Modelling sex-specific crossover patterning in Arabidopsis

Andrew Lloyd*† and Eric Jenczewski†‡

* Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, SY23 3EB, UK
† Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay, 78000, Versailles, France
‡ INRA UMR1349 Institut de Génétique, Environnement et Protection des Plantes, France
Short Title: Modelling crossovers in Arabidopsis

Keywords: recombination, crossovers, interference, beam-film, sex-specific

Corresponding author: Andrew Lloyd

Address: Institute of Biological, Environmental and Rural Sciences, Plas Gogerddan, Aberystwyth University, Aberystwyth, SY23 3EB, UK

Phone: +44 (0) 1970823229

Email: anl50@aber.ac.uk
ABSTRACT

Interference is a major force governing the patterning of meiotic crossovers. A leading model describing how interference influences crossover-patterning is the beam film model, a mechanical model based on the accumulation and redistribution of crossover-promoting stress along the chromosome axis. We use the beam-film model in conjunction with a large Arabidopsis reciprocal back-cross data set to gain mechanistic insights into the differences between male and female meiosis and crossover patterning. Beam-film modelling suggests that the underlying mechanics of crossover patterning and interference are identical in the two sexes, with the large difference in recombination rates and distributions able to be entirely explained by the shorter chromosome axes in females. The modelling supports previous indications that fewer crossovers occur via the class II pathway in female meiosis and that this could be explained by reduced DNA double strand breaks in female meiosis, paralleling the observed reduction in synaptonemal complex length between the two sexes. We also demonstrate that changes in the strength of suppression of neighboring class I crossovers can have opposite effects on effective interference depending on the distance between two genetic intervals.

INTRODUCTION

Meiotic crossovers shuffle parental genetic information generating new combinations of alleles. In most species the presence of one crossover inhibits nearby crossover formation so that inter-crossover distances are greater and more uniform, than if placed at random. This phenomenon, crossover interference, was first noted in genetic studies over a century ago (1,
2), however it is only in the last few years that insights into its mechanistic basis have begun to surface.

Studies in Arabidopsis (3), yeast (4) and humans have shown that the inhibitory effect of interference spreads across a defined physical chromosome distance. This is usually measured in µm synaptonemal complex (SC) although there is some evidence that it is mediated by the chromosome axes (whose lengths are proportional to that of the SC) prior to synapsis (reviewed 5). In yeast, interference is, at least in part, mediated by Topoisomerase II (4) and wt levels of interference require SUMOylation and subsequent ubiquitin-mediated removal of TopoII and also of the axis component Red1/Asy3 (4). These findings are consistent with suggested roles for the chromosome axis and local stress relief via DNA remodeling, in mediating interference.

Several approaches have been used to model CO patterning, the most notable being the gamma model and the beam-film model. The gamma model is a statistical model based on the observation that the distances between two crossovers are relatively uniform, following a gamma distribution. Under this model “effective interference strength” is highest when distances between crossovers show the least variation. This results in a large value of the gamma shape parameter.

In contrast, the beam-film model is a mechanistic model whose various parameters have biological correlates. In the beam-film model, each bivalent has a number of precursor sites (DSBs) that are subject to mechanical stress. CO maturation at precursor sites is promoted by stress and this stress is relieved locally following CO-designation. As stress promotes COs, stress relief propagating out from crossover sites inhibits the formation of additional COs.
nearby. In the beam-film model, interference strength is highest when stress relief propagates furthest from designated crossover sites.

In most species, there are multiple crossover pathways. The majority of crossovers occur via the interference sensitive class I pathway and are dependent on the ZMM group of proteins identified initially in yeast (Zip1, Zip2, Zip3, Zip4, Mer3, Msh4, and Msh5) (6). Crossovers occurring via this pathway are specifically marked by Zip3/Hei10 and MLH1 foci at late pachytene (7–9). A number of secondary “clean-up” pathways repair DSBs not metabolized by the class I pathway (10, 11). These clean-up pathways mostly repair DSBs as non-crossovers, but also contribute a smaller number of crossovers (i.e. class II crossovers) which are insensitive to interference (12) and usually make up 10-15% of total recombination events (e.g. 7, 12, 13).

In their simplest forms, the gamma and beam-film models of crossover patterning deal exclusively with class I crossovers and therefore experimental data used in comparisons with simulated data must be class I specific, e.g. Zip3/Hei10 foci distributions. Crossover patterning in yeast has been extensively explored using the single-pathway beam-film model and the effects of different parameters on crossover distributions are now well understood. Consideration of class I crossovers alone, however, is not always possible or appropriate. For example, cytological analyses of female meiosis in Arabidopsis remain challenging and, as the number of crossovers per chromosome is low, well over a thousand cells would need to be analyzed to achieve the same number of inter-interval distances (the limiting factor for analyses) commonly reported for yeast chromosomes (4, 14). Cytological data are also not appropriate when wanting to consider patterning of both class I and class II crossovers, as there is as yet no robust cytological marker for class II crossovers. In these instances, beam-film
best-fit simulations must be used in combination with genetic data, which include the result of both class I and class II crossovers. Backcross populations, for example, are particularly well suited to this kind of analysis.

Using a large Arabidopsis reciprocal backcross recombination data set (~1500 individuals, ~380 markers for both male and female) it has been shown using the two-pathway gamma model that crossover interference is higher in female meiosis than in male meiosis, that male meiosis has a higher proportion of non-interfering COs and that class-II COs are non-uniformly distributed (15).

We use the two-pathway beam-film model to analyze the same large Arabidopsis dataset, identifying parameter best fits to gain mechanistic insights into the differences between male and female crossover patterning. Using the best-fit parameter values identified as a starting point we then model the influence of various parameters on crossover patterning under the two-pathway beam-film model.

