

# Crossover Patterning by the Beam-Film Model: Analysis and Implications

Liangran Zhang<sup>1</sup>, Zhangyi Liang<sup>1</sup>, John Hutchinson<sup>2</sup>, Nancy Kleckner<sup>1\*</sup>

**1** Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, United States of America, **2** School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, United States of America

## Abstract

Crossing-over is a central feature of meiosis. Meiotic crossover (CO) sites are spatially patterned along chromosomes. CO-designation at one position disfavors subsequent CO-designation(s) nearby, as described by the classical phenomenon of CO interference. If multiple designations occur, COs tend to be evenly spaced. We have previously proposed a mechanical model by which CO patterning could occur. The central feature of a mechanical mechanism is that communication along the chromosomes, as required for CO interference, can occur by redistribution of mechanical stress. Here we further explore the nature of the beam-film model, its ability to quantitatively explain CO patterns in detail in several organisms, and its implications for three important patterning-related phenomena: CO homeostasis, the fact that the level of zero-CO bivalents can be low (the “obligatory CO”), and the occurrence of non-interfering COs. Relationships to other models are discussed.

**Citation:** Zhang L, Liang Z, Hutchinson J, Kleckner N (2014) Crossover Patterning by the Beam-Film Model: Analysis and Implications. *PLoS Genet* 10(1): e1004042. doi:10.1371/journal.pgen.1004042

**Editor:** R. Scott Hawley, Stowers Institute for Medical Research, United States of America

**Received:** May 10, 2013; **Accepted:** November 5, 2013; **Published:** January 30, 2014

**Copyright:** © 2014 Zhang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Funding to LZ, ZL and NK provided by a grant to NK from the National Institutes of Health: NIH/NIGMS R01 GM044794 (<http://nih.gov>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: kleckner@fas.harvard.edu

## Introduction

Crossover (CO) recombination interactions occur stochastically at different positions in different meiotic nuclei. Nonetheless, along a given chromosome, COs tend to be evenly spaced. This interesting phenomenon implies the existence of communication along chromosomes, the nature of which is not understood. CO patterning, commonly known as “CO interference”, was originally detected from genetic studies in *Drosophila* [1,2]. It was found that the frequency of meiotic gametes exhibiting two crossovers close together along the same chromosome (“double COs”) was lower than that expected for their independent occurrence. The implication was that occurrence of one CO (or more correctly one CO-designation) “interferes” with the occurrence of another CO (CO-designation) nearby.

We previously proposed a model for CO patterning in which macroscopic mechanical properties of chromosomes play governing roles via accumulation, relief and redistribution of stress (Figure 1A) [3,4]. In that model, a chromosome with an array of precursor interactions comes under mechanical stress along its length. Eventually, a first interaction “goes critical”, undergoing a stress-promoted molecular change which designates it to eventually mature as a CO. By its intrinsic nature, this change results in local relief of stress. That local relaxation then redistributes outward in the immediate vicinity of its nucleation point, in both directions, dissipating with distance. A new stress distribution is thereby produced, with the stress level reduced in the vicinity of the CO-designation site, to a decreasing extent with increasing distance from its nucleation point. This effect disfavors occurrence of additional (stress-promoted) CO designations in the affected region. The spreading inhibitory signal

comprises “CO interference”. More such CO-designations may then occur, sequentially, each accompanied by spreading interference. Each subsequent event will tend to occur in a region where the stress level remains higher, which will necessarily tend to be regions far away from prior CO-designated sites. Thus, as more and more designation events occur, they tend to fill in the holes between prior events, ultimately producing an evenly-spaced array. The most attractive feature of this proposed mechanism is the fact that redistribution of stress is an intrinsic feature of any mechanical system, thus comprising a built-in communication network as required for spreading CO interference.

CO-designated interactions then undergo multiple additional biochemical steps to finally become mature CO products [5]. Precursors that do not undergo CO-designation mature to other fates, predominantly inter-homolog non-crossovers (NCOs).

CO patterning by the above stress-and-stress relief mechanism can be modeled quantitatively by analogy with a known physical system that exhibits analogous behavior, giving the beam-film (BF) model [3].

We note that BF model simulations can be applied to any mechanism whose effects are described by the same mathematical expressions as the beam-film case. In such a more general formulation (Figure 1B), there is again an array of precursor interactions. That array would be acted upon by a “Designation Driving Force” (DDF). Event-designations would occur sequentially (or nearly so). Each designation would set up a spreading inhibitory effect that spreads outward in both directions, decreasing in strength with increasing distance, thereby decreasing the ability of the affected precursors to respond to the DDF. When

**Author Summary**

Spatial patterning is a common feature of biological systems at all length scales, from molecular to multi-organismic. Meiosis is the specialized cellular program in which a diploid cell gives rise to haploid gametes for sexual reproduction. Crossing-over between homologous maternal and paternal chromosomes (homologs) is a central feature of this program, playing a role not only for increasing genetic diversity but also for ensuring regular segregation of homologs at the first meiotic division. The distribution of crossovers (COs) along meiotic chromosomes is a paradigmatic example of spatial patterning. Crossovers occur at different positions in different meiotic nuclei but, nonetheless, tend to be evenly spaced along the chromosomes. We previously described a mechanical “stress and stress relief” model for CO patterning with an accompanying mathematical description (the “beam-film model”). In this paper we explore the roles of mathematical parameters in this model; show that it can very accurately describe experimental data sets from several organisms, in considerably quantitative depth; and discuss implications of the model for several phenomena that are directly related to crossover patterning, including the features which can ensure that every chromosome always acquires at least one crossover.

multiple designation/interference events occur, they would produce an evenly-spaced array. Maturation of CO-designated and not-CO-designated interactions ensues.

The present study adds several new features to the BF simulation program and explores in further detail the predictions and implications of the BF model (whether mechanical or general). We evaluate the ability of the model to quantitatively explain experimental CO pattern data sets in budding yeast, tomato,

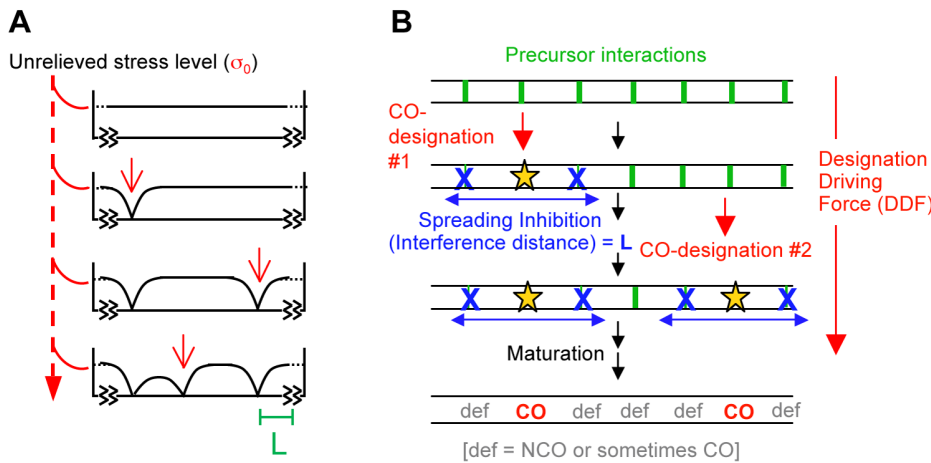
grasshopper and *Drosophila*. Our results show that the logic and mathematics of the BF model are remarkably robust in explaining experimental data. New information of biological interest also emerges. We then present detailed considerations of three phenomena of interest, the so-called “obligatory CO” and “CO homeostasis”, and the nature of “non-interfering COs”. We discuss how these phenomena are explained by the BF model and show that BF predictions can very accurately explain experimental data pertaining to these effects. Overall, the presented results show that BF simulation analysis is a useful approach for exploring experimental CO patterns. Other applications of this analysis are presented elsewhere. The current study has also provided new criteria for characterization of CO patterns using Coefficient of Coincidence analysis and illustrates both short-comings and useful applications of gamma distribution analysis. Relationships of the BF model to other models are discussed.

**Results**

**Part I. Coefficient of Coincidence (CoC) Relationships and the Event Distribution (ED)**

CO data sets, whether experimental or from BF simulations, comprise descriptions of the positions of individual COs along the lengths of each of a large number of different chromosomes (“bivalents”). Each bivalent represents the outcome of CO-designation in a single meiotic nucleus; the entire data set comprises the outcomes of CO patterning for a particular chromosome in many nuclei.

**CoC relationships.** The classical description of CO interference relationships is Coefficient of Coincidence (CoC) analysis [1,2]. For this purpose (Figure 2A), the chromosome of interest is divided into a number of intervals and for each interval the total frequency of COs in the many chromosomes examined is observed. Intervals are then considered in pairs, in all possible pairwise combinations. For each pair, the observed frequency of



**Figure 1. The beam-film model.** (A) Beam-film model [3]. CO-designation is promoted by stress. Each stress-promoted CO-designation reduces the stress level to zero at the designation point. That effect redistributes in the vicinity, decreasing exponentially with distance, correspondingly reducing the probability of subsequent designation(s) in the immediate vicinity. Subsequent CO-designations will tend to occur in regions with higher remaining stress levels, thus giving an even distribution. More specifically: with a film attached to a beam, if the beam expands relative to the film, stress arises along the film, causing it to crack at the positions of flaws. A crack at one position will release the stress nearby (with a distance  $L$  that is characteristic of the materials) thus reducing the probability that another crack occurs nearby. Assuming a maximal possible stress level of  $\sigma_0$ , if a crack occurs at an isolated position that is unaffected by any prior cracks, then the stress level at any distance “ $x$ ” to either side is  $\sigma = \sigma_0 (1 - e^{-x/L})$ . If two cracks occur near one another, the stress levels at positions between them is the sum of their individual effects, with additional considerations also coming into play at the ends of the beam as described [3]. (B) A generalized version of the beam-film model involving sequential rounds of event designation and spreading interference as described by the mathematical expressions of the BF model.

doi:10.1371/journal.pgen.1004042.g001

bivalents exhibiting a CO in both intervals (“double COs”) is compared with the frequency expected if events occurred independently in the two intervals. The latter frequency is given by the product of the total frequencies of events in the two intervals, each considered individually. For any pair of intervals, the ratio of the frequency of observed double COs to the frequency of expected double COs (Observed/Expected) is the Coefficient of Coincidence (CoC). If events occur independently in the two intervals, the CoC for that pair of intervals is one. If (positive) CO interference is present for the two intervals, the CoC is less than one (some expected COs have been inhibited). CoC values for all interval pairs are then plotted as a function of the distance between the corresponding intervals (defined as the distance between the centers of the two intervals).

The classical resulting CO interference CoC pattern is illustrated by an appropriate set of BF simulations (Figure 2B). When intervals are close together (short inter-interval distances), the frequency of observed COs is much less than that expected from independent occurrence ( $CoC \ll 1$ ), reflecting “interference”. The CoC increases with increasing inter-interval distance to a value of  $\sim 1$ . Additionally, because COs tend to be evenly spaced, the CoC value rises above  $\sim 1$  specifically at the average distance between adjacent COs (or multiples thereof): at these particular spacings, the probability of a double CO is higher than that predicted by random occurrence. This tendency is increasingly pronounced as interference extends over longer and longer fractions of total chromosome length.

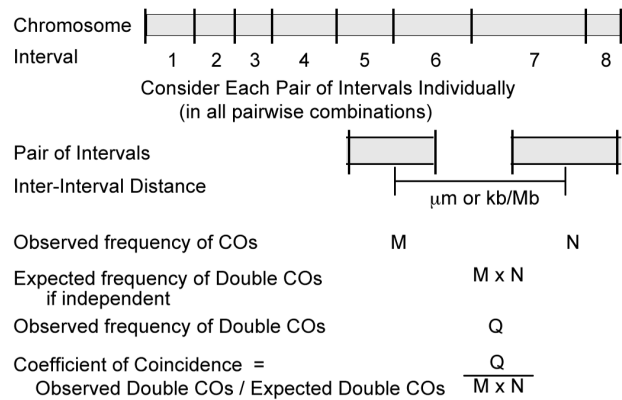
BF simulations specify a parameter for interference distance, denoted “L” ([3]; below). Figure 2B illustrates CoC curves for simulations at varying values of L. For any actual CoC curve, whether experimental or simulated, a useful parameter for describing the strength of CO interference is the inter-interval distance at which the  $CoC = 0.5$ . We define this parameter as  $L_{CoC}$  (Figure 2B; vertical arrows). Where appropriate, the value of the BF-specified parameter “L” is denoted alternatively as  $L_{BF}$  to distinguish it from  $L_{CoC}$  (Figure 2B). Interestingly, the values of  $L_{CoC}$  and  $L_{BF}$  are always quite similar (e.g. Figure 2B).

CoC analysis provides a very accurate and reproducible description of CO patterns for experimental data sets as long as two requirements are met (Figure S1, Protocol S1). First, chromosomes must be divided into a large enough number of intervals that double COs within an interval are rare. If this requirement is not met, closely-spaced double COs will be missed. In general, interval size should be less than  $\sim 1/4$  the average distance between adjacent COs. Second, the data set must be large enough to give significant numbers of double COs. As a practical matter, where possible (e.g. for cytological markers of CO positions), interval size should be decreased progressively until the CoC curve no longer changes.

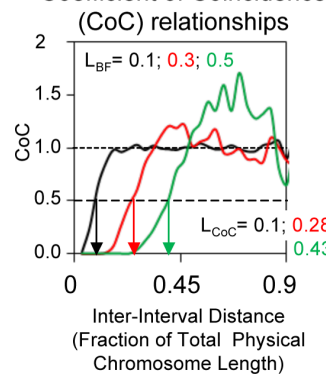
We further note that the appropriate metric for CO interference is physical distance along the chromosome. This has been shown to be the case for mouse and Arabidopsis [6–8]; for tomato (as described below); and for budding yeast (L.Z., unpublished). Accordingly, disruption of chromosome continuity abolishes the transmission of interference in *C.elegans* [9]. Experimentally, “interference distance” is defined in units of  $\mu m$  pachytene (synaptonemal complex; SC) length. In reality, SC length is often (or always) a proxy for chromosome length at the preceding stage (leptotene): in yeast, *Sordaria* and likely other organisms, CO patterning occurs at the leptotene stage and nucleates SC formation (e.g., [4]).

**The ED.** CO patterns are also reflected in the average number of COs per bivalent and the fractions of chromosomes exhibiting different numbers of COs, which we refer to as the

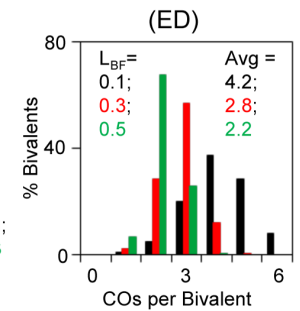
**A Coefficient of Coincidence (CoC) analysis**



**B Coefficient of Coincidence (CoC) relationships**



**C CO Event Distribution (ED)**



**Figure 2. Descriptors of CO patterns: Coefficient of Coincidence (CoC) and Event Distribution (ED).** (A) Determination of CoC. Interval sizes can be identical or different (Figure S1 for more details). (B,C) Data sets were generated by BF simulations at the indicated varying values of parameter (L), also called  $L_{BF}$ . Other specified parameter values for the simulations are:  $S_{max} = 3.5$ ,  $A = 1$ ,  $cL = cR = 0.85$ ,  $N = 13$ ,  $B = 1$ ,  $E = 0$ ,  $M = 1$ .

(B) CoC curves. Inter-interval distances given as fractions of total physical chromosome length in  $\mu m$ . The inter-interval distance at which  $CoC = 0.5$  (vertical arrows) is defined as  $L_{CoC}$ .  $L_{CoC}$  and  $L_{BF}$  are always quite similar in magnitude. (C) EDs. Fraction of bivalents exhibiting different numbers of COs and average values.

doi:10.1371/journal.pgen.1004042.g002

“Event Distribution” or “ED”. As the “interference distance” increases, the distribution of COs per bivalent shifts to lower numbers with a corresponding decrease in the average number of COs per bivalent (Figure 2C).

**Part II. Parameters of the BF Model and Their Roles for CO Patterns**

BF simulations require specification of three types of parameters (Table 1). One set describes the nature of the precursor array upon which CO-designation acts; a second set describes features of the patterning process *per se*; and a third precursor specifies the efficiency with which a designated event matures into a detectable CO or CO-correlated signal.

**The precursor array.** The precursors for CO patterning are generally assumed to be the total array of double strand break (DSB)-initiated interactions between homologs. Several BF

**Table 1.** BF parameters.

Precursor parameters	
<b>N</b>	Average number of precursors per bivalent
<b>E</b>	Precursor distribution along chromosomes. 0 = random; 1 = even; $0 < E < 1$
<b>B</b>	Precursor distribution among chromosomes 0 = Poisson; 1 = constant; $0 < B < 1$
<b>s</b>	Distribution of precursor sensitivities to DDF (currently random)
<b>Density</b>	Precursor density can be varied along the chromosome as desired
Patterning parameters	
<b>S (Smax)</b>	Designation driving force; increased in simulations to maximum level; usually $S > 1$
<b>L = L<sub>BF</sub></b>	Distance over which interference signal spreads (total for both directions)
<b>A</b>	Reactivity of precursor to local value of (s); (A 1 = s <sup>2</sup> ; A2 = s; A3 = 5/s; A4 = 1/s)
<b>c</b>	End effects on interference (cL and cR) 0 = unclamped; $\geq 1$ = clamped
Other parameters	
<b>M</b>	Efficiency with which a CO-designated interaction matures to a detectable CO

doi:10.1371/journal.pgen.1004042.t001

parameters describe the nature of this array. (N) specifies the average number of precursors per bivalent. (E) specifies the extent to which the precursors along a given bivalent are evenly or randomly spaced. (B) specifies the extent to which precursors occur at a constant value along a given bivalent in different meiotic nuclei or are randomly (Poisson) distributed among different nuclei. Also, different precursors will naturally exhibit a range of intrinsic sensitivities to the DDF. The parameter(s) specifies the distribution of those sensitivities as specified by the Matlab function “rand”.

The original BF model included (N) and (s) and assumed a given chromosome has the same number of precursors in different nuclei but assumed that precursors are distributed randomly along a given chromosome (B = 1, E = 0). The latter assumption is likely not the case *in vivo*. Experimental evidence in several organisms shows that precursors tend to be evenly spaced, sometimes dramatically (e.g. [10–14]). And anecdotal evidence further suggests that the number of precursors tends to be quite constant for a given bivalent in different nuclei (e.g. [10,12,15,16]). We further note that evenly-spaced precursors have not been taken into account in any previous quantitative model for CO patterning (e.g. [17–20]). Variations in the nature of the precursor array can affect CoC relationships, with interactive effects, particularly at low values of (N) (Figure 3ABC; Figure S2).

The BF simulation program now also includes a feature which permits the density of precursors to be varied along the chromosome in a desired pattern. This feature is useful for modeling effects such as the paucity of DSBs in centromeric regions, or other regional and domain variations in DSB levels along chromosomes (e.g. [5,6,10,21–26]). Application to grasshopper CO patterns is described below.

The BF model assumes that precursors do not turn over, i.e. that a precursor either develops into a CO or into some other type of product, without being recycled to serve again as a precursor in another position. This assumption has not been directly tested. However, it seems reasonable because precursors are known to be highly evolved multi-protein complexes whose numbers can be constant over long periods of time (e.g. [12]).

The BF model also assumes that the entire precursor array is established prior to CO-designation (or essentially so). This is

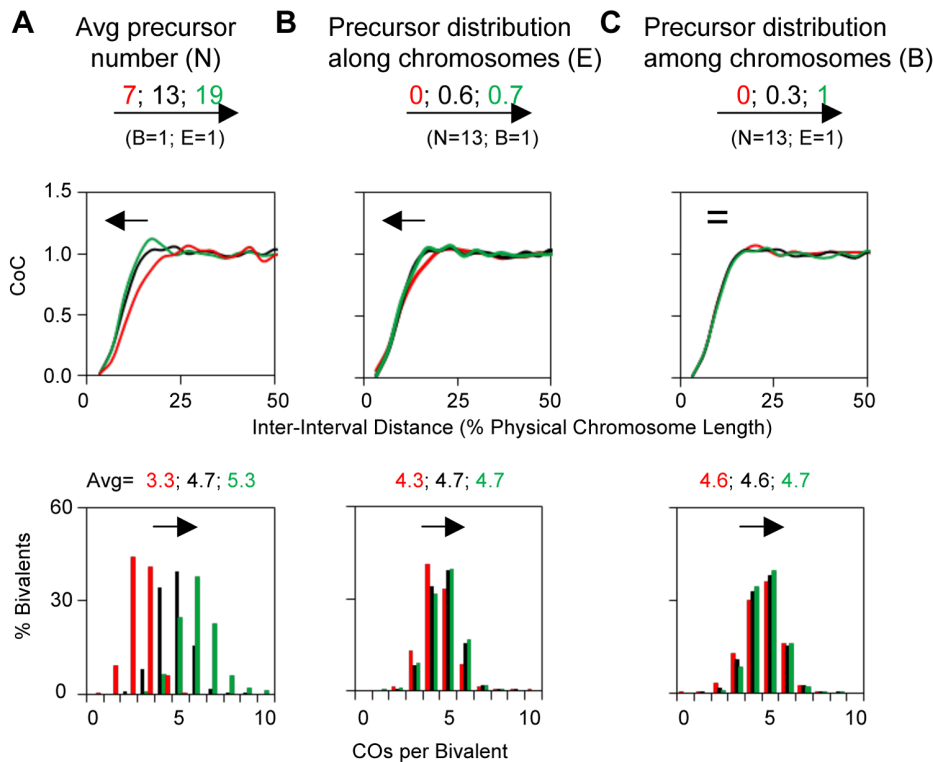
clearly true in some organisms (e.g. [12]). It is not so clear in other organisms, where different regions of the genome can be at significantly different stages within a single nucleus (e.g. [27]). However, BF simulations will still pertain in the latter case if the effects of CO-designation at earlier-evolved positions can be “stored” within the chromosomes and exert their effects upon nearby positions when the appropriate precursors do finally evolve.

Finally, it has sometimes been considered that CO patterns evolve in two stages (e.g. [11,15]). In such a case, one round of event-designation is imposed on total DSB-mediated interactions, giving a set of intermediate designated sites. That intermediate set then undergoes a second round of designation. BF simulations can directly model this situation. DSB-mediated interactions are used as a first set of precursors for a first simulation to give the intermediate array of events. That intermediate array is then used as a second set of precursors for a second simulation. The predicted outcomes from one- and two-round scenarios for recombination-related markers in *Sordaria* meiosis are presented elsewhere. A useful feature in distinguishing between the two scenarios is whether closely-spaced COs ever do, or do not, occur at the specific spacing characteristic of the first precursor array. If COs arise in a single step, closely-spaced events can occur at the positions of adjacent precursors. If COs arise in two steps, then this will not occur; instead, closely-spaced COs can only occur at the spacing of adjacent precursors in the intermediate array (Figure S3).

**CO patterning parameters: DDF (S), interference distance (L), precursor reactivity (A) and end effects (cL, cR).** All of these patterning parameters are present in the original BF model. Detailed explanations of their significance are as follows:

(S,Smax) versus (L): The outcome of the patterning process is determined primarily by two basic parameters: the strength of the (CO)-designation driving force (global stress or the DDF), as given by parameter (S); and the distance over which interference spreads, given by parameter (L).

For simulation purposes, the value of (S) is progressively increased to a specified maximum (Smax). A first, most reactive, precursor goes critical to give a CO-designation with



**Figure 3. Variations in the precursor array can alter CO patterns.** BF simulations were carried out under a set of “standard conditions” except for variations in the parameter of interest as indicated. Panels (A), (B) and (C) illustrate the effects of variations in N, E and B. Standard simulation parameter values are the same as the best-fit values for yeast Chromosome XV (Table 2). Parameter (E) is the standard deviation of the average inter-precursor distance and the corresponding evenness level is also shown by the shape ( $\nu$ ) of gamma distribution in the program output. For E=0, 0.6 and 0.7, the corresponding  $\nu=1, 2.4$  and  $4.2$ , respectively. Parameter (B) is set by using the binomial distribution, in which with an average number of precursors on each bivalent (a constant mean), the distribution of the number of precursors among bivalents (probability of success for each trial) can be adjusted by changing the number of the trials.  
doi:10.1371/journal.pgen.1004042.g003

accompanying interference; the level of (S) is then further increased, giving a next designation at the next most reactive position. This process is increased up to a final desired level. This procedure gives sequential CO-designations. The higher the final maximum value, (Smax), the more CO-designations. Interference arises instantaneously after each designation, reducing the probability that affected precursors can respond to the driving force over the specified distance (L).

The final overall pattern of COs reflects the balance between the CO-designation driving force (DDF) and the interference distance, i.e. the values of (Smax) and (L). Correspondingly, a change in the value of either parameter can confer a similar alteration in CoC relationships and the ED (Figure 4AB). Higher (Smax) or lower (L) permits more COs to occur at shorter inter-interval distances, thus shifting the CoC curve to the left. Concomitantly, the overall level of COs increases. Lower (Smax) or higher (L) has the opposite effects. The ED changes commensurately.

