIRE *z* scores above the indicated threshold, expressed as likelihood ratios. Values in the table are TPF/FPF at decreasing *z* score thresholds, listed vertically on the left. Timepoints are organized on the horizontal axis. Predictive value is expressed with respect to two DSB hotspot datasets [12, 15]. Because this analysis is probe-based, multiple true positives may be derived from a single DSB hotspot.

Figure 5.5. Regions proximally adjacent to the subtelomere borders exhibit meiosis-specific open chromatin. (A) Examples of meiosis-specific open chromatin at the subtelomeric borders. Distance from the subtelomeric border (coordinate 0) is plotted on the x-axis and FAIRE enrichment is plotted on the y. Subtelomeres (orange bars) are shown. FAIRE data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted. Asterisks indicate sites significantly open ($P < 0.01$) **(B)** Analysis of telomeric regions. The x-axis reports the region analyzed and the y-axis reports the ratio of observed sites of open chromatin to the expected number from random sampling. Category I (red), category II (maroon), category III (light blue) and category IV sites (navy) are shown. The telomere and subtelomeres do not deviate from the genomic averages of expected/observed sites of open chromatin. Meiosis-specific open chromatin is most prominent in the 10 kb adjacent to the subtelomeres. In the 10-20 kb and 30-40 kb windows, only category IV sites are significantly above expected values. **(C)** The ratio of observed to expected number of DSB hotspots as plotted as a function from distance from the subtelomeric border. Data from Buhler *et al.* [12] (blue) and Blitzblau *et al*. [15] (red) is shown. **(D)** Recombination in the subtelomeric regions as measured by linkage mapping [51]. Distance from the subtelomeric border is plotted on the x-axis and recombination rate (cM/kb) is shown on the y. Recombination rates are presented as the average of the 15 reported rates closest to each position plotted. Subtelomeric regions are defined as having either > 50% DNA sequence identity to another region of the genome or as having < 60% non-dubious ORFs using a 20 kb sliding window [51]. The subtelomeric border is the point at which neither of these

criteria is met. The gray line represents the genomic average recombination rate of 0.37 cM/kb [51].

Figure 5.6. Centromeres and pericentromeric regions exhibit constitutively open chromatin. (A) FAIRE profiles of each of the 16 *S. cerevisiae* centromeres. Asterisks indicate sites significantly open at $P < 0.01$. Chromosomal coordinates are plotted on the x-axis and FAIRE enrichment is plotted on the y. FAIRE data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted. Centromeres (black circles) are shown. **(B)** Analysis of centromere-associated regions. The x-axis reports the region analyzed and the y-axis reports the ratio of observed sites of open chromatin to the expected number from random sampling. Category I (red), category II (maroon), category III (light blue) and category IV sites (navy) are shown.

Figure 5.7. A 15-bp motif is overrepresented in meiosis-specific open chromatin surrounding tRNA genes. The 15-bp motif identified from **(A)** FAIRE-enriched sites isolated from any meiotic sample but not YPD, and FAIRE-enriched sites isolated from all meiotic samples but not YPD are represented [79]. **(B)** Open chromatin increases during meiosis at sites associated with tRNA genes. The position relative to tRNA annotated start sites (bp) is plotted on the x-axis and FAIRE enrichment is plotted on the y. Cytoplasmic tRNA genes associated with a site of open chromatin were aligned at their annotated transcription start. Shown is a moving average of the average FAIRE-enrichment *z* score (window size = 15 probes). FAIRE data from YPD (red), 0 hr (dashed

green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted. Regions corresponding to the functional tRNA have a median length of 72 bp (minimum of 70 bp and a maximum of 132 bp). The graph includes the 98 tRNA-associated sequences that overlap a site of open chromatin ($P < 0.01$).

Table 5.1. **Sites of open chromatin are preferentially associated with DSB hotspots.** Shown are the percentages of bp from each sample (top) that overlap with DSB hotspots [12, 15]. Permuted values are the average derived from 10 permutations of FAIRE site positions (**Methods**). Observed/expected ratios indicate the strength of association between the sample and DSB hotspots.

SUPPLEMENTARY FIGURE AND TABLE LEGENDS

Figure 5.S1. DSB hotspots can be divided into those that occur in sites of open chromatin and those that do not. DSB hotspots were aligned at their annotated start. Hotspot position is plotted on the x-axis and FAIRE enrichment is plotted on the y. Shown are plots of the moving averages of *z* scores of the closest 15 probes at each position. Data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), 3 hr (purple and dotted), meiotic histone H3 ChIP-chip from 0 hr (orange), 1 hr (brown) and 2 hr (dark brown and dotted) are plotted. **(A)** The 167 DSB hotspots included in both datasets that overlap a site of open chromatin ($P < 0.01$). **(B)** The 129 DSB hotspots included in both datasets [12,15] that do not overlap a site of open chromatin.

Figure 5.S2. Sites identified as mitosis-specific open chromatin are not just under the significance threshold in meiosis. (A) Sites of open chromatin identified as mitosis-specific (Category V). Relative position (scaled to one) within sites is plotted on the x-axis and FAIRE enrichment is plotted on the y. Shown are plots of the moving averages of *z* scores of the closest 15 probes at each relative position. FAIRE data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted. **(B)** Sites of open chromatin identified as constitutive (Category II). FAIRE data from YPD (red), 0 hr (dashed green), 1.5 hr (blue fine dashes), and 3 hr (purple and dotted) samples are plotted.

Table 5.S1. **The association between sites of open chromatin and DSB hotspots is significant even when permutations are restricted to intergenic regions.** Shown are the numbers of DSB hotspots (top) [12,15] that overlap with a site of open chromatin from the corresponding sample (left). Intergenic permuted values are the average number of DSB hotspots that overlap after 10,000 permutations of the open chromatin sites, in which the placement was restricted to intergenic space (**Methods**).

Table 5.S2. **Crossover and non-crossover hotspots are associated with open chromatin to the same degree.** Shown are the percentages of bp from each sample (top) that overlap with total recombination hotspots, crossover hotspots and non-crossover hotspots [11]. Permuted values are the average bp overlap of 10 permutations of the sample (**Methods**). Observed/expected ratios indicate the strength of association between the sample and DSB hotspots.

Table 5.S3. **Enrichment of open chromatin at subtelomeric border regions and pericentromeric regions.** Shown are the total counts of sites of open chromatin from each sample (top) that overlap with the listed genomic region (left). Permuted values are the average number of regions of open chromatin that overlap after 10,000 permutations of the FAIRE sample (**Methods**). Observed/expected ratios indicate the strength of association between the sample and the genomic region.

Table 5.S4. Coordinates of proximal subtelomeric borders. Shown are the coordinates for the left and right proximal subtelomeric borders on each chromosome from Barton *et al.* [51]. See figure 5D legend for the criteria used to define these borders.

Table 5.S5. **DSB hotspots are not enriched at regions associated with tRNA genes.** Shown are the total counts of DSB hotspots (top) [12,15] that overlap with either tRNA coding regions or tRNA coding regions ± 500 bp. Permuted values are the average number of DSB hotspots that overlap after 10,000 permutations of the DSB hotspots ± 1 standard deviation (**Methods**). None of the associations are significant ($P > 0.05$).

FIGURES

Figure 5.1.

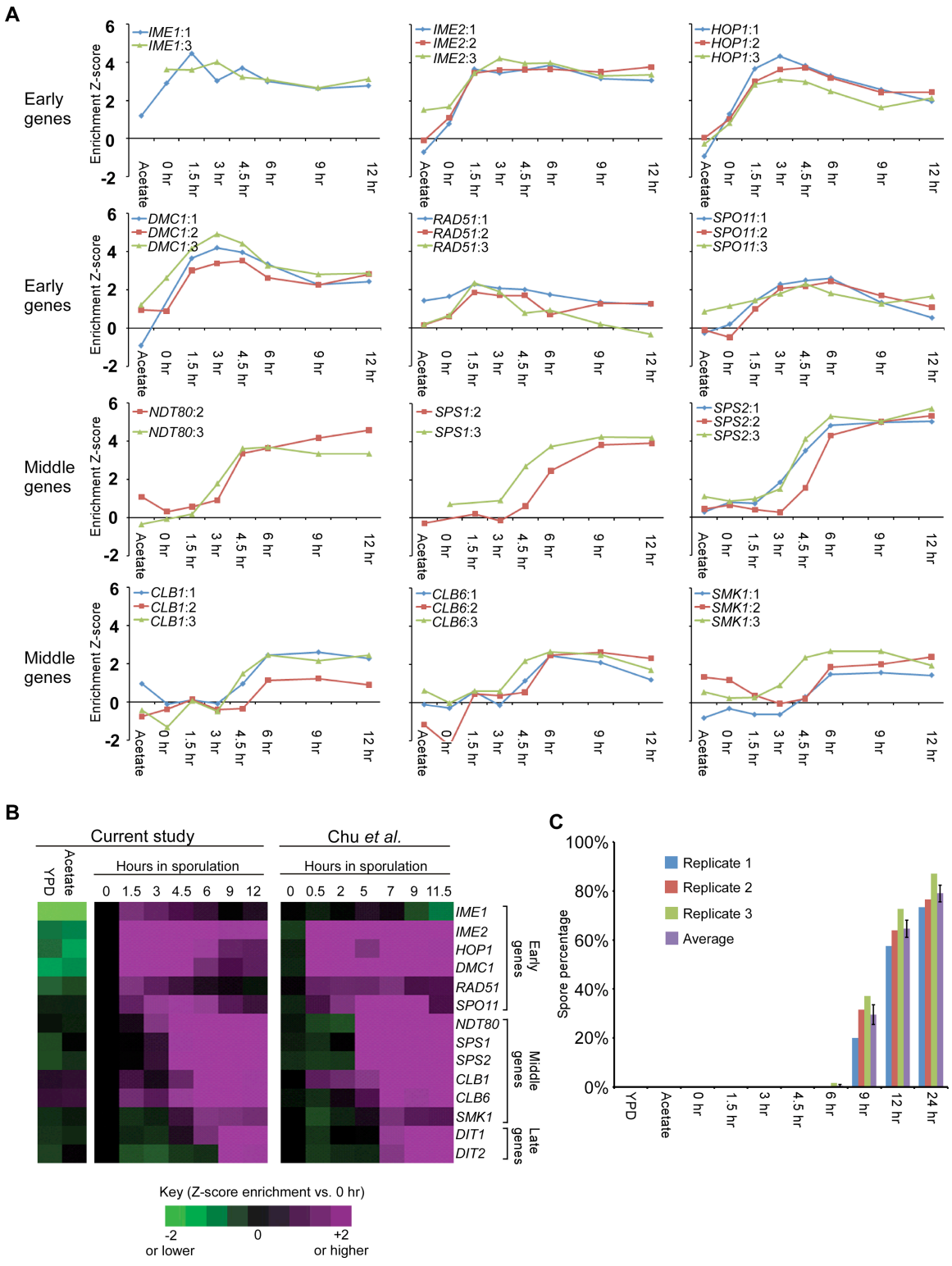


Figure 5.2.