**MATERIALS AND METHODS**

**Experimental data:** Experimental dataset used has been previously published (16) and was derived from large Arabidopsis reciprocal backcross populations. On average, 1,505 plants were genotyped for 380 SNPs in the male population and 1,507 plants genotyped for 386 SNPs in the female population (380 in common). As the average distance between markers is small in this data set – 316 kb in male, 311 kb in female – the number of double crossovers (DCOs) in a single interval are expected to be negligible. It was therefore assumed during analysis that all recombination events were identifiable.
Beam film parameter optimization: Beam-film simulations were performed and best-fit parameters determined using MADpatterns (17) and custom perl scripts with an approach based on that described in (14). For each chromosome and each sex at least three rounds of analysis were undertaken. In each round of each analysis 30,000 bivalents were simulated for a range of parameter values. In the first round, relatively broad value ranges of optimised parameters ($S_{\text{max}}$: 2 - 10; $T_2$Prob: 0.002 - 0.008; $c_L$: 0.3 - 1.3 and $c_R$: 0.3 - 1.3) were chosen based on values described in Zhang et al (2014) and comparison of ad hoc simulations with analysis of experimental datasets (15). Parameters $N$, $B$, $E$, $B_s/B_e/B_d$, $A$ and $M$ were set at appropriate default values (see below). In the next two rounds, progressively smaller step-sizes between values were used to arrive at the final parameter values.

For each round of analysis, the crossover distributions, CoC curves and event distributions (distribution of number of COs per gamete) simulated for each chromosome were determined using MADpatterns (17) and compared to those obtained for the relevant sex and chromosome from the experimental data set; number of intervals = 13. Importantly the experimental data are gamete data, while the MADpatterns program simulates (and outputs) bivalent data (i.e. all crossovers on a pair of homologous chromosomes). Therefore, all simulated bivalent crossover frequencies were halved to convert to gamete crossover frequencies. Bivalent event distributions were also converted to gamete event distributions, assuming random assignment of each crossover to two of the four chromatids. Parameter sets were ranked based on the difference between simulated and experimentally determined CoC distributions [$\text{Score}_{\text{CoC}} = \Sigma_{\text{IID}} \text{abs}(\log_2(CoC_{\text{sim}}/CoC_{\text{exp}}))$], CO distributions [$\text{Score}_{\text{CO}} = \Sigma_{\text{Int}} (CO_{\text{sim}} - CO_{\text{exp}})^2$] and event distributions [$\text{Score}_{\text{ED}} = \Sigma_{\text{Enum}} (ED_{\text{sim}} - ED_{\text{exp}})^2$]. Simulations were ranked for each score and final parameter values chosen were those with the lowest rank-sum.
**Fixed Parameter Values:** Beam-film model parameters $N$, $B$, $E$, $Bs/Be/Bd$, $A$ and $M$ were fixed based on known values of the biological correlates, parameters that tend to be stable between species (14), or suggested default values (18). A description of each of these parameters and values used is given below, further explanations can be found in refs. (14, 17).

$N$ – *Precursor sites per bivalent:* Each meiotic double strand break (DSB) corresponds to a recombination precursor site. In Arabidopsis the total number of DSBs per meiosis is estimated at around 250 (19). For beam-film modelling, the number of DSBs for each chromosome was calculated as 250 multiplied by the proportion of the total genome size (in Mb) contributed by that chromosome (Table S1.). To simulate reduced DSB formation in female meiosis, total DSB number was set at 150 and the per-chromosome number of precursors calculated as above.

$B$ – *Similarity in total precursor number between bivalents:* $B$ sets the similarity of precursor number between the multiple bivalents simulated for a given chromosome in each round of analysis: $0 = $ Poisson, $1 = $ constant. When $B \neq 1$, i.e. precursor number is not constant, parameter $N$ specifies the average number of precursors per bivalent. $B$ was set to 1 for all simulations.

$E$ – *Evenness of precursor spacing:* $E$ sets the evenness of precursor spacing: $0 = $ random, $1 = $ even. There is considerable experimental evidence that DSB spacing is non-random (20, 21). For numerous organisms a parameter value of 0.6 has been found appropriate (14). $E$ was set to 0.6 for all simulations.

$Bs/Be/Bd$ – *Recombination “black hole” start/end/precursor density:* Recombination black hole start and end points correspond to the start and end of the recombination suppressed
centromeric region. These values were determined based on the recombination frequencies observed in the backcross data (Table S1, Figure S1) and correspond to regions with high DNA methylation and low H3K4me3 indicative of heterochromatin (21).

**Bmax** – *Similarity in maximum stress levels between bivalents*: *Bmax* sets the similarity of *Smax* between simulated bivalents: 0 = Poisson, 1 = constant. When *Bmax* ≠ 1, i.e. maximum stress levels are not constant, parameter *Smax* specifies the average maximum stress per bivalent. *Bmax* was set to 1 for all simulations.

**A** – *Intrinsic precursor sensitivities*: Parameter *A* determines how precursor sensitivities are assigned. For all simulations *A* was set to 1 – sensitivities assigned from a uniform distribution.

**M** – *Crossover maturation efficiency*: In the beam film model, it is possible to model failure of crossover maturation. If failure occurs, the CO-designated site inhibits nearby crossovers but does not itself develop into a crossover. 0 = no maturation, 1 = 100% maturation. *M* set to 1 for all simulations.