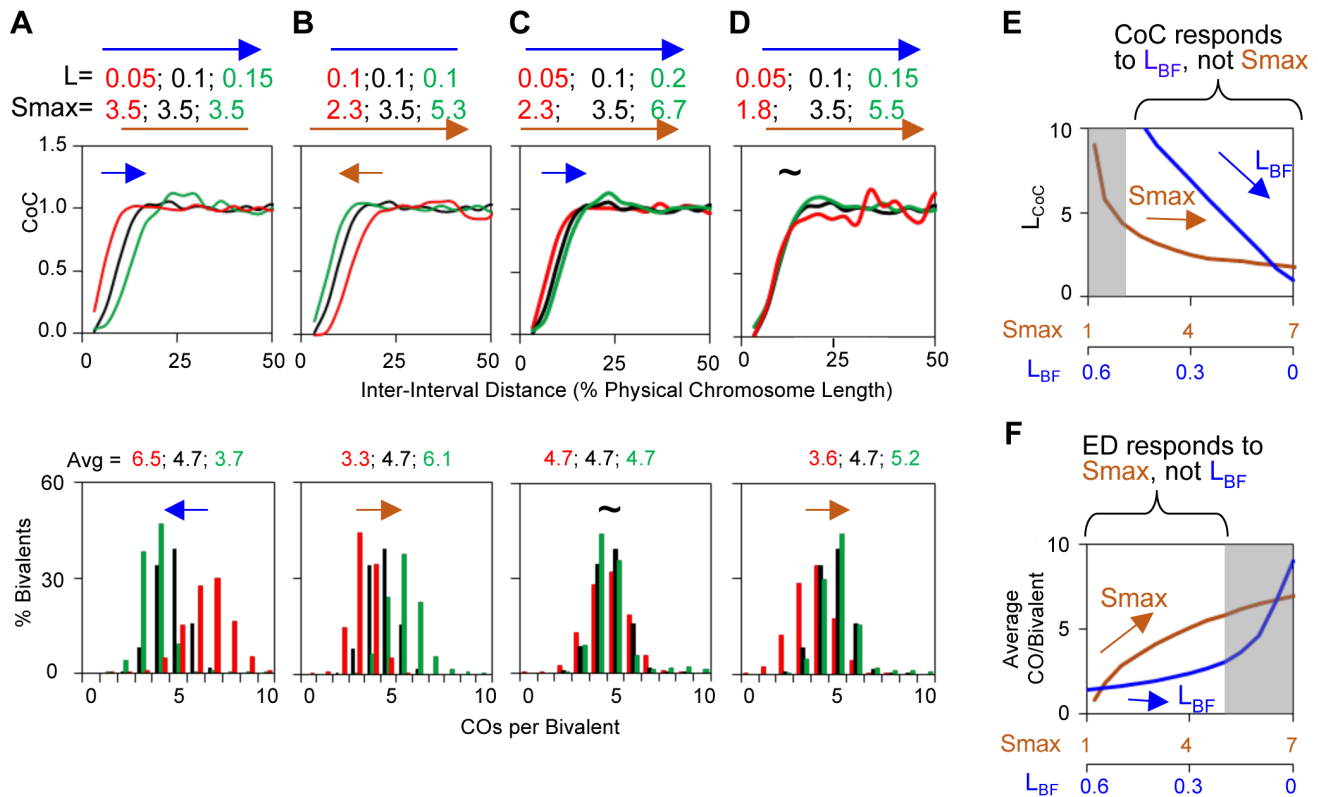
To a considerable extent, opposing variations in the two parameters can compensate for one another (Figure 4CD). Nonetheless, in most cases, the effects of variations in (L) and (Smax) can be distinguished. The primary target of (L) is inter-CO communication, with the number of COs affected as a secondary consequence. The primary target of the DDF (Smax) is the number of COs, not inter-CO communication, with inter-CO relationships affected as a secondary consequence. Correspondingly, variations in (L) primarily affect CoC relationships

whereas variations in (Smax) primarily affect the ED (Figure 4CDEF). The practical implication for best-fit BF simulations is that the values of these two parameters can be specified independently (Figure S4).

We note that, *in vivo*, (S) and (Smax) could take a variety of forms. (i) The value of Smax could potentially be defined by the time available for CO-designation. Interestingly, in *Drosophila*, the presence of a structural chromosome heterozygosity (deletion or inversion) results in a delay in meiotic progression and an increase in the number of COs without loss of CO interference [28,29]. This constellation of phenotypes could be modeled by an increase in Smax. (ii) CO designation could occur sequentially without any progressive increase in the DDF, simply because different precursor complexes will tend to undergo designations sooner or later in relation to their intrinsic reactivities, up to the maximum number specified by Smax.

Also: in the BF model, the strength of the interference signal decays exponentially with distance from the CO-designation site (Figure 1A) in accord with the way in which stress redistributes in the beam-film system upon which the mathematical expressions are based [3,4] (Figure 1A legend). This decay relationship can be altered in the simulation program. However, we have found no need to do so thus far (e.g. below).

Finally, the value of (Smax) actually incorporates the combined effects of the driving force and the sensitivity of precursors to that force. Similarly, the value of (L) incorporates the combined effects of the strength of the interference signal (as it dissipates with



**Figure 4. Variations in the L and/or Smax can alter CO patterns.** All panels: BF simulations as in Figure 3. Panels (A) and (B) illustrate the effects of variations in L or Smax, respectively. Panels (C) and (D) illustrate the fact that very similar CoC and ED relationships can be achieved by appropriate different combinations of L and Smax, but with a differential response of CoC relationships to changes in L (C) and of ED relationships to changes in Smax (D). These differential responses are further documented in Panels (E) and (F).  
 doi:10.1371/journal.pgen.1004042.g004

distance) and the sensitivity of precursors to that signal. Put another way: any difference that can be modeled by a change in ( $S_{max}$ ), or a change in ( $L$ ) could, potentially, reflect a change either an actual change in the altered feature or a change in the ability of precursors to sense that feature. Other information must be brought to bear to distinguish the two types of effects.

(A): A third patterning parameter, ( $A$ ), describes the dose/response relationship between precursor sensitivity ( $s$ ) and the local stress/DDF level at the corresponding position (i.e. the value of  $s$  as modified by the effects of any interference signals that have emanated across that position). Parameter ( $A$ ) can have one of four possible values. In two cases ( $A = 1, 2$ ), reactivity varies directly with ( $s$ ); in the other two cases ( $A = 3, 4$ ), reactivity varies inversely with ( $s$ ) (Table 1). Variations in ( $A$ ) can affect CO patterns (Figure 5A).

Clamping:  $c(L)$  and  $c(R)$ : Special considerations apply to interference at chromosome ends. These effects are incorporated into BF simulations by “clamping” parameters ( $c(L)$  and  $c(R)$ ). In the absence of any other consideration, a terminal region will behave the same way as any other region of the chromosome with respect to its response to the DDF ( $S_{max}$ ), the interference signal ( $L$ ) and precursor reactivity ( $A$ ). The interesting consequence of this effect, not regularly appreciated, is that there will automatically be intrinsic tendency for the ends to exhibit higher frequencies of COs, because these regions will be subjected to interference signals emanating in from only one direction (i.e. from internal regions of the chromosomes and not from regions “beyond” the end of the chromosome) (Figure 5B top). In a mechanical model, this “default” situation is achieved by “clamping” the end of the

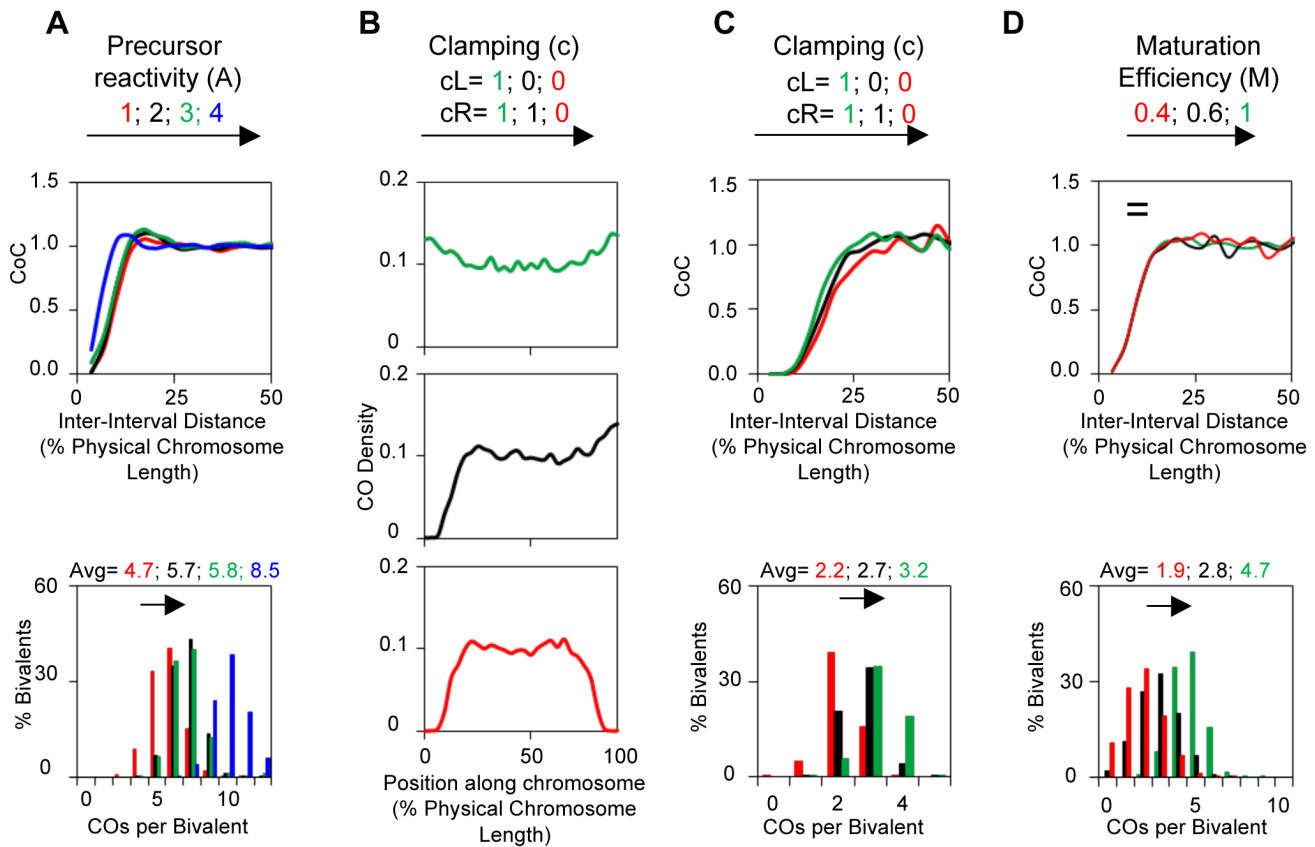
chromosome to some object. In BF simulations, clamping is defined by parameter ( $c$ ), which can be specified individually at each chromosome end ( $c(L)$  and  $c(R)$ ). The default case, fully clamped, is  $c = 1$ .

In a mechanical model, the chromosome end could alternatively be free in space, i.e. would be “unclamped”. Since such a free end cannot support stress, it would behave as if it already had experienced a CO, i.e. with an interference signal having spread inward. The result would be a decreased probability that COs will occur near that end. In this case,  $c = 0$ . Intermediate situations can also occur. Thus,  $c(L)$  and  $c(R)$  can take any value between 0 and 1.

As a practical matter, specification of ( $c(L)$ ) and ( $c(R)$ ) permits more accurate modeling of *in vivo* patterns where end effects are prominent. For example, many organisms exhibit a tendency for COs to occur near the ends of chromosomes whereas others do not. Such effects tend to emerge when chromosome ends are clamped (e.g. Figure 5B middle, bottom). [Notably, however, genetic variations that result in paucities or excesses of DSBs near ends (e.g. [6,10,21–26,30]) should be modeled by variations in the precursor density (above) rather than as effects on interference].

Variations in ( $c(L)$ ) and ( $c(R)$ ) primarily affect the distribution of COs along the length of a bivalent but also have secondary effect on overall CoC and ED relationships (Figure 5BC).

**Maturation efficiency (M).** A precursor that undergoes designation may not mature efficiently into the signal used to define designations experimentally. This situation, defined by the value of the parameter ( $M$ ), occurs in diverse mutant situations. If maturation is less than 100% efficient, the initial array of CO-



**Figure 5. Variations in A, c or M can alter CO patterns.** All panels: BF simulations as in Figure 3. (A) Effects of variations in precursor reactivity relationships. (B, C) Effects of variations in end clamping status on the distribution of COs along the chromosome (Panel B) and on CoC and ED relationships (Panel C). (D) Variations in maturation efficiency (M) do not affect CoC relationships but do affect the ED.  
doi:10.1371/journal.pgen.1004042.g005

designated events undergoes random subtraction such that the final array of detected events reflects only a subset of the original designation array. Variations in (M) do not affect CO patterns. Since maturation efficiency only affects CO status after patterning is established, a decrease in (M) only shifts ED relationships to lower CO numbers, with no/little effect on CoC relationships (Figure 5D).

### Part III. BF Simulations Accurately Describe Experimental Data Sets

Application of the BF model to an experimental data set permits the identification of a set of parameter values for which simulated CO patterns most closely match those observed experimentally (general strategy described and illustrated in Figure S4).

Best-fit simulation analysis for data sets from yeast, *Drosophila*, tomato and grasshopper demonstrates that the logic and mathematics of the BF model can describe experimental CO patterns with a high degree of quantitative accuracy. This conclusion is evident in descriptions of CoC and ED patterns as described in this section (III). Additional evidence is provided by applications and extensions of BF simulation analysis to CO homeostasis, the obligatory CO and non-interfering COs as described in sections IV–VI. Inspection of experimental CoC relationships has also provided new information regarding the metric of CO interference in tomato and the fact that interference spreads across centromere regions (in grasshopper, as previously described, and also in tomato and yeast).

**Budding yeast.** Yeast provides a favorable system for analysis of CO patterning in general, and for application and evaluations of BF modeling in particular, for several reasons. First, in this organism, the sites of patterned (“interfering”) COs are marked by foci of ZMM proteins Zip2 or Zip3 along pachytene chromosomes ([31]; Materials and Methods; Figure 6A). Zip2/3 foci mark CO sites very soon after they are designated (and independent of the two immediate downstream consequences of CO-designation, i.e. formation of the first known CO-specific DNA intermediate and nucleation of SC formation). Thus, effects of CO maturation defects on CO patterns are minimized. Second, the positions of these foci, and thus of CO-designations, can be determined along any specifically-marked chromosome to the resolution of fluorescence microscopy (e.g. Figure 6B). Also, Zip2/3 focus positions can be determined even in the absence of SC [31]. Third, the average number of Zip3 foci (COs) per bivalent varies over a significant range (e.g. Table 2). The lowest value described thus far, ~2 for Chromosome III, is close to the number of COs seen in some other organisms (e.g. mouse or human or grasshopper) and thus provide appropriate models for such cases. At the highest values, ~7, multiple foci (COs) occur quite evenly along the length of the chromosome, which is very useful for revealing general patterns. Fourth, analysis of chromosomes in hundreds of nuclei is readily achievable, thus readily providing sufficiently large data sets for both wild-type and diverse mutant cases. Correspondingly, CoC and ED relationships can be determined extremely accurately. For a given data set, CoC values at each inter-interval distance can vary over a significant range (Figure 6C). This

**Table 2.** Experimental characteristics and best-fit simulation parameters.

Organisms\ Chromosomes	Chrom. length <sup>#</sup>	Experimental data				Best-fit simulation values									
		CO/Biv		% Zero- CO Biv	L <sub>CoC</sub>	L <sub>BF</sub>	Smax	A	cL/cR	N <sup>@</sup>	E	B	M		
		Mbp	μm			Fraction	μm								
<i>S.cerevisiae</i>	Chr III	0.32	1.2	1.8	1	0.3	0.25	0.3	3.5	1	0.85/0.85	6	0.6	1	1
	Chr XV	1.05	3.2	4.7	0	0.3	0.1	0.3	3.5	1	0.85/0.85	13	0.6	1	1
	Chr IV	1.53	4.8	6.8	0	0.3	0.075	0.3	3.5	1	0.85/0.85	19	0.6	1	1
	Chr XIV (BR)	0.79	2.1	3.7	0	0.3	0.15	0.3	3.5	1	0.85/0.85	11	0.6	1	1
<i>Chorthippus</i>	L3	1800	94	2.2	0	28	0.3	28	2.3	2	1.1/0.96	14	0	1	1
<i>D.Melanogaster</i>	X	22	15	1.4	5	6	0.2	6	2.8	1	0.65/0.5	7	0.4	0.5	1
<i>S. lycopersicum</i>	Chr 2–4	70	22	1.4	0	11	0.65	14	1.05	3	1.7/1.1	25	0.6	1	1
		70	17	1.2	0	11	0.8	14	1.05	3	1.7/1.1	20	0.6	1	1

<sup>#</sup>The genomic chromosome length and SC length are from the following sources. (1) *S.cerevisiae*: Saccharomyces Genomic Database (<http://www.yeastgenome.org>) and SC length are from this study and [31]. (2) *Chorthippus*: its C value = ~10 pg (<http://www.genomesize.com/results.php?page=1>), and 1 pg = 978 Mb [70]; thus its genome size is ~9780 Mb, chromosome L3 SC = 94 μm (0.189 of total SC length [71]), thus the predicted L3 is ~1800 Mb assuming the SC length is proportional to genome size in this organism. (3) *D. Melanogaster*: Fly Database (<http://flybase.org/>). SC length is based on [72,73]. (4) *S. lycopersicum*: Sol Genomics Network (<http://solgenomics.net>). SC length is from [38].  
<sup>@</sup>The number of precursors is based on the following. (1) *S. cerevisiae*: Spo11 oligos and also microarray data [21,52]. (2) *Chorthippus*: from [3] (3) *D. Melanogaster*: based on [3,34,35] (4) *S. lycopersicum*: based on [38].  
doi:10.1371/journal.pgen.1004042.t002

variation is largely due to sampling variation because it is significantly reduced in BF simulation data sets that involve 5000 chromosomes rather than the 300 that usually comprises a typical data set. Despite this variation, which is present in all experimental data sets (below), the average CoC curve obtained from such an experiment is highly reproducible, as illustrated by the results of four independent experiments (Figure 6D–G).

Experimental CO patterns in wild-type SK1: We have defined Zip3 focus patterns along WT chromosomes III, IV and XV in the SK1 strain background. These chromosomes range in size from 330 to 1530 kb (Table 2). For CoC analysis, each chromosome was divided into 100 nm interval, a size that provides maximal accuracy (Figure S1). Inter-interval distances for CoC analysis are expressed in units of μm SC length (rationale above). The three analyzed chromosomes exhibit virtually overlapping CoC curves, with L<sub>CoC</sub> = ~0.3 μm (Figure 6F; Table 2). Along a pachytene chromosome of SK1, this distance corresponds to ~100 kb. The CoC curve remains less than 1 up to inter-interval distances corresponding to ~150 kb, in accord with the maximal distance over which interference is detected by genetic analyses (e.g. [32]).

The value of Zip3 foci as a marker for CO patterns is further confirmed by analysis of an *mlh1Δ* mutant. Since Zip3 foci mark CO sites shortly after they are designated, and long before the late step at which Mlh1 is thought to act (above), the *mlh1Δ* mutation should not reduce total Zip3 focus levels and should have no effect on CoC relationships for Zip3 foci. This expectation is confirmed (Figure 6H).

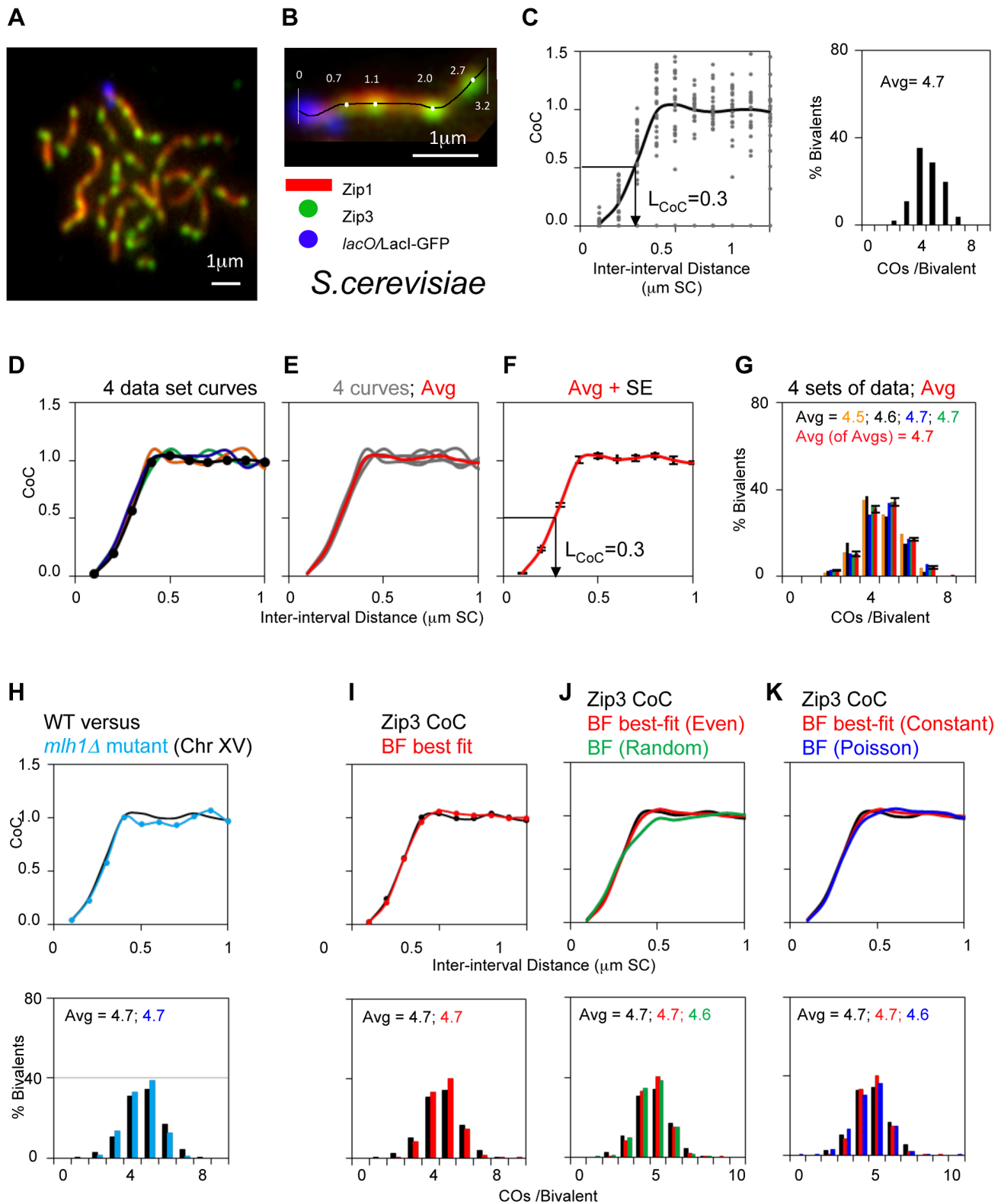
BF analysis: CoC and ED data for all three analyzed chromosomes can be very closely matched by BF best-fit simulations (e.g. Figure 6I; best-fit simulations for SK1 Chromosome IV and III and for chromosome XIV in BR are shown in Figure S5 A–D; parameter values in Table 2). Points of note include:

- (1) Despite differences in absolute chromosome length, and numbers of Zip3 foci, all three chromosomes are described

by the same set of optimal parameter values with the exception of the predicted number of precursors, which increases with chromosome length as could be expected (Table 2).

- (2) The thus-defined best-fit value of (L), i.e. the distance over which the inhibitory interference signal spreads, is ~300 nm (L<sub>BF</sub> = ~0.3 μm) for all analyzed chromosomes (Table 2). This value of (L<sub>BF</sub>) turns out to correspond closely to “interference distance” as defined experimentally by CoC analysis in all cases (L<sub>CoC</sub> = ~0.3 μm; Table 2).
- (3) For each chromosome, the number of precursors used for the best-fit simulation corresponds well to that described experimentally by analysis of DSBs and is approximately proportional to chromosome length [21].
- (4) An optimal match between best-fit CoC curves and experimental data requires that precursors be relatively evenly spaced along the chromosomes (E = 0.6; Figure 6I; Figure S5A–E), thus confirming and extending experimental evidence that yeast DSBs are evenly spaced (discussion in [13,14]; see also below).
- (5) Notably, also, for the shortest chromosome (III), an optimal match is obtained only if precursors are also assumed to occur in a relatively constant number along a given chromosome in different nuclei (B = 1; Figure S5F). At lower values of B, the frequency of zero-Zip3 focus chromosomes is higher than that observed experimentally because a significant fraction of chromosomes fail to acquire enough precursors to give at least one focus. This finding suggests that a given chromosome usually acquires the same (or nearly the same) number of precursors in every meiotic nucleus. For longer chromosomes, the value of B is not very important (e.g. Figure 6K; further discussion below).
- (6) Zip3 focus analysis reveals that the shortest yeast chromosome (III) has a significant number of zero-event chromosomes (1%) whereas the two longer chromosomes have much lower





**Figure 6. Experimental and BF analysis of CO patterns in budding yeast.** Panels (A, B): Experimental System. (A) Spread yeast pachytene chromosomes fluorescently labeled for SC component Zip1, CO-correlated foci of ZMM protein Zip3, and terminally labeled at the end of Chromosome XV by a *lacO/LacI-GFP* array. (B) Positions of Zip3 foci along a single Chromosome XV bivalent were defined as shown. Panels (C–H): Experimental CO patterns for Chromosome XV. (C) CoC and ED relationships for a single representative Chromosome XV data set reflecting CO positions defined along >300 bivalents (as in (B)). Average CoC curve (black line) shows  $L_{CoC} = 0.3 \mu\text{m}$ . (D–G) CoC curves and EDs for four independent experiments like that in (C). (D) shows the four individual average CoC curves; data set from panel (C) in black. (E) shows the four curves from four independent data sets and their average (in red). (F) shows the average of the four average CoC curves with the standard error at each

inter-interval distance. (G) Shows the EDs for four independent experiments in (D–F) and the average (in red) with standard error. (H) Compares the average CoC curve and ED for an *mlh1Δ* mutant (blue) with those for WT (black; average of averages from panels (F) and (G)). Both ED relationships and CoC relationships in the mutant are as WT since Mlh1 acts very late (text). Panels (I–K): BF simulations of CO patterns for Chromosome XV (data from average of averages in panels (F) and (G)). (I) Best-fit simulation (red) versus experimental data (black). Best-fit simulation specifies relatively even spacing of precursors ( $E = 0.6$ ) and a constant number of precursors along the chromosome in all nuclei ( $B = 1$ ). Other parameter values are in Table 1. (J, K). Experimental data and best-fit simulation data (black and red, from panel (I)) are compared with simulation using the same parameter values as the best-fit simulation except that precursors are either randomly spaced (Panel J;  $E = 0$ ; green) or Poisson distributed among chromosomes in different nuclei (Panel K;  $B = 0$ ; blue). Best-fit simulations for Chromosome IV and III data and for chromosome XIV in BR are shown in Figure S5 A–D; parameter values in Table 1. Importantly, even spacing is important for the best fit in all cases. In contrast, constant and Poisson distributions give very similar matches to experimental data except for the case of Chromosome III, where constant distribution must be required to ensure a sufficiently low number of zero-CO chromosomes (Figure S5 EF).  
doi:10.1371/journal.pgen.1004042.g006

numbers (no zero-focus chromosome has been detected among >1000 chromosomes analyzed) (e.g. Figure 6G). These values are recapitulated by BF simulations (Figure 6I; Figure S5EF; further discussion below).

The ability of BF simulations to accurately describe yeast data is further supported by analysis of CO homeostasis and the “obligatory CO” as described below.