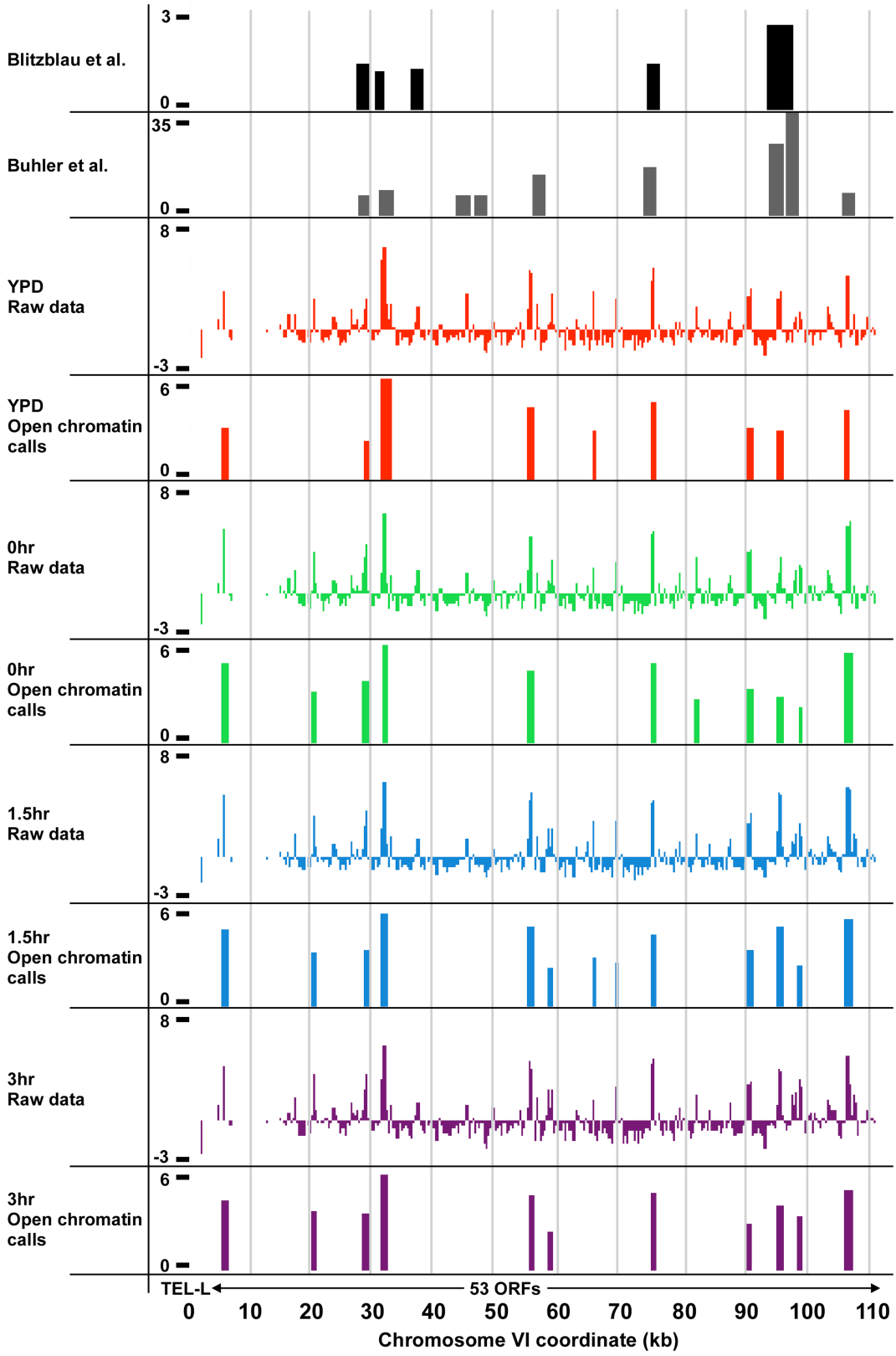

Blitzblau et al.
3
0
Buhler et al.
35
0
YPD
Raw data
8
-3
YPD
Open chromatin
calls
6
0
0hr
Raw data
8
-3
0hr
Open chromatin
calls
6
0
1.5hr
Raw data
8
-3
1.5hr
Open chromatin
calls
6
0
3hr
Raw data
8
-3
3hr
Open chromatin
calls
6
0
TEL-L
53 ORFs
0
10
20
30
40
50
60
70
80
90
100
110
Chromosome VI coordinate (kb)

Figure 5.3.

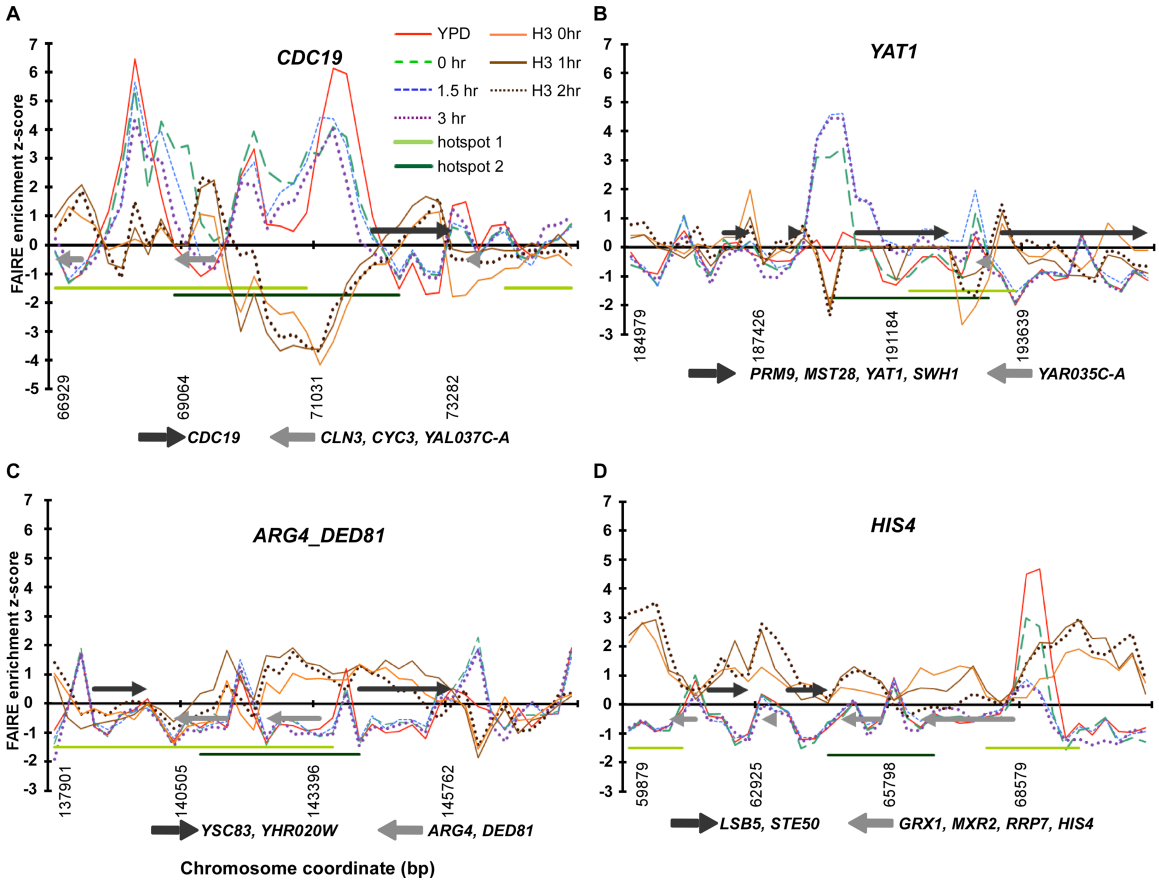
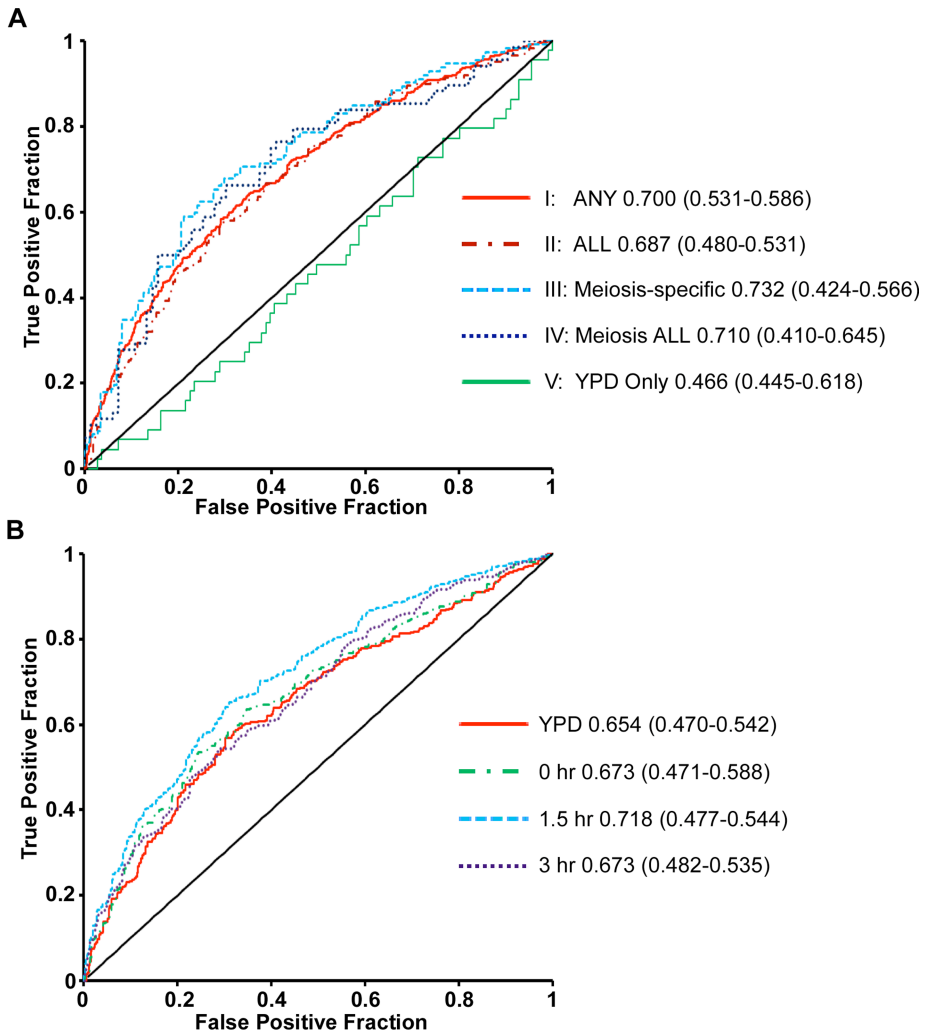

A
CDC19
YPD
0 hr
1.5 hr
3 hr
H3 0hr
H3 1hr
H3 2hr
hotspot 1
hotspot 2
FAIRE enrichment z-score
66929
69064
71031
73282
CDC19
CLN3, CYC3, YAL037C-A
B
YAT1
184979
187426
191184
193639
PRM9, MST28, YAT1, SWH1
YAR035C-A
C
ARG4_DED81
137901
140505
143396
145762
YSC83, YHR020W
ARG4, DED81
D
HIS4
59879
62925
65798
68579
LSB5, STE50
GRX1, MXR2, RRP7, HIS4
Chromosome coordinate (bp)

Figure 5.4.



C

min FAIRE enrichment z-score	YPD		0 hr		1.5 hr		3 hr	
	Buhler et al	Blitzblau et al	Buhler et al	Blitzblau et al	Buhler et al	Blitzblau et al	Buhler et al	Blitzblau et al
5	7.11	6.42	5.89	4.74	7.88	6.46	6.79	6.02
4	6.88	5.46	5.39	5.15	5.71	5.60	5.72	5.29
3	4.68	4.24	4.92	4.40	5.27	4.90	4.89	4.90
2	3.56	3.26	3.67	3.56	3.70	3.76	3.33	3.68
1	2.44	2.35	2.33	2.35	2.47	2.43	2.16	2.23
0	1.26	1.24	1.24	1.29	1.31	1.36	1.18	1.21

Figure 5.5.