**Double crossover class determination**: The proportion of each class of DCO for a given inter-interval distance was determined from simulations modelling the formation of class I crossovers only (*T2Prob* = 0), class II crossovers only (*Smax* = 0), or both class I and II crossovers. For each simulation, numbers of DCOs were tallied for each inter-interval distance (IID, the distance between a pair of genetic intervals). For each IID, numbers of DCOs involving two class I COs (*DCO_{I,I}*) , two class II COs (*DCO_{II,II}*) or all DCOs (*DCO_{ALL}*) were calculated from the respective simulations. *DCO_{I,II} = DCO_{ALL} – (DCO_{I,I} + DCO_{II,II}).*
Response of model to parameters $L$, $S_{\text{max}}$, $T_2\text{Prob}$ and $N$: To investigate the response of the model to parameters $L$, $S_{\text{max}}$ and $T_2\text{Prob}$ we simulated 30000 bivalents for an “idealized” male Arabidopsis chromosome ($N = 60$, $B = 1$, $E = 0.6$, $B_s = 0.45$, $B_e = 0.55$, $B_d = 0.01$, $S_{\text{max}} = 9$, $B_{s\text{max}} = 1$, $A = 1$, $L = 0.7$, $c_L = 0.8$, $c_R = 0.8$, $M = 1$, $T_2\text{Prob} = 0.004$) as described above, varying one specified parameter.

RESULTS

Beam-Film Parameter Best-Fit Simulations: We used the MADpatterns program (18) and custom scripts to simulate crossover patterning and identify parameter best-fits for each of the five Arabidopsis chromosomes for both male and female meioses. Parameters $N$, $B$, $E$, $B_s/B_e/B_d$, $B_{s\text{max}}$, $A$ and $M$ were fixed for a given set of simulations (see methods), while $S_{\text{max}}$, $L$, $T_2\text{Prob}$ and $c_L/c_R$ were explicitly optimized. Best-fit values were identified by comparing crossover number and distribution, and interference relationships (CoC curves) of simulated and experimental datasets (Figure 1 & S1-2). Simulations using best-fit parameters gave good approximations of the experimental data (Figure 1 & S1-2).
Figure 1. Crossover analysis for Arabidopsis chromosome 5. Each analysis includes experimental (solid lines) and simulated (dashed lines) data from male (blue) and female (orange) meioses. A Crossover distributions for Arabidopsis chromosome 5. Dashed lines represent the limits of the centromeric region over which precursor number is markedly reduced during simulations. B-C CoC curves for chromosome 5 with inter-interval distance (IID, the distance between a pair of genetic intervals) measured in Mb (B) or µm SC (C). L_{CoC} positions for male and female are indicated by blue and orange arrows respectively. D Event distribution for chromosome 5.

Optimized Parameters:

S_{max} - Maximum stress level per bivalent: In the beam-film model, whether or not a precursor site (DSB) develops into a crossover is determined by the crossover-promoting “stress” (S) experienced by that precursor as well as the precursor’s sensitivity (determined by parameter A, see methods). The biological correlate of S_{max} i.e. the nature of the crossover promoting stress is not precisely defined but may relate to the expansion of chromatin during early prophase (22). When simulating each bivalent, the value of S is progressively increased until S = S_{max}. At some point S reaches a critical value at which the most sensitive precursor undergoes CO-designation and stress-relief extends out from that position. As S increases to S_{max}, additional precursors usually experience sufficient stress to promote the designation of further COs. At any point during the simulation, the actual stress experienced by a precursor is the current value of S, minus the sum of any stress-relief due to interference from nearby COs.

Optimum values of S_{max} for the five chromosomes were similar in males and females: male 7 ± 1.9 and female 6.9 ± 0.7, p = 1 (Bonf. corrected) (Figure 2, Table S1).

L_{BF} – Stress relief distance: The parameter L_{BF} corresponds to the length of the chromosomal interval over which a CO relieves stress i.e. stress-relief propagates out from COs a distance of \( \frac{1}{2} L_{BF} \) in either direction. The magnitude of the stress-relief decreases exponentially with
distance from the CO, such that there is maximal stress-relief in the middle of the interval (i.e. immediately surrounding the CO) and almost no stress-relief at either end of the interval (14, 22).

When running simulations $L_{BF}$ is specified as the proportion of total chromosome length (i.e. chromosome length is set to 1), but is converted to length in Mb or $\mu$m SC to enable comparisons between chromosomes of different lengths. When measured in Mb ($L_{BF, Mb}$) best-fit estimates of $L_{BF}$ were significantly higher in females: $L_{BF, Mb}$ – male 17.1 ± 3.5 Mb and female 28.8 ± 3.1 Mb, $p = 0.0095$ Bonf. corr. (Figure 2, Table S1). When the distance metric was converted to $\mu$m SC ($L_{BF, SC}$), using the best available estimates of SC length in the two sexes (3), there was no-longer any difference in $L_{BF}$ between the two sexes: $L_{BF, CoC}$ – male 27.7 ± 5.6 $\mu$m and female 23.7 ± 2.5 $\mu$m, $p = 1$ Bonf. corr. (Figure 2). For some chromosomes, optimum values of $L_{BF}$ were larger than the chromosome in question. This is expected if one CO suppresses the formation of additional COs more than half the length of the chromosome away. An example can be seen for chromosome 2 in females which has an estimated SC length of 16.2 $\mu$m and an $L_{BF, CoC}$ estimate of 25.9 $\mu$m. As can be seen from the CoC curve for this chromosome (Figure S2) it is clear that the observed number of DCOs are less than expected (i.e. CoC < 1) even when intervals are at opposite ends of the chromosome (e.g. inter-interval distance ~14 $\mu$m).

cL/R – Left and Right end clamping: In the beam-film model, “clamping” at chromosome ends determines how stress is supported in terminal regions. Unclamped chromosome ends (cL = 0; cR = 0) cannot support stress and so locally relieve stress, behaving as if there were a crossover at the chromosome end. Clamped chromosome ends (cL = 1; cR = 1) experience stress as elsewhere along the bivalent. As the interference signal cannot come from beyond the end of the chromosome, recombination frequencies will tend to be higher at the end of chromosomes.
than for internal regions when clamped. Total clamping averages (cL/R) for male and female were calculated from the estimated values of cL and cR for each sex. Clamping values were variable between chromosomes but there was no significant difference between the average clamping values, male 0.78 ± 0.16 and female 0.69 ± 0.13, p = 1 Bonf. corr. (Figure 2).