**Drosophila melanogaster.** CO interference was first discovered, and CoC analysis developed, by genetic analysis of *Drosophila* X chromosome (Introduction). These classical genetic data do provide for a quantitatively accurate CoC curve with  $L_{CoC} = \sim 6 \mu\text{m}$  (Figure 7A), because the interval sizes are small enough and the data set is large enough [33]. The average number of COs per bivalent defined by experimental analysis is 1.44 with an unusually high level of zero-CO bivalents (5%) (Figure 7B). Recent studies of  $\gamma$ -His2Av foci have defined a total of  $\sim 24$  DSBs for the entire genome in *Drosophila* female meiosis [34]. Assuming that DSBs are proportional to genome length, this implies 6 DSBs for the X-chromosome ( $N = 6$ ). A BF simulation with  $N = 6$  can quantitatively match the classical *Drosophila* X chromosome data for all descriptors (Figure 7AB). Best-fit simulation requires differential clamping at the two ends, with less clamping at the centromeric end (Table 2). This feature provides for the experimentally-observed tendency for CO distributions to be shifted away from that end (illustrated in [35]).

The ability of BF simulations to accurately describe *Drosophila* data is further supported by analysis of CO homeostasis and the “obligatory CO” as described below.

**Chorthippus brunneus (grasshopper).** CO sites along the L3 bivalent of grasshopper have been defined by analysis of chiasmata in 1466 diplotene nuclei [36,37]. These data yield a CoC curve with  $L_{CoC} = 28 \mu\text{m}$  and an ED with an average of 2.2

COs/bivalent and no detected zero-CO chromosomes (<0.07%; Figure 8A).

A prominent feature of the L3 bivalent is a severe paucity of COs in the centromere region (Figure 8B left panel). This feature presumably reflects a defect in occurrence of precursors (DSBs) in centromeric heterochromatin. Thus, for BF best-fit simulations, the precursor array was adjusted accordingly, to give a strong paucity of precursors in the centromere region (a “black hole”; Figure 8B left; C).

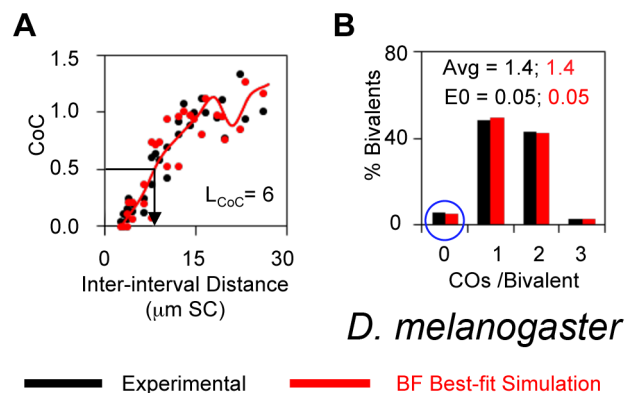
BF best-fit simulations can accurately describe L3 CoC and ED relationships, with or without inclusion of the black hole (Figure 8D, E); however, inclusion of the centromere precursor defect dramatically improves CO distributions along the chromosome, not only for total COs but for bivalents with two or three COs (Figure 8F versus 8G). Unlike all of the other cases analyzed above, there is no information for grasshopper regarding the number of precursors per bivalent. In BF simulations, the best fit between experimental and simulated data sets is provided when  $N = \sim 14$  (Figure 8H).

**Solanum lycopersicum (tomato).** CO patterns in tomato have been defined by analysis of Mlh1 foci ([38]; Lhuissier F.G. personal communication). Chromosomes in this organism exhibit a range of different pachytene bivalent lengths [38]. Bivalents comprise two groups, chromosomes 2–4 and 5–11, on the basis of longer and shorter SC lengths respectively (Figure 9A). We find that experimental CoC curves are significantly offset for the two groups when the metric of inter-interval distance is Mb (Figure 9A); in contrast, CoC curves for the two groups are superimposable when inter-interval distance is  $\mu\text{m}$  SC length (physical distance), with  $L_{CoC} = 11 \mu\text{m}$  (Figure 9B).

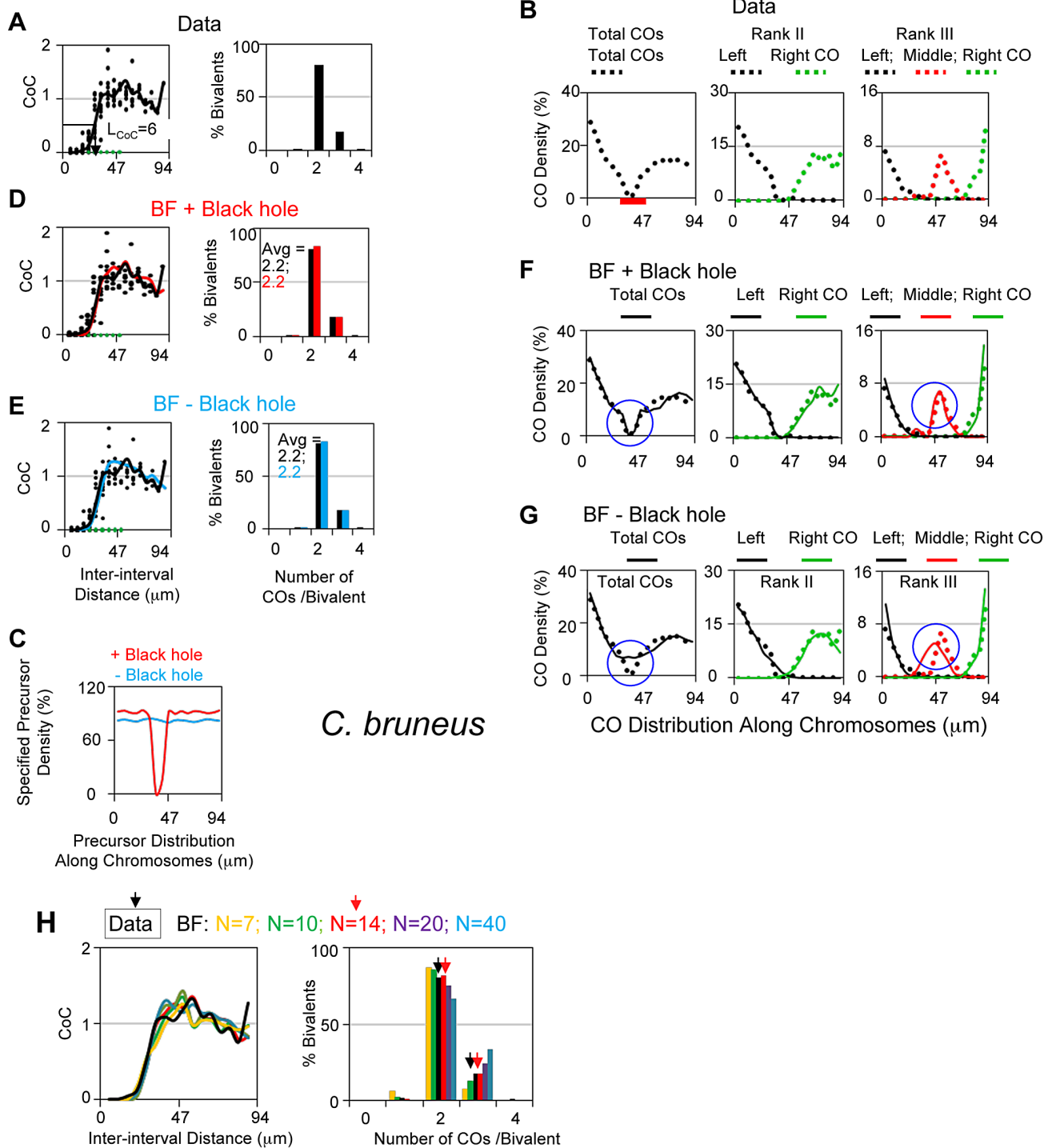
Physical distance has shown to be the appropriate metric for CO interference in mouse and *Arabidopsis* (above). The experimental results for tomato described above (Figure 9 AB) imply that this is also true in tomato. In addition, these data imply that the ratio of Mb to  $\mu\text{m}$  SC length is higher for shorter chromosomes than for longer chromosomes. This difference is explained, quantitatively, by two facts that: (i) heterochromatic DNA is much more densely packed along the SC than euchromatin and (ii) shorter chromosomes have a higher proportion of heterochromatin than euchromatin (and thus shorter SC lengths) (Figure S6).

CoC and ED patterns for both groups of chromosomes are well-described by BF simulations (Figure 9CD). All parameters have the same values in both cases, including the interference distance ( $L$ ) when expressed in  $\mu\text{m}$  SC length, except that precursor number varies with CO number/SC length (Figure 9CD legend).

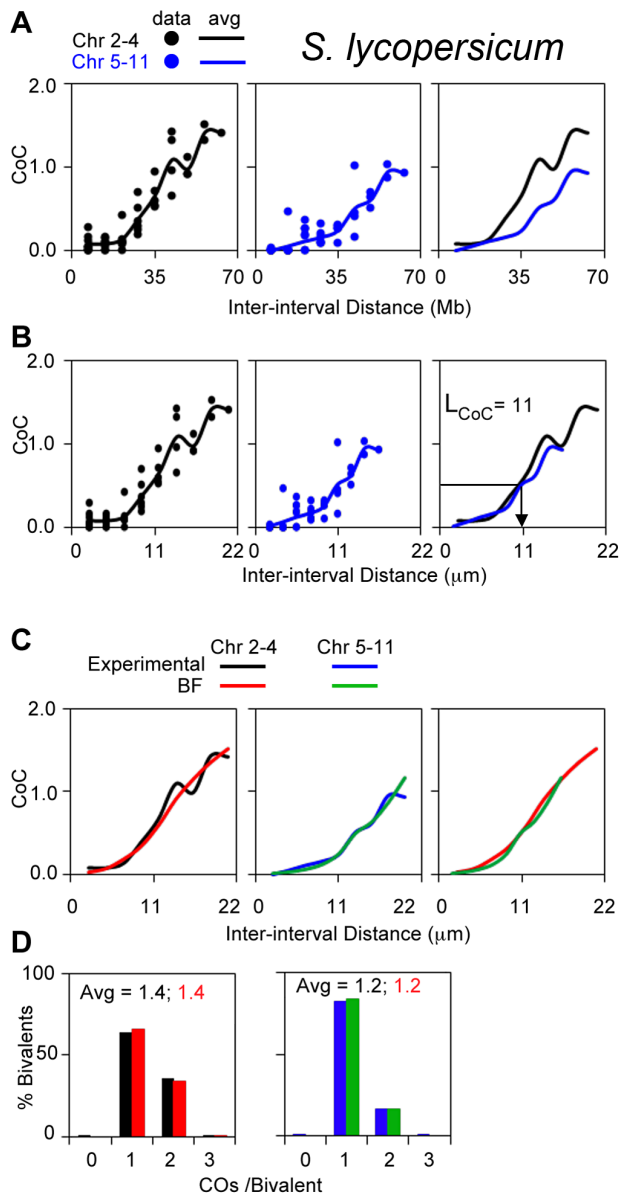
**Interference spans centromeres in grasshopper, yeast and tomato.** Previous analyses of interference have shown that CO interference is transmitted across centromeric regions [6,37–42]. Correspondingly, CoC values for interval pairs that span centromeres are almost indistinguishable from those for pairs separated by the same distances that do not span centromeres



**Figure 7. Experimental and BF analysis of CO patterns in *Drosophila*.** (A, B). Experimental data (black) and BF simulations (red) for the *D. melanogaster* X chromosome.  
doi:10.1371/journal.pgen.1004042.g007



**Figure 8. Experimental and BF analysis of CO patterns in grasshopper (*Chorthippus bruneus*).** (A, B) Experimental data. (A) CoC and ED relationships; (B) Distribution of COs along the bivalent for total COs (left) and for bivalents with either two or three COs with different colors for first, second (and third) COs from the left end of the bivalent. The centromeric region is labeled by a red bar. (C–H) BF simulation analyses. (C) Precursor density (frequency of precursors per interval specified for simulations, where the number of intervals = 17 as for CoC analysis) for BF simulations that used either an even distribution along the chromosome (blue; – Black hole) or a distribution where precursor levels decrease to zero over a region corresponding to the paucity of COs in the centromeric region (red; + Black hole; centromeric region defined in Panel B). (D) BF simulations with best-fit parameter values (Table 2) using the two precursor distributions defined in (C), i.e. with or without the centromere region black hole. CoC and ED relationships are the same in both cases (D and E); the distribution of COs along the chromosomes are well-fit when the black hole is included (F) as compared to when it is not (G). (H) BF simulations were used to estimate the likely value of N. CoC relationships seen experimentally (black) were compared by those given by BF simulations that use all best-fit parameter values except that the value of N (precursors per bivalent), which was varied from N = 7 to N = 40 (colors). Left panel: CoC relationships match the experimental curve for any  $N \geq 14$  (i.e. all curves except gold and green which are N = 7 and N = 10). Right panel: ED relationships are best fit by N = 14 (compare red and black), with less good fits at lower and higher values (right side). doi:10.1371/journal.pgen.1004042.g008



**Figure 9. Experimental and BF analysis of CO patterns in tomato.** (A, B) Experimental CoC data for chromosomes with shorter or longer SC lengths (2-4 and 5-11, respectively). Curves for the two groups are offset when inter-interval distance is expressed in units of Mb (genomic distance) (Panel A) and are overlapping when distance is expressed in units of  $\mu\text{m}$  SC (physical distance) (Panel B). (C, D) BF best-fit simulations for both groups of chromosomes using the physical distance metric ( $\mu\text{m}$  SC length) and adjusting precursor number as required to give  $\sim 20$  precursors per CO for each type of chromosome. This relationship is based on observation of  $\sim 280$  Rad51 foci [69] and 15 Mlh1 foci genome wide [38]. Interestingly, this relationship further implies that  $N$  is proportional to SC length. doi:10.1371/journal.pgen.1004042.g009

(Figure 10B). This is remarkable given the manifestly different structure in these regions. We further find that the same is true for budding yeast, based on Zip3 focus analysis (Figure 10A) and for tomato, from Mlh1 focus analysis (Figure 10C).

#### Part IV. CO Homeostasis

Experimental evidence has revealed that variations in the level of recombination-initiating double-strand breaks (DSBs) are not

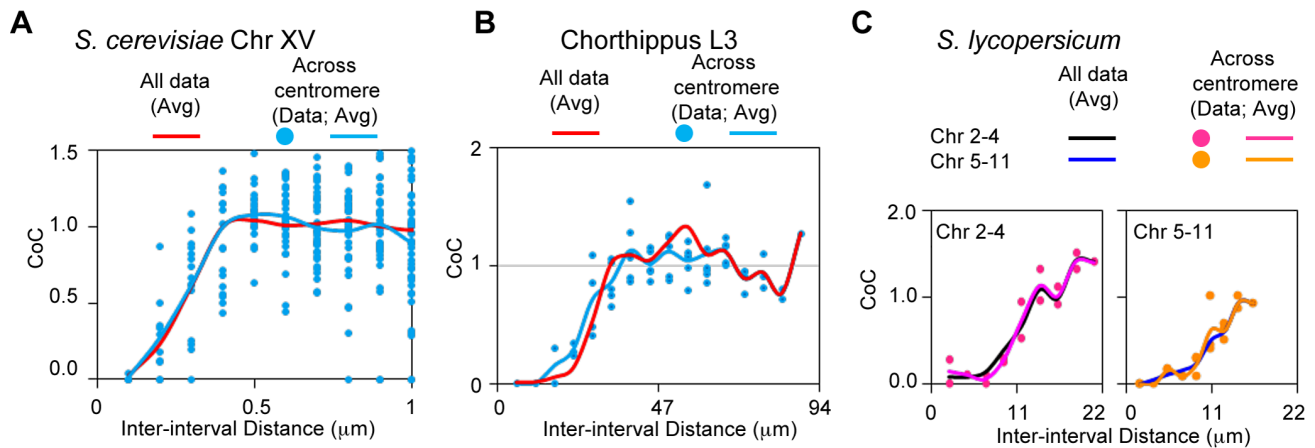
accompanied by corresponding variations in the number of COs. When DSB levels are either reduced or increased, CO levels are not reduced or increased commensurately [15,34,43–46]. This phenomenon is referred to as CO homeostasis [43].

According to the BF model, CO homeostasis is dependent upon, and in fact is a direct consequence of, CO interference (Figure 11A), as proposed [43,46]. In the absence of interference, the probability that a precursor will give rise to a CO is a function only of its own intrinsic properties, independent of the presence/absence of other precursors nearby. Thus, as the number of precursors decreases, the number of COs will decrease proportionately. In contrast, if interference is present, each individual precursor is subject to interference that emanates across its position from CO-designation events at neighboring positions. The lower the number of precursors, the less this effect will be. Thus, assuming a fixed level of CO interference, the frequency of COs per precursor will increase as the number of precursors decrease. Put another way: as the density of precursors decreases, the ratio of COs to precursors increases, even though there is no change in CO interference. Importantly, since CO homeostasis requires CO interference, its magnitude will also depend on the strength of CO interference as discussed below.

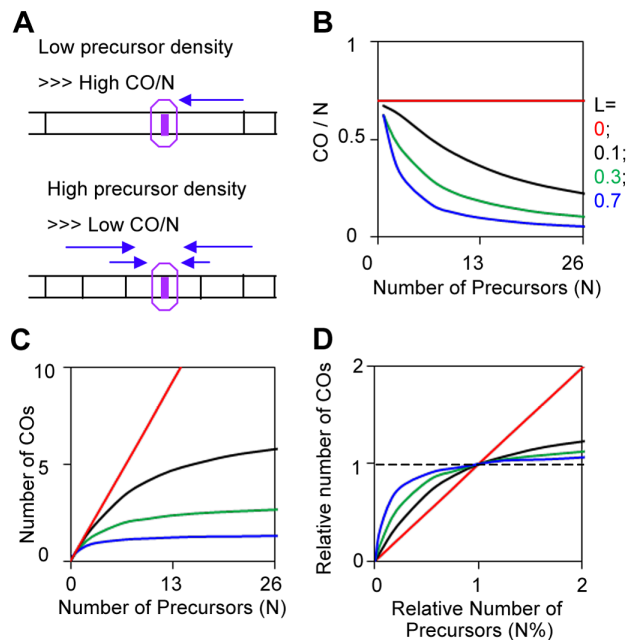
**In the BF model, CO homeostasis involves interplay between  $N$  and patterning parameters, e.g.  $L$  and  $S_{\text{max}}$ .** CO homeostasis for a given condition can be defined quantitatively by BF simulations in which the value of  $N$  (which is the number of precursors per bivalent and thus corresponds to precursor density) is varied, with all other parameter values remaining constant. Homeostasis can be described by plotting, as function of ( $N$ ), either the frequency of COs per precursor ( $\text{CO}/N$ ) or the total number of COs per bivalent (which corresponds to CO density) (Figure 11BC). Alternatively, CO homeostasis can be seen from the perspective of a starting wild-type situation with the number of COs per bivalent at the wild type precursor level taken as the point of reference and variations in the number of COs and precursors expressed relative to those reference values (Figure 11D).

The magnitude of CO homeostasis, i.e. the extent to which CO levels fail to respond to changes in precursor levels, will vary with the level of CO interference (above). This relationship can be described quantitatively by carrying out simulations for different values of ( $N$ ) at different values of patterning parameter(s), illustrated here for variations in the interference distance ( $L$ ) (Figures 11B–D). In the absence of interference, CO levels vary directly with precursor levels; the greater the interference distance, the less the change in CO levels with precursor levels. Notably, at very long interference distances, CO levels do not change at all with precursor levels. This is basically because interference precludes CO-designation at all other precursor sites, thus rendering variations in CO density irrelevant. This latter situation has recently been documented for *C.elegans* [44,45]. Analogous effects can be seen for variations in  $S_{\text{max}}$  (not shown), in accord with the fact that both  $L$  and  $S_{\text{max}}$  play important roles for  $L_{\text{CoC}}$  (above).

**BF simulations of CO homeostasis in experimental data sets can be used to evaluate the validity of the BF model.** BF best-fit simulations for wild type experimental data sets provide specific predicted values of all BF parameters, including  $N$ ,  $L$  and  $S_{\text{max}}$  (above). Given this starting point, BF simulations can then specifically predict how CO levels will vary if the level of DSBs (precursors) is decreased or increased. If the BF model accurately describes CO patterning, the best-fit simulation will accurately describe the experimental data. Application of this approach to budding yeast and *Drosophila* shows that the BF



**Figure 10. Interference spreads across centromeres in yeast, grasshopper and tomato.** CoC values for interval pairs that span centromere regions, and the corresponding average CoC curves, closely match the average CoC curves for all interval pairs for yeast (Panel A), grasshopper (Panel B) and tomato, both bivalent groups as described in Figure 9B (Panel C). doi:10.1371/journal.pgen.1004042.g010



**Figure 11. CO homeostasis and quantification by BF simulations.** (A) CO homeostasis is the phenomenon that, as the level of DSBs (precursors) increases or decreases, there is a less than proportional change in the frequency of COs. The basis for this effect is illustrated. At lower (higher) precursor density, a given precursor will be less (more) likely to be subject to interference and thus more (less) likely to give a CO. These relationships further imply that the extent of CO homeostasis at a given precursor density will also vary with interference distance and the strength of the DDF ( $L$  and  $S_{max}$ ) and other patterning features as reflected in  $L_{CoC}$  (text). (B, C, D) CO homeostasis was modeled by BF simulations using standard parameter values (Figure 3) with the indicated variations in the number of precursors ( $N$ ), and varying values for the interference distance ( $L$ ). CO homeostasis can be viewed as a function of  $N$  and  $L$ , as an effect on the probability that a single precursor will become a CO (B); or as an effect on the total number of COs along a chromosome (C). Also, the values of  $N$  and of COs can be defined as 1 for the wild-type reference situation and the effects of varying  $N$  (as a percentage of the reference value) can be seen directly as effects on CO levels (also as a percentage of the reference value). This approach considers the “densities” of precursors and COs rather than the absolute levels. doi:10.1371/journal.pgen.1004042.g011

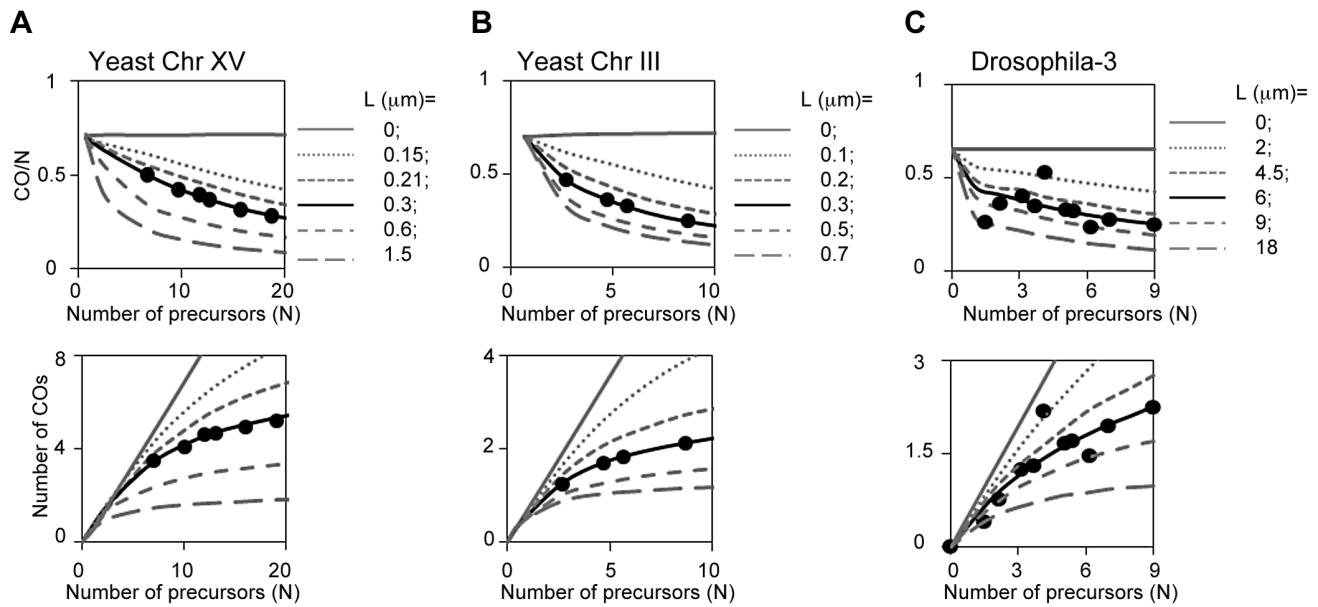
model can quantitatively predict experimental CO homeostasis patterns in both organisms.

The BF model quantitatively describes CO homeostasis in yeast: We asked whether the BF model could quantitatively explain CO homeostasis along chromosomes XV and III as defined by Zip3 focus analysis (above). We determined experimentally the number of Zip3 foci that occur along the two test chromosomes in a series of mutants that are known to exhibit particular, defined decreases or increases in DSB levels. Reductions in DSB levels were provided by the hypomorphic alleles of DSB transesterase Spo11 used to originally define CO homeostasis [43]. An increase in DSB levels was provided by a *tel1D* mutation, which increases DSB levels without significantly altering CO interference ([13]; unpublished). The observed experimental relationships are described by the BF-predicted relationships (Figures 12AB,  $L = 0.3 \mu\text{m}$ ).

The robustness of BF simulations is further supported by analysis of CoC and ED relationships for Zip3 foci in each of the DSB mutant strains. Best-fit simulations of these data sets should occur at exactly the same values of all parameters except for the number of precursors ( $N$ ), which should match that defined by experimental analysis of DSBs. Both of these predictions are fulfilled (Figure S7). These comparisons also reveal some interesting subtleties to DSB formation in the mutants (Figure S7).

The BF model quantitatively explains CO homeostasis in *Drosophila*: We analogously evaluated whether BF simulations accurately predict CO homeostasis relationships in *Drosophila*, which were defined experimentally by analysis of a fragment of chromosome 3 [34]. In *Drosophila*, as in yeast, CO-versus-DSB experimental data exhibit a very good match to the CO-versus- $(N)$  relationships predicted using BF best-fit parameter values, uniquely and specifically at the best-fit value of ( $L$ ) (Figure 12C).

BF simulations should (and do) accurately predict CO homeostasis in interference-defective mutants: CO homeostasis relationships should be altered in mutants where interference is defective, in a predictable way according to the magnitude of the reduction (Figure 12ABC, grey lines above the curves describing the wild-type relationships). We will present elsewhere data showing that BF best-fit simulations for a yeast mutant specifically defective in interference accurately predict CO homeostasis relationships in that mutant, thus further supporting the validity of BF simulation analysis.