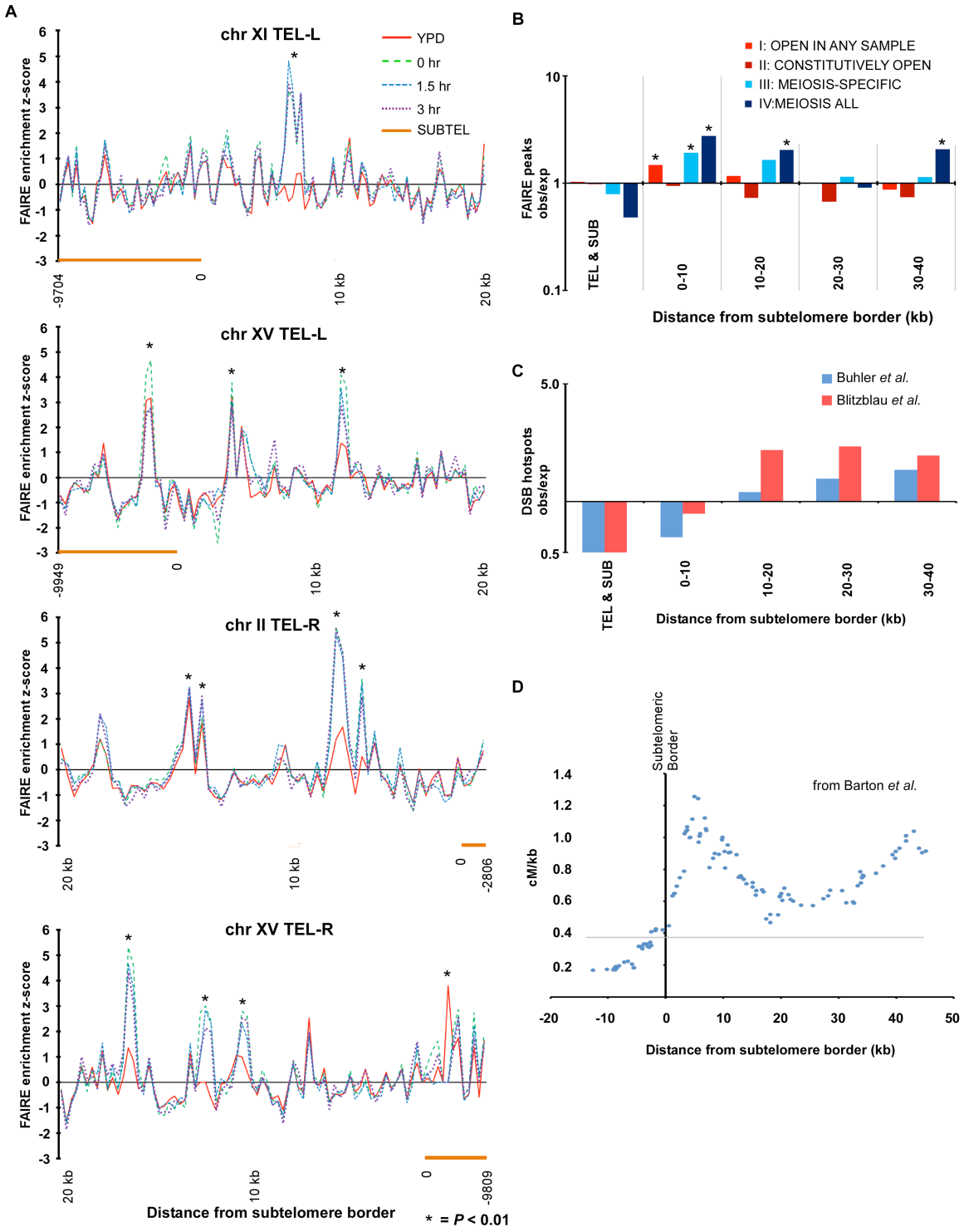

A
chr XI TEL-L
YPD
0 hr
1.5 hr
3 hr
SUBTEL
FAIRE enrichment z-score
chr XV TEL-L
chr II TEL-R
chr XV TEL-R
-9704
-9949
-2806
-9809
0
10 kb
20 kb
Distance from subtelomere border
* = P < 0.01
B
I: OPEN IN ANY SAMPLE
II: CONSTITUTIVELY OPEN
III: MEIOSIS-SPECIFIC
IV:MEIOSIS ALL
FAIRE peaks obs/exp
TEL & SUB
0-10
10-20
20-30
30-40
Distance from subtelomere border (kb)
C
Buhler et al.
Blitzblau et al.
DSB hotspots obs/exp
Distance from subtelomere border (kb)
D
Subtelomeric Border
from Barton et al.
cM/kb
Distance from subtelomere border (kb)

Figure 5.6.

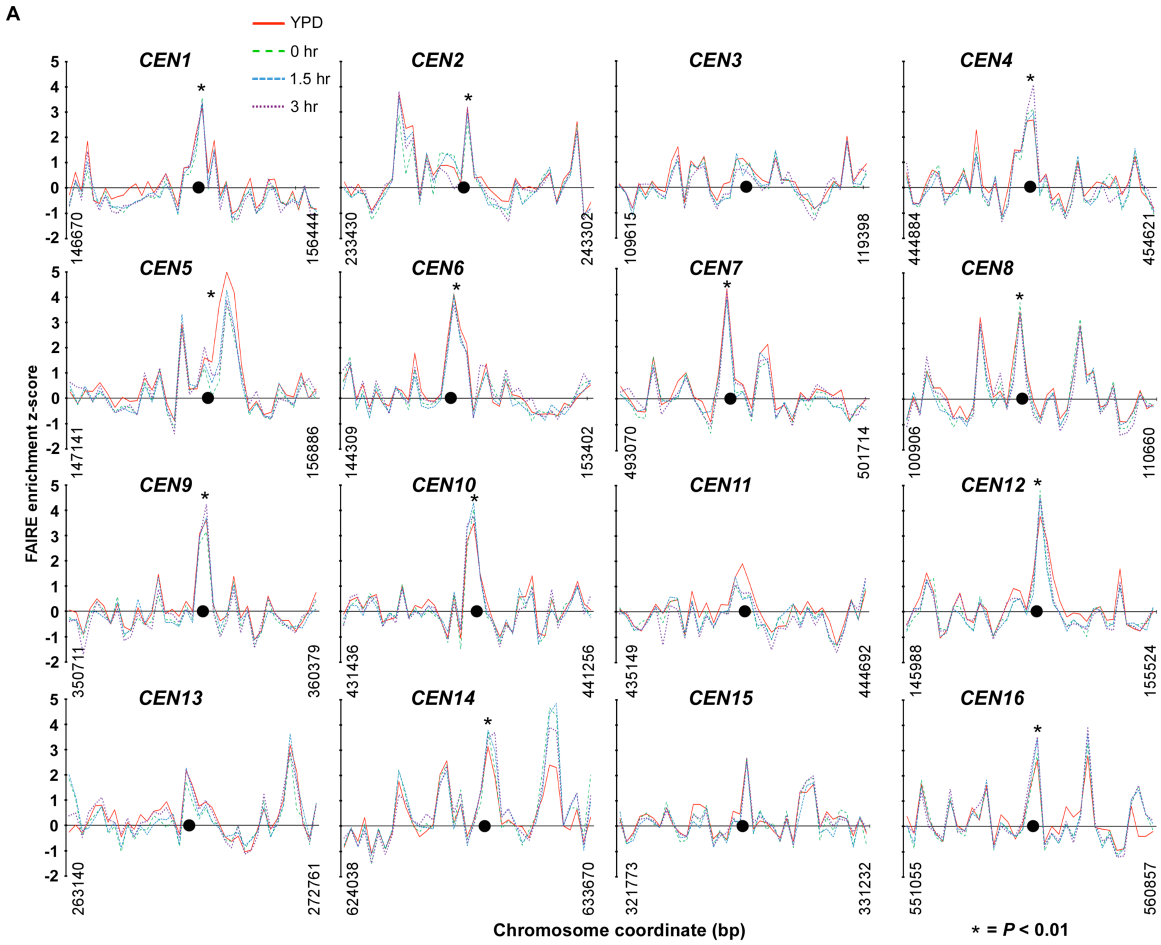
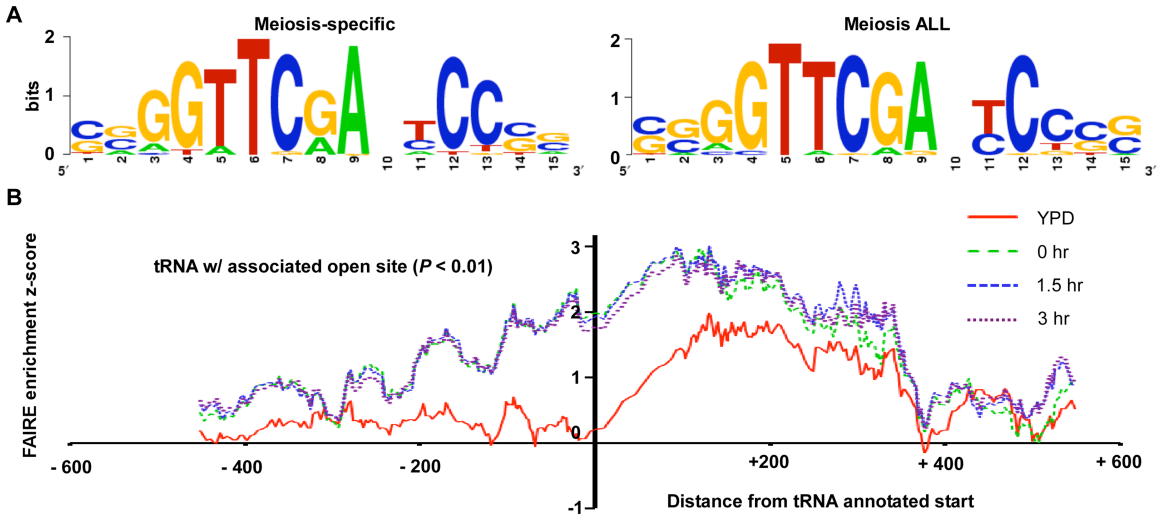

A
YPD
0 hr
1.5 hr
3 hr
CEN1
CEN2
CEN3
CEN4
CEN5
CEN6
CEN7
CEN8
CEN9
CEN10
CEN11
CEN12
CEN13
CEN14
CEN15
CEN16
FAIRE enrichment z-score
Chromosome coordinate (bp)
* = P < 0.01
B
I: OPEN IN ANY SAMPLE
II: CONSTITUTIVELY OPEN
III: MEIOSIS-SPECIFIC
IV: MEIOSIS ALL
FAIRE peaks obs/exp
20-30 L
10-20 L
0-10 L
CEN
0-10 R
10-20 R
20-30 R
Distance from centromere (kb)

Figure 5.7.


A
Meiosis-specific
Meiosis ALL
bits
B
tRNA w/ associated open site (P < 0.01)
FAIRE enrichment z-score
YPD
0 hr
1.5 hr
3 hr
- 600
- 400
- 200
+200
+ 400
+ 600
Distance from tRNA annotated start

TABLES

Table 5.1.

	What percentage of open chromatin bases in the samples…					
	YPD (689 kb)	**0hr (667 kb)**	**1.5hr (738 kb)**	**3hr (585 kb)**	**ANY (961 kb)**	**ALL (507 kb)**
	... overlap with the 2,299 kb (1179 hotspots) identified by Buhler *et al* (Threshold 5)					
Experiment	46.6	48.7	47.9	49.8	44.6	51.9
Permuted (SD)	19.6 (2.1)	20.1 (0.9)	20.7 (1.5)	19.9 (1.5)	20.4 (1.1)	21.2 (1)
obs/exp	**2.4***	**2.4***	**2.3***	**2.5***	**2.2***	**2.5***
	… overlap with the 673 kb (258 hotspots) identified by Blitzblau *et al*					
Experiment	16.8	18.3	18.6	20	16.1	20.5
Permuted (SD)	5.7 (0.8)	6.1 (0.8)	6 (0.8)	6.1 (0.9)	6 (0.6)	5.9 (0.6)
obs/exp	**3.0***	**3.0***	**3.1***	**3.3***	**2.7***	**3.5***

* $= P < 0.0001$

SUPPLEMENTARY FIGURES

Figure 5.S1.