**T2Prob - Probability that a non-crossover designated precursor will form a Type II crossover:**
The average optimized T2Prob values were significantly higher in male than female meiosis, 0.0063 ± 0.0010 and 0.0036 ± 0.0008 respectively (p = 0.026, Bonf. corrected, Table 1, Figure 2). This value corresponds to the proportion of non-crossover designated precursor sites that become Type II COs and assumes equal numbers of precursor sites i.e. meiotic DSBs in male and female meiosis. We also determined what proportion of the total number of crossovers occur via the class II pathway (i.e. p = CO_{II} / (CO_{I} + CO_{II}). These values were equivalent for the two sexes: 0.14 ± 0.02 male, 0.14 ± 0.03 female, p = 1 Bonf. corr.

**Figure 2. Beam-film best-fit parameter values.** Standard simulations for male (M) and female (F) assumed 250 DSBs per meiosis. Simulations for female were run a second time assuming 150 DSBs (F_{DSB}).

* p < 0.05, ** p < 0.01, after Bonferroni multiple comparison correction.

**Beam-Film Parameter Best-Fit Simulations – Modelling reduced DSB formation in females:** While there are relatively good estimates for the number of DSBs in male meiosis in Arabidopsis, cytological analyses of female meiosis are more challenging and there are no reliable estimates of DSB numbers. Thus, while we have assumed equal numbers of DSBs in...
male and female meiosis, it is possible that DSB numbers differ between the two sexes. Meiotic DSBs occur in loop DNA that has been recruited to the chromosome axis (23). In Arabidopsis female meiosis there are fewer (albeit larger) chromatin loops and the chromosome axis is 40% smaller than in male meiosis (3) which could feasibly result in a similar reduction in DSBs (24, 25). To understand whether reduced DSB numbers would have any effect on crossover patterning and/or estimates of parameter values in female meiosis, we repeated the best fit simulations assuming a reduction in DSBs equal to the reduction in SC length i.e. approx. 40% reduction, or 150 (rather than 250) DSBs per meiosis.

Optimised values of $S_{max}$, $L_{BF Mb}$, $L_{BF SC}$ and cL/cR were identical for both sets of simulations (F and F$_{DSB}$ Figure 2). Estimates of $T_{2Prob}$ were higher for simulations of female meiosis with reduced DSB numbers, and the optimized value no-longer differed from the value of $T_{2Prob}$ estimated for male meiosis (Figure 2). In the beam film model, the parameter $T_{2Prob}$ is the number of non-crossover designated DSBs that become class II COs. Thus, despite the difference in the parameter value estimates of $T_{2Prob}$ for female meiosis with 250 or 150 DSBs, the total number of class II COs estimated using both approaches is the same: 250 DSBs $\times T_{2Prob} 0.0036 \pm 0.0006 = 0.87 \pm 0.15$ COs; 150 DSBs $\times T_{2Prob} 0.0056 \pm 0.0010 = 0.90 \pm 0.17$ COs.

**Crossover Distribution Analysis**: Simulated CO distributions closely matched those observed in the experimental data with highest recombination in males in distal regions and highest recombination in females closer to the centromere. (Figure 1, Figure S1). The exception was the short arms of chromosomes 2 and 4 in males which have high experimental recombination rates but had low levels of recombination when simulated using the global best-fit parameters
It is possible that this is related to the presence of the NORs on the short arms of chromosomes of these two chromosomes.

**CoC Analysis:** The coefficient of coincidence (CoC) is the ratio of the observed and expected numbers of double crossovers (DCOs) for a given pair of intervals, given the rates of single COs in the two intervals. When interference strength is high, CoC values tend to be low as there are fewer DCOs observed than expected (one CO suppresses the occurrence of nearby COs). CoC shows a characteristic curve when plotted against inter-interval distance (Figure 1B-C), with low CoC for small inter-interval distances and CoC approximating 1 for large inter-interval distances. A useful measure when analyzing such curves is $L_{CoC}$, the inter-interval distance at which the curve crosses the line CoC = 0.5 (dashed line, Figure 1B-C). For all analyses the simulated data gave $L_{CoC}$ values that were no different from those determined from the equivalent experimental data (Table 1.). For both experimental and simulated data, $L_{CoC}$ was significantly smaller in males than in females if measured in Mb but showed no difference between when measured in µm SC (Table 1, Figure 1 and Figure S2).

<table>
<thead>
<tr>
<th></th>
<th>Mb</th>
<th>µm SC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>male</td>
<td>female</td>
</tr>
<tr>
<td>experimental</td>
<td>7.05 ± 0.50</td>
<td>12.84 ± 1.50</td>
</tr>
<tr>
<td>simulated</td>
<td>6.30 ± 1.05</td>
<td>11.60 ± 0.83</td>
</tr>
<tr>
<td>p value*</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

*Bonferroni corrected

**Response of the two-pathway beam-film model to parameters $L$, $Smax$ and $T2Prob$:** To further investigate how changes in individual parameters affect the two-pathway beam-film model we simulated crossovers using a range of parameter values.
**Crossover Distributions:** As the distance over which crossover interference is propagated ($L_{BF}$) decreased, crossover frequencies increased. The change in frequency was particularly marked in distal regions although all regions experienced some increase in crossovers (Figure 3A). When $L_{BF}$ was reduced below half of chromosome length, increases in crossovers was also strong adjacent to the centromere (Figure 3A).