**Figure 12. CO homeostasis in yeast, *Drosophila*, tomato and grasshopper.** (A, B) CO homeostasis relationships were determined experimentally for yeast chromosomes XV and III (text; Figure S7); plotted (filled circles); and compared with the values predicted from BF simulations based on best-fit parameter values for the two chromosomes ( $L = 0.3 \mu\text{m}$ , black lines). Predictions for other values of  $L$  are shown for comparison (grey lines). Experimental data precisely match BF simulation predictions. (C) CO homeostasis relationships determined experimentally for *Drosophila* chromosome 3 were compared with the predictions of the corresponding BF simulation ( $L = 6 \mu\text{m}$ , black line). Predictions for other values of  $L$  are shown for comparison (grey lines). Experimental data precisely match BF simulation predictions. [Note: *Drosophila* analysis was carried out as follows. Variation in CO number as a function of DSB level was determined experimentally for a fragment of chromosome 3 [34]. We first defined the theoretical CO homeostasis curve (black line) for full length chromosome 3 by BF simulations using the same set of parameters defined for chromosome X except that the number of precursors was adjusted in proportion to relative chromosome length (chromosome 3 is  $1.5 \times$  the length of chromosome X). The CO levels observed experimentally on the chromosome 3 fragment at the different DSB levels were then adjusted to those expected for the full length chromosome under the assumption that CO frequency is proportional to chromosome length, as in all other cases (above). The resulting experimental values were then compared with the theoretical curve and its relatives constructed at varying values of  $L$  (grey lines).] doi:10.1371/journal.pgen.1004042.g012

**The strength of CO homeostasis reflects the ratio of inter-precursor distance and interference distance ( $L_{CoC}$ ).** It would be convenient to have a standard way of comparing different situations (e.g. different chromosomes, mutants or organisms) with respect to the “strength” of CO homeostasis. In principle, the strength of homeostasis should vary according to the ratio between the CO interference distance ( $L_{CoC}$ ) and the distance between adjacent precursors (roughly given by the number of precursors divided by  $\mu\text{m}$  SC length). If this ratio is higher homeostasis will be stronger because a greater fraction of precursors are within the interference distance and thus can be eliminated without effect; and in the limit, reduction of precursor density will have no effect whatsoever. Oppositely, if this ratio is lower; homeostasis will be weaker because a greater fraction of precursors are outside of the interference distance and thus, when eliminated, will directly reduce CO levels; and in the limit, when the interference distance is zero, CO homeostasis is absent. By this criterion, i.e. [ $L_{CoC}$ /average inter-precursor distance from the data in Table 2], the strength of CO homeostasis is the same for all yeast chromosomes ( $0.3 \mu\text{m}/0.3 \mu\text{m} = 1$ ); are essentially the same for yeast chromosomes as for the *Drosophila* X chromosome ( $6 \mu\text{m}/5 \mu\text{m} = 1.1$ ); is significantly greater for grasshopper ( $28 \mu\text{m}/2.6 \mu\text{m} = 10.8$ ); and is even greater for tomato chromosomes, both groups ( $11 \mu\text{m}/0.888 \mu\text{m} = 12.5$ ). Systematic exploration of such relationships by BF simulations remains for future studies.

#### Part V. The “Obligatory CO”

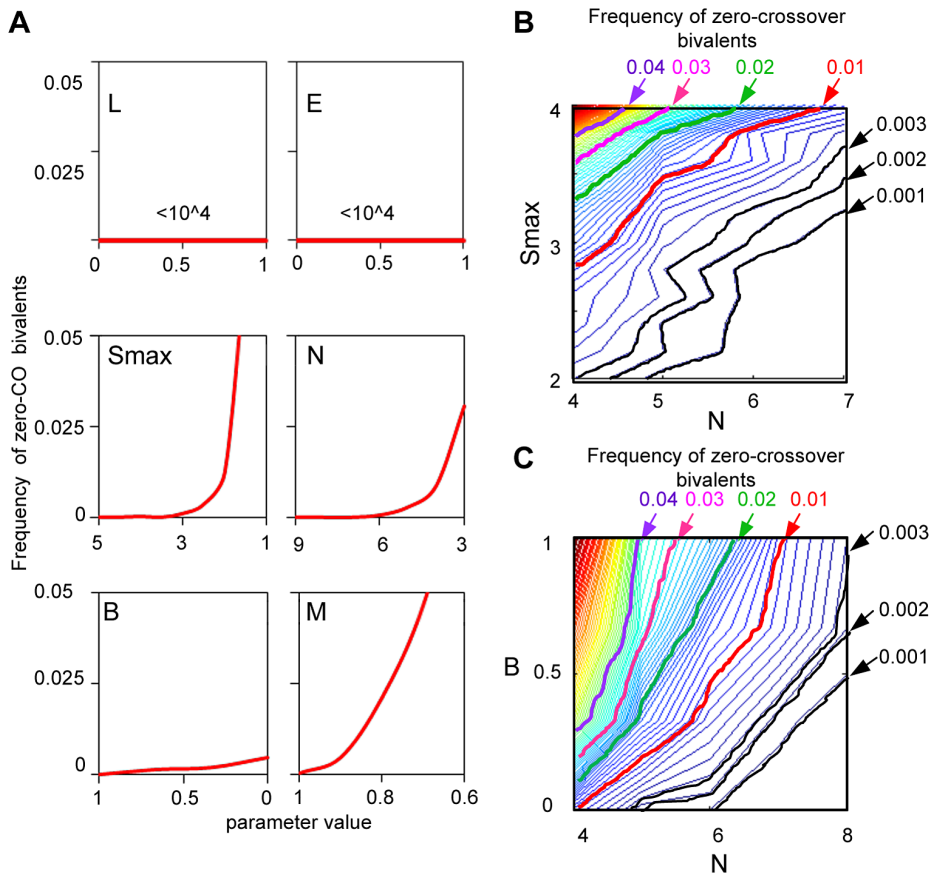
Regular segregation of homologs to opposite poles at the first meiotic division requires that they be physically connected. During

meiosis in all organisms, in at least one sex and usually both, the requisite physical connection is provided by the combined effects of a crossover between non-sister chromatids of homologs and connections between sister chromatids along the chromosome arms. Correspondingly, in such organisms, in wild-type meiosis, every bivalent almost always acquires at least one CO [47]. This first CO that is essential for homolog segregation is often referred to as the “obligatory CO”. In fact, the obligatory CO is simply a biological imperative: the level of zero-CO chromosomes should be low. The CO patterning process, by whatever mechanism, must somehow explain this feature.

In most situations, the frequency of zero-CO bivalents is extremely low ( $<10^{-3}$ ), but higher frequencies also occur in certain wild-type situations as well as in certain mutants (below). In some models for CO patterning, the obligatory CO is ensured by a specific “added” feature of the patterning process (e.g. the King and Mortimer model; Discussion). In contrast, in the beam-film model, the requirement for one CO per bivalent is satisfied as an intrinsic consequence of the basic functioning of the process, as follows.

**In the BF model, the “obligatory CO” is independent of ( $L$ ) and ( $E$ ) and requires an appropriate combination of values for ( $S_{max}$ ), ( $N$ ), ( $B$ ) and ( $M$ ).** In the BF model, the obligatory CO is ensured as an intrinsic consequence of all of the features that ensure occurrence of a first event; features that act later in the process are not relevant (Figure 13A).

$L$  and  $E$ : Variations in ( $L$ ) and ( $E$ ) have no effect on the level of zero-CO chromosomes because: (i) spreading inhibition of CO-designation (“interference”) only affects the number of COs after



**Figure 13. How is a low level of zero-CO bivalents ensured by the beam-film model?** (A) BF simulations were carried out under a set of “standard conditions” (as described in Figure 3), except that  $S_{max} = 5$ ,  $N = 8$ , while the value of one parameter was systematically varied as indicated in each panel. The frequency of zero-CO chromosomes as a function of the value of the varied parameter is plotted. Variations in L and E have no effect; variations in  $S_{max}$ , N, B and M all have effects. (B,C). A given frequency of zero-CO chromosomes can be achieved by diverse constellations of parameter values that play off against one another. This situation is illustrated by BF simulations under above conditions where the frequency of zero-CO chromosomes is determined over a range of combinations of values of two parameters. This interplay is illustrated for combinations of precursor number (N) and either parameter  $S_{max}$  (Panel B) or parameter B (Panel C). The lower the number of precursors, (N), the higher the  $S_{max}$  needed to ensure that at least one will be sensitive enough to undergo CO designation and the more important it will be for precursors to occur at the same average number along the bivalent in every nucleus, thus minimizing the probability that a bivalent will have too-few precursors (higher B) (text).

doi:10.1371/journal.pgen.1004042.g013

the first CO designation; and (ii) the distribution of multiple events along a chromosome also comes into play only after the first designation event has occurred. In contrast, essentially all other basic BF parameters are important in ensuring a low level of zero-CO chromosomes:

**$S_{max}$ , N and B:** The frequency of zero-CO chromosomes will be minimized if every chromosome has at least one precursor that is adequately sensitive to the DDF. This effect can be favored by either (i) higher  $S_{max}$ ; (ii) higher N; or (iii) higher B. Higher  $S_{max}$  means that a higher fraction of sites in a particular precursor array will be adequately sensitive. Higher N means that there will be more sites and there is a higher chance that an adequately-sensitive precursor will be present. Higher B means a reduced probability that a bivalent will have a lower-than-average number of precursors and thus a lower-than-average chance for an adequately-sensitive precursor to be present.

(M): Even if the CO patterning process ensures the occurrence of at least one CO per chromosome, a defect in maturation of CO-designated interactions to detectable COs will tend to counteract that effect, converting chromosomes with one (or a few) COs to chromosomes with zero-COs.

(A): The more likely a precursor is to give a CO-designation in response to a particular local level of interference, the lower will be the frequency of zero-CO chromosomes. The frequency of zero-CO chromosomes is lowest for  $A = 4$  and increases progressively for  $A = 3, 2$  and 1 (not shown).

**Evolutionary implications.** The above considerations imply that, according to the BF model, the “obligatory CO requirement” will be met in any given organism because the relevant features have been coordinately tuned by evolution into a combination that ensures a low level of zero-CO chromosomes. That is: a suitably low level of zero-CO chromosomes can be achieved by a variety of combinations, with more and less favorable values of different parameters in different cases. Interplay between pairs of parameters is illustrated for various combinations of ( $S_{max}$ ) and (N) (Figure 13B) and for various combinations of (N) and (B) (Figure 13C). The lower the number of precursors, (N), the higher the  $S_{max}$  needed to ensure that at least one will be sensitive enough to undergo CO designation and the more important it will be for precursors to occur in a constant level along the bivalent in every nucleus (higher B).

**Application to experimental data: When the obligatory CO appears to be “missing”, in wild-type meiosis.** In most organisms, for most chromosomes, in wild-type meiosis the observed frequency of zero-CO chromosomes is  $<0.1\%$  (e.g. Table 1). However, zero-CO chromosomes occur at significantly higher levels on the *Drosophila* X chromosome (5%) and on yeast chromosome III (1% by Zip3 foci).

Yeast: the importance of N: BF analysis suggests that, in yeast, the feature responsible for the high level of zero-Zip3 focus chromosomes along chromosome III is simply a paucity of precursors and thus is a function of its diminutive size, *per se*. In this organism, patterns on all analyzed chromosomes can be described by the same set of BF parameters with the exception of (N), which varies roughly in proportion to chromosome length as defined in  $\mu\text{m}$  SC (above). The high level of zero-Zip3 focus chromosomes along chromosome III, relative to other chromosomes, is thus solely a reflection of the fact that it is much shorter than other chromosomes (N = 6 versus  $N \geq 13$ ; Table 2). This conclusion is directly and strongly supported by experimental analysis of Zip3 foci along chromosome III in mutants where precursor levels are gradually decreased or increased (by alterations in DSB levels; described below). In such mutants, the frequencies of zero-Zip3 focus chromosomes are commensurately increased or decreased. In mutants with relative DSB levels of 1.5, 1 (WT), 0.8 and 0.7, the frequencies of zero-CO chromosomes are, respectively,  $<0.003$ ; 0.01; 0.02; and 0.08 (Figure S7B).

Analysis of yeast chromosome III further shows that the best-fit simulation requires that precursors occur in a constant number per bivalent in all nuclei (B = 1; frequency of zero-Zip3 focus bivalents = 0.01). If precursors are Poisson distributed among bivalents in different nuclei (B = 0), the frequency of zero-Zip3 focus bivalents increases to 0.04 (Figure S5). This comparison not only suggests that DSBs/precursors always occur at the same number along a different chromosome but provide a rationale for the existence of this feature which, in the general case, is essential to minimize the frequency of zero-CO bivalents along short chromosomes.

*Drosophila* X chromosome: The high level of zero-CO bivalents for the *Drosophila* X chromosome is recapitulated in BF simulations without adding any unusual features (Figure 7B), suggesting that there is nothing remarkable about this chromosome. Further, this chromosome the same number of precursors as yeast chromosome III (N = 6), suggesting that here, too, the fact that the chromosome is “too short” could be an important factor for the high level of zero-CO chromosomes. In accord with this possibility, *Drosophila* chromosome 3 is 50% longer than chromosome X; and BF simulation analysis shows that if X chromosome length is increased by 50%, with a proportional increase in precursors (to N = 9), and without any change in any other parameter, the frequency of zero-CO would decrease to 0.008.

However: the number of precursors cannot be the only relevant feature, because the level of zero-CO chromosomes in *Drosophila* is higher than that for yeast chromosome III even though they both have N = 6. This difference could be attributable in part to less regular distribution of precursors among chromosomes in *Drosophila* (B = 0.5 versus B = 1 in yeast) and to a lower DDF level (Smax = 2.8 versus 3.5 in yeast [Table 2]).

Implications: The general implication of these considerations is that CO patterning features have evolved to give a very low level of zero-CO bivalents, in accord with the biological imperative. However, in certain organisms, the constellation of features may be tuned to just such a level that the number of precursors along shorter chromosomes is just at the limit of the necessary threshold.

Notably, also, in both *Drosophila* and yeast, additional mechanisms exist which complement the patterned CO system to ensure regular homolog segregation. *Drosophila* exhibits robust CO-independent segregation [48] and, in budding yeast, the significant level of non-interfering COs also ensure disjunction. For example: 1% of chromosome III's exhibit zero Zip3 foci whereas only 0.1% of chromosome III's exhibit no COs as defined genetically (Hunter N and Bishop D, personal communication).

**Application to experimental data: When the obligatory CO appears to be “missing” in mutant meiosis.** Mutant phenotypes that affect interference and/or the “obligatory CO” could fall into three different categories [47]:

IF– OC+: The BF model predicts that mutants with decreased CO interference, as defined specifically by decreased L, will show no defect in formation of the first (obligatory) CO, i.e. will show the same level of zero-CO chromosomes as that seen in WT meiosis. This phenotype, “IF– OC+”, not previously reported, is now described by observations in yeast to be presented elsewhere.

IF+ OC–: In certain mutants, the level of zero-CO bivalents is increased but interference is unaltered. In the context of the BF model, this phenotype could arise from several types of defects. For example: this phenotype is expected for mutants with altered recombination biochemistry such that CO designation is normal but CO maturation is inefficient, i.e.  $M < 1$  (Figure 5A). For example, mutants lacking Mlh1 exhibit reduced levels of COs and significant levels of zero-CO chromosomes but relatively normal CO interference, as shown by genetic/chiasma analysis for several organisms [49–51]. On the other hand, by Zip3 focus analysis, both CoC and ED in a yeast *mlh1*Δ are essentially the same as in WT (Figure 6H), implying that CO-designation is normal. Thus the obligatory CO defect seen by genetic/chiasma analysis is specifically attributable to a maturation defect ( $M < 1$ ). Correspondingly, there are multiple indications that Mlh1 acts very late in recombination, for maturation of dHJs to COs [50,51].

This phenotype could also be conferred by genetic shortening of a chromosome. A reduction in the Mb length of the chromosome will decrease the precursor number (N) without alternation of any other properties and thus could push N below the minimum necessary threshold. Such an effect could explain the high level of zero-CO chromosomes seen when plant chromosomes are shortened by centric fission [47] and when yeast chromosome III is reduced in length (Hunter N., and Bishop, D.K., personal communication).

IF– OC–: This phenotype could be conferred in several ways. One example would be a defect early in the recombination process that eliminate the ability of recombination precursors to both generate COs and generate/respond to the CO interference signal.

Implications: The existence of IF+ OC– mutants has sometimes been cited as evidence for the existence of a specific feature, separate from interference, that “ensures the obligatory CO”. The existence of IF– OC– mutants has sometimes been cited as evidence that interference is required for the “obligatory CO”. Both of these phenotypes have alternative explanations in the context of the beam-film model. In contrast, the existence of IF– OC+ mutants is specifically predicted by the beam-film model.

## Part VI. Non-interfering COs

In some organisms, a significant fraction of COs arises outside of the patterning process. The existence of these “non-interfering” COs is most rigorously documented for budding yeast, where the number of “non-interfering” COs is  $\sim 30\%$  among total COs (by compassion the number of patterned COs defined by analysis of CO-correlated Zip2/Zip3 foci with the number of



total COs from genetic and microarray analyses) (e.g. [22,31,52,53]; below).

The origin of non-interfering COs is unknown. One possibility is that they arise from the majority subset of interactions that do not undergo CO-designation [5]. By this model (“Scenario 1”; Figure 14A left), not-CO-designated interactions would mostly mature to NCOs but sometimes would mature to COs, analogously to the situation in mitotic DSB-initiated recombinational repair [54]. Alternatively, such COs might arise from some other set of DSBs that arise outside of the normal process, e.g. because they occur later in prophase after CO-designation is completed or earlier in prophase before patterning conditions are established (“Scenario 2”; Figure 14A right).

Both scenarios can be examined using the BF simulation program. To simulate the outcome of Scenario 1, where non-patterned COs arise from non-designated interactions left over after patterning, a standard CO-designation BF simulation is performed to define the interfering COs; the precursors that have *not* undergone CO-designation are then used as the starting array of precursors for a second round of CO-designation. In this second round, COs are randomly selected from among the precursors remaining after the first round of designation. The COs resulting from the two simulations are then combined and the total pattern is analyzed.

To model Scenario 2, in which non-patterned COs arise from an unrelated set of precursors, a standard CO-designation BF simulation is performed to define interfering COs. Then a second, independent simulation is performed using a specified number of precursors that are unrelated to the first set and random selection of COs from among that precursor set. COs generated by the two types of simulations are then again combined and analyzed.

CoC relationships for total COs (interfering plus non-interfering) will depend significantly on whether the precursors that give rise to the “non-interfering” COs are evenly or randomly spaced along the chromosomes. CoC curves for total COs reflect the combined inputs of CoC relationships for interfering COs and non-interfering COs. CoC curves for interfering COs are affected only modestly by even-versus-random spacing due to the overriding effects of CO interference (above; e.g. Figure 14B left). However, non-interfering CO relationships are a direct reflection of precursor relationships, which differ dramatically in the two cases. For precursors, CoC = 1 for random spacing and significant “interference” for even spacing; Figure 14B second from left). CoC relationships for non-interfering COs alone exhibit the same features (Figure 14B, rightmost two panels). These differences are directly visible in CoC curves for total COs, with greater or lesser prominence according to the relative abundance of non-interfering COs versus interfering COs (Figure 14C). Notably, CoC relationships for Scenario 1, where precursors exhibit the even spacing defined by BF best-fit simulations ( $E = 0.6$ ), show a qualitatively different shape than CoC relationships under Scenario 2.

Given this framework, we defined CoC curves for total COs along yeast chromosomes IV and XV as defined by microarray analysis (Figure 14D left panel). The general shapes of these experimental curves correspond qualitatively to those predicted for emergence of non-interfering COs from an evenly-spaced precursor array, with a closer correspondence to those predicted for Scenario 1 than to those predicted for Scenario 2 (compare Figure 14D left panel with Figure 14C).

This impression is further supported by BF simulations. To model Scenario 1, we began with the set of best-fit parameters defined for interfering COs (Zip3 foci) above (Figure 6I) and generated predicted total CoC curves, assuming that non-interfering COs comprise 30% of the total (above), for each

of the three possible case of non-interfering COs: Scenario 1 (where precursors are assumed to be evenly spaced as for interfering COs); and Scenario 2 with precursors assumed to be either evenly or randomly spaced (Figure 14D, second panel from left). The CoC curve for the first of these three cases has the same shape as the experimental CoC curves for total COs (compare Figure 14D left and second from left panels) and direct comparison shows that it gives a quite good quantitative match with the experimental curves (Figure 14D third panel from left). Scenario 2 with evenly-spaced precursors is a less good match (Figure 14D, right panel). Scenario 2 with randomly-spaced precursors (Figure 14D, second panel from left, red) is a quite poor match (not shown).

These analyses suggest that, in yeast, non-interfering COs arise from the not-CO-designated precursors as a minority outcome of the “NCO” default pathway (Figure 14A, Scenario 1).

## Part VII. Variations in Diverse BF Parameters Alter the Value of the Gamma Shape Parameter

Many studies of CO interference characterize CO patterns by defining a gamma distribution that best describes an experimentally observed distribution of the distances between adjacent COs, often with the assumption (implicit or explicit) that a higher value of the gamma shape parameter ( $\nu$ ) corresponds to “stronger” CO interference (e.g. [11]). We have examined the way in which ( $\nu$ ) varies as a function of changes in the values of several BF parameters. Variations in  $L$  or  $S_{max}$  increase or decrease the value of ( $\nu$ ) in correlation with increased or decreased  $L_{CO}$ ; and in opposition to the average number of COs per bivalent (Figure 15AB, compare green line and blue/pink distributions with red and black lines). This is the pattern expected for a change in the “strength of interference”. In contrast, the value of ( $\nu$ ) is also altered by variations in  $M$  or  $N$ , which have little or no effect on  $L_{CO}$ ; moreover, the change in ( $\nu$ ) co-varies with the change in the average number of COs per bivalent (Figure 15CD, compare green line and blue/pink distributions with red and black lines). The BF model thus implies that a change in the value of ( $\nu$ ), e.g. in a mutant as compared to wild type, may or may not imply a change in the patterning process per se. However, comparison of the variation in ( $\nu$ ) with the variation in average COs per bivalent can distinguish between the two possibilities, with opposing variation implying a patterning difference and co-variation implying a difference in some other feature.

## Discussion

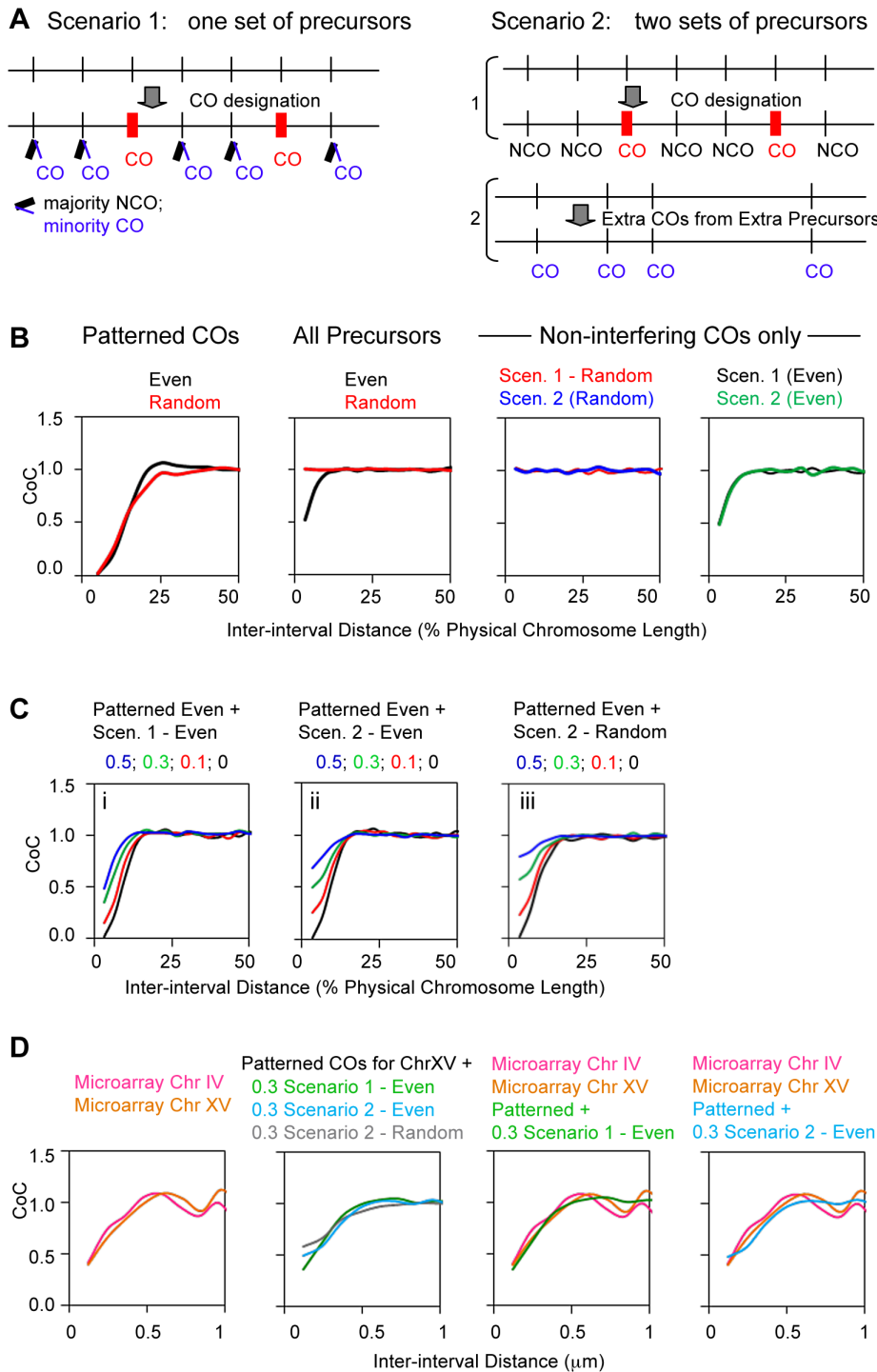
The presented analysis has provided new information on CO patterning from several different perspectives.

### BF Simulations Accurately, Quantitatively Describe Experimental Data Sets

This is true not only with respect to CoC and ED relationships but with respect to more detailed effects such as CO homeostasis and the obligatory CO. These matches, and the information that emerges there-from, support the notion that the basic logic of the BF model provides a robust and useful way of thinking about CO patterning. These matches are also specifically supportive of the proposed mechanical stress-and-stress relief mechanism.