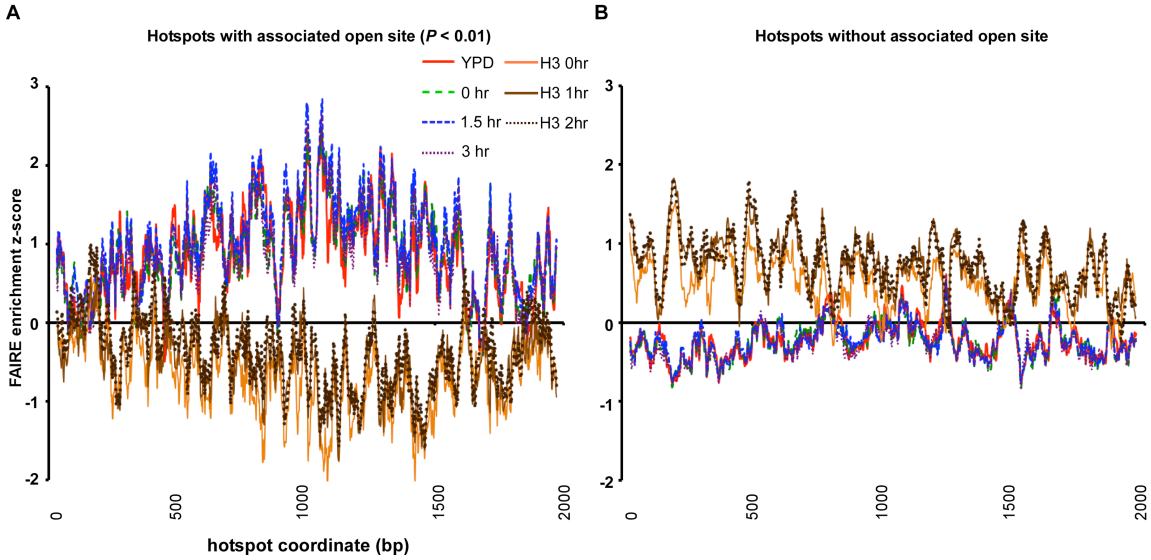
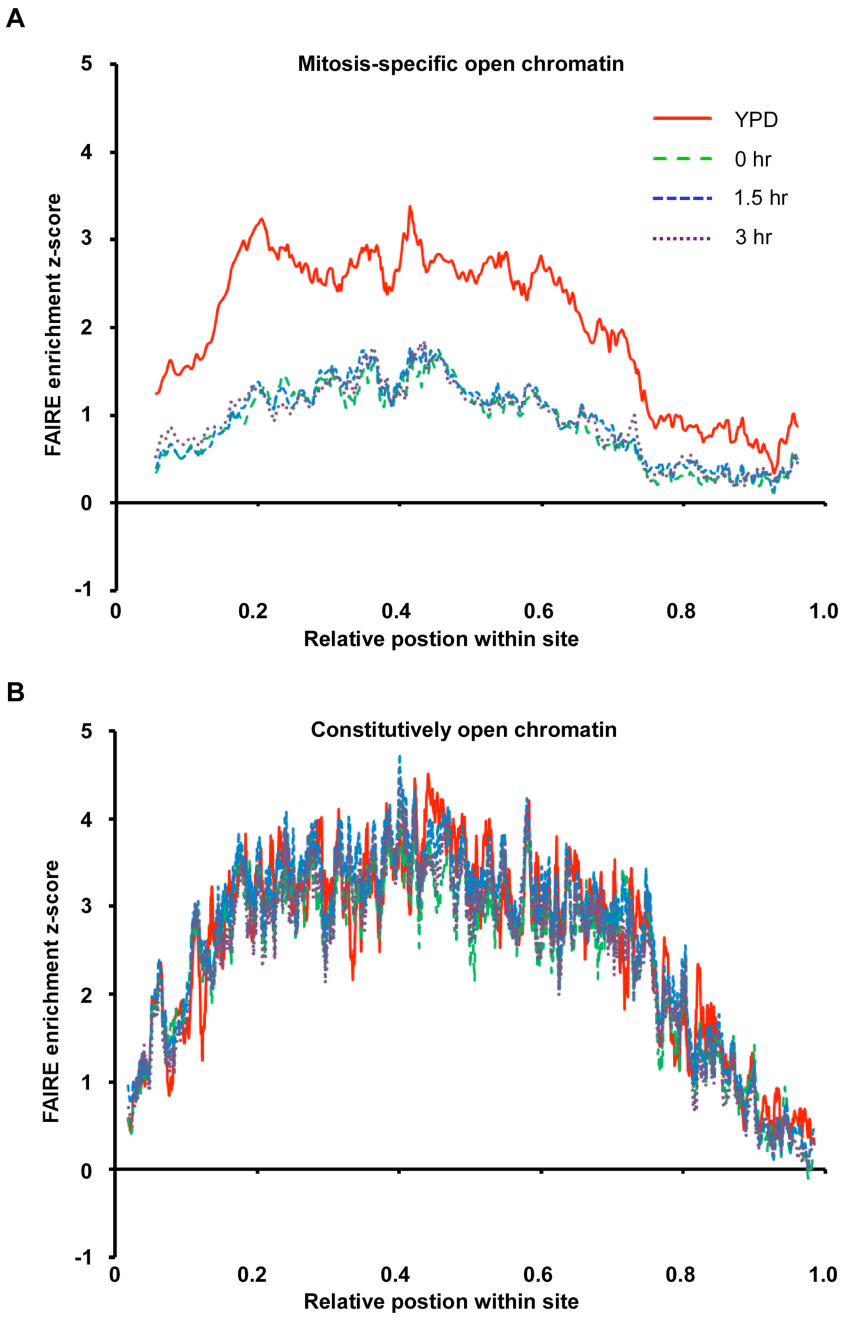


Figure 5.S2.


A
Mitosis-specific open chromatin
YPD
0 hr
1.5 hr
3 hr
FAIRE enrichment z-score
Relative postion within site
B
Constitutively open chromatin
FAIRE enrichment z-score
Relative postion within site

SUPPLEMENTARY TABLES

Table 5.S1.

	Buhler *et al.* 1179 DSB hotspots (T5)	**Blitzblau *et al.* 258 DSB hotspots**
YPD (857 regions)	475	129
intergenic permuted (SD)	270 (12.5)	74 (6.8)
obs/exp	1.8*	1.7*
0 hr (871 regions)	467	132
intergenic permuted (SD)	271 (12.5)	74 (6.8)
obs/exp	1.7*	1.8*
1.5 hr (937 regions)	484	142
intergenic permuted (SD)	290 (13.0)	79 (6.8)
obs/exp	1.7*	1.8*
3 hr (738 regions)	408	130
intergenic permuted (SD)	237 (12.1)	65 (6.6)
obs/exp	1.7*	2.0*
ANY (1204 regions)	591	157
intergenic permuted (SD)	356 (13.9)	95 (7.2)
obs/exp	1.7*	1.7*
ALL (510 regions)	327	101
intergenic permuted (SD)	180 (10.8)	49 (5.9)
obs/exp	1.8*	2.0*

* = $P < 0.0001$

Table 5.S2

	What percentage of open chromatin bases in the samples…					
	YPD (689 kb)	**0hr (667 kb)**	**1.5hr (738 kb)**	**3hr (585 kb)**	**ANY (961 kb)**	**ALL (507 kb)**
	… overlap with the 236 kb (136 total hotspots) identified by Mancera *et al*					
Experiment	5.1	4.4	4.8	4.8	4.5	5.1
Permuted (SD)	2.2 (0.6)	2.0 (0.4)	2.0 (0.4)	2.3 (0.5)	2.1 (0.2)	2.3 (0.3)
obs/exp	**2.4***	**2.2***	**2.4***	**2.0***	**2.1***	**2.3***
	…overlap with the 156 kb (92 crossover hotspots) identified by Mancera *et al*					
Experiment	4.1	3.5	3.7	3.6	3.5	4.1
Permuted (SD)	1.3 (0.5)	1.5 (0.4)	1.4 (0.3)	1.7 (0.6)	1.6 (0.2)	1.7 (0.2)
obs/exp	**3.0***	**2.4***	**2.6***	**2.1***	**2.3***	**2.4***
	... overlap with the 84 kb (74 non-crossover hotspots) identified by Mancera *et al*					
Experiment	1.8	1.7	1.9	2.1	1.7	1.8
Permuted (SD)	0.8 (0.1)	0.6 (0.3)	0.9 (0.3)	0.9 (0.3)	0.7 (0.2)	0.7 (0.4)
obs/exp	**2.2***	**2.7***	**2.3***	**2.3***	**2.4***	**2.5***

* = $P < 0.01$

Table 5.S3

	ANY (1204 regions)	ALL (510)	MEI-SPEC (285)	MEI ALL (151)
tel & subtel	50	21	9	3
Permuted (SD)	48.35 (5.41)	21.49 (4.50)	11.42 (3.44)	6.24 (2.54)
obs/exp	1.03	0.98	0.79	0.61
0-10 kb from subtel border	49	14	15	12
Permuted (SD)	33.01 (4.43)	14.93 (3.74)	7.82 (2.83)	4.33 (2.01)
obs/exp	1.48 *	0.94	1.92 *	2.77 *
10-20 kb from subtel border	39	11	13	9
Permuted (SD)	33.21 (4.42)	15.02 (3.74)	7.86 (2.81)	4.37 (2.10)
obs/exp	1.17	0.73	1.65	2.05 *
20-30 kb from subtel border	33	10	9	4
Permuted (SD)	33.09 (4.50)	14.97 (3.71)	7.86 (2.83)	4.38 (2.10)
obs/exp	1.00	0.67	1.15	0.91
30-40 kb from subtel border	29	11	9	9
Permuted (SD)	33.49 (4.55)	14.96 (3.73)	7.88 (2.83)	4.35 (2.16)
obs/exp	0.87	0.74	1.14	2.07 *
20-30 kb L of centromere	13	6	5	1
Permuted (SD)	16.71 (3.20)	7.56 (2.66)	3.94 (2.03)	2.19 (1.54)
obs/exp	0.78	0.79	1.27	0.46
10-20 kb L of centromere	17	6	3	3
Permuted (SD)	16.77 (3.21)	7.52 (2.68)	3.94 (2.02)	2.19 (1.52)
obs/exp	1.01	0.80	0.76	1.37
0-10 kb L of centromere	23	13	4	0
Permuted (SD)	16.75 (3.23)	7.56 (2.67)	3.94 (2.01)	2.19 (1.51)
obs/exp	1.37	1.72 *	1.02	0.00
CEN	12	10	0	0
Permuted (SD)	1.29 (0.88)	0.80 (0.87)	0.32 (0.55)	0.20 (0.43)
obs/exp	9.30 *	12.5 *	0	0
0-10 kb R of centromere	26	15	3	1
Permuted (SD)	16.57 (3.18)	7.54 (2.63)	3.92 (2.03)	2.15 (1.49)
obs/exp	1.57 *	1.99 *	0.77	0.47
10-20 kb R of centromere	24	12	8	3
Permuted (SD)	16.77 (3.22)	7.49 (2.62)	3.98 (2.03)	2.21 (1.52)
obs/exp	1.43	1.60	2.01	1.36
20-30 kb R of centromere	15	11	2	0
Permuted (SD)	16.89 (3.25)	7.49 (2.63)	3.95 (2.05)	2.19 (1.53)
obs/exp	0.89	1.47	0.51	0.00

* = $P < 0.05$

Table 5.S4.

Chromosome #	**Left subtelomeric border coordinate**	**Right subtelomeric border coordinate**
1	27969	208077
2	8482	807099
3	12283	304358
4	18566	1523235
5	23231	566225
6	15431	253579
7	11730	1076606
8	12847	525392
9	21220	430494
10	19497	727396
11	14485	662927
12	12076	1062917
13	8383	917577
14	8330	772657
15	26975	1071791
16	19079	941132

Table 5.S5.