Changes in $S_{max}$ – the maximum crossover promoting force (or “stress”) experienced by recombination pre-cursors – had similar effects on crossover distribution as those observed for changes in $L_{BF}$. Higher values of $S_{max}$ (like low values of $L_{BF}$) resulted in large increases in recombination in distal regions, but recombination rates were largely unchanged in proximal regions (Figure 3D). For low $S_{max}$, crossover rates were very low near the ends of the chromosomes and resulted in the most proximal crossover distribution (Figure 3D). Increases in $T2Prob$ resulted in a small uniform increase in COs across the chromosome (Figure 3G).

**CoC:** Increases in $L_{BF}$ shifted the CoC curve to the right resulting in increased $L_{CoC}$ values (Figure 3B). For small inter-interval distances increases in $L_{BF}$ resulted in no change or a small increase in CoC (Figure 3C).

For low values of $S_{max}$, CoC curves were shifted more to the right and for higher values of $S_{max}$ CoC curves were shifted to the left, however the effect on $L_{CoC}$ was not as pronounced as observed for changes in $L_{BF}$ (Figure 3B & 3E). Again, for small inter-interval distances decreases in $S_{max}$ resulted in no change or an increase in CoC (Figure 3F). As has been observed previously in yeast (14), $S_{max}$ had less influence on CoC than $L$. 
Changes in $T2Prob$ had limited effect on CoC for intermediate and large inter-interval distances, with no broad effects on CoC curves and no change in $L_{CoC}$ values (Figure 3H). For small inter-interval distances, however, the level of class II crossovers had a marked effect, with higher number of class II crossovers resulting in higher values of CoC (Figure 3I).

Thus for smaller inter-interval distances the main determinant of CoC is the number of class II crossovers, while for intermediate and large inter-interval distances CoC values are mainly determined by $L_{BF}$ and to a lesser extent $S_{max}$ i.e. factors affecting (interference sensitive) class I crossovers.

**Figure 3.** Effect of beam-film parameters on crossover patterning in Arabidopsis. The effect of altering a single beam film parameter – $L$ (A-C), $S_{max}$ (D-F) or $T2Prob$ (G-I) – on crossover distribution (A, D & G) and CoC (B-C, E-F and H-I). Red vertical lines in C, F and I represent IID = 0.1.

Differing classes of double COs at small and large IIDs cause opposing effects of interference on CoC: One of the less intuitive outputs of the analysis above was that for small inter-interval distances, an increase in the distance over which the interference signal is propagated can result in increased values of CoC (Figure 3C). To further understand this...
phenomenon, we sought to identify how changes in $L_{BF}$ might differentially affect the expected and observed number of double COs (the determinants of CoC) for different IIDs. Beam film simulations demonstrated that increased $L_{BF}$ resulted in a small decrease in the expected number of double COs (DCOs) for small and large IIDs (IID = 0.1 and 0.5; Figure 4A). This is unsurprising given that the expected number of DCOs for a pair of intervals is purely based on the respective rates of COs in the two intervals. In contrast, the observed number of DCOs changed dramatically for IID = 0.5, but only marginally for IID = 0.1 (Figure 4A) in response to changes in $L_{BF}$. As a result, CoC dramatically decreased for IID = 0.5 with increased $L_{BF}$ but increased slightly for IID = 0.1 (Figure 4B).

We reasoned that the difference in behavior might be due to the nature of the DCOs formed at smaller and larger IIDs. For example, DCOs can occur between two class I COs, two class II COs or between a class I and a class II CO but interference only directly suppresses those involving two class I COs. We therefore ran beam film simulations with class I COs only (T2Prob = 0), class II COs only (Smax = 0), or both class I and class II COs and determined numbers of the different classes of DCOs formed for each set of simulations (Figure 4C). From these numbers we determined the proportions of the different classes of DCOs (Figure 4D) that occur for different IIDs under standard conditions (i.e. when simulating both class I and class II COs). For small IIDs DCOs are almost exclusively formed between a class I CO and a class II CO (Figure 4D). In contrast, for larger IIDs ($\geq 0.4$) the majority of DCOs are formed between two class I COs (Figure 4D). Cytological observations in tomato reporting the same phenomenon (26) suggest this is a general feature of meiosis. As interference only suppresses DCOs involving two class I COs, changes in $L_{BF}$ will only directly affect DCO formation at larger IIDs. This pattern holds when the proportion of class II crossovers falls within the range normally observed (5-20%), although when class II crossovers are absent or make up the
majority of crossovers then most DCOs involve two class I or two class II COs respectively for all IID (Figure S3).

Both the expected number of DCOs and observed DCOs at small IIDs are indirectly affected by increased $L_{BF}$ due to the associated decrease in the frequency of class I COs. The magnitude of the change is greater for the expected number of DCOs, which can be seen from the equations below. Here $CI$ and $CII$ are the rates of class I and class II crossovers respectively in the two intervals:

$$DCO_{exp} = (CI_{int1} + CII_{int1}) \times (CI_{int2} + CII_{int1})$$

$$DCO_{obsSmall,IID} \sim (CI_{int1} \times CII_{int2}) + (CI_{int2} \times CII_{int1})$$

For small IIDs, while $CI >> CII$, the reduction in the expected number of DCOs is approximately twice that of the observed reduction in DCOs, resulting in an increase in CoC.