### New Information about CO Patterning in Several Organisms

In budding yeast: (i) CO patterning has the same basic features for shorter and longer chromosomes; (ii) Mlh1 is required



**Figure 14. Modeling of non-interfering COs with the example of budding yeast.** (A) Non-interfering COs could arise from the same array of precursors as designated COs by low probability of CO formation at remaining not-CO-designated sites (left); alternatively, non-interfering COs could arise from a different set of precursors (right). (B) BF simulations. (Far left) Simulations under standard parameter conditions (Figure 3) illustrate the facts that CoC relationships for interfering COs vary relatively little according to whether precursors are randomly or evenly spaced; see also Figure 3 and Figure S2). (Second from left) An array of precursors was defined under standard parameter conditions with  $E=0.6$  (relatively even spacing) or random spacing ( $E=0$ ) and subjected to a simulation at very high  $S_{\text{max}}$  such that all precursors were manifested as “events”, for which CoC curves were then constructed. Precursor CoC relationships are dramatically different for random spacing (CoC = 1 for all inter-interval distances) or even spacing (significant interference at small inter-interval distances). (Right two panels) CoC curves for non-interfering COs generated by Scenario 1 or Scenario 2 with even or random precursor spacing ( $E=0.6$  or  $E=0$ ) directly match the CoC curves for the precursors from which they arose. (C) BF simulations for mixtures of interfering and non-interfering COs, at the different indicated proportions of non-interfering COs. Interfering COs were generated under standard parameter conditions which correspond to yeast chromosome XV (Figure 3;  $E=0.6$ ). Non-interfering COs were generated according to Scenario 1 (left) or by Scenario 2 from either even or random precursors ( $E=0.6$ ;  $E=0$ ) (middle; right). (D) Left: Experimental CoC curves for CO data from microarray analysis of yeast Chromosomes IV and XV [22,52]. Data from the two studies were combined and the positions of COs

defined under the assumption, as made in both studies, that the CO position is at the site of the exchanged polymorphism. CoC analysis of these COs was then performed as for Zip3 foci. Second from left: CoC curves for yeast Chromosome XV corresponding to the combination of interfering COs as described by the best-fit simulation (Figure 6;  $E=0.6$ ) plus non-interfering COs at a level comprising 30% of total COs as described by best-fit simulations under the indicated Scenario and precursor array conditions. Two right-most panels: Experimental CoC curves for Chromosomes IV and XV (from left panel) plus best-fit simulation CoC curves for the indicated mixtures of interfering and non-interfering COs (from the second from left panel). The best match is provided if non-interfering COs arise from precursors left over after CO-designation (Scenario 1 with relatively evenly-spaced precursors;  $E=0.6$ ) as shown in second panel from right.  
doi:10.1371/journal.pgen.1004042.g014

specifically for CO maturation not for CO patterning; and (iii) Precursors are evenly spaced, as shown by both CoC analysis and analysis of total (interfering-plus-non-interfering) COs.

In tomato (and, to be described elsewhere, in budding yeast), the metric of CO interference is physical chromosome length ( $\mu\text{m}$ ) not genomic length (Mb). In the case of tomato, differences in CoC relationships expressed in the two different metrics is attributable to differential packaging of heterochromatin versus euchromatin along the chromosome plus differential proportions of heterochromatic versus euchromatic regions among different chromosomes.

In tomato and yeast, as previously described for grasshopper, human and several other organisms, crossover interference spreads across centromeres with the same metric as along chromosome arms.

In budding yeast, non-interfering COs arise from evenly-spaced precursors, most probably by occasional resolution of NCO-fated precursors to the CO fate.

### New Insights into CO Homeostasis and the “Obligatory CO”

With respect to CO homeostasis, the importance of CO interference as a determinant in the strength of homeostasis is emphasized and BF simulations are shown to permit accurate quantitative descriptions of homeostasis. Also, the strength of homeostasis can be seen to reflect the ratio of interference distance ( $L_{\text{CoC}}$ ) to the distance between adjacent precursors.

With respect to the obligatory CO, the general logic of the BF model (Figure 1) suggests that occurrence of a low level of zero-CO chromosomes is independent of CO interference (and precursor spacing) and is achieved by an appropriate evolved constellation of all other parameters. Explanations can also be provided for several known cases where the level of zero-CO chromosomes is unusually high, but interference is robust, and potential explanations for other mutant phenotypes are suggested. Importantly, the logic of the beam-film model predicts the existence of mutants that lack interference but still exhibit the obligatory CO, evidence for which will be presented elsewhere.

### Models and Mechanisms of CO Interference

The central issue for CO patterning is how information is communicated along the chromosomes. Three general types of mechanisms have been envisioned. (1) A molecular signal spreads along the chromosomes, e.g. as in the polymerization model of King and Mortimer [55] or the “counting model” of Stahl and colleagues [20,35]. (2) A biochemical reaction/diffusion process surfs along the chromosomes [56], as recently described in detail for bacterial systems [57,58]. (3) Communication occurs via redistribution of mechanical stress, as in the beam-film model [3,4] or via other mechanical mechanisms (e.g. [59]).

The counting model can provide good explanations of experimental data; however, the underlying mechanism is contradicted by experimental findings ([43]; but see [60]). No specific reaction/diffusion mechanism has been suggested thus far for CO interference. The King and Mortimer model and the

beam-film model are significantly different, in three respects. First, in the King and Mortimer model, the final array of COs reflects the relative rates of CO-designation and polymerization. Thus it is the kinetics of the system that governs its outcome. In the beam-film model, where interference arises immediately after each CO-designation, kinetics does not play a role. Second, in the King and Mortimer model, the interference signal continues to spread until it runs into another signal approaching from the opposite direction. In the beam-film model, the interference signal is nucleated and spreads for an intrinsically limited distance, with an intrinsic tendency to dissipate with distance from its nucleation site. Third, the King and Mortimer model envisioned that precursors were Poisson distributed among chromosomes. As a result, significant numbers of chromosomes would initially acquire no precursors if the average number of precursors is low and thus would never give a CO, thereby giving an unacceptably high level of zero-CO chromosomes. To compensate for this effect, the model proposed that the effect of interference was to release encountered precursors, which then rebound in regions that were not yet affected by interference (and thus on chromosomes with no precursors). This precursor turnover would ensure that all chromosomes achieved a precursor that could ultimately give a CO. Because of this feature, the King and Mortimer model envisions that interference is required to ensure a low level of zero-CO chromosomes (i.e. to ensure the “obligatory CO”). By the beam-film model, instead, precursors do not turn over and interference is not required to ensure a low level of zero-CO chromosomes, which results instead from an appropriate constellation of other features, as described above. The beam-film model predicts the existence of mutants that are defective in interference but do not exhibit an increase in the frequency of zero-CO chromosomes.

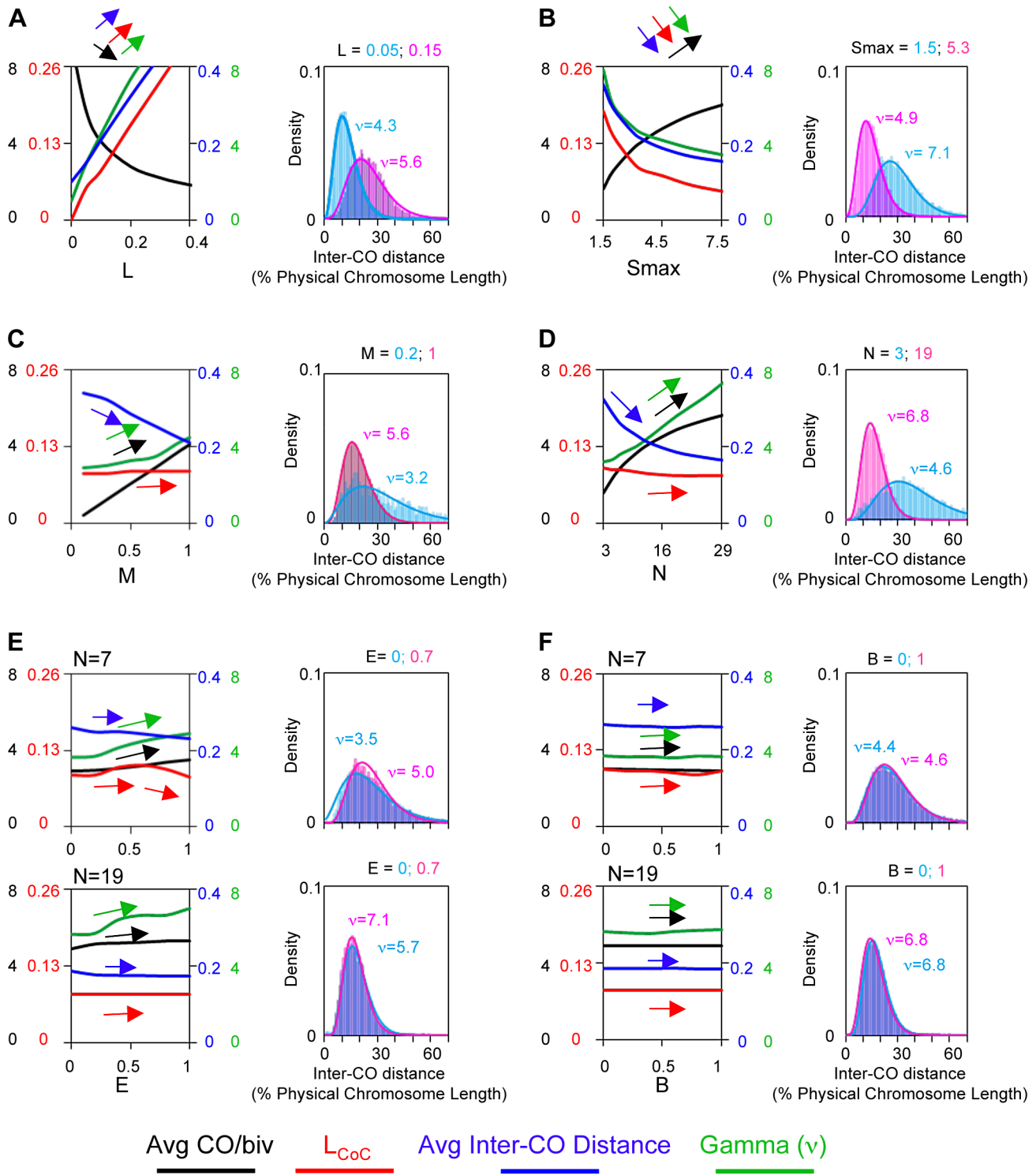
## Materials and Methods

### Yeast Strains

Yeasts SK1 strains (Figure 6 and S7) are described in Table S1. In all strains, ZIP3 carries a MYC epitope tag; a construct expressing LacI-GFP and is integrated at either *LEU2* or *URA3*, and a *lacO* array [61] is inserted at HMR (chromosome III), *Scp1* (Chromosome XV) or Chromosome IV telomere (SGD1522198) to specifically label each chromosomes by binding of LacI-GFP.

### Zip2/Zip3 Foci on Yeast Pachytene Chromosomes Correspond to Programmed (“Interfering”) COs

Pachytene chromosomes exhibit  $\sim 65$  foci of Zip2, Zip3 and Msh4, with strong colocalization of Zip3 and Msh4 foci ([31,62]; this work). Zip2 foci [63] exhibit interference as defined by CoC relationships for random adjacent pairs of intervals [31]. We further show here that Zip2 and Zip3 foci exhibit interference as defined by full CoC relationships along specific individual chromosomes (Figure 6 and Figure S5). Zip2 and Zip3 foci also both occur specifically on association sites of *zip1Δ* chromosomes [31,64]. The total number of COs per yeast nucleus as defined by microarray and genetic analysis is  $\sim 90$  [22,52,65] implying that



**Figure 15. Application of the gamma distribution to CO patterns: Risks and rewards.** The value of the gamma distribution shape parameter ( $\nu$ ) varies not only with variations in interference distance ( $L$ ) but with variations in other parameters. Notably, the value of  $\nu$  and the value of the average number of COs per bivalent vary inversely with changes in patterning parameters  $L$  and  $S_{max}$  but vary directly, or very little, with changes in precursor and maturation parameters. This is shown by BF simulations performed under standard conditions (Figure 3) except that individual parameters were systematically varied. The values of each of the four indicated descriptors (average CO/bivalent;  $L_{CoC}$ ; average distance between adjacent COs (“inter-CO distance”) and the gamma distribution shape parameter  $\nu$ ) were plotted as a function of the value of the varied parameter (left). Also, the distribution of distances between adjacent COs for representative values of the varied parameter are shown (right). (A–D) Effects of systematic variation of parameters  $L$ ,  $S_{max}$ ,  $N$  and  $M$ . (E, F) Effects of variation of  $E$  and  $B$  at two different values of  $N$ .

doi:10.1371/journal.pgen.1004042.g015

Zip2/Zip3/Msh4 foci represent  $65/90 = 70\%$  of the total. Correspondingly, mutant analysis suggests that “non-interfering” COs comprise  $\sim 30\%$  of total COs (e.g. [50]). Additionally, BF analysis accurately explains CoC relationships for total COs on the assumption of 70% patterned COs and 30% “non-interfering” COs (Figure 14, Results).

### Cytological Mapping Zip3 Foci in Yeast

Synchronous meiotic cultures (SPS sporulation procedure from [66]) were prepared and harvested at a time when pachytene nuclei are most abundant ( $\sim 4\text{--}5$  hours). Cells were spheroplasted and chromosomes spread on glass slides according to Loidl et al. and Kim et al. [67,68]. Primary antibodies were mouse monoclonal anti-myc, goat polyclonal anti-Zip1 (Santa Cruz) and rabbit polyclonal anti-GFP (Molecular Probes). Each was diluted appropriately in the above BSA/TBS blocking buffer. Secondary antibodies were donkey anti-mouse, donkey anti-goat, and donkey anti-rabbit IgG labeled with Alexa488, Alexa645 or 594 and Alexa555 (Molecular Probes), respectively. Stained slides were mounted in Slow Fade Light or Prolong Gold Antifade (Molecular Probes). Spread chromosomes were visualized on an Axioplan IEmot microscope (Zeiss) with appropriate filters. Images were collected using Metamorph (Molecular Devices) image acquisition and analysis software. Acquired images were then analyzed with Image J software (NIH), with total SC length and positions of Zip3 foci for the specifically labeled bivalent were measured from the *lacO/LacIGFP*-labeled end to the other end (Figure 6B bottom). For each type of chromosome analyzed (III, IV and XV) in each experiment, measurements were made for  $>300$  bivalents, one from each of a corresponding number of spread nuclei. Resulting data were transferred into an EXCEL worksheet for further analysis.

### CoC Calculations for Yeast Zip3 Foci

Coefficient of coincidence (CoC) curves were generated from SC length and Zip3 focus positions determined as described above. Each analyzed bivalent was divided into a series of intervals of  $0.1\ \mu\text{m}$  in length (corresponding to the resolution with which adjacent Zip3 foci can be resolved). Chromosome III, IV and XV were thus usually divided into 9, 42 and 30 intervals with equal size, respectively. Each chromosome length was normalized to 100% and each Zip3 focus position was also normalized correspondingly. Each Zip3 focus was then assigned to a specific interval according to its coordinate. The total frequency of bivalents having a Zip3 focus in each interval was calculated. For each pair of intervals, the frequency of bivalents having a Zip3 focus in both intervals was determined to give the “observed” frequency of double COs. For each pair of intervals, the total CO frequencies for the two intervals were multiplied to give the frequency of double COs “expected” on the hypothesis of independent occurrence. The ratio of these two values is the CoC. Thus in each pair of intervals,  $\text{CoC} = (\text{Obs DCO})/(\text{Pred DCO})$ . CoC values for all pairs of intervals can be plotted as a function of the distance between the midpoints of the two involved intervals (“inter-interval distance”). However, for all of the data shown here, the CoC values from all pairs of intervals having same inter-interval distance were averaged and this average CoC was plotted as a function of inter-interval distance (e.g. Figure 6C and others).

### BF Software and Simulations

The previous Beam Film program [3] was rewritten in MATLAB (R2010a) for easy use and modified to include more features as described in the text. Extensive details regarding

program structure and application are provided in the Protocol S1 section. However, briefly, there are three options in the software that serve three different purposes:

- (1) Analyze existing CO data. This option allows the user to process an experimental CO data set. Outputs include a variety of different CO distribution descriptors including CoC curves, ED relationships including the average number of COs/bivalent, average inter-CO distance and evaluation of the gamma distribution shape parameter ( $v$ ).
- (2) Do a single BF simulation with a particular set of specified parameter values. This option gives the same outputs as for analysis of an existing CO data set. It can help the user to understand how the BF model works (e.g. by extension of examples presented in the Results). It also enables a skilled user to do a single round of BF simulation at some single particular parameter condition.
- (3) Scan a range of parameters to get a BF best-fit simulation of an experimental data set. This option automatically scans all parameter combinations (over specified ranges of each parameter) and outputs the results in rank order according to the goodness of fit levels calculated based on PLS (Projected Likelihood Score) as defined by Falque et al. [18,19]. However, the rank order defined by PLS is not a maximum likelihood method. As a result, the best fit judged by PLS is not always the actual best fit. To overcome this drawback, the software outputs the results for all parameter combinations scanned, which the user can further evaluate to select the actual best fit. We normally choose the best fit by comparing experimental and simulated data sets with respect to the CoC curve, the average number of COs per bivalent and the distribution of CO number per bivalents (the ED distribution).

### Other Data Used in This Study

The *Chorthippus L3* chiasmata data were generously provided by Gareth Jones (University of Birmingham, UK). The *Drosophila* X-chromosome crossover data are from [33]. The tomato (*S. lycopersicum*) Mlh1 foci data are from [38] (generously provided by F. Lhuissier). Zip2 data in the *S.cerevisiae* BR background are from [31] (generously provided by J. Fung).

### Supporting Information

**Figure S1** Determination of interval sizes required for accurate CoC curves. Bivalents must be divided into a sufficiently large number of intervals that few if any closely-spaced COs are missed. A general rule is that the interval size should be less than  $1/4$  the average distance between COs. Operationally, where possible, interval size should be progressively decreased until the position of the CoC curve no longer changes. The simulations presented were performed under standard parameter conditions (text Figure 3) except that the number of intervals (and thus the inter-interval distance) was progressively increased. For this particular case, the CoC curves do not change significantly once the number of intervals is at or above 20 (interval size = 5% total chromosome length in  $\mu\text{m}$ ). (TIF)

**Figure S2** Interplay among precursor parameter values at low precursor number ( $N$ ). Variations in the distributions of precursors along or between bivalents (parameters E and B) have more significant effects at lower average precursor numbers ( $N$ ). Panels (A) and (B): Effects of variations in B and E are illustrated by

simulations using the same parameter values as for text Figure 3 except that  $N = 7$ .

(TIF)

**Figure S3** CoC curves can have signatures that reflect inter-precursor spacing. BF simulations show that if precursors are very evenly spaced ( $E \geq 0.8$ , the corresponding  $v > 10$ ), and if the interference distance is relatively long as compared to the average distance between precursors (e.g.  $L = 0.15$  versus  $0.06$ – $0.17$ ), CoC curves can exhibit “humps” corresponding to the average inter-precursor distance. These humps reflect the fact that closely-spaced double COs will tend to occur specifically at adjacent precursors, and when those precursors are evenly spaced, there is an elevated probability of double CO occurrence at that particular inter-interval distance. This feature is not apparent in standard simulation conditions (text Figures 3–5) because, in those conditions, precursors are less evenly spaced ( $E = 0.6$ , the corresponding  $v = 2.4$ ). Other BF parameter values for the simulations show in this figure are:  $L = 0.15$ ,  $S_{\max} = 1.8$ ,  $A = 2$ ,  $cL = cR = 1$ ,  $E = 0.8$ ,  $B = 0.9$ .

(TIF)

**Figure S4** How to obtain the best-fit BF simulation for an experimental data set. For each data set, the constellation of BF parameters that provides the best fit to the data set was obtained in three stages, as illustrated for data from yeast Chromosome XV (text). (1) A simulation was carried out at probable approximate values of  $N$  and  $L$  (Panel A). The range of sensible values of ( $N$ ) is suggested by total DSB levels, total levels of inter-homolog events (COs+NCOs), numbers of inter-axis bridges, immunofluorescent foci and/or EM-defined SC-associated recombination nodules, all of which approximately reflect total precursor interactions. The initial value of  $L$  ( $L_{BF}$ ) was generally set at  $L_{CoC}$ . With respect to other parameters: values of ( $cL$ ) and ( $cR$ ) were selected based on the distribution of COs along the chromosome; the value of  $M$  was assumed to be 100% for wild-type meiosis; the value of  $A = 1$  was selected as a reasonable first approximation. (2) The value of  $S_{\max}$  was then refined so as to optimize the fit between experimental and predicted ED arrays with respect to both the average number of COs per bivalent, including the probability of zero-CO chromosomes (Panel B). (3) The values of all parameters were then further refined by empirical trial-and-error, guided by knowledge as to the general effects of changes in each parameter on CoC and ED outputs as described above. Panel C describes initial refinements with respect to  $L$  and  $S_{\max}$ ; Panel D describes subsequent refinements of these two parameters. The final selected best-fit simulation parameters are those described in Panel D,  $L = 0.1$ ;  $S_{\max} = 3.5$  (other parameter values in text Table 2).

(TIF)

**Figure S5** Experimental data and BF simulations for yeast chromosomes. Panels A–C: CoC and ED relationships, and best-fit BF simulations, for SK1 Chromosomes XV (from text Figure 6I) and analogously analyzed Chromosomes IV and III. Panel D: CoC and ED relationships and BF simulation for Chromosome XIV in the BR background ([31]; J. Fung, personal communication). Parameter values for simulations in A–D in text Table 2. Note that all chromosomes, in both strain backgrounds, have the same CoC relationships when inter-interval distance is expressed in  $\mu\text{m}$  SC length. Further, pachytene SC length is  $\sim 10\%$  less in BR than in SK1. This comparison, along with other comparisons (L.Z., unpublished), shows that the metric for interference is physical distance in yeast as in other analyzed organisms (text). Panels E and F: the average CoC curve and the ED relationships for Chromosome III observed experimentally (black) and BF best-fit simulations using optimal parameter values (including  $E = 0.6$

and  $B = 1$ , which give relatively even spacing and a constant number of COs per bivalent; Table 2) (red) are compared with BF simulations using the same parameter values except that precursors were considered to be randomly spaced along chromosomes III in different nuclei ( $B = 0$ ; Panel E, green) or Poisson-distributed among chromosomes III in different nuclei ( $B = 0$ ; Panel F, blue). Even-versus-random spacing affects CoC relationships, confirming that precursors are evenly spaced (text), but does not affect ED relationships. Oppositely, constant-versus-Poisson distribution among chromosomes does not affect CoC relationships but significantly alters ED relationships, with a decrease in the average number of COs per bivalent overall but, more importantly, a significant increase in the frequency of zero-CO bivalents, from 1% to 4%. This effect strongly suggests that a given chromosome always acquires the same/similar number of precursors in every meiotic nucleus.

(TIF)

**Figure S6** Heterochromatin/euchromatin ratios underlie differences in SC length in tomato. CoC analysis suggests that the metric for CO interference is physical chromosome length (manifested as SC length) and that the density of DNA, i.e. Mb per  $\mu\text{m}$  SC length, is greater for short chromosomes than for longer chromosomes (text). The latter feature reflects the combination of two effects: (i) short chromosomes have a higher percentage of their DNA in heterochromatin and (ii) there is more DNA per  $\mu\text{m}$  SC length in heterochromatin versus euchromatin. This conclusion requires knowing: (i) the “packing ratio” of Mb/ $\mu\text{m}$  SC length for heterochromatin versus euchromatin; and (ii) the fraction of each pachytene SC length that underlies heterochromatin versus euchromatin. This information is available from published data (Panel A; <http://solgenomics.net>; [38,74,75]. Panel B: (i) Genome-wide, total DNA in heterochromatin and euchromatin are 693 Mb and 203 Mb, respectively and total SC length in heterochromatin and euchromatin are 78  $\mu\text{m}$  and 152  $\mu\text{m}$ , respectively. Thus, Mb/ $\mu\text{m}$  SC is 8.9 for heterochromatin and 1.3 for euchromatin, respectively. (ii) For shorter and longer chromosomes, the fractions of SC length in euchromatin and heterochromatin in the two groups are 66% and 34% respectively. This information implies that the group of shorter chromosomes will comprise 70 Mb, with an average Mb/ $\mu\text{m}$  SC length of 3.35, while the group of longer chromosomes will comprise 73 Mb, with an average Mb/ $\mu\text{m}$  SC length of 4.12. This difference quantitatively explains the offset in CoC curves when inter-interval distance is expressed in Mb versus  $\mu\text{m}$  SC length.  $L_{CoC}$  (and  $L_{BF}$ ) are 14  $\mu\text{m}$  SC length, which thus corresponds to 47 Mb and 58 Mb, respectively, which is a difference of 11 Mb. The offset in the experimental CoC curves for the two groups when inter-interval distance is expressed in Mb exactly matches this difference:  $L_{CoC}$  occurs at inter-interval distances of 32 Mb and 43 Mb respectively, which is a difference of 11 Mb. Notably, these considerations also show that, despite their differences in SC lengths, all of the chromosomes in tomato contain about the same total amount of DNA.