	Buhler *et al.*	**Blitzblau *et al.***
tRNA coding sequence	49	14
Permuted (SD)	51.3 (6.4)	15.9 (3.9)
tRNA sequence ± 500 bp upstream & downstream	77	21
Permuted (SD)	71.6 (7.1)	21.5 (4.5)

REFERENCES

1. Zickler, D. and N. Kleckner, *Meiotic chromosomes: integrating structure and function.* Annu Rev Genet, 1999. **33**: p. 603-754.

2. Felsenstein, J., *The evolutionary advantage of recombination.* Genetics, 1974. **78**(2): p. 737-56.

3. Gerton, J.L. and R.S. Hawley, *Homologous chromosome interactions in meiosis: diversity amidst conservation.* Nat Rev Genet, 2005. **6**(6): p. 477-87.

4. Keeney, S., C.N. Giroux, and N. Kleckner, *Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family.* Cell, 1997. **88**(3): p. 375-84.

5. Schwacha, A. and N. Kleckner, *Identification of joint molecules that form frequently between homologs but rarely between sister chromatids during yeast meiosis.* Cell, 1994. **76**(1): p. 51-63.

6. Allers, T. and M. Lichten, *Differential timing and control of noncrossover and crossover recombination during meiosis.* Cell, 2001. **106**(1): p. 47-57.

7. Bishop, D.K., et al., *DMC1: a meiosis-specific yeast homolog of E. coli recA required for recombination, synaptonemal complex formation, and cell cycle progression.* Cell, 1992. **69**(3): p. 439-56.

8. Szostak, J.W., et al., *The double-strand-break repair model for recombination.* Cell, 1983. **33**(1): p. 25-35.

9. Sun, H., D. Treco, and J.W. Szostak, *Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site.* Cell, 1991. **64**(6): p. 1155-61.

10. McMahill, M.S., C.W. Sham, and D.K. Bishop, *Synthesis-dependent strand annealing in meiosis.* PLoS Biol, 2007. **5**(11): p. e299.

11. Mancera, E., et al., *High-resolution mapping of meiotic crossovers and non-crossovers in yeast.* Nature, 2008. **454**(7203): p. 479-85.

12. Buhler, C., V. Borde, and M. Lichten, *Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in Saccharomyces cerevisiae.* PLoS Biol, 2007. **5**(12): p. e324.

13. Baudat, F. and A. Nicolas, *Clustering of meiotic double-strand breaks on yeast chromosome III.* Proc Natl Acad Sci U S A, 1997. **94**(10): p. 5213-8.

14. Gerton, J.L., et al., *Inaugural article: global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2000. **97**(21): p. 11383-90.

15. Blitzblau, H.G., et al., *Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres.* Curr Biol, 2007. **17**(23): p. 2003-12.

16. Lichten, M. and A.S. Goldman, *Meiotic recombination hotspots.* Annu Rev Genet, 1995. **29**: p. 423-44.

17. Petes, T.D., *Meiotic recombination hot spots and cold spots.* Nat Rev Genet, 2001. **2**(5): p. 360-9.

18. Wu, T.C. and M. Lichten, *Meiosis-induced double-strand break sites determined by yeast chromatin structure.* Science, 1994. **263**(5146): p. 515-8.

19. Kirkpatrick, D.T., et al., *Control of meiotic recombination and gene expression in yeast by a simple repetitive DNA sequence that excludes nucleosomes.* Mol Cell Biol, 1999. **19**(11): p. 7661-71.

20. Fan, Q.Q. and T.D. Petes, *Relationship between nuclease-hypersensitive sites and meiotic recombination hot spot activity at the HIS4 locus of Saccharomyces cerevisiae.* Mol Cell Biol, 1996. **16**(5): p. 2037-43.

21. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

22. Han, M. and M. Grunstein, *Nucleosome loss activates yeast downstream promoters in vivo.* Cell, 1988. **55**(6): p. 1137-45.

23. Han, M., et al., *Depletion of histone H4 and nucleosomes activates the PHO5 gene in Saccharomyces cerevisiae.* Embo J, 1988. **7**(7): p. 2221-8.

24. Ohta, K., T. Shibata, and A. Nicolas, *Changes in chromatin structure at recombination initiation sites during yeast meiosis.* Embo J, 1994. **13**(23): p. 5754-63.

25. Fan, Q., F. Xu, and T.D. Petes, *Meiosis-specific double-strand DNA breaks at the HIS4 recombination hot spot in the yeast Saccharomyces cerevisiae: control in cis and trans.* Mol Cell Biol, 1995. **15**(3): p. 1679-88.

26. Cao, L., E. Alani, and N. Kleckner, *A pathway for generation and processing of double-strand breaks during meiotic recombination in S. cerevisiae.* Cell, 1990. **61**(6): p. 1089-101.

27. Borde, V., et al., *Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites.* Embo J, 2009. **28**(2): p. 99-111.

28. Hirota, K., et al., *Distinct Chromatin Modulators Regulate the Formation of Accessible and Repressive Chromatin at the Fission Yeast Recombination Hotspot ade6-M26.* Mol Biol Cell, 2008. **19**(3): p. 1162-1173.

29. Hirota, K., et al., *Multiple modes of chromatin configuration at natural meiotic recombination hot spots in fission yeast.* Eukaryot Cell, 2007. **6**(11): p. 2072-80.

30. Ohta, K., et al., *Competitive inactivation of a double-strand DNA break site involves parallel suppression of meiosis-induced changes in chromatin configuration.* Nucleic Acids Res, 1999. **27**(10): p. 2175-80.

31. Shenkar, R., M.H. Shen, and N. Arnheim, *DNase I-hypersensitive sites and transcription factor-binding motifs within the mouse E beta meiotic recombination hot spot.* Mol Cell Biol, 1991. **11**(4): p. 1813-9.

32. Borde, V., T.C. Wu, and M. Lichten, *Use of a recombination reporter insert to define meiotic recombination domains on chromosome III of Saccharomyces cerevisiae.* Mol Cell Biol, 1999. **19**(7): p. 4832-42.

33. Chu, S., et al., *The transcriptional program of sporulation in budding yeast.* Science, 1998. **282**(5389): p. 699-705.

34. Hogan, G.J., C.K. Lee, and J.D. Lieb, *Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters.* PLoS Genet, 2006. **2**(9): p. e158.

35. Nagy, P.L., et al., *Genomewide demarcation of RNA polymerase II transcription units revealed by physical fractionation of chromatin.* Proc Natl Acad Sci U S A, 2003. **100**(11): p. 6364-9.

36. Giresi, P.G., et al., *FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin.* Genome Res, 2007. **17**(6): p. 877-85.

37. Eeckhoute, J., et al., *Cell-type selective chromatin remodeling defines the active subset of FOXA1-bound enhancers.* Genome Res, 2009. **19**(3): p. 372-80.

38. Polach, K.J. and J. Widom, *Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation.* J Mol Biol, 1995. **254**(2): p. 130-49.

39. Lee, C.K., et al., *Evidence for nucleosome depletion at active regulatory regions genome-wide.* Nat Genet, 2004. **36**(8): p. 900-5.

40. Zweig, M.H. and G. Campbell, *Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine.* Clin Chem, 1993. **39**(4): p. 561-77.

41. Malone, R.E., et al., *A meiotic gene conversion gradient opposite to the direction of transcription.* Nature, 1992. **359**(6391): p. 154-5.

42. Goldway, M., et al., *A short chromosomal region with major roles in yeast chromosome III meiotic disjunction, recombination and double strand breaks.* Genetics, 1993. **133**(2): p. 159-69.

43. Wu, T.C. and M. Lichten, *Factors that affect the location and frequency of meiosis-induced double-strand breaks in Saccharomyces cerevisiae.* Genetics, 1995. **140**(1): p. 55-66.

44. Kugou, K., et al., *Rec8 Guides Canonical Spo11 Distribution Along Yeast Meiotic Chromosomes.* Mol Biol Cell, 2009.

45. White, M.A., et al., *DNA-binding protein RAP1 stimulates meiotic recombination at the HIS4 locus in yeast.* Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9755-9.

46. White, M.A., M. Dominska, and T.D. Petes, *Transcription factors are required for the meiotic recombination hotspot at the HIS4 locus in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1993. **90**(14): p. 6621-5.

47. Kirkpatrick, D.T., Q. Fan, and T.D. Petes, *Maximal stimulation of meiotic recombination by a yeast transcription factor requires the transcription activation domain and a DNA-binding domain.* Genetics, 1999. **152**(1): p. 101-15.

48. Zickler, D. and N. Kleckner, *The leptotene-zygotene transition of meiosis.* Annu Rev Genet, 1998. **32**: p. 619-97.

49. Rockmill, B. and G.S. Roeder, *Telomere-mediated chromosome pairing during meiosis in budding yeast.* Genes Dev, 1998. **12**(16): p. 2574-86.

50. Barton, A.B., et al., *A function for subtelomeric DNA in Saccharomyces cerevisiae.* Genetics, 2003. **165**(2): p. 929-34.

51. Barton, A.B., et al., *Meiotic recombination at the ends of chromosomes in Saccharomyces cerevisiae.* Genetics, 2008. **179**(3): p. 1221-35.

52. Su, Y., A.B. Barton, and D.B. Kaback, *Decreased meiotic reciprocal recombination in subtelomeric regions in Saccharomyces cerevisiae.* Chromosoma, 2000. **109**(7): p. 467-75.

53. Lambie, E.J. and G.S. Roeder, *Repression of meiotic crossing over by a centromere (CEN3) in Saccharomyces cerevisiae.* Genetics, 1986. **114**(3): p. 769-89.

54. Chen, S.Y., et al., *Global Analysis of the Meiotic Crossover Landscape.* Dev Cell, 2008.

55. Lamb, N.E., et al., *Susceptible chiasmate configurations of chromosome 21 predispose to non-disjunction in both maternal meiosis I and meiosis II.* Nat Genet, 1996. **14**(4): p. 400-5.

56. Koehler, K.E., et al., *Spontaneous X chromosome MI and MII nondisjunction events in Drosophila melanogaster oocytes have different recombinational histories.* Nat Genet, 1996. **14**(4): p. 406-14.

57. D'Ambrosio, C., et al., *Identification of cis-acting sites for condensin loading onto budding yeast chromosomes.* Genes Dev, 2008. **22**(16): p. 2215-27.

58. Goyon, C. and M. Lichten, *Timing of molecular events in meiosis in Saccharomyces cerevisiae: stable heteroduplex DNA is formed late in meiotic prophase.* Mol Cell Biol, 1993. **13**(1): p. 373-82.

59. Liu, J., T.C. Wu, and M. Lichten, *The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate.* Embo J, 1995. **14**(18): p. 4599-608.

60. de Massy, B., V. Rocco, and A. Nicolas, *The nucleotide mapping of DNA double-strand breaks at the CYS3 initiation site of meiotic recombination in Saccharomyces cerevisiae.* Embo J, 1995. **14**(18): p. 4589-98.

61. Kundu, T.K., Z. Wang, and R.G. Roeder, *Human TFIIIC relieves chromatin-mediated repression of RNA polymerase III transcription and contains an intrinsic histone acetyltransferase activity.* Mol Cell Biol, 1999. **19**(2): p. 1605-15.

62. Scott, K.C., S.L. Merrett, and H.F. Willard, *A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains.* Curr Biol, 2006. **16**(2): p. 119-29.

63. Schmidt, C.K., N. Brookes, and F. Uhlmann, *Conserved features of cohesin binding along fission yeast chromosomes.* Genome Biol, 2009. **10**(5): p. R52.

64. Wahls, W.P., E.R. Siegel, and M.K. Davidson, *Meiotic recombination hotspots of fission yeast are directed to loci that express non-coding RNA.* PLoS ONE, 2008. **3**(8): p. e2887.