**Figure 4. Influence of IID on CoC response to changes in $L_{BF}$**. A The expected (solid line) and observed (dashed line) proportion of interval pairs receiving a double crossover (DCO) for two different inter-interval distances (IIDs); calculated from simulations with varying values of $L_{BF}$. B CoC values for two IIDs calculated from simulations with varying values of $L_{BF}$. C The number of DCOs observed for different IIDs from simulations involving class I and class II crossovers (CI & CII), class I crossovers only (CI) or class II
crossovers only (CII). D The proportions of DCOs formed between two class I crossovers (CI_CI), two class II crossovers (CII_CII), or a class I and a class II CO (CI_CII) for different IIDs.

Response of the two-pathway beam-film model to pre-cursor (DSB) numbers: As described above a 40% decrease in DSB numbers had few effects on estimates of beam-film parameters. To further investigate the effect of DSBs on crossover patterning in Arabidopsis we simulated crossovers for a range of DSB numbers. As well as simulating wt meiosis, we also modelled the effect of altered DSB numbers in situations where a large proportion of DSBs become class II crossovers, such as is observed in some mutant contexts (27).

Altered DSB number had relatively little effect on crossover distributions or CoC curves in a wt context (Figure 5A-B). For CO distribution, the main change was a general flattening of the distribution with lower DSB number. The only clear difference in CoC was for small inter-interval distances, where higher DSB numbers resulted in higher values of CoC (Figure 5C). In contrast, altering the number of DSBs in a context where a high proportion become class II crossovers had a dramatic effect on crossover patterning. Here increased DSBs resulted in proportionate increases in crossovers (Figure 5D). Regardless of the number of DSBs, CoC values were approximately 1 for all inter-interval distances (Figure 5E-F).

We next modelled how DSB number affects the total number of crossovers for male and female meiosis in both contexts. In wild type, doubling the number of DSBs resulted in ~15% increase in crossovers in male and female (Figure 5G). In a context with a high number of class II crossovers, doubling the number of DSBs resulted in almost doubling the number of crossovers (Figure 5G). For a given number of DSBs the modelling predicts ~65% more crossovers in wt male than wt female, but essentially equal numbers of crossovers when the probability of class II crossovers is high (Figure 5G).
**Figure 5. Influence of DSB number on crossover patterning.** Crossover distributions (A & D) and CoC relationships (B-C and E-F) for beam-film simulations with varying numbers of DSBs. A-C show results for simulations of wt meiosis, D-F show results for simulation of meiosis with increased class II crossover formation (T2Prob). (G) Total crossover number for genome-wide simulations using best-fit parameters for male and female meiosis and varying numbers of DSBs.

**DISCUSSION**

Crossover interference is a well-known genetic phenomenon, however its mechanistic basis is only just now coming to light. The interference signal is thought to propagate a set physical distance ($L_{BF}$, usually measured in µm SC) from designated crossover sites (4, 14), and analyses commonly use cytological observations and simulations of class I crossover positions along the length of a synapsed bivalent (4, 14).

To gain insights into the differences between female and male meiosis in Arabidopsis, we analyzed a large Arabidopsis reciprocal backcross data set (16) and performed two-pathway (i.e. both class I and class II COs) beam film best-fit simulations. Our modelling suggests that the major differences in crossover number, crossover distribution and interference relationships between the sexes can all be explained by the observed difference in SC length between male and female meiosis. The relationship between genome size and SC length is governed by the size/number of chromatin loops, which occur at a conserved density of ~20 per µm SC across
a wide range of organisms (28). As genome size is identical for both sexes in Arabidopsis, we
would expect loop size in male meiocytes to be about 60% of that found in female meiocytes.
Exactly how chromatin loop size is determined remains unclear but this decision occurs very
early in, or prior to, meiosis (5, 29). It is probable, therefore, that the cause of differences in
crossover patterning also occurs very early in, or prior to, meiosis. Interestingly humans also
display a sex-specific differences in chromatin loop-size and SC length, although in this case
female meiocytes have shorter loop-size, longer SC and more crossovers (30).

It has been reported previously that effective crossover interference is stronger in females than
in males in Arabidopsis (15). Our analyses indicate that the interference signal is propagated
over the same physical distance (µm SC) in both male and female meiosis, and thus from a
mechanistic standpoint interference is identical in the two sexes. The higher effective
interference (i.e. the effect on the inheritance of two linked genetic loci) observed in females
can be entirely explained by the difference in SC length between the two sexes, as a given
distance in µm SC corresponds to a greater length in Mb. It is worth noting that our estimates
of $L_{BF}$ for male (27.7 ± 5.6 µm) and female (23.7 ± 2.5 µm) Arabidopsis are similar to estimates
for tomato (14 µm, ref 14) but are 80 to 90-fold larger than for yeast (0.3 µm, ref 14). This vast
difference in the distance across which interference propagates in different taxa, as estimated
by the beam-film model, remains challenging to explain biologically.

In addition to explaining differences in effective interference, SC length also explained the
differences in CO distribution observed between the sexes. In male meiosis, crossovers are
high adjacent to the peri-centromeres and in the distal regions, while in female meiosis
crossovers are high adjacent to the peri-centromeres but low in the distal regions (3, 16). Our
modelling shows that increases in the proportion of the chromosome over which interference
spreads (either through a reduction in SC length, or an increase in $L_{BF}$) reduces crossovers particularly in distal regions. The lower SC length in females can therefore account for the observed differences in crossover distribution.