(TIF)

**Figure S7** BF simulations of CO patterns in yeast mutants with altered DSB levels. CoC and ED relationships for yeast mutants with changed DSB levels and BF best-fit simulations (Panels A, B). Also indicated are the number of DSBs predicted from experimental analysis (relative DSBs levels from pulse-field gels along chromosome III, VII and VIII in a *rad50S* strain background; [43]) and the number of precursors required to give a best fit simulation. Predicted values of ( $N$ ) and observed levels of DSBs match very precisely for most of the mutants; however,

predicted values are slightly but significantly higher than experimental values at the very lowest DSB levels. This could mean that *rad50S* DSB levels are underestimated at low DSB levels; that DSB levels are auto-catalytic such that occurrence of DSBs above a threshold level tends to promote the formation of additional DSBs; or that best-fit simulations do not give precisely the correct values at low DSB levels.  
(TIF)

**Protocol S1** Instructions for the BF program.  
(DOCX)

**Table S1** Strains used in this study.  
(DOCX)

## References

- Muller HJ (1916) The Mechanism of Crossing Over. *Am Nat* 50: 193–434.
- Sturtevant AH (1915) The behavior of the chromosomes as studied through linkage. *Z indukt Abstamm-u VererbLehre* 13: 234–287.
- Kleckner N, Zickler D, Jones GH, Dekker J, Padmore R, et al. (2004) A mechanical basis for chromosome function. *Proc Natl Acad Sci U S A* 101: 12592–12597.
- Borner GV, Kleckner N, Hunter N (2004) Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* 117: 29–45.
- Hunter N (2006) Meiotic recombination. In: Aguilera A, Rothstein R editors. *Molecular Genetics of Recombination*. pp 381–442.
- Drouaud J, Mercier R, Chelysheva L, Berard A, Falque M, et al. (2007) Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. *PLoS Genet* 3: e106.
- Petkov PM, Broman KW, Szatkiewicz JP, Paigen K (2007) Crossover interference underlies sex differences in recombination rates. *Trends Genet* 23: 539–542.
- Billings T, Sargent EE, Szatkiewicz JP, Leahy N, Kwak IY, et al. (2010) Patterns of recombination activity on mouse chromosome 11 revealed by high resolution mapping. *PLoS One* 5: e15340.
- Hillers KJ, Villeneuve AM (2003) Chromosome-wide control of meiotic crossing over in *C. elegans*. *Curr Biol* 13: 1641–1647.
- Oliver-Bonet M, Campillo M, Turek PJ, Ko E, Martin RH (2007) Analysis of replication protein A (RPA) in human spermatogenesis. *Mol Human Reprod* 13: 837–844.
- de Boer E, Stam P, Dietrich AJ, Pastink A, Heyting C (2006) Two levels of interference in mouse meiotic recombination. *Proc Natl Acad Sci U S A* 103: 9607–9612.
- Storz A, Gargano S, Ruprich-Robert G, Falque M, David M, et al. (2010) Recombination proteins mediate meiotic spatial chromosome organization and pairing. *Cell* 141: 94–106.
- Zhang L, Kim KP, Kleckner NE, Storlazzi A (2011) Meiotic double-strand breaks occur once per pair of (sister) chromatids and, via Mec1/ATR and Tel1/ATM, once per quartet of chromatids. *Proc Natl Acad Sci U S A* 108: 20036–41.
- Berchowitz LE, Copenhaver GP (2010) Genetic interference: don't stand so close to me. *Curr Genomics* 11: 91–102.
- Cole F, Kauppi L, Lange J, Roig I, Wang R, et al. (2012) Homeostatic control of recombination is implemented progressively in mouse meiosis. *Nat Cell Biol* 14: 424–430.
- Mets DG, Meyer BJ (2009) Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure. *Cell* 139: 73–86.
- McPeck S, Speed TP, Mor H (1995) Modeling Interference in Genetic Recombination. *Genetics* 1044: 1031–1044.
- Falque M, Anderson LK, Stack SM, Gauthier F, Martin OC (2009) Two types of meiotic crossovers coexist in maize. *Plant Cell* 21: 3915–3925.
- Gauthier F, Martin OC, Falque M (2011) CODA (crossover distribution analyzer): quantitative characterization of crossover position patterns along chromosomes. *BMC Bioinformatics* 12: 27.
- Foss E, Lande R, Stahl FW, Steinberg CM (1993) Chiasma interference as a function of genetic distance. *Genetics* 133: 681–691.
- Pan J, Sasaki M, Kniewel R, Murakami H, Blitzblau HG, et al. (2011) A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* 144: 719–731.
- Chen SY, Tsubouchi T, Rockmill B, Sandler JS, Richards DR, et al. (2008) Global analysis of the meiotic crossover landscape. *Dev Cell* 15: 401–415.
- Borde V, Lin W, Novikov E, Petrini JH, Lichten M, et al. (2004) Association of Mre11p with double-strand break sites during yeast meiosis. *Mol Cell* 13: 389–401.
- Gerton JL, Derisi J, Shroff R, Lichten M, Brown PO, et al. (2000) Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 97: 11383–11390.
- Blat Y, Protacio RU, Hunter N, Kleckner N (2002) Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. *Cell* 111: 791–802.
- Buhler C, Borde V, Lichten M (2007) Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*. *PLoS Biol* 5: e324.
- Kauppi L, Barchi M, Lange J, Baudat F, Jasin M, et al. (2013) Numerical constraints and feedback control of double-strand breaks in mouse meiosis. *Genes Dev* 27: 873–886.
- Joyce EF, McKim KS (2010) Chromosome axis defects induce a checkpoint-mediated delay and interchromosomal effect on crossing over during *Drosophila* meiosis. *PLoS Genet* 6: e1001059.
- Joyce EF, McKim KS (2011) Meiotic checkpoints and the interchromosomal effect on crossing over in *Drosophila* females. *Fly (Austin)* 5: 134–140.
- Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, et al. (2012) Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley. *Plant Cell* 24: 4096–4109.
- Fung JC, Rockmill B, Odell M, Roeder GS (2004) Imposition of crossover interference through the nonrandom distribution of synapsis initiation complexes. *Cell* 116: 795–802.
- Malkova A, Swanson J, German M, McCusker JH, Housworth EA, et al. (2004) Gene conversion and crossing over along the 405-kb left arm of *Saccharomyces cerevisiae* chromosome VII. *Genetics* 168: 49–63.
- Charles DR (1938) The spatial distribution of cross-overs in X-chromosome tetrads of *Drosophila melanogaster*. *J Genetics* 36: 103–126.
- Mehrotra S, McKim KS (2006) Temporal analysis of meiotic DNA double-strand break formation and repair in *Drosophila* females. *PLoS Genet* 2: e200.
- Lande R, Stahl FW (1993) Chiasma interference and the distribution of exchanges in *Drosophila melanogaster*. *Cold Spring Harb Symp Quant Biol* 58: 543–552.
- Laurie DA, Jones GH (1981) Inter-individual variation in chiasma distribution in *Chorthippus brunneus* (Orthoptera: Acrididae). *Heredity* 47: 409–416.
- Jones GH (1984) The control of chiasma distribution. *Symp Soc Exp Biol* 38: 293–320.
- Lhuissier FG, Offenberg HH, Wittich PE, Vischer NO, Heyting C (2007) The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. *Plant Cell* 19: 862–876.
- Colombo PC, Jones GH (1997) Chiasma interference is blind to centromeres. *Heredity (Edinb)* 79 (Pt 2): 214–227.
- Saintenac C, Falque M, Martin OC, Paux E, Feuillet C, et al. (2009) Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). *Genetics* 181: 393–403.
- Lian J, Yin Y, Oliver-Bonet M, Liehr T, Ko E, et al. (2008) Variation in crossover interference levels on individual chromosomes from human males. *Hum Mol Genet* 17: 2583–2594.
- Broman KW, Rowe LB, Churchill GA, Paigen K (2002) Crossover interference in the mouse. *Genetics* 160: 1123–1131.
- Martini E, Diaz RL, Hunter N, Keeney S (2006) Crossover homeostasis in yeast meiosis. *Cell* 126: 285–295.
- Rosu S, Libuda DE, Villeneuve AM (2011) Robust crossover assurance and regulated interhomolog access maintain meiotic crossover number. *Science* 334: 1286–1289.
- Yokoo R, Zawadzki KA, Nabeshima K, Drake M, Arur S, et al. (2012) COSA-1 reveals robust homeostasis and separable licensing and reinforcement steps governing meiotic crossovers. *Cell* 149: 75–87.
- Globus ST, Keeney S (2012) The joy of six: how to control your crossovers. *Cell* 149: 11–12.
- Jones GH, Franklin FC (2006) Meiotic crossing-over: obligation and interference. *Cell* 126: 246–248.
- Hawley RS, Theurkauf WE. (1993) Requiem for distributive segregation: achiasmate segregation in *Drosophila* females. *Trends Genet* 9: 310–317.
- Argueso JL, Kijas AW, Sarin S, Heck J, Waase M, et al. (2003) Systematic mutagenesis of the *Saccharomyces cerevisiae* MLH1 gene reveals distinct roles

## Acknowledgments

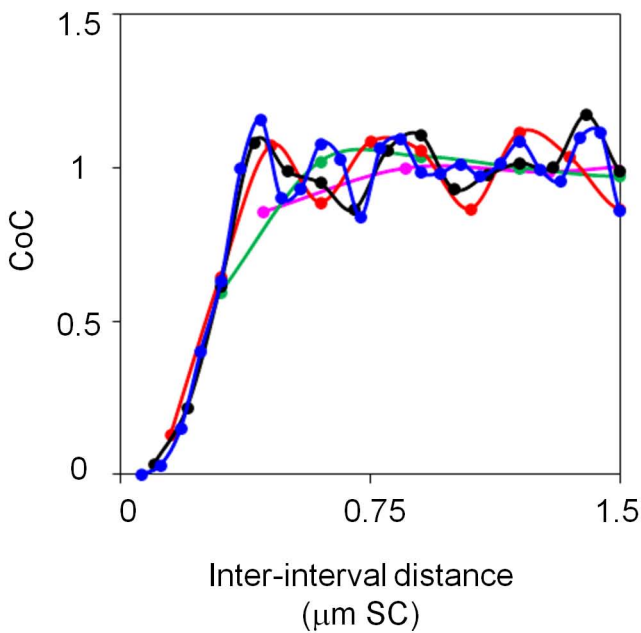
We thank Gareth Jones (University of Birmingham, UK) for grasshopper chiasmata data, Jennifer Fung (UCSF) for Zip2 data in the *S.cerevisiae* BR background and Franck G.P. Lhuissier (Netherlands) for tomato Mlh1 foci data. We thank Scott Hawley for provoking deeper thinking about the obligatory CO; Aurora Storlazzi and Denise Zickler for comments; and Guillaume Witz and Frederick Chang for help with mathematical analyses.

## Author Contributions

Conceived and designed the experiments: LZ JH NK. Performed the experiments: LZ. Analyzed the data: LZ NK. Contributed reagents/materials/analysis tools: LZ ZL JH. Wrote the paper: LZ NK.

- for Mlh1p in meiotic crossing over and in vegetative and meiotic mismatch repair. *Mol Cell Biol* 23: 873–886.
50. Argueso JL, Wanat J, Gemici Z, Alani E (2004) Competing crossover pathways act during meiosis in *Saccharomyces cerevisiae*. *Genetics* 168: 1805–1816.
  51. Zakharyevich K, Tang S, Ma Y, Hunter N (2012) Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* 149: 334–347.
  52. Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM (2008) High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature* 454: 479–485.
  53. Franklin WS, Foss HM (2010) A two-pathway analysis of meiotic crossing over and gene conversion in *Saccharomyces cerevisiae*. *Genetics* 186: 515–536.
  54. Bzymek M, Thayer NH, Oh SD, Kleckner N, Hunter N (2010) Double Holliday junctions are intermediates of DNA break repair. *Nature* 464: 937–941.
  55. King JS, Mortimer RK (1990) A polymerization model of chiasma interference and corresponding computer simulation. *Genetics* 126: 1127–1138.
  56. Fujitani Y, Mori S, Kobayashi I (2002) A reaction-diffusion model for interference in meiotic crossing over. *Genetics* 161: 365–372.
  57. Han YW, Mizuuchi K (2010) Phage Mu transposition immunity: protein pattern formation along DNA by a diffusion-ratchet mechanism. *Mol Cell* 39: 48–58.
  58. Vecchiarelli AG, Hwang LC, Mizuuchi K (2013) Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proc Natl Acad Sci U S A* 110: E1390–1397.
  59. Hulten MA (2011) On the origin of crossover interference: A chromosome oscillatory movement (COM) model. *Mol Cytogenet* 4: 10.
  60. Getz TJ, Banse SA, Young LS, Banse AV, Swanson J, et al. (2008) Reduced mismatch repair of heteroduplexes reveals “non”-interfering crossing over in wild-type *Saccharomyces cerevisiae*. *Genetics* 178: 1251–1269.
  61. Straight AF, Belmont A S, Robinett CC, Murray AW (1996) GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. *Curr Biol* 6: 1599–1608.
  62. Hollingsworth NM, Ponte L, Halsey C (1995) MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes Dev* 9: 1728–1739.
  63. Agarwal S, Roeder GS (2000) Zip3 provides a link between recombination enzymes and synaptonemal complex proteins. *Cell* 102: 245–255.
  64. Chua PR, Roeder GS (1998) Zip2, a meiosis-specific protein required for the initiation of chromosome synapsis. *Cell* 93: 349–359.
  65. Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, et al. (1997) Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* 387: 67–73.
  66. Koszul R, Kleckner N (2009) Dynamic chromosome movements during meiosis: a way to eliminate unwanted connections? *Trends Cell Biol* 19: 716–724.
  67. Loidl J, Klein F, Engebrecht J (1998) Genetic and morphological approaches for the analysis of meiotic chromosomes in yeast. *Methods Cell Biol* 53: 257–285.
  68. Kim PK, Weiner BM, Zhang L, Jordan A, Dekker J, et al. (2010) Sister cohesion and structural axis components mediate homolog bias of meiotic recombination. *Cell* 143: 924–937.
  69. Qiao H, Lohmiller LD, Anderson LK (2011) Cohesin proteins load sequentially during prophase I in tomato primary microsporocytes. *Chromosome Res* 19: 193–207.
  70. Dolezel J, Bartos J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. *Cytometry A* 51: 127–128.
  71. Santos J, del Cerro A, Diez M (1993) Spreading synaptonemal complexes from the grasshopper *Chorthippus jacobsi*: pachytene and zygotene observations. *Hereditas* 118: 235–241. *Hereditas* 118: 235–241.
  72. Page SL, Hawley RS (2001) c(3)G encodes a *Drosophila* synaptonemal complex protein. *Genes Dev* 15: 3130–3143.
  73. Carpenter AT (1979) Recombination nodules and synaptonemal complex in recombination-defective females of *Drosophila melanogaster*. *Chromosoma* 75: 259–292.
  74. Peterson DG, Stack SM, Price JH, Johnston JS. DNA content of heterochromatin and euchromatin in tomato (*Lycopersicon esculentum*) pachytene chromosomes. *Genome* 39: 77–82.
  75. Sherman JD, Stack SM (1995) Two-dimensional spreads of synaptonemal complexes from solanaceous plants. VI. High-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics* 141: 683–708.





Number of Intervals

7

10

20

30

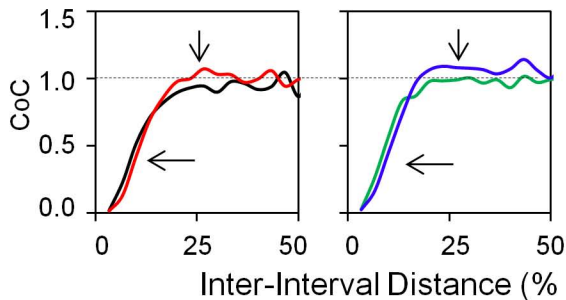
50

**A**

Along Chromosomes (E)

Even &gt;&gt; Random

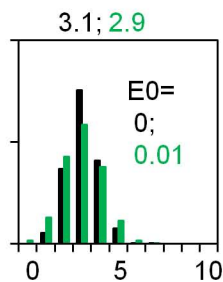
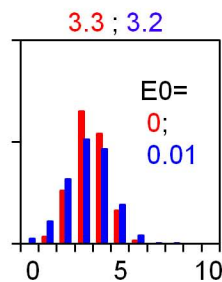
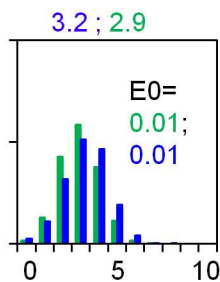
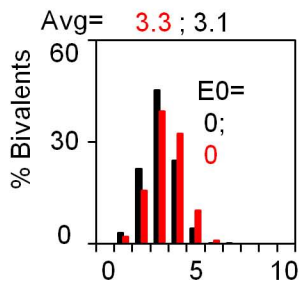
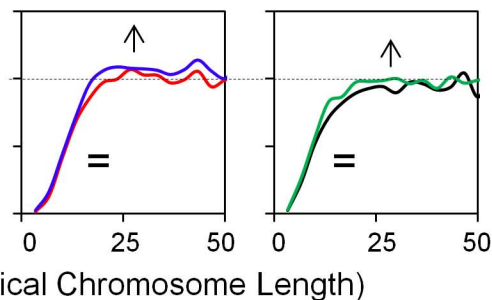

 (B) Constant  
 (B) Poisson

**B**

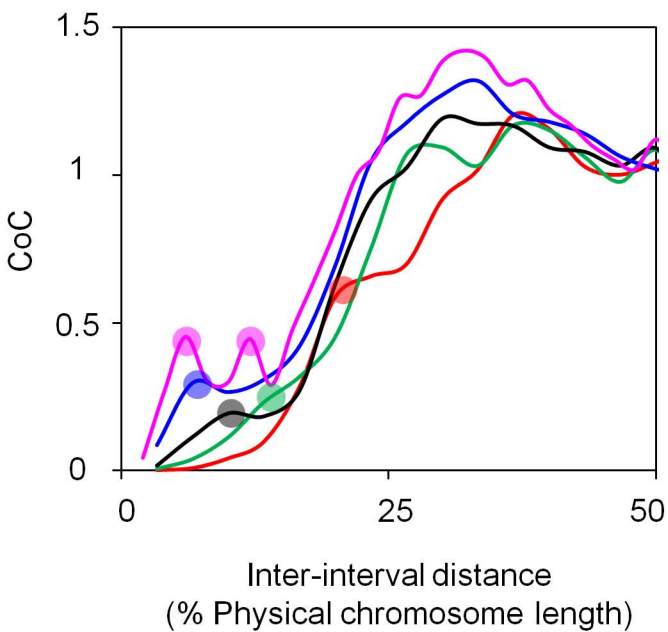
Among chromosomes (B)

Constant &gt;&gt; Poisson


 (E) Even  
 (E) Random



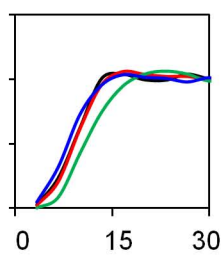
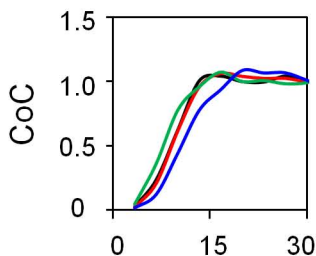
COs per Bivalent



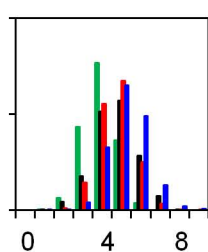
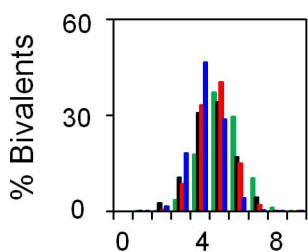
N	inter-N distance	Small Hump position
5	17	17
7	13	13
9	10	1
13	7	7
17	6	6

**A** Data;  
L = 0.08; 0.1; 0.12  
Smax = 3.5 for all

**B** Data;  
Smax = 2.8; 3.5; 4.2  
L = 0.1 for all



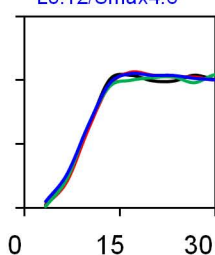
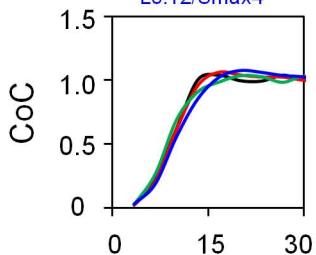
Inter-interval distance (% Physical chromosome length)



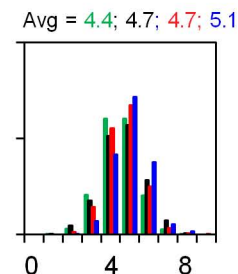
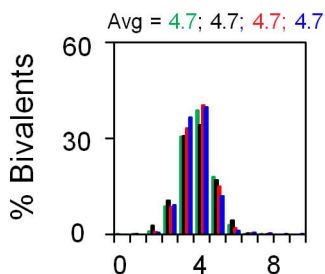
COs / Bivalent

**C** Chr XV Data;  
L0.08/Smax3;  
L0.1/Smax3.5;  
L0.12/Smax4

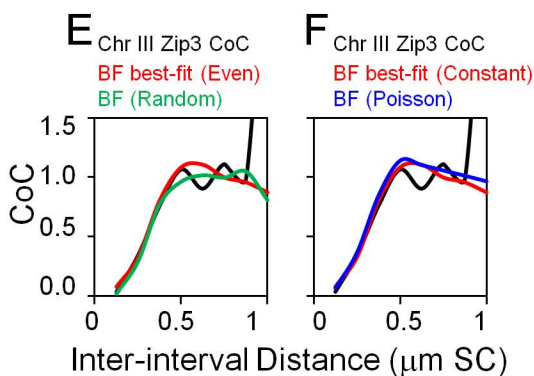
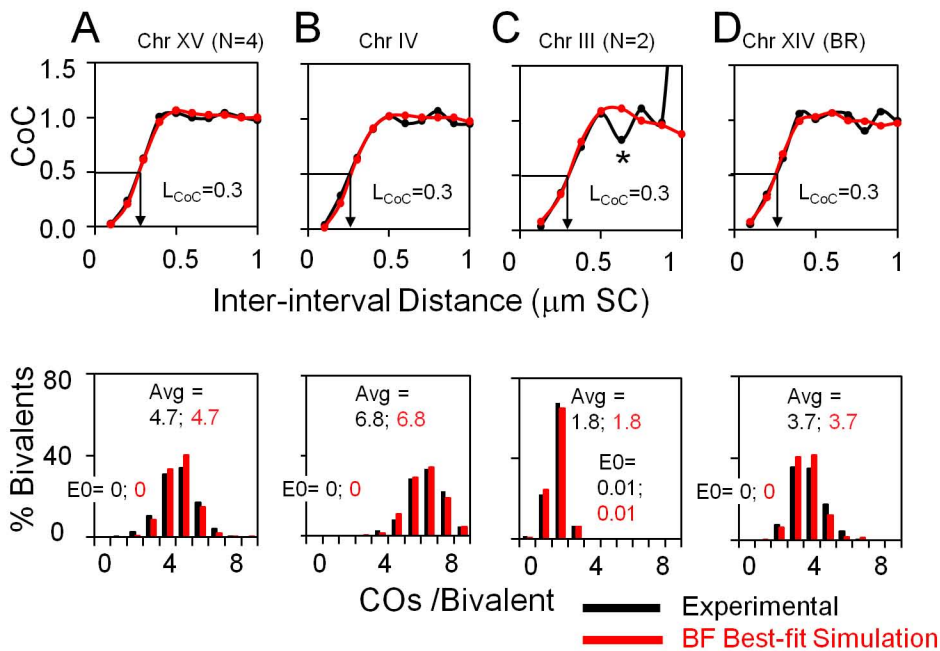
**D** Chr XV Data;  
L0.08/Smax2.8;  
L0.1/Smax3.5;  
L0.12/Smax4.6



Inter-interval distance (% Physical chromosome length)



COs / Bivalent



## A Chromosome Properties\*\*

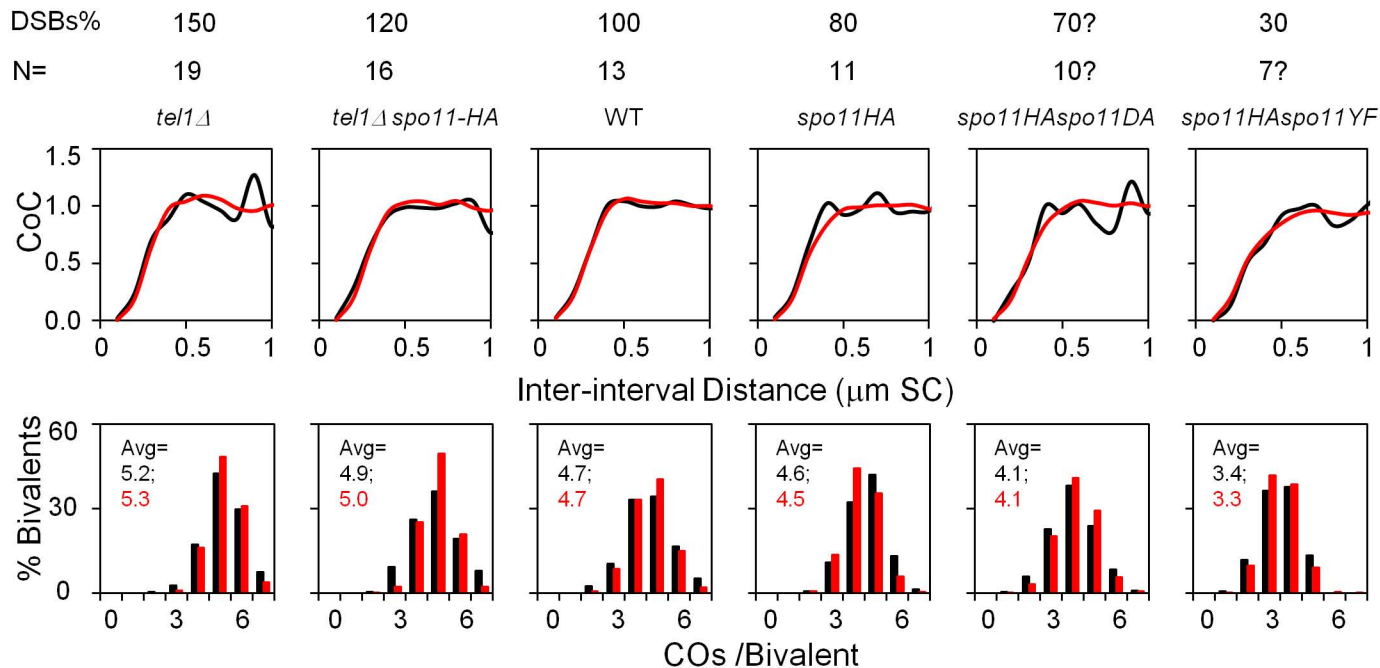
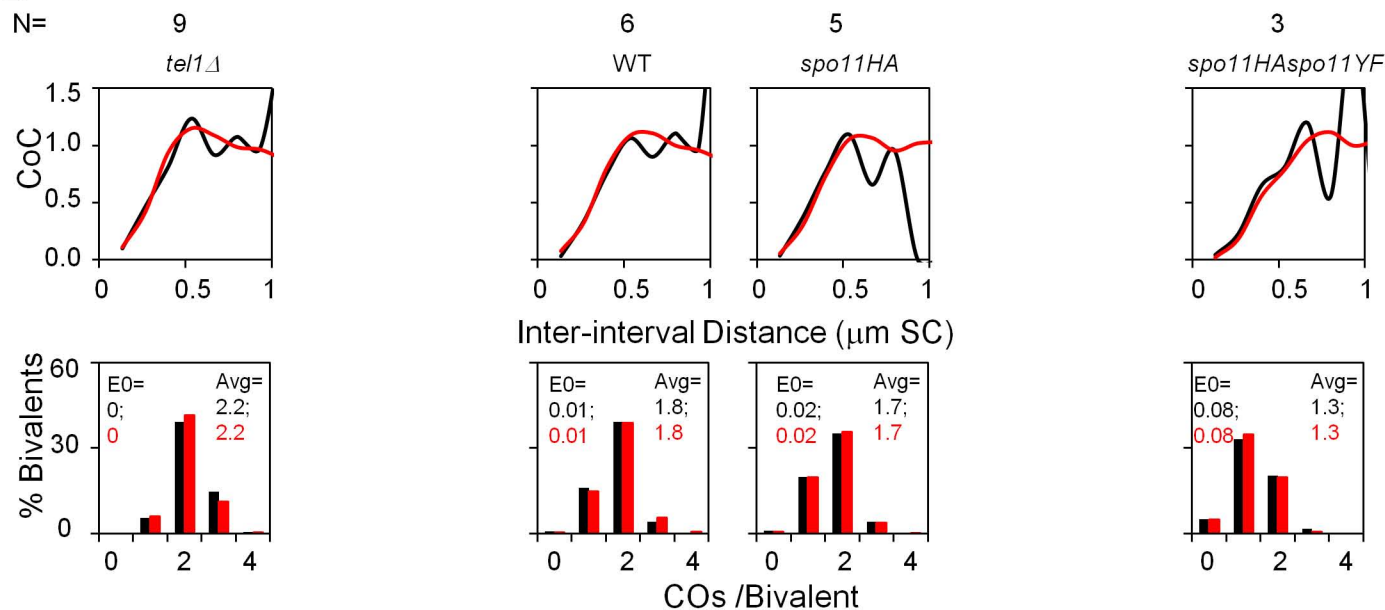
Chromosome number	1	2	3	4	5	6	7	8	9	10	11	12	total
Chromosome length (Mb)	107	59	77	76	77	55	77	75	80	77	63	77	900
SC length ( $\mu\text{m}$ )	30	21	23	21	16	19	19	19	16	16	16	14	230
Euchromatin (% $\mu\text{m}$ SC)	72	80	69	66	59	70	63	64	62	62	60	59	66
Heterochromatin (% $\mu\text{m}$ SC)	28	20	31	34	41	30	37	36	38	38	40	41	34

>> average for Chr 2-4 = 22  $\mu\text{m}$  = 71 Mb; average for Chr 5-11 = 17  $\mu\text{m}$  = 72 Mb

## B

	Heterochromatin	Euchromatin	
SC length ( $\mu\text{m}$ )	0.34 of 230 $\mu\text{m}$ = 78 $\mu\text{m}$	0.66 of 230 = 152 $\mu\text{m}$	
Genome (Mb)**	0.77 of genome = 693Mb	0.23 of genome = 207Mb	
1 $\mu\text{m}$ =?Mb	8.9Mb/ $\mu\text{m}$	1.3 Mb/ $\mu\text{m}$	
1Mb=? $\mu\text{m}$	0.12 $\mu\text{m}$ /Mb	0.7096 $\mu\text{m}$ /Mb	
Chr2-4= 22 $\mu\text{m}$	0.28 SC=6.1 $\mu\text{m}$ =51.2Mb	0.72 SC=15.696 $\mu\text{m}$ =22.13Mb	total=73 Mb
Chr5-11= 17 $\mu\text{m}$	0.37 SC=6.45 $\mu\text{m}$ =54.13Mb	0.63 SC=10.987 $\mu\text{m}$ =15.49Mb	total=70 Mb

\*\* All data are from <http://solgenomics.net>; Sherman 1995; Lhuissier 2007; Peterson 1996.