65. Donze, D. and R.T. Kamakaka, *RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae.* Embo J, 2001. **20**(3): p. 520-31.

66. Biswas, M., et al., *Limiting the extent of the RDN1 heterochromatin domain by a silencing barrier and Sir2 protein levels in Saccharomyces cerevisiae.* Mol Cell Biol, 2009. **29**(10): p. 2889-98.

67. Borde, V., et al., *Association of Mre11p with double-strand break sites during yeast meiosis.* Mol Cell, 2004. **13**(3): p. 389-401.

68. Robine, N., et al., *Genome-wide redistribution of meiotic double-strand breaks in Saccharomyces cerevisiae.* Mol Cell Biol, 2007. **27**(5): p. 1868-80.

69. Yamada, T., et al., *Roles of histone acetylation and chromatin remodeling factor in a meiotic recombination hotspot.* Embo J, 2004. **23**(8): p. 1792-803.

70. Zalevsky, J., et al., *Crossing over during Caenorhabditis elegans meiosis requires a conserved MutS-based pathway that is partially dispensable in budding yeast.* Genetics, 1999. **153**(3): p. 1271-83.

71. Chen, C.* and H. Ma, *Plant meiocyte collection for genome research*, in *Manuscript; * correspondence for material distribution*. 2007.

72. Ausubel, F.M., *Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology*. 3rd ed. 1997, New York ; Chichester: Wiley. 1 v. (various pagings).

73. O'Geen, H., et al., *Comparison of sample preparation methods for ChIP-chip assays.* Biotechniques, 2006. **41**(5): p. 577-80.

74. Buck, M.J., A.B. Nobel, and J.D. Lieb, *ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data.* Genome Biol, 2005. **6**(11): p. R97.

75. Giardine, B., et al., *Galaxy: a platform for interactive large-scale genome analysis.* Genome Res, 2005. **15**(10): p. 1451-5.

76. Shapiro, S. and M. Wilk, *An analysis of variance test for normality (complete samples).* Biometrika, 1965. **52**(3 & 4): p. 591-611.

77. Sprinthall, R.C., *Basic Statistical Analysis*. 8 ed. 2003: Allyn & Bacon.

78. Liu, X., D.L. Brutlag, and J.S. Liu, *BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes.* Pac Symp Biocomput, 2001: p. 127-38.

79. Crooks, G.E., et al., *WebLogo: a sequence logo generator.* Genome Res, 2004. **14**(6): p. 1188-90.

CHAPTER 5 ADDENDUM

Specific contributions by L.E.B.

S.E.H. and myself conducted the experiments that led to the identification of regions of open chromatin in a meiotic time course in *S. cerevisiae*. S.E.H. conducted the time-course and formaldehyde-treated each sample before storage. I prepared and fluorescent-labeled the sample and reference DNA for microarray competitive hybridization. S.E.H. performed the hybridization, photographed the chips, and applied algorithms resulting in FAIRE ratios at each probe and significant regions of FAIRE enrichment. S.E.H. and myself conducted data analysis. I categorized all peaks into the groupings used throughout the paper and performed overlap analysis with published DSB hotspots. I also conducted the centromere and telomere analysis. S.E.H conducted motif analysis and produced all moving average plots. S.E.H and myself jointly designed and performed ROC curve and tRNA analysis. I prepared the manuscript, which was reviewed and edited by S.E.H, G.P.C. and J.D.L.

Chapter 6

Conclusions and Future Study

Luke E. Berchowitz

I. Using the FTL system to study recombination in *A. thaliana*

By monitoring the segregation of cell autonomous fluorescent proteins within pollen tetrads of *A. thaliana*, our lab is able to quickly generate large datasets that can be used to analyze many different facets of meiotic recombination. My early work in the lab, along with several other lab members, focused on the development of the FTL system to study CO rates, CO interference, and GCs. We have mapped and quality-tested a collection of transgenic markers encoding fluorescent proteins with a density of ~ 1 insertion/1.14 Mb in the *A. thaliana* genome which is equivalent to ~ 1 insertion every 5.7 cM (assuming an genomic average of 5 cM/Mb). The FTL system is currently the most time-efficient method to generate CO frequency datasets in any higher eukaryotic model system. Below I discuss the main conclusions regarding the study of meiotic recombination using the FTL system and offer insights into future applications.

A. The effects of development and environment on recombination

Using the FTL system, we demonstrated that developmental and environmental factors have a significant effect on CO frequency. We found that CO frequency increased with increasing temperature and that these differences were significant in the two genetic intervals tested [1]. This reinforced studies done in both plants and animals that also demonstrated increasing temperature increases CO frequency [2, 3]. Additionally, we showed that flowers collected from the secondary or tertiary bolts have higher CO rates than those collected from the primary bolt. However, along the primary bolt, flower age did not affect CO frequency.

This set of observations allowed us to establish standardized procedures for data collection so that data collected in different experiments can be compared. It is necessary to keep growth conditions, particularly temperature, constant among populations that are to be compared. It is also necessary to score only flowers from the same (typically the primary) bolt and although it is unlikely that flower number affects CO frequency, we prefer only to score flowers 5-30 as we have clearly demonstrated that flowers in this range have comparable CO frequencies.

Moving forward, it will be interesting to determine how temperature and developmental stage affect recombination frequency. An experiment that would address this is a forward genetic screen for mutants that do not affect the recombination rates of untreated plants but either a) do not have altered recombination rates at elevated temperatures or b) do not have altered recombination rates on secondary/tertiary bolts. One hypothesis to explain increased rates of recombination due to high temperature and flower architecture is that the plant senses extrinsic signals such as temperature and these signals are transduced to elicit a response in the meiotic cells. The response could take the form of transcription, translation, RNA turnover, protein turnover or protein modifications in the recombination machinery. If this is the case, our screen would likely pull out genes involved in the signal and response, which could encompass a diverse group. Another experiment that would help understand how the plant conditionally alters its recombination rate would be to conduct expression microarray analysis of meioses at different temperatures (or on different bolts). Although many of the changes that plants undergo with these treatments will be unrelated to recombination, this experiment could

be used in conjunction with the screen described above and identify genes that could be involved in regulation of recombination.

B. Further study of interference using the FTL system

Using the FTL system we developed a three-color assay for CO interference in *A. thaliana*. By crossing multiple lines together, it is possible to calculate CO interference almost anywhere in the *A. thaliana* genome. Initially, we used the COI [4] and NPD ratio [5] methods to quantify CO interference [1], but subsequently we have used the method of Malkova *et al.* in which the map distance of one interval is measured when there are CO(s) in a second adjacent interval and compared to the map distance calculated when there are no adjacent COs [6]. This method is advantageous because it can take into account triple and quadruple COs, which the previous methods cannot. Additionally, interference ratios generated by the Malkova *et al.* method can be compared using relatively simple statistics [7].

To study certain aspects of interference such as interference over large distances and decay pattern without the complications of ≥ triple exchanges, it is useful to analyze intervals that are not adjacent to one another. To this end, we are in the process of developing a fourth color in our FTL system that can be differentiated from the other three currently available. Four segregating markers allow for a type of interference calculation called S_4 analysis, which allows the quantification of interference across intervals that are not necessarily adjacent [8]. Triple exchanges can be kept to a minimum using S_4 analysis since the genetic distance of the test intervals can be small while the distance separating the two test intervals can be any size. Defining the decay pattern of

interference will give insight into how the interference signal is imposed. This is could be an important piece of the puzzle while developing and testing interference models because most models make predictions as to the strength to distance ratio of the interference signal. Finally, defining the decay pattern will give clues as to where, relative to CO sites, on the chromosomes to look for the interference signal by cytological methods.

C. Measuring GC using the FTL system

A modified version of the FTL system can be used to quantify GC frequency by monitoring the frequency of 3:1 segregation patterns at a single locus [1]. We plan to use the FTL system to determine the ratio of GCs that are associated with COs to those that are not. Meiotic DSBs can be repaired by either the DSBR pathway, which (given polymorphisms at the site of exchange) is capable of producing GCs associated with COs, assuming that *A. thaliana* is like *S. cerevisiae* in that the DSBR pathway produces COs rather than NCOs [9, 10], or the SDSA pathway, which only is capable of producing GCs without associated COs [11]. Defining the ratio of GCs with and without associated COs will give insight into the relative use of the two repair pathways in *A. thaliana*. To do this, we will create two small (less than 3 cM) adjacent intervals delineated by three colors with the middle locus being GC test locus (**Figure 1a**). The frequency of outer marker recombination accompanied by GC at the central locus will serve as a measure of CO with GC. Conversely, GC at the central locus without exchange of the outer makers is an indication of a NCO GC. Measurements of this type at multiple sites will provide a genome-wide assessment. NCO GCs with another CO within one of the two intervals not

related to the GC repair would complicate the analysis, but this is a rare event if the two intervals are very small.

Once we determine the ratio of NCO GCs to CO GCs, we can test whether NCO GCs and COs exhibit positive interference by constructing two adjacent intervals delineated by three colors with the most distal marker being a GC test locus (**Figure 1b**). Whether or not NCO GCs and COs exhibit positive interference is a controversial topic (reviewed in Chapter 1) and many interference models make predictions related to this phenomenon. This arrangement of markers will enable us to assay whether or not a GC at the distal locus interferes with CO frequency in the proximal adjacent interval. Since the goal is to determine if NCO GCs exhibit positive interference, a complication of this assay is that a portion of the interference that we will observe between GCs and adjacent COs will be due to CO GCs. However, determination of the ratio of NCO GC to CO GC will allow us to correct our experimentally derived interference value by removing the interference expected from CO GCs.

II. *A. thaliana* has at least two pathways for producing COs, one of which is interference-insensitive.

In *S. cerevisiae* and mammals, genetic and biochemical experiments have shown that there are at least two pathways for producing COs: a primary interference-sensitive pathway and a secondary interference-insensitive pathway [12-14]. This is different from meiosis in *D. melanogaster* and *C. elegans* where all COs are interference-sensitive or *S. pombe* where all COs are not subject to interference. The idea that *A. thaliana* has at least two pathways had been proposed based on three lines of evidence: 1. mathematical

modeling of CO distribution in *A. thaliana* on all chromosomes suggests that a distribution that includes 15-20% non-interfering COs fits the experimental data best [15, 16]. 2. *msh4* mutants severely reduce, but do not eliminate COs and the remaining chiasmata appeared to be distributed in a non-interference pattern [17]. 3. BLAST analysis showed that *A. thaliana* has at least one *MUS81* homolog [18, 19]. However, the first published report on the genetic analysis of the *Atmus81* mutant suggested that *AtMUS81* was not involved in meiotic recombination [18, 19].

A. *AtMUS81* mediates the secondary CO-insensitive pathway

My research shows that *Atmus81* mutants are, in fact, significantly defective for CO formation (10-15%) and that the remaining COs had significantly stronger interference than wild-type, which would be expected when interference-insensitive COs are diminished [20]. Additionally, the reduction of COs in *Atmus81* was additive with the CO defect in *Atmsh4*, indicating that the two genes work in different pathways, a result which was later confirmed by cytology [21]. Interestingly, the *Atmus81*; *Atmsh4* double mutant does not abolish COs completely, suggesting that *A. thaliana* has yet a third mechanism for generating COs outside of the primary and secondary pathways. Holloway *et al.* demonstrated a similar result in mouse [22]. No genes that mediate the tertiary COs present in the *Atmus81*; *Atmsh4* double mutant have been identified in any organism. A screen designed to target these genes could be done by EMS mutagenizing seeds produced from *Atmus81*/*Atmus81*; *Atmsh4*/+ individuals. We would screen for M2 plants that are *Atmus81*/*Atmus81*; *Atmsh4*/*Atmsh4* and have no residual COs. Since this phenotype would be laborious to score in large mapping population, whole genome

sequencing of candidates would likely be a more efficient means of identifying mutations, as opposed to standard F2 mapping.