In mammals, SC length is correlated with the number of DSBs (24, 25, 31). If the same holds true in plants, then we might expect fewer DSBs in female meiosis. Our analysis revealed that while the number of DSBs had very little influence on crossover distributions and CoC curves, a decrease in the number of DSBs resulted in an increase in the estimated proportion of DSB sites that become class II crossovers ($T2Prob$). Thus, the reduction in SC length observed for females, if accompanied by an equivalent reduction in DSBs, can also account for proposed differences in the number of class II crossovers between male and female meiosis. At least one line of evidence suggests this question may not be fully resolved however. In mutant lines with large numbers of additional class II crossovers, the recombination landscape of male and female meiosis are roughly equivalent with even a slightly higher number of crossovers in female (27). This suggests the possibility of similar numbers of DSBs in male and female meiosis. Further comparative cytological studies of male and female meiosis will be required to fully answer these questions.

Given the substantial differences in crossover patterning between female and male meiosis it is striking that they can all be accounted for by the difference in SC length. It is similarly striking that despite the differences in crossover patterning there are also no significant differences between the sexes in the estimated beam-film model parameters. This gives us good confidence in our approach, and suggests that similar investigations, in different contexts (e.g. mutants, over expression lines, environmental conditions), could provide further mechanistic insights into the factors governing crossover patterning in Arabidopsis.
The model can also be used to make predictions about how important agricultural goals such as heightened recombination rates could be achieved. For example, with the development of CRISPR and related technologies, it is possible to modulate the number or location of DSBs in early meiosis and there is interest in using this approach to alter recombination rates in plant breeding programs (32–34). In most organisms, crossover numbers are thought to be maintained independently from the number of DSBs through crossover homeostasis (35–37). Our modelling suggests that the extent to which homoeostasis maintains crossover numbers is determined by the proportion of DSBs that become class II crossovers: The higher the proportion of class II crossovers, the more DSB number will affect crossover number. Thus, we predict that combining the knock out of class II CO suppressing proteins (e.g. RECQ4, FANCM, FIGL1, 38–40) with approaches to increase meiotic DSBs could maximize increases in recombination and the associated benefit to breeding programs.

One of the surprising findings of our analysis is that for small inter-interval distances, an increase in the distance over which the interference signal is propagated can result in increased values of CoC (Figure 3) i.e. decreased effective interference. This behavior is not specific to the beam-film model but is expected whenever both class I and class II crossovers occur, and there is a change in the strength of suppression of closely spaced class I crossovers. This finding highlights the need for caution when interpreting interference data and particularly in the distinction between mechanistic (e.g. $L_{BF}$) and effective (e.g. CoC from genetic data) measurements of interference. It should also be noted that at small inter-interval distances the magnitude of the predicted change in CoC is small, and that for specific interval pairs the effect of the local chromosomal landscape (e.g. recombination hotspots etc) may out-weigh the effect predicted by the model. Despite these caveats, it is clear that an increase or decrease in
mechanistic interference strength ($L_{BF}$) is not expected to result in an equivalent increase or decrease respectively in effective interference for small IIDs. Given the widespread use of reporter lines that determine recombination rates and CoC values for closely linked intervals (41) it is important to realize that these lines give little to no insight into any change in the mechanics of crossover interference.

As an example, two recent papers investigated altered recombination rates at temperature extremes in Arabidopsis (42, 43). In both cases, increased temperature gave rise to more class I COs, but the increased COs were associated with no change, or a decrease in genetic measurements of CoC (i.e. effective interference). In the studies, CoC (or interference ratio) was measured by tracking the inheritance of closely linked fluorescent reporter genes in pollen, and thus combined both class I and class II crossovers measured at a small inter-interval distance. While it could be concluded from these studies that temperature increases class I crossovers without any effect on interference, these results are also consistent with an alternative hypothesis i.e. that increased temperature decreases the distance over which interference is propagated, resulting in increased class I COs, but with no effect on genetic measurements of interference at small-inter-interval distances. Or to put it another way, high temperature might decrease mechanistic interference, but result in an increase (or no change) in effective interference for small IIDs. There is good evidence that heightened temperature might have such a mechanistic effect, given that the chromosome axis is thought to mediate interference (5) and the synaptonemal complex / axis structure is sensitive to temperature (44, 45) but this remains to be experimentally validated.

While the beam-film model was able to reliably model genetic recombination data, there are several ways in which the model might further be improved with increased understanding of
the underlying biology. For example, when calculating $L_{BF}$ and $L_{CoC}$ in µm SC using backcross data, we assume a direct relationship between SC length and Mb. In Arabidopsis the relationship between SC length and Mb is constant between whole chromosomes ($R^2 = 0.99$, based on data from (46)), however the relationship may not be constant within a chromosome (47). Establishing how the relationship between Mb and µm SC changes for different chromosomal domains would provide one means to improve models of crossover patterning when using genetic data. Another question is whether DSB density is constant along the length of the chromosome? If so, is it constant relative to SC length or length in Mb? Recently Spo11-oligo sequencing has demonstrated relatively constant DSB formation along the length of the chromosome, although there are clearly regions of higher and lower DSB density, particularly the centromeres where DSB formation is strongly suppressed (21). Incorporating such data into the beam-film model would be an interesting future avenue.

Despite these possible improvements, it is clear that we can gain novel insights into crossover patterning using genetic recombination data in combination with beam-film simulations. These are particularly powerful when, as for this study, we have good estimates of SC length for all chromosomes, circumventing the need for cytological determination of crossover locations. This enables us to take advantage of the main benefit of genetic data, that it incorporates all crossover events, and thus enables us to develop a more nuanced understanding of the interplay between the mechanistic determinants of crossover-interference and the final effect on patterns of inheritance.