**A****B**

— Experimental  
 — BF Best-fit Simulation

Strains	Genotype
NKY4146	ho::hisG <sup>+</sup> , leu2 <sup>+</sup> , ura3 <sup>+</sup> , HMR::LacO-URA3 <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup>
NKY4147	ho::hisG <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup>
NKY4148	ho <sup>+</sup> , ura3 <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , leu2::Lacl-GFP::Clonat <sup>+</sup> , tel4::226xLacO::Kan <sup>+</sup>
NKY4149	ho::hisG <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , ZIP3-13myc::hygromycin B/+, URA3::CYC1p-Lacl-GFP <sup>+</sup> , scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup> , mlh1::KanMX <sup>+</sup>
NKY4152	ho <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , arg4-nsp/?, ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup> , tel1D::KanMX4 <sup>+</sup>
NKY4153	ho <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup> , spo11-HA3His6::KanMX4 <sup>+</sup> , tel1D::KanMX4 <sup>+</sup>
NKY4154	ho <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP/ura3, scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup> , spo11-HA3His6::KanMX4 <sup>+</sup>
NKY4155	ho::hisG <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , arg4-nsp?, ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup> , spo11-HA3His6::KanMX4/spo11(D290A)-HA3His6::KanMX4, Rec8-3HA::URA3/+
NKY4156	ho::hisG <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , arg4-nsp?, ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , scp1(Ch XV telomere)::LacO-LEU2 <sup>+</sup> , spo11-HA3His6::KanMX4/spo11(Y135F)-HA3His6::KanMX4, Rec8-3HA::URA3/+
NKY4157	ho::hisG <sup>+</sup> , leu2 <sup>+</sup> , ura3 <sup>+</sup> , pSE350 xHpal, HMR::LacO Sherret-URA3 <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , tel1::KanMX4 <sup>+</sup>
NKY4158	ho::hisG <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , Rec8-3HA::URA3 <sup>+</sup> , spo11-HA3His6::KanMX4 <sup>+</sup> , HMR::LacO-URA3 <sup>+</sup>
NKY4159	ho::hisG <sup>+</sup> , leu2::hisG <sup>+</sup> , ura3(Smal-PstI) <sup>+</sup> , ZIP3-13myc::Hygromycin B <sup>+</sup> , URA3::CYC1p-Lacl-GFP <sup>+</sup> , Rec8-3HA::URA3/+, spo11-HA3His6::KanMX4/spo11-HA3His6::KanMX4/spo11(Y135F)-HA3His6::KanMX4, HMR::LacO-URA3 <sup>+</sup>



# Instructions for BF program

The new version of BF (Beam Film) code is written in MATLAB (R2010a) for easy use (both the original code and the executable program are downloadable with the following link: <https://app.box.com/s/hv91q2nrtq0cp9n8iy9m>). Users can use, distribute or modify it freely.

The purpose of this software is to help researchers quantitatively analyze experimental CO data sets and to simulate CO patterns according to the predictions of the BF model.

## 1. User interface of the software: general instructions

Three analysis options are available: (A) Analyze an existing CO data set; (B) Carry out a single round of BF simulation under a specific set of BF parameters; and (C) Scan a range of parameters to get the best fit of the CO data. These three options appear in the main window (a simple graphical user interface; Protocol Figure 1A). After one option is selected, a corresponding next window will appear (Protocol Figure 1BCD).

After selecting one of the three options, a job is initiated by specifying appropriate parameters and, for option A, to input an experimental data set. In the job window that appears upon selection of a particular option (above), when the mouse cursor is moved onto each item, a short explanatory message pops out (e.g. Protocol Figure 1A, mouse cursor symbol and pop out message). For further detailed explanation of these items, consult the remainder of this Instruction (below) and also the main paper.

Once all necessary inputs have been entered and the "RUN" button is pushed, the program will carry out the requested analyses for each of 5000 independent chromosomes. This value is more than sufficient for all uses we have encountered thus far and thus is embedded in the program; however, it can be changed if needed by entering and altering the software itself.

**Note:** Once a particular job has been run, the results can be saved for later use before to running another job, otherwise they will be lost. Results can be preserved by copying them into another location (e.g. another file or a new folder). If results of a job are not saved in this way, they will be substituted by results of the next job in the current work folder.

## **2. Option (A): Analyze CO data.**

**Inputs.** The user should first specify the values for two input parameters, specifically:

1. *Interval number of CoC.* This input defines the intervals used for CoC analysis. The chromosome can be divided either into a particular number of intervals of identical length or a particular array of intervals, of any number, and position along the chromosome can be specified. Same-length intervals are more appropriate for simulation of cytological data where the analyzed chromosome can be divided into intervals as desired. However, for genetic data, intervals are usually specified by particular genetic markers and thus are not all of the same size, so individual specification is required.

In both cases, the necessary specification is entered into the box "Number of intervals for CoC analysis" (Protocol Figure 1B).

For same-length intervals, all that is required is to specify the desired total interval number per chromosome. The default value is 5.

To specify particular intervals, first the genetic marker positions must be normalized. There are two ways to do this normalization according to nature of the data set. (a) If the markers span the entire length of the chromosome, and their positions relative to the chromosome ends are known, the entire chromosome length is defined as "1" and all marker positions correspondingly normalized. (b) If markers define only a continuous internal segment of a chromosome, then the first marker position is defined as "0" and the last marker position is defined as "1" and other marker positions correspondingly normalized. Then, in either case, a list of the normalized genetic marker positions, framed by a pair of square brackets ("[" ]"), is entered into the

appropriate box. For example in text Figure 7, the 9 genetic marker positions would be specified as: [0.013, 0.165, 0.249, 0.335, 0.483, 0.58, 0.765, 0.869, 0.958].

Note: BF simulations include parameters that define specialized effects at chromosome ends (cL) and (cR) (text). For simulations, the values of these parameters must be specified appropriately. In the case of option (b), the optimal cL and cR mean effects either from the chromosome region(s) outside the fragment analyzed or from ends (or combined effects) since the ends positions are unknown.

**Note re choice of interval number.** For CO data sets based on cytological analysis of CO-correlated immuno-stained foci or nodules, the following criteria should be considered in order to set the interval number properly. (a) The interval size should be greater than the microscope resolution ( $\geq 0.1$  micron for standard epifluorescence or confocal microscopy). (b) Along an individual chromosome, the inhibitory effect of CO interference decays with increasing distance from the site of a CO. As a result, the CoC (classical indicator of CO interference) decreases with increasing inter-interval distance. Thus, a smaller interval size (larger interval number) is preferred because it permits more sensitive detection of interference over smaller inter-interval distances. Also, when precursor interactions are evenly spaced, the value of the CoC tends to rise up to a value larger than one before fluctuating around one (e.g. text Figure 6J) and this feature is more sensitively revealed with smaller inter-interval sizes (text Figure S1). On the other hand, with a smaller inter-interval size, more data (for more bivalents) are needed to get a statistically meaningful result. Thus, a big interval size may require less data but may miss the chance to see short inter-interval interference and CoC rise above one (e.g. text Figure S1). (c) As a general guideline, if possible, interval size should be less than  $\sim 1/4$  the average distance between adjacent COs. Fewer intervals would tend to place more than one COs into a single interval on a same chromosome which would artificially reduce the CoC. (d) A very large interval number (e.g. 50) will make the calculation process more time-consuming. Conclusion: where possible, the number of intervals should usually be set at five to ten times the CO number.

As an example of these effects, text Figure S1 shows CoC curves with different number of intervals, respectively. The hump where the CoC curve rises above one can be seen with 20 or more interval number, but it is barely seen with  $\leq 10$  intervals. Also, the

position of this feature and its prominence are different with different interval number. A bigger interval size (smaller interval number) would average the details from several smaller intervals (bigger interval number), thus obscuring the effects of interest.

*2. Interval number for rank plots.* Rank plot analysis defines the positions of successive COs along sets of bivalents exhibiting 1, 2 or 3 COs (example in text Figure 8). Analysis of bivalents exhibiting more than 3 COs is not informative because patterns are difficult to discern. For rank plot analysis, chromosomes (bivalents) are divided into intervals and the frequencies of COs of each type are plotted as a function of the interval. If no number is entered, the default value is 0, which means that no rank plots analysis will be performed.

*3. Entering experimental data set.* After input of the two above values, the user should push the “RUN” button. The program will now ask the user to choose the file and the worksheet containing the data to be analyzed. Data must be in an EXCEL file with proper format as described in Protocol Figure 2. For virtually all analyses, chromosomes of a particular genetic identity are first normalized to length 100% and each CO position is then normalized correspondingly.

**Outputs.** The results of running this option are presented in output 3 files:

(1) “output\_CO” file. This file includes most of the output information including CO number per chromosome, average CO number, CoC, etc, as detailed in Protocol Figure 3A.

(2) “gamdata\_CO” file. This file contains three worksheets. (a) All inter-CO distances for all adjacent pairs of COs are in a worksheet named “gamdata”. (b) Cumulative curve data, which describe the cumulative number of COs as a function of distance from the defined "initial" end, are in worksheet “cumulative”. The primary data are in the second column. The first column is the data for normalized X-axis plot data. (c) The distribution of inter-CO distances, sorted into a specified number of bins, is in worksheet “bin”. The default is 20 bins, but the user can change that number in the bin function.

(3) “coc\_CO” file. This file contains the all data relating to CoC relationships which are, for all pairs of intervals: observed double COs (Obs DCOs); predicted double COs (Pred DCOs); and the CoC (= Obs DCO/ Pred DCOs).

### **3. Option (B): A single BF simulation.**

This option provides many different outputs for a single set of specified BF parameters. This is useful in helping to understand how the BF model works and, once the set of possible parameter values can be narrowed down by experience or by application of Option (C) below, for doing a single round of BF simulation with all possible outputs. Users may also need to refer to the fuller discussions in the main text (not referenced specifically below); guidelines concerning how to set the number of intervals along a bivalent presented above under “Analyze CO data” and Section 6 below “More hints for obtaining the best fit between a BF simulation and an experimental data set.”

#### **Inputs**

(1) *CO interference distance (L)*. All bivalent length are normalized to “1”, and usually  $0 \leq L \leq 1$ . An intrinsic consequence of this feature is that even when different chromosomes of a given organism (and sex) have the same actual value of L in a biological metric (e.g.  $\mu\text{m}$  SC), the best-fit value of L will be different for different chromosomes, in proportion to chromosome length by that metric. In some organisms with very strong CO interference, L could be set as  $> 1$ . *C.elegans* could be such a case since chromosome fusion studies indicate that the CO interference extends beyond the length of a normal whole chromosome [1]. In organisms with where L is large relative to total chromosome length, the CoC curve rises prominently to values greater than one. As the value of L increases, this feature becomes more prominent and shifts to larger inter-interval distances (to the right on the CoC curve). These effects are illustrated in text Figure 2B and 4A. In arriving at a best-fit simulation it is useful to keep in mind that the inter-interval distance at which the CoC curve exceeds one is generally greater than the value of L. When the average number of COs per bivalent is greater than two, the CoC curve usually reaches a value of one at an inter-interval distance that is about half of the average inter-CO distance. If the average COs per bivalent is less than two, the corresponding point in the CoC curve occurs at an inter-interval distance that is closer to

the average inter-CO distance. As a good estimation of the interference distance  $L$ , use the inter-interval distance where  $CoC=0.5$  ( $L_{CoC}$ ) from the CoC curve (text for more details).

(2) *CO-designation driving force (DDF, Smax)*.  $S_{max}$  is the maximum strength of driving force for CO designation and could correspond to both magnitude and/or time (text for more details). *In vivo*, the parameter  $S_{max}$  could represent a driving force that increases over time, or equivalently, represents increased time duration for CO-designation as proposed for the "interchromosomal effect" in *Drosophila* ([2,3] ; text). Increased ( $S_{max}$ ) has effects on CO number and patterning that resemble those conferred by decreased ( $L$ ) (see text Figure 4). For organisms examined thus far, the best-fit value of  $S_{max}$  is between one and five. The actual value is easily determined after a range is established for the values of ( $L$ ) and ( $N$ ) (below). The BF program also includes an item "Smaxstd" below the  $S_{max}$  item that allows the user to specify different  $S_{max}$  values for different nuclei.

(3) *End conditions (cL and cR)*. This parameter can be used to modulate the different end effects for CO formation in different organisms and/or different chromosomes. The two ends can be set differently. The default value is 0 for both ends ( $cL=0$ ,  $cR=0$ ). This default equates to defining the end as a pre-existing CO, with a corresponding reduction in the probability of CO-designation. The effect is to reduce the frequency of COs in the immediate vicinity of the end (over some distance given by the value of ( $L$ )). At the other extreme, the value 1 ( $cL=1$ ,  $cR=1$ ) means that the end is no different from any other region. However, since CO interference will enter a terminal region from only one side, the consequence is an increase in the number of COs in near-terminal regions relative to internal regions. For example, in some organisms such as *S.cerevisiae*, the frequency of COs (but not NCOs) is suppressed preferentially in telomere regions [4- 8]. This effect is simulated in the BF model by using a low clamping level. However, in other organisms/ chromosomes with a tendency for COs to occur sub-telomerically (e.g. in grasshopper; main text Figure 8), the clamping level could be set at a high level (and may even above 1) to mimic these special terminal effects.

(4) *Precursor parameters*. Five parameters are used to describe the precursor array: ( $N$ ), ( $E$ ), ( $B$ ), ( $A$ ) and "black holes". ( $N$ ) is the average number of precursors per

bivalent. (E) is the extent to which precursors are evenly-versus-randomly distributed along each given chromosome. (B) is the extent to which precursors are distributed at a constant number, or are Poisson-distributed, among different chromosomes. (A) specifies the distribution of sensitivities of precursors to the CO driving force. "Black holes" describe cases in which the average number of precursors per unit length varies along the length of a chromosome.

- (N) can be experimentally determined or estimated (at least in some organisms) either by physical DSBs mapping (e.g. [7]) or cytological markers such as Rad51/Dmc1 foci, H2AX foci, RPA foci (e.g. [9-11]), or genetically counted CO and NCO numbers especially from microarray experiments (e.g. [12]). In *S.cerevisiae* the precursor number on each chromosome can be more accurately determined based on comprehensive evaluation results from DSBs mapping (e.g. [7]), microarray (e.g. [12]) and classical genetic measurements (<http://www.yeastgenome.org>). Thus the number of precursors on chromosome III, IV and XV are defined as 6, 19, and 13 respectively. This is about 2~3 times the number of interference COs on these chromosomes (text; Table 2). In mammals, the number of precursors is about 10 times the number of COs (e.g. [9]). In higher plants this ratio is higher, perhaps up to 20-30 folds (e.g. in maize ~20 CO/nucleus and ~500 Rad51 foci/nucleus [13,14]). Due to intrinsic technical limitations, any single assay may not give the accurate precursor number and in some organisms there is absence of data to determine the number of precursors. However, in lots of organisms, particularly those with a high ratio of precursors to COs, a rough estimate of N can work very well. For example, in simulating results for grasshopper (text Figure 8H), the value of N was set at 14; however, we did not see obvious difference for N = 8 – 25. On the other hand, for organisms with a low number of precursors and/or a low N/CO ratio, variations in N have more dramatic effects on CoC and ED (CO number/distribution) values (see Figure S2). Thus, specifying (N) as accurately as possible helps to accurately define the strength of interference.

- The parameter (B) makes it possible to model the extent to which a given (genetic) chromosome gets different number of precursors in different nuclei. By changing the value of (B), the user can vary this condition from a random distribution (Poisson distribution; B = 0) to a constant number (B = 1), where the number of precursors for a given chromosome is the same in all nuclei. When the number of

precursors is high, the value of (B) has little effect on ED (CO number/distribution) or on CoC relationships. However, at low values of N, more random distribution of precursors among different chromosomes (nuclei) will increase the fraction of chromosomes that fail to get even one precursor (or, more generally, even one precursor that is adequately sensitive to the CO driving force defined by  $S_{max}$ ). This will lead to an increase in the frequency of chromosomes that fail to acquire even a single CO (zero-CO chromosomes or E0's; text Figure 13, Figure S2B and S5F). Correspondingly, there are indications that the precursors tend to occur at a relatively constant level along a given chromosome in different nuclei. For example, in *C. elegans*, there are only a small number average of precursors, but every chromosome in every nucleus gets at least one DSB [15]. In *Sordaria*, precursor numbers defined by Rad51 and Mer3 foci are relatively similar for each chromosome among different nuclei [11]. Thus usually  $B \gg 0$ .

- The parameter (E) sets the precursor distribution along each given chromosome. The value of E is between 0 and 1. The default value is 0, i.e. randomly distributed precursors, which is set by the MATLAB function "rand" based on Poisson distribution in [0, 1] interval. E = 1 means absolutely evenness. Between these two extremes, evenness levels are adjusted by varying the standard deviation relative to the mean of [the distance between adjacent precursors], which in turn is given by the average precursor number N. A commonly-used metric for defining "evenness of spacing" is the shape parameter of the gamma distribution (" $\nu$ "). To relate values of (E) to this metric, the distribution of inter-precursor distances at different values of (E) were subjected to gamma distribution analysis. Roughly, E = 0 corresponds to  $\nu=1$ ; E = 0.6 corresponds to  $\nu = 2.4$ ; E = 0.7 corresponds to  $\nu = 4.5$  and at E=1,  $\nu$  is infinite. Users can obtain the values of  $\nu$  corresponding to any specific value of (E) from the BF output file under the "Precursor Gamma" item.

Specification of the value of (E) is important. In most cases, the expectation is that  $E \gg 0$  because, recently, there is more and more evidence, from studies in various organisms, that precursors are distributed with some level of evenness. For example: (i) In *Sordaria*, total precursors marked by Mer3 or Msh4 foci are arranged along chromosomes like peas in a pod, with a  $\nu$  value  $>200$  ([11] and discussion therein). (ii) In mouse meiosis, Msh4 and RPA foci, which are also believed to label total inter-



homolog recombination intermediates, are distributed non-randomly ( $v = \sim 4 - 5$ ; [10,16]).  
(iii) In *S.cerevisiae*, there is interference not only between COs but between COs and NCO [12]. Moreover, evidence for evenly distributed precursors can be detected as early as DSBs stage in other types of analysis [17].

Correspondingly, implementation of evenly distributed precursors in our model, which is never considered/ included in any other model, is very helpful for researchers to understand/distinguish the interplay between precursor "interference" and CO interference. Simulations illustrate the fact that effects of even spacing of precursors can be seen in CO patterns when the level of N is relatively low (as seen e.g. for *Drosophila* X or *S. cerevisiae* chromosome III; text Table 2; text Figure S2A). When the value of N is relatively high, the effects of CO patterning override the effects of even spacing of precursors (text Figure 3B).

- (A). The sensitivity (s) of each precursor to the CO driving force is randomly assigned with the MATLAB function "rand". Our original model assumed that the value of (A) is the square root of this random number [18]. In the BF model, 4 preset amplitude distributions with different average sensitive levels are ordered from lowest to highest based on this with little modification. Different values of (A) are required for best-fit simulations in different organisms, in the following order. A=1 (*S.cerevisiae*, *Drosophila*); A= 2 (grasshopper); A=3 (tomato and probably also for other higher plants) (text Table 2).

- For specifying a "black hole" the BF default is [0, 0, 1] which implies no difference in the average frequency of precursors along the chromosome. In this option, the first two numbers define the black hole interval and the third number defines precursor density in the black hole interval compared to the average precursor density along the whole chromosome. Multiple intervals with different average precursor densities can be specified although not incorporated in this program.

Application of this option is most prominently useful to accommodate the fact that, in some organisms or for some chromosomes of a given organism, the centromere appears to suppress DSB formation in a large region around it, which contributes to a reduced COs density in this region (see text for more details). Thus, the black hole

option has been set up as a low precursor density in the region surrounding the centomere (e.g. text Figure 8B). Necessarily, and as documented by simulations (e.g. text compare Figure 8F with 8G), introduction of a black hole has no effect on CO interference (for interval pairs in which both intervals actually acquire a significant number of COs).

Application of the black hole and ends-clamping options are very useful features of BF simulations. This is particularly obvious in some organisms such as grasshopper. On the other hand, variations in these features are of negligible significance in some organisms and/or particularly on the long chromosomes of such organisms where ends comprise a relatively small fraction of the total chromosome length.

(5) “Non-interfering” COs. Genetic studies indicate that there are two different pathways for CO formation in most organisms. The major pathway in most organisms is subject to patterning via the CO interference process. This pathway is sometimes referred to as the "ZMM" pathway (in *S.cerevisiae* it includes at least Zip1-4, Mer3, Msh4/5, Mlh1/3). In some organisms, e.g. *C.elegans* this is likely the only pathway, at least in WT meiosis. A secondary pathway also appears to give rise to COs, outside of the normal patterning process, resulting in what are referred to as "non-interfering COs" (although this may not be a correct characterization in the strict sense). These COs are thought to comprise 5-30% of total COs among several different species as defined by biological or computational analysis [19]. In some organisms, COs do not exhibit interference (e.g. *S. pombe* and *Aspergillus nidulans*). In such organisms, all COs might arise by the same "non-interfering" pathway present in most other cases, or by some different route.

In the BF model, the effects of non-interfering COs on the total array of COs can be modeled. The nature of the non-interfering CO contribution is specified in three steps. First, one of two possible sources of such COs is selected, either Scenario 1 or 2 (further discussion in text and text Figure 14). In Scenario 1, non-interfering COs arise from the same array of precursors as the patterned COs, by random selection among the set of precursors that is left behind after CO-designation. In Scenario 2, non-interfering COs arise from an independently-established array of precursors which might appear either before or after the precursors that give rise to patterned COs. Second, the final fraction

of total COs that arise from that source is specified. For both Scenarios, the default option is zero, i.e. no non-interfering COs. Third, for Scenario 1, the precursor array is automatically defined by the parameters used for CO patterning. However, for Scenario 2, the nature of the precursor array can be set either randomly or evenly distributed along chromosomes (i.e.  $E=0$  or  $E>0$ ) and the number is either Poisson distributed or more constantly among chromosomes (i.e.  $B=0$ , or  $B>0$ ).

(6) *CO maturation efficiency (M)*. The default value of M is 1, which implies that 100% of CO-designated interactions actually finally mature into COs. This condition applies in most cases. However, for some chromosomes in some organisms in wild-type meiosis, and in several mutants, the probability that a CO-designated interaction will emerge as a mature CO/chiasma is not 100% because of some inefficiency or defect in events after the designation step. We describe this situation as a reduction in the efficiency of "CO maturation", the magnitude of which is given by a value of  $M<1$ . In the BF program, a defect in maturation is simulated simply by randomly removing some CO designation sites from the primary output array of programmed COs. Necessarily, and as confirmed by BF simulations for an *mlh1D* mutant in budding yeast, CO interference as defined by the CoC curve is unaltered by reductions in the value of M (except at very low frequencies of COs/bivalent).