My work on *AtMUS81* clearly demonstrates the presence of the hypothesized secondary interference-insensitive CO pathway in Arabidopsis that is mediated by *AtMUS81*. Subsequent *in vitro* biochemical research has shown that AtMus81-AtEme1 has the ability to resolve HJs [23]. I am currently working on the genetic characterization of the two *EME1* homologs in *A. thaliana*: *AtEME1a* and *AtEME1b*, both of which produce a protein that heterodimerizes with AtMus81 [23].

B. What is the mechanistic difference between primary Msh4-Msh5 COs and secondary Mus81-Eme1 COs?

It has been demonstrated that there are at least two CO pathways, but the question remains as to the mechanistic difference between the two. Of particular interest is why one pathway is subject to interference while the other is not. Evidence suggests that the 'decision' to repair a DSB as a CO or NCO is made directly after break formation because single-end intermediates accumulate in *zmm* mutants, which is the step just after resection of the breaks [24]. Thus, both Msh4-Msh5 and Mus81-Eme1 likely act after strand invasion in the CO designated pathway. Interestingly, interference-sensitive COs in *S. cerevisae* arise from double HJ intermediates which are likely mediated by Msh4-Msh5 [12], while COs in *S. pombe* arise from single HJ intermediates, which are very likely to be resolved by Mus81-Eme1 [25]. One hypothesis to explain the mechanistic difference between primary and secondary COs is that primary COs are produced from Msh4-Msh5 mediated resolution of DSBs that are processed via a dHJ repair pathway

while secondary COs are produced from Mus81-Eme1 mediated resolution of DSBs that are processed via a single HJ pathway.

Recent evidence from *A. thaliana* suggests that Mus81-Eme1 is present at all DSB sites [21]. Furthermore, it has been shown that Mus81-Eme1, in addition to its role as a 3' flap endonuclease and nicked HJ resolvase, can resolve aberrant multi-strand CO intermediates [14, 26]. Taken together, these results suggest a 'toolbox' hypothesis for CO formation in *S. cerevisiae* and *A. thaliana* in which all CO-related proteins are recruited to all DSBs. The majority of CO designated DSBs follow the primary double HJ pathway, which are resolved by ZMM proteins. Secondary COs arise from Mus81-Eme1 CO resolution of aberrant multi-strand or single HJ molecules that cannot be resolved by the ZMM pathway. I am postulating that all Mus81-Eme1 COs are produced from aberrant and single HJs, however, it is possible that some Mus81-Eme1 COs are produced from dHJ structures or non-HJ intermediates.

The mechanistic difference between primary and secondary COs may be explained by supposing that primary CO reactions occur first and initiate an 'interference signal' (discussed below), while secondary COs take longer to process, as they must first be bypassed by the primary pathway. In this model, secondary COs do not produce an interference signal, and are resolved after the interference signal has been imposed. Additionally, aberrant joint molecules resolved by Mus81-Eme1 would occur at a low frequency and their probability would be influenced by many factors including sequence and chromatin context at each CO designated DSB, and would explain the observed fraction of COs that are essentially randomly distributed.

III. How is interference imposed?

Many models (discussed in chapter 1) have been proposed to explain the mechanism that results in interference, but the answer to this question remains unanswered. Although it remains unclear whether or not and to what degree DSBs interfere with one another, some evidence suggests that there are at least two levels of interference: one at the DSB-DSB level and another at the CO (and possibly NCO)-CO level [27-30].

I propose that DSB-DSB interference acts on a smaller scale and is mediated by two main factors. First, the pattern of Spo11 distribution may be established by the meiosis-specific cohesin subunit Rec8. Rec8 has been shown to colocalize with Spo11 in *S. cerevisiae*, and *rec8*Δ mutants exhibit a drastic alteration in Spo11 distribution [31]. It has been speculated that Rec8 not only provides landmarks along the chromosomal axes that guide the distribution of Spo11, but it is also responsible for the transition of Spo11 from the axes to the loops, where breaks are subsequently formed [31]. Rec8 could preferentially bind certain locations within the chromatin context of the chromosomal axis with intervening DNA loops, resulting in a minimum physical distance between subunits. This idea requires that only one DSB could be formed at each site of Rec8 localization, which results in the minimization of closely spaced DSBs. This creates a pattern of DSB-DSB interference on a small scale. Secondly, DSBs could recruit limiting break forming factors away from nearby sites [28].

To explain CO-CO interference in *S. cerevisiae* and *A. thaliana*, I propose a signal propagation model [32] with added elements from the counting model [8, 33]. In this model, stabilization of SEIs by ZMM proteins results in an interference signal that propagates down the chromosomal axial elements. A signal propagated down the axis

rather than the DNA itself is attractive because it allows interference to act over very large distances of linear DNA by varying the DNA loop/axis ratio. Importantly, the SC itself is not required for interference in this model, just the axis.

What is the nature of the interference signal? A model based on protein modification that results in the inhibition of primary-pathway crossing over at sites where the signal spreads is intriguing. The modified protein could be a cohesin, histone, or even an axial element protein, but Rec8 is a particularly interesting target since phosphorylation of Rec8 is important for many of its meiotic roles. Two *rec8* mutants with mutations at multiple phosphorylation sites exhibit disrupted synapsis have delayed production of mature recombinants [34]. Since Rec8 phosphorylation is required for recombination in *S. cerevisiae*, the protein modification could be a de-phosphorylation of Rec8 that disables crossing over. Another possibility is de-methylation of H3K4me3, which has been shown to mark sites of meiotic recombination [35].

In this model, there are a number of secondary-pathway COs that are created that are randomly distributed and not affected by the interference signal. These COs are described by the parameter 'm' in the counting model [36]. These COs would be dependent on Mus81-Eme1 and as described above, could be produced by the resolution of aberrant JMs that escaped the ZMM machinery. Mus81-Eme1 would act later in meiotic prophase to 'clean up' these aberrant JMs and CO resolution would be independent of CO interference. Furthermore, it is possible the residual COs in the *zmm*, *mus81* double mutants are produced by a backup mechanism for DSB repair that is not active in wild type cells.

IV. Using FAIRE to study the relationship between open chromatin and recombination in *S. cerevisiae*

We used FAIRE to investigate the patterns of chromatin organization in a synchronized early meiotic time course in *S. cerevisiae*. FAIRE samples were competitively hybridized on a tiling microarray with a genomic DNA reference to determine regions of the genome that exhibited significant open chromatin. Below I discuss the main conclusions of this study as well as follow-up experiments to further understand how chromatin structure influences meiotic processes.

A. The association between DSB hotspots and open chromatin

Several locus-by-locus studies in *S. cerevisiae* and *S. pombe* have demonstrated that meiotic recombination hotspots occur in regions of open chromatin, but the degree to which this relationship held at the genomic level had not been determined. We found that although open chromatin exhibited a significant association with DSB hotspots, some hotspots occur in open chromatin (class I) while a separate class of hotspots occur in regions that are not more open than the genomic average (class II). To refine our understanding of the factors that are required for hotspot activity, it will be of interest to determine features that distinguish class I hotspots from class II hotspots. To compare histone modifications between the two classes of hotspots, one could use ChIP-chip with antibodies that recognize different modified histones. This experiment has been conducted for H3K4me3, which marks most hotspots, but does not distinguish the two classes from one another [35]. Another possibility is that genes within class II hotspots exhibit higher rates of transcription than genes not occurring in hotspots but with similar

levels of chromatin openness. Using expression microarrays to compare transcription rates of genes occurring in class II hotspots to genes that do not occur in hotspots but have similar FAIRE profiles could test this hypothesis.

We found that the degree of openness was predictive of open chromatin, but only if the region was open in at least one meiotic sample. Degree of openness was not predictive of hotspots in mitosis-specific regions. Although degree of openness is not predictive in mitosis-specific regions, these regions do exhibit a significant association with DSB hotspots. We interpret this to mean that some hotspots that are established during mitotic growth, which could be due to transcription factor binding, do not require large amounts of open chromatin and the degree to which those regions are open is immaterial. However, regions that stay open during meiosis are more likely to recruit break-forming complexes with frequency proportional to degree of openness. To test this idea, one could create a series of *dmc1* strains harboring 2 kb insertions (at the same locus) that exclude nucleosomes to different degrees. In each of these strains, one would test DSB activity in the inserted regions by isolating ssDNA five hours into meiosis (as in Buhler et al. [37]) and then using qPCR to quantify ssDNA enrichment at the sites. Our idea predicts that ssDNA enrichment will be proportional to nucleosome exclusion.

Mancera *et al.* used a genome tiling approach to create high-resolution recombination maps in *S. cerevisiae* [38]. They presented evidence that some hotspots can be categorized into those that are more likely to form COs and those that are more likely to form NCOs. We did not find evidence that open chromatin was associated to a greater degree with either of the classes of hotspots. We interpret this to mean that open chromatin is important for establishment of DSB hotspots, but does not influence the

repair pathways. However, the list of CO and NCO hotspots is incomplete and this question could be answered more thoroughly as more hotspot classification data becomes available.

B. Analysis of chromatin profiles at particular genomic features

We found that both regions associated with tRNA genes and subtelomeric border regions exhibited significant enrichment of meiosis-specific open chromatin. The biological significance of meiosis-specific open chromatin at tRNA genes is not obvious to us, but non-coding RNAs have been shown to be associated with DSB hotspots in *S. pombe* [39]. We proposed that meiosis-specific opening of chromatin at the subtelomeric border regions could play a role in homolog pairing. Pairing has been hypothesized to initiate at telomeric regions, but telomeric and subtelomeric regions contain mostly repetitive and non-unique sequences making a homology search using these regions difficult. Meiosis-specific opening at the unique euchromatic border regions could provide a substrate for an effective homology search. To test this hypothesis one could analyze pairing in a mutant that was defective for meiosis-specific open chromatin at subtelomeric border regions. Since the open chromatin in these regions is induced rather than constitutive, a good candidate for this would be a temperature conditional *swi1* mutant since *swi1* mutants are defective for meiotic entry. If a conditional *swi1* mutant can undergo meiosis but cannot open the chromatin at subtelomeric borders, our hypothesis predicts that this mutant will be pairing defective.

We found that pericentromeric and centromeric DNA exhibited constitutive enrichment of open chromatin. 12/16 centromeres exist within a region of open

chromatin and 10/16 exist in a region open in all samples. This result was contrary to our expectations because many proteins associate with centromeres and these regions have drastically reduced rates of recombination [40]. However, recent DSB mapping has shown that pericentromeric regions do not have decreased DSB formation. The enrichment of centromere-associated regions by FAIRE could be due to FAIRE's propensity to remove histone-associated DNA to a greater degree than protein-DNA interactions that do not include histones. Although we know that centromere-associated proteins are abundant, nucleosome occupancy in these regions could be sparse. Another explanation could be that the histone variant found at the centromere is somehow less crosslinkable.

To resolve the apparent paradox that DSBs are normal yet recombination is suppressed at the 50 kb regions surrounding the centromeres, it has been suggested that meiotic DSBs in pericentromeric regions are repaired using sister chromatids rather than by homologous recombination [41]. While it has been historically difficult to study recombination between sister chromatids since sisters have identical DNA, an assay to study this phenomenon has been developed [42]. This assay relies on the incorporation of BrdU specifically during pre-meiotic S-phase to differentiate sisters. Then after meiosis, one can denature the sisters and treat with fluorescent antibodies to BrdU to detect sister chromatid exchanges. In this context the immunostaining procedure will include labeled antibodies that recognize centromeres. To test if sister chromatid exchanges occur at increased rates in close proximity to the centromeres, one compares the rate of sister chromatid exchange near centromeres to that of a control genomic region of comparable

size. This assay may be more feasible in organisms that have larger chromosomes such as hamster ovary cell lines or lily.