**ACKNOWLEDGEMENTS**

We thank Christine Mezard for providing the experimental Arabidopsis recombination dataset and for critical reading of the manuscript. A.L. was funded by a Marie Curie fellowship (PIOF-
GA-2013-628128) and the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 663830. The IJPB benefits from the support of the LabEx Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS).
REFERENCES


23. Panizza S, et al. (2011) Spo11-accessory proteins link double-strand break sites to the


44. Loidl J (1989) Effects of elevated temperature on meiotic chromosome synapsis in


Table S1. Chromosome metrics and beam-film parameters

<table>
<thead>
<tr>
<th>Chr</th>
<th>Sex</th>
<th>Mb</th>
<th>µm SC</th>
<th>N</th>
<th>B</th>
<th>E</th>
<th>Bs</th>
<th>Be</th>
<th>Bd</th>
<th>Smax</th>
<th>Bsmax</th>
<th>A</th>
<th>Lp</th>
<th>Lm</th>
<th>cL</th>
<th>cR</th>
<th>M</th>
<th>T2prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>30.4</td>
<td>49.2</td>
<td>64</td>
<td>1</td>
<td>0.6</td>
<td>0.475</td>
<td>0.5</td>
<td>0.01</td>
<td>8.5</td>
<td>1</td>
<td>1</td>
<td>0.65</td>
<td>19.8</td>
<td>32.0</td>
<td>0.8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>19.7</td>
<td>31.9</td>
<td>41</td>
<td>1</td>
<td>0.6</td>
<td>0.175</td>
<td>0.225</td>
<td>0.01</td>
<td>7.5</td>
<td>1</td>
<td>1</td>
<td>0.85</td>
<td>16.7</td>
<td>27.1</td>
<td>0.3</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>23.5</td>
<td>37.9</td>
<td>49</td>
<td>1</td>
<td>0.6</td>
<td>0.5</td>
<td>0.65</td>
<td>0.01</td>
<td>5.5</td>
<td>1</td>
<td>1</td>
<td>0.7</td>
<td>16.4</td>
<td>26.6</td>
<td>0.4</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>18.6</td>
<td>30.1</td>
<td>39</td>
<td>1</td>
<td>0.6</td>
<td>0.125</td>
<td>0.225</td>
<td>0.01</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0.6</td>
<td>11.2</td>
<td>18.0</td>
<td>0.6</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>27.0</td>
<td>43.6</td>
<td>56</td>
<td>1</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.01</td>
<td>9.5</td>
<td>1</td>
<td>1</td>
<td>0.8</td>
<td>21.6</td>
<td>34.9</td>
<td>1.1</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>F</td>
<td>30.4</td>
<td>25.0</td>
<td>64</td>
<td>1</td>
<td>0.6</td>
<td>0.475</td>
<td>0.5</td>
<td>0.01</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>30.4</td>
<td>25.0</td>
<td>0.7</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>19.7</td>
<td>16.2</td>
<td>41</td>
<td>1</td>
<td>0.6</td>
<td>0.175</td>
<td>0.25</td>
<td>0.01</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1.6</td>
<td>31.5</td>
<td>25.9</td>
<td>1.2</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>23.5</td>
<td>19.3</td>
<td>49</td>
<td>1</td>
<td>0.6</td>
<td>0.5</td>
<td>0.65</td>
<td>0.01</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>23.5</td>
<td>19.3</td>
<td>0.5</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>18.6</td>
<td>15.3</td>
<td>39</td>
<td>1</td>
<td>0.6</td>
<td>0.125</td>
<td>0.225</td>
<td>0.01</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td>31.6</td>
<td>26.0</td>
<td>0.8</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>27.0</td>
<td>22.2</td>
<td>56</td>
<td>1</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.01</td>
<td>6.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>27.0</td>
<td>22.2</td>
<td>0.7</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>Foa</td>
<td>30.4</td>
<td>25.0</td>
<td>38</td>
<td>1</td>
<td>0.6</td>
<td>0.475</td>
<td>0.525</td>
<td>0.01</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0.9</td>
<td>27.4</td>
<td>22.5</td>
<td>0.5</td>
<td>0.7</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Foa</td>
<td>19.7</td>
<td>16.2</td>
<td>38</td>
<td>1</td>
<td>0.6</td>
<td>0.475</td>
<td>0.525</td>
<td>0.01</td>
<td>6.5</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>29.6</td>
<td>24.3</td>
<td>1</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Foa</td>
<td>23.5</td>
<td>19.3</td>
<td>29</td>
<td>1</td>
<td>0.6</td>
<td>0.5</td>
<td>0.65</td>
<td>0.01</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>23.5</td>
<td>19.3</td>
<td>0.5</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Foa</td>
<td>18.6</td>
<td>15.3</td>
<td>23</td>
<td>1</td>
<td>0.6</td>
<td>0.125</td>
<td>0.225</td>
<td>0.01</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td>31.6</td>
<td>26.0</td>
<td>0.8</td>
<td>0.4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Foa</td>
<td>27.0</td>
<td>22.2</td>
<td>34</td>
<td>1</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.01</td>
<td>8.5</td>
<td>1</td>
<td>1</td>
<td>1.1</td>
<td>29.7</td>
<td>24.4</td>
<td>0.6</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

* Optimised parameter

* Fixed parameter

* Calculated based on $L_p$
**Figure S1.** Experimental and simulated crossover distributions for Arabidopsis chromosome 1-5.
**Figure S2.** CoC curves for simulated and experimental recombination data.
Figure 3S. Proportions of DCOs formed between two class I crossovers (CI_CI), two class II crossovers (CII_CII), or a class I and a class II CO (CI_CII) for different IIIDs and different values of T2Prob (grey bars). Total proportion of class II crossovers are as follows: T2Prob = 0, 0% class II COs; T2Prob = 0.001, 3% class II COs; T2Prob = 0.005, 13% class II COs; T2Prob = 0.01, 23% class II COs; T2Prob = 0.025, 43% class II COs; T2Prob = 0.125, 81% class II COs.