C. Future meiotic applications of FAIRE

FAIRE is a low-cost system for identifying regions of open chromatin in a meiotic time course. Regions identified by FAIRE are good candidates for hotspots of meiotic recombination. In *S. cerevisiae*, regions most open during meiosis are most likely to predict the site of a DSB hotspot and it will be interesting to see if this holds true in other systems. FAIRE is best used in systems in which tiling microarrays are available, but organisms with a published genomic sequence can also be used by sequencing DNA isolated after crosslinking and phenol extraction. FAIRE can be used to predict hotspots in organisms where ssDNA mapping is not available *i.e.* those without mutations that arrest the cell after ssDNA has been made.

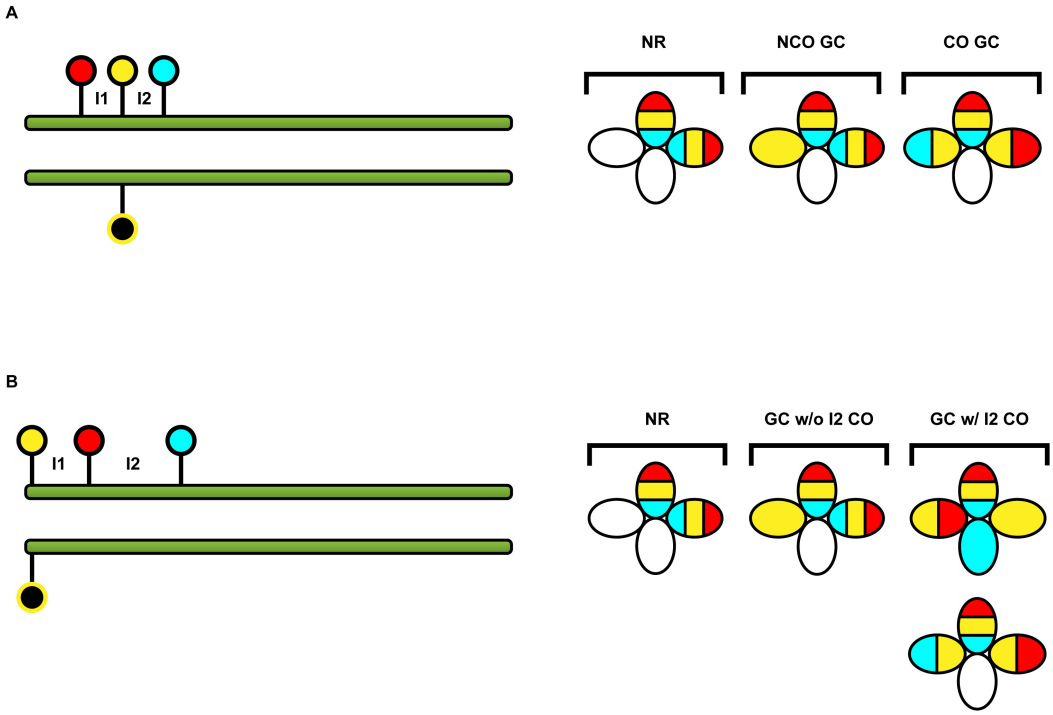
In organisms such as *C. elegans*, in which cells at each meiotic stage can be identified and isolated [43], FAIRE could be used to create stage-specific meiotic prophase open chromatin maps. However, it will be difficult to isolate the amount of tissue required at each stage. FAIRE could also be used to determine meiotic open chromatin in *A. thaliana*, where meiocytes can be quickly and cheaply isolated by the thousands [44].

FIGURE LEGENDS

Figure 6.1. Using the FTL system to assay properties of GCs **A.** Marker setup used to determine the ratio of GCs produced as the result of CO repair to GCs produced as the result of NCO repair at a particular locus. In this genotype, scorable individuals have one chromosome containing three closely linked FTL markers and the other chromosome containing a point-mutated (GC scorable) allele of the middle marker. GCs at the middle marker can be segregated into those that occur with or without outer marker recombination. Note that not all possibilities of fluorescent proteins within the tetrads are present in this figure. I1 = interval 1 and I2 = interval 2. **B.** Marker setup used to determine if GCs not associated with a CO interfere with an adjacent interval. In this genotype, scorable individuals have one chromosome containing three linked FTL markers and the other chromosome containing a point-mutated (GC-scorable) allele of the distal marker. GCs are assayed at the distal locus and simultaneous COs are monitored at the adjacent interval (I2). Note that not all possibilities of fluorescent proteins within the tetrads are present in this figure. I1 = interval 1 and I2 = interval 2.

FIGURES

Figure 6.1.



REFERENCES

1. Francis, K.E., et al., *A pollen tetrad-based visual assay for meiotic recombination in Arabidopsis.* PNAS, 2007. **104**(10): p. 3913-3918.

2. Berkemeier, J. and G. Linnert, *Effects of High and Low Temperature on Meiotic Recombination in Normal and Oligochiasmatic Lines of Vicia-Faba.* Biol Zentralb, 1987. **106**: p. 219-230.

3. Stern, C., *An Effect of Temperature and Age on Crossing-Over in the First Chromosome of Drosophila Melanogaster.* Proc Natl Acad Sci U S A, 1926. **12**(8): p. 530-2.

4. Sturtevant, A.H., *The behavior of chromosomes as studied through linkage.* Z. Induct. Abstammungs-Vererbungsl, 1915. **13**: p. 234-287.

5. Papazian, H.P., *The Analysis of Tetrad Data.* Genetics, 1952. **37**(2): p. 175-88.

6. Malkova, A., et al., *Gene conversion and crossing over along the 405-kb left arm of Saccharomyces cerevisiae chromosome VII.* Genetics, 2004. **168**(1): p. 49-63.

7. Getz, T.J., et al., *Reduced mismatch repair of heteroduplexes reveals "non"-interfering crossing over in wild-type Saccharomyces cerevisiae.* Genetics, 2008. **178**(3): p. 1251-69.

8. Foss, E., et al., *Chiasma interference as a function of genetic distance.* Genetics, 1993. **133**(3): p. 681-91.

9. Szostak, J.W., et al., *The double-strand-break repair model for recombination.* Cell, 1983. **33**(1): p. 25-35.

10. Sun, H., D. Treco, and J.W. Szostak, *Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site.* Cell, 1991. **64**(6): p. 1155-61.

11. Paques, F. and J.E. Haber, *Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae.* Microbiol Mol Biol Rev, 1999. **63**(2): p. 349-404.

12. de los Santos, T., et al., *The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast.* Genetics, 2003. **164**(1): p. 81-94.

13. de los Santos, T., et al., *A role for MMS4 in the processing of recombination intermediates during meiosis in Saccharomyces cerevisiae.* Genetics, 2001. **159**(4): p. 1511-25.

14. Jessop, L. and M. Lichten, *Mus81/Mms4 endonuclease and Sgs1 helicase collaborate to ensure proper recombination intermediate metabolism during meiosis.* Mol Cell, 2008. **31**(3): p. 313-23.

15. Lam, S.Y., et al., *Crossover interference on nucleolus organizing region-bearing chromosomes in Arabidopsis.* Genetics, 2005. **170**(2): p. 807-12.

16. Copenhaver, G.P., E.A. Housworth, and F.W. Stahl, *Crossover interference in Arabidopsis.* Genetics, 2002. **160**(4): p. 1631-9.

17. Higgins, J.D., et al., *The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in Arabidopsis.* Genes Dev, 2004. **18**(20): p. 2557-70.

18. Hartung, F., et al., *The role of AtMUS81 in DNA repair and its genetic interaction with the helicase AtRecQ4A.* Nucleic Acids Res, 2006. **34**(16): p. 4438-48.

19. Hartung, F., S. Suer, and H. Puchta, *Two closely related RecQ helicases have antagonistic roles in homologous recombination and DNA repair in Arabidopsis thaliana.* Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18836-41.

20. Berchowitz, L.E., et al., *The role of AtMUS81 in interference-insensitive crossovers in A. thaliana.* PLoS Genet, 2007. **3**(8): p. e132.

21. Higgins, J.D., et al., *Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over.* Plant J, 2008. **54**(1): p. 152-62.

22. Holloway, J.K., et al., *MUS81 generates a subset of MLH1-MLH3-independent crossovers in mammalian meiosis.* PLoS Genet, 2008. **4**(9): p. e1000186.

23. Geuting, V., et al., *Two Distinct MUS81-EME1 Complexes from Arabidopsis Process Holliday Junctions.* Plant Physiol, 2009. **150**(2): p. 1062-71.

24. Bishop, D.K. and D. Zickler, *Early decision; meiotic crossover interference prior to stable strand exchange and synapsis.* Cell, 2004. **117**(1): p. 9-15.

25. Cromie, G.A., et al., *Single Holliday junctions are intermediates of meiotic recombination.* Cell, 2006. **127**(6): p. 1167-78.

26. Oh, S.D., et al., *RecQ helicase, Sgs1, and XPF family endonuclease, Mus81-Mms4, resolve aberrant joint molecules during meiotic recombination.* Mol Cell, 2008. **31**(3): p. 324-36.

27. Mezard, C., et al., *The road to crossovers: plants have their say.* Trends Genet, 2007. **23**(2): p. 91-9.

28. Ohta, K., et al., *Competitive inactivation of a double-strand DNA break site involves parallel suppression of meiosis-induced changes in chromatin configuration.* Nucleic Acids Res, 1999. **27**(10): p. 2175-80.

29. Anderson, L.K., K.D. Hooker, and S.M. Stack, *The distribution of early recombination nodules on zygotene bivalents from plants.* Genetics, 2001. **159**(3): p. 1259-69.

30. de Boer, E., et al., *Two levels of interference in mouse meiotic recombination.* Proc Natl Acad Sci U S A, 2006. **103**(25): p. 9607-12.

31. Kugou, K., et al., *Rec8 Guides Canonical Spo11 Distribution Along Yeast Meiotic Chromosomes.* Mol Biol Cell, 2009.

32. King, J.S. and R.K. Mortimer, *A polymerization model of chiasma interference and corresponding computer simulation.* Genetics, 1990. **126**(4): p. 1127-38.

33. Foss, E.J. and F.W. Stahl, *A test of a counting model for chiasma interference.* Genetics, 1995. **139**(3): p. 1201-9.

34. Brar, G.A., et al., *The multiple roles of cohesin in meiotic chromosome morphogenesis and pairing.* Mol Biol Cell, 2009. **20**(3): p. 1030-47.

35. Borde, V., et al., *Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites.* Embo J, 2009. **28**(2): p. 99-111.

36. Stahl, F.W., et al., *Does crossover interference count in Saccharomyces cerevisiae?* Genetics, 2004. **168**(1): p. 35-48.

37. Buhler, C., V. Borde, and M. Lichten, *Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in Saccharomyces cerevisiae.* PLoS Biol, 2007. **5**(12): p. e324.

38. Mancera, E., et al., *High-resolution mapping of meiotic crossovers and non-crossovers in yeast.* Nature, 2008. **454**(7203): p. 479-85.

39. Wahls, W.P., E.R. Siegel, and M.K. Davidson, *Meiotic recombination hotspots of fission yeast are directed to loci that express non-coding RNA.* PLoS ONE, 2008. **3**(8): p. e2887.

40. Lambie, E.J. and G.S. Roeder, *Repression of meiotic crossing over by a centromere (CEN3) in Saccharomyces cerevisiae.* Genetics, 1986. **114**(3): p. 769-89.

41. Chen, S.Y., et al., *Global Analysis of the Meiotic Crossover Landscape.* Dev Cell, 2008.

42. Latt, S.A. and R.R. Schreck, *Sister chromatid exchange analysis.* Am J Hum Genet, 1980. **32**(3): p. 297-313.

43. Zalevsky, J., et al., *Crossing over during Caenorhabditis elegans meiosis requires a conserved MutS-based pathway that is partially dispensable in budding yeast.* Genetics, 1999. **153**(3): p. 1271-83.

44. Chen, C.* and H. Ma, *Plant meiocyte collection for genome research*, in *Manuscript; * correspondence for material distribution*. 2007.