Experimentally, a defect in CO maturation could be detected by comparing the number and distribution of sites defined by very early marker for CO-designation (e.g. Zip3 foci in *S. cerevisiae*) with those defined by a later marker, e.g. Mlh1/3 foci detected by immunostaining; chiasmata detected at diakinesis or metaphase I; or events detected genetically or by microarray analysis of marker/sequence heterozygosities.

Non-interfering COs (section (5) above) could potentially also exhibit inefficient maturation. This possibility can be mimicked by simply adjusting the fraction of these types of COs specified for the simulation (above). Thus, variations in CO maturation efficiency are specifically applied, as such, only to the programmed COs.

(7) *Number of intervals for CoC analysis*. Please refer the Note in section "Analyze CO data" above for details and guidelines.

(8) *Interval number for rank plot analysis.* Please refer to corresponding section above in Part 2. Option (A): Analyze CO data.

### **Outputs.**

The user will be presented with four output files after running a job according to this Option:

- (1) "output\_BF" contains essentially the same outputs obtained from running a job in Option A "Analyze CO data" (example in Protocol Figure 3A).
- (2) "CoC\_BF" contains observed DCOs, predicted DCOs and CoC value for all pairs of intervals as described above for the "coc\_CO" output file of Option A "Analyze CO data".
- (3) "precursor\_BF" shows the array of precursors.
- (4) "co\_BF" contains the CO array generated by the simulation run. The array is presented in essentially the same format as that described above of CO data sets to be input into "Analyze CO data" in Option A above except that all chromosome lengths have been normalized to one and the "zeros" are used to designate sites of precursors that did not undergo CO-designation.

This data, after add number "1" (the normalized total bivalent length) as the last column, can be used as a data set for analysis of a simulation output by Option A. Such analysis can confirm the validity of the precursor distribution and provide additional further information (e.g. inter-precursor distances, gamma distribution, etc).

## **4. Option (C): Automatically search the best fit**

This option automatically scans all combinations of parameter values over particular specified ranges for goodness-of-fit to an input data set. The results of a scan are presented in an ordered list, ranked according to goodness of fit levels based on a PLS (Projected Likelihood Score) derived from that of Falque et al. [20] but with some improvements (including information concerning events at ends and giving more weight to the importance of the ED (average CO number/distribution)). Importantly, however: this weighting method is not a maximum likelihood method and thus the "best fit" as

judged by PLS may not be the true best fit. To overcome this drawback, the software output presents all parameter combinations scanned so that the user can check particular cases individually to get the best fit based on all kinds of criteria. We usually choose the best fit by comprehensively comparing simulated and actual data for average number of COs per bivalent, the CoC curve, and the distribution of CO number per bivalent. As discussed in the main paper, the gamma distribution is not a very good indicator of the strength of CO interference level but can be useful in distinguishing whether a given variation in CO patterning reflects a change in a patterning parameter ( $L$ ,  $S_{max}$ ) or some other parameter (for the precursor array or maturation efficiency). Thus we also use the values of average inter-CO distance and the shape parameter  $\nu$  as additional inputs when deciding on a best fit.

The scanning provided by Option C is useful, especially for users without extensive prior experience with BF simulations, to provide rough estimates of reasonable parameter values. However, to run a job by this Option is very time consuming. We therefore suggest that users first run a job in which parameters are varied over a large range, but with large step sizes between the evaluated values. This can be followed by a second (and maybe also a third) round in which the range of values is narrowed and the step sizes decreased based on the results of the prior scan(s).

Because this scan is so time consuming, readers must be well-informed as to likely ranges of parameter values (as discussed above and in the main text) before initiating a job.

### **Inputs.**

All parameters are same as those in Option B “A single BF simulation”. In addition, the user will have to specify a range and a step size for parameters  $L$ ,  $N$  and  $S_{max}$ . Once these values are specified and the job is started, the user will be asked to select the specific CO data file to be used as the reference (formatting as described for Option A “Analyze CO data”).

**Outputs.** The user will be presented with a single output file named “BFscan”. This file contains three worksheets:

(1) *"Zip3" sheet* is the experimental data set to which the program seeks a best fit. This sheet presents: (a) average CO number per bivalent, (b) CO events distribution (ED) and (c) average CoC from the experimental data set.

(2) *"BFcopattern" sheet* includes average CO number and CO number per bivalent (ED) from BF simulation (see Figure 3C for an example output).

(3) *"BFcoc" sheet* includes CoC data from BF simulation (see Protocol Figure 3B for an example output).

## 5. Inner structure of the software

Once all input information has been entered, the program carries out the following tasks in a specific order (which depends somewhat on the option(s) selected, Figure S5 and S6). The details of these progressions are very useful if the user wishes to modify the program.

(1) *Precursor array*. The program first sets up the precursor array from the input values of N, B, E and A as described above. If a black hole for the precursor is set up, the precursor array is then adjusted according to the black hole setting.

(2) *Raw CO array*. Once the precursor array is defined, the program will apply the previously-described beam-film function for CO-designation and interference [18] to each of the specified number of bivalents (beams). This operation defines, for each bivalent, which precursor will be a CO (crack) based on the values of the distance (L) over which inhibiting interference (stress) spreads, the maximal level of CO-designation (stress) (Smax) and the distribution of precursor sensitivities to CO-designation (flaw amplitude) (A). The process reaches its limit when no remaining precursors have an amplitude high enough to respond to the specified maximum CO-designation force (Smax). The final outcome is an array of COs and an array of "leftover" un-reacted precursors, output as described above.

(3) *Adjusted CO array*. The raw CO array is then adjusted according to additional inputs such as maturation efficiency (M) and/or the level of non-interfering COs (above). From this reorganized array, program calculates ED relationships, i.e. the average

number of COs per bivalent and the frequencies of bivalents with different numbers of COs; the average inter-CO distance, the distribution of distances between adjacent COs and, based on the latter distribution, shape parameter of the Gamma distribution ( $\nu$ ).

(4) *Finalized CO array.* COs in the reorganized, adjusted CO array are now sorted into the intervals specified for CoC analysis. Each CO is assigned to a specific interval, thus defining a "finalized" CO array.

(5) *CoC.* Given the finalized CO array in (4), the program now calculates, for each pair of intervals: (i) the actual frequency of double COs (Obs DCO); the predicted frequency of double COs (Pred DCO) based on the assumption that COs occur independently in the two intervals; and, finally, the ratio of (Obs DCO)/ (Pred DCO), i.e. the CoC.

(6) *Outputs.* Output files are created, according to the option selected, as outlined above.

## **6. More hints for obtaining the best fit between a BF simulation and an experimental data set.**

There are two approaches for arriving at the best-fit BF simulation of an experimental data set. One approach is to use the best-fit scan described above. However, a user with more experience regarding the effects of changes in various parameters should be able to arrive at the best fit more directly by trial-and-error with a constrained set of possible parameter values (see main text for more details).

Additional practical tips for arriving at the best fit:

(a) *Interplay between  $L$ ,  $S_{max}$ .* Both increased  $L$  (or decreased  $L$ ) and decreased  $S_{max}$  (or increased  $S_{max}$ ) can decrease (or increase) CO number and shift CoC curve to right (or left). However the two parameters have distinguishable effects for CoC and CO number as discussed in the text and text Figure 4. Furthermore, as shown in Figure 4 and Figure S4, if both  $L$  and  $S_{max}$  are changed coordinately while keeping Event Distribution (ED) relationships (CO number/ distribution) fixed, the CoC curve changes; and if both  $L$  and  $S_{max}$  are changed coordinately while keeping the CoC curve is kept

constant, the CO number/distribution changes. In summary, it is always possible to find specific optimal values of L and Smax.

(b) *L, N and physical chromosome length.* In a given organism, different chromosomes are of different lengths, both genetically and physically (e.g. at the pachytene stage when CO-correlated foci are examined for patterning). When comparing chromosomes of different lengths in BF simulations, two different considerations must be kept in mind.

First: in the BF model, chromosome length is always defined as 1. Therefore, if interference spreads over a particular physical distance along the chromosome (as appears to be the case; [21,22,23]), and assuming that distance is the same for all chromosomes, the best-fit value of (L) in the BF model will have to be normalized such that the absolute length in physical chromosome distance is the same for all chromosomes. That is, for longer chromosomes, the value of L must be a smaller percent of total chromosome length, and *vice versa* for shorter chromosomes. If this normalization is not performed, different chromosomes will exhibit different values of L (as % total chromosome length) whereas the biological reality is that the interference distance is the same for all chromosomes.

Second: experimental data suggest that the number of precursors along a chromosome is proportional to chromosome length ([15]; L.Z. and N.K. in preparation). For BF simulations, this fact implies that for best fit simulations of chromosomes of different lengths, the value of (N) must be scaled such that (N)/length is a constant. This is important because CoC curves and the number/distribution of COs per bivalent are both sensitive to the number of precursors (text, Figure 3).

When these two considerations are taken together, they imply that correct modeling of genetically different chromosomes of different lengths, in the same organism, requires that both (L) and (N) be appropriately normalized to chromosome length. An example is shown in Protocol Figure 4. Simulations were carried out for a hypothetical set of four chromosomes whose lengths vary over a four-fold range. From largest to smallest, the BF value of (L) is increased four-fold (to provide a constant absolute value per unit actual chromosome length) while the value of (N) decreases four-fold in accord with a constant density of precursors per unit length. All other parameters are kept



constant. The average number of COs/bivalent varies in direct proportion to the chromosome length, just as does the average number of precursors per bivalent, as shown in Protocol Figure 4A over a wider range of values of chromosome length. However, the CoC curves for the four types of chromosomes are exactly overlapping (Protocol Figure 4B).

There are also situations in which a chromosome of a given genetic composition exhibits two different physical lengths in two different situations, where the interference distance extends over the same physical length in both cases. Modeling of these situations requires the same two adjustments, for analogous reasons (see text Figure 9).

## 7. Running environments

The following apply to both Windows and Mac operating systems.

- (a) .exe (large file) can be run in computers without MATLAB.
- (b) Original codes (MATLAB code) works only in a MATLAB window.

## References

1. Hillers KJ, Villeneuve AM (2003) Chromosome-wide control of meiotic crossing over in *C. elegans*. *Curr Biol* 13: 1641-1647.
2. Joyce EF, McKim KS (2010) Chromosome axis defects induce a checkpoint-mediated delay and interchromosomal effect on crossing over during *Drosophila* meiosis. *PLoS Genet* 6: e1001059.
3. Joyce EF, McKim KS (2011) Meiotic checkpoints and the interchromosomal effect on crossing over in *Drosophila* females. *Fly (Austin)* 5: 134-140.
4. Gerton JL, DeRisi J, Shroff R, Lichten M, Brown PO, et al. (2000) Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 97: 11383-11390.
5. Buhler C, Borde V, Lichten M (2007) Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*. *PLoS Biol* 5: e324.
6. Blitzblau HG, Bell GW, Rodriguez J, Bell SP, Hochwagen A (2007) Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres. *Curr Biol* 17: 2003-2012.

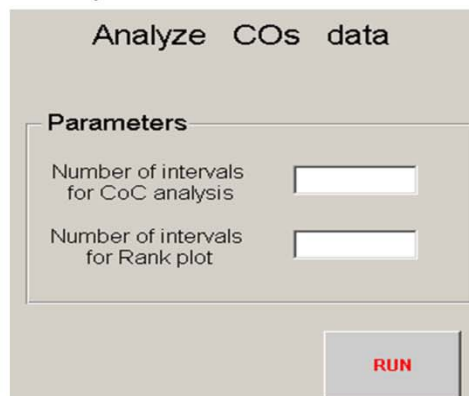
7. Pan J, Sasaki M, Kniewel R, Murakami H, Blitzblau HG, et al. (2011) A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* 144: 719-731.
8. Chen SY, Tsubouchi T, Rockmill B, Sandler JS, Richards DR, et al. (2008) Global analysis of the meiotic crossover landscape. *Dev Cell* 15: 401-415.
9. Cole F, Kauppi L, Lange J, Roig I, Wang R, et al. (2012) Homeostatic control of recombination is implemented progressively in mouse meiosis. *Nat Cell Biol* 14: 424-430.
10. de Boer E, Stam P, Dietrich AJ, Pastink A, Heyting C (2006) Two levels of interference in mouse meiotic recombination. *Proc Natl Acad Sci U S A* 103: 9607-9612.
11. Storlazzi A, Gargano S, Ruprich-Robert G, Falque M, David M, et al. (2010) Recombination proteins mediate meiotic spatial chromosome organization and pairing. *Cell* 141: 94-106.
12. Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM (2008) High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature* 454: 479-485.
13. Pawlowski WP, Golubovskaya IN, Cande WZ (2003) Altered nuclear distribution of recombination protein RAD51 in maize mutants suggests the involvement of RAD51 in meiotic homology recognition. *Plant Cell* 15: 1807-1816.
14. Anderson LK, Doyle, G.G., Brigham, B., Carter, J., Hooker, K.D., Lai, A., Rice, M., and Stack, S.M. (2003) High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics* 165: 849-865.
15. Mets DG, Meyer BJ (2009) Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure. *Cell* 139: 73-86.
16. Oliver-Bonet M CM, Turek PJ, Ko E, Martin RH (2007) Analysis of replication protein A (RPA) in human spermatogenesis. *Mol Human Reprod* 13: 837-844.
17. Zhang L, Kim KP, Kleckner NE, Storlazzi A (2011) Meiotic double-strand breaks occur once per pair of (sister) chromatids and, via Mec1/ATR and Tel1/ATM, once per quartet of chromatids. *Proc Natl Acad Sci U S A* 108:20036-41.
18. Kleckner N, Zickler D, Jones GH, Dekker J, Padmore R, et al. (2004) A mechanical basis for chromosome function. *Proc Natl Acad Sci U S A* 101: 12592-12597.
19. Berchowitz LE, Copenhaver GP (2010) Genetic interference: don't stand so close to me. *Curr Genomics* 11: 91-102.
20. Petkov PM, Broman KW, Szatkiewicz JP, Paigen K (2007) Crossover interference underlies sex differences in recombination rates. *Trends Genet* 23: 539-542.

21. Petkov PM, Broman KW, Szatkiewicz JP, Paigen K (2007) Crossover interference underlies sex differences in recombination rates. *Trends Genet* 23: 539-542.
22. Billings T, Sargent EE, Szatkiewicz JP, Leahy N, Kwak IY, et al. (2010) Patterns of recombination activity on mouse chromosome 11 revealed by high resolution mapping. *PLoS One* 5: e15340.
23. Drouaud J, Mercier R, Chelysheva L, Berard A, Falque M, et al. (2007) Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. *PLoS Genet* 3: e106.

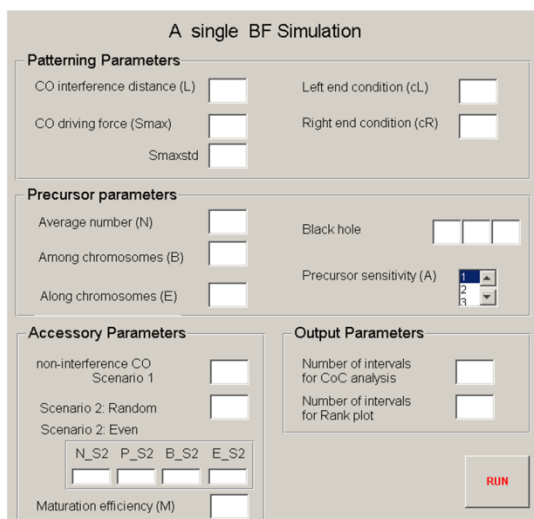
A. Main Window (3 options)



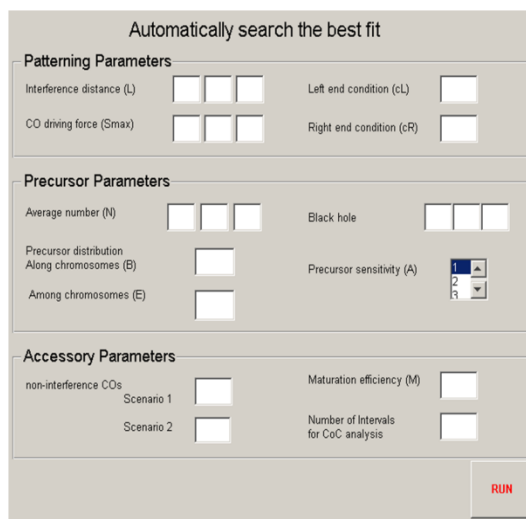
B. "Analyze COs Data" Window



C. "A single BF Simulation" Window



D. "Automatically search the best fit" Window



**Protocol Figure 1. BF Software User Interface.**

(A) The main window includes three options (B-D). (B) The sub-window of Option A: "Analyze CO data"; (C) The sub-window of Option B: "A single BF simulation". (D) The sub-window of Option C: "Automatically search the best fit".

An example for formatted CO data to be analyzed

Bivalent	Ordered CO position ( $\mu\text{m}$ )							Bivalent Length
	a	b	c	d	e	f	g	
1	0.29	0.8	1.14	1.55	2.11	2.83	3.26	3.45
2	0.13	0.47	0.95	1.17	2.34	0	0	3.3
3	0.13	0.84	1.64	2.94	0	0	0	3.15
4	1	1.8	2.57	2.99	0	0	0	3.33
5	0.19	0.75	1.75	2.13	2.6	3.07	0	3.3

**Protocol Figure 2. An example of the proper format for a data set to be subjected to for CO analysis (Option A) or BF scan (Option C).**

For each bivalent in the data set, CO positions are ordered, beginning at one end defined as 0, from smaller to larger. The last column is the total bivalent length. For bivalents with fewer COs than the maximum number present in the data set, the positions beyond the last CO are filled with zeros, e.g. for the second bivalent, columns f and g are filled with “0”.

**A**

```
CO/biv      interCO-gam
0           5.8904
0           Stahl gam
0.006      5.6049
0.0926     0.053
           L
0.331     0.1
0.41      Smax
0.141     3.5
0.0174    smaxstd
0.001     0
           cL
Avg CO     0.85
4.6492    cR
Avg CoC   0.85
0.5579    N
1.0315    13
0.9978    Actual N
           12.679
1.0045    Precursor
0.9956    gamma
0.99      2.4375
1.0096    0.0312
           E
0.9835    0.6
0.9974    B
hump      0
0.1929    A
CO/interval 3
0.3258    Black hole
0.5014    0
0.4848    1
0.4986    nonIF S1
0.4862    0
0.5014    nonIF S2
0.488     0
0.4834    Maturation
0.5036    1
0.3108
```

**B**

Average CoC between pairs of intervals with "X" of intervals spaced. X =

Parameter values				X =									
				0	1	2	3	4	5	6	7	8	
L	N	Smax	PLS	CoC									
0.1	12	3.5	13	0.539	1.030	0.993	1.012	0.993	1.013	0.973	1.030	0.983	
0.1	12	4	14	0.610	1.038	1.001	0.997	0.990	1.009	0.995	0.983	0.994	
0.1	13	3.5	18	0.555	1.043	0.986	1.004	0.999	0.987	1.016	0.999	1.016	
0.1	13	3	24	0.464	1.016	1.022	1.003	0.987	1.004	1.023	0.969	1.008	
0.1	12	3	25	0.450	1.018	0.998	1.007	0.980	1.015	0.995	0.990	1.002	
0.1	14	3	25	0.466	1.037	1.015	0.999	0.999	0.995	0.996	1.005	0.998	
0.1	13	4	27	0.635	1.027	0.999	1.001	0.999	0.992	0.993	0.989	1.012	
0.1	14	3.5	29	0.570	1.033	1.002	0.991	1.014	0.987	0.993	1.008	0.978	
0.1	14	4	47	0.636	1.029	1.003	0.997	1.009	1.003	0.986	1.011	1.003	

**C**

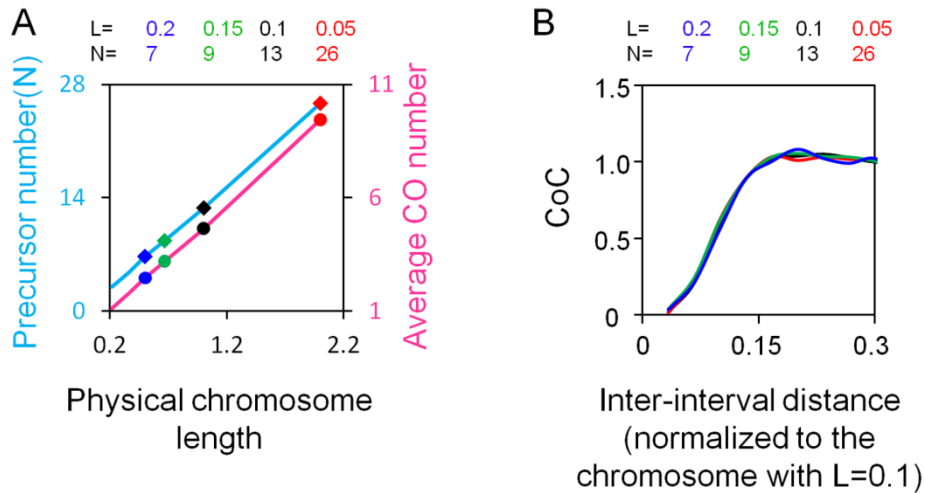
Average Fraction of "X" COs per bivalents. X =

Parameter values				Average CO#	X =									
				CO#	0	1	2	3	4	5	6	7	8	
L	N	Smax	PLS	AveCO#	COpatn									
0.1	12	3.5	13	4.49	0.000	0.000	0.012	0.113	0.377	0.370	0.111	0.013	0.002	
0.1	12	4	14	4.89	0.000	0.000	0.002	0.060	0.270	0.396	0.219	0.040	0.007	
0.1	13	3.5	18	4.66	0.000	0.000	0.008	0.081	0.342	0.400	0.146	0.018	0.004	
0.1	13	3	24	4.16	0.000	0.001	0.024	0.192	0.437	0.295	0.049	0.002	0.000	
0.1	12	3	25	4.01	0.000	0.001	0.033	0.232	0.452	0.246	0.035	0.000	0.000	
0.1	14	3	25	4.24	0.000	0.000	0.016	0.162	0.447	0.310	0.062	0.002	0.000	
0.1	13	4	27	5.05	0.000	0.000	0.001	0.039	0.212	0.426	0.255	0.052	0.007	
0.1	14	3.5	29	4.79	0.000	0.000	0.003	0.059	0.304	0.430	0.178	0.022	0.002	
0.1	14	4	47	5.18	0.000	0.000	0.002	0.026	0.183	0.419	0.288	0.070	0.006	

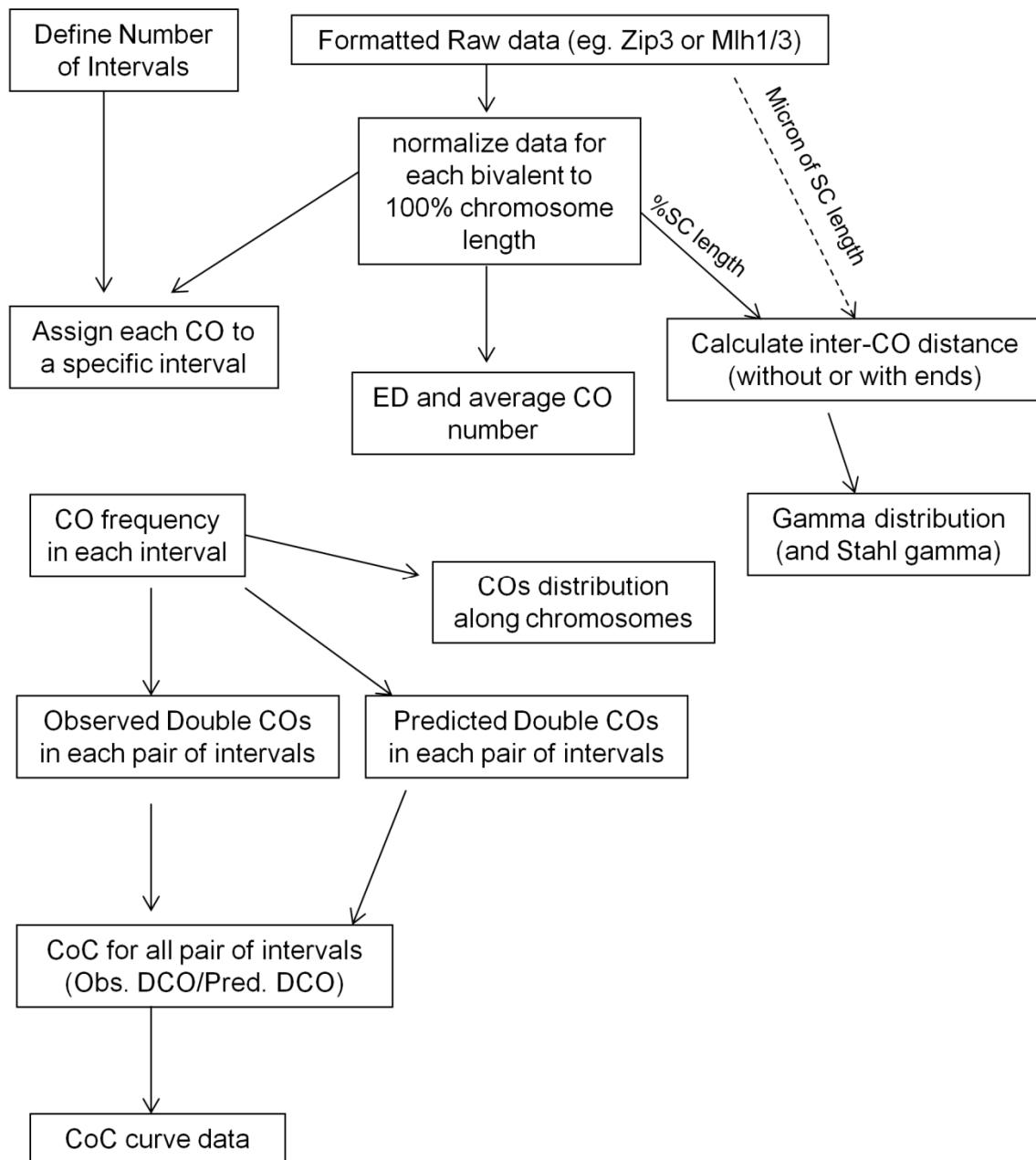
**Protocol Figure 3. Output examples.**

(A) An example of the output from Option B: "a single BF simulation" output.

(B, C) Examples of the outputs from Option C: "BFscan", for worksheets named "BFcoc" and "BFcopattern" respectively. The parts below the dash lines are the actual outputs, and above it are annotations added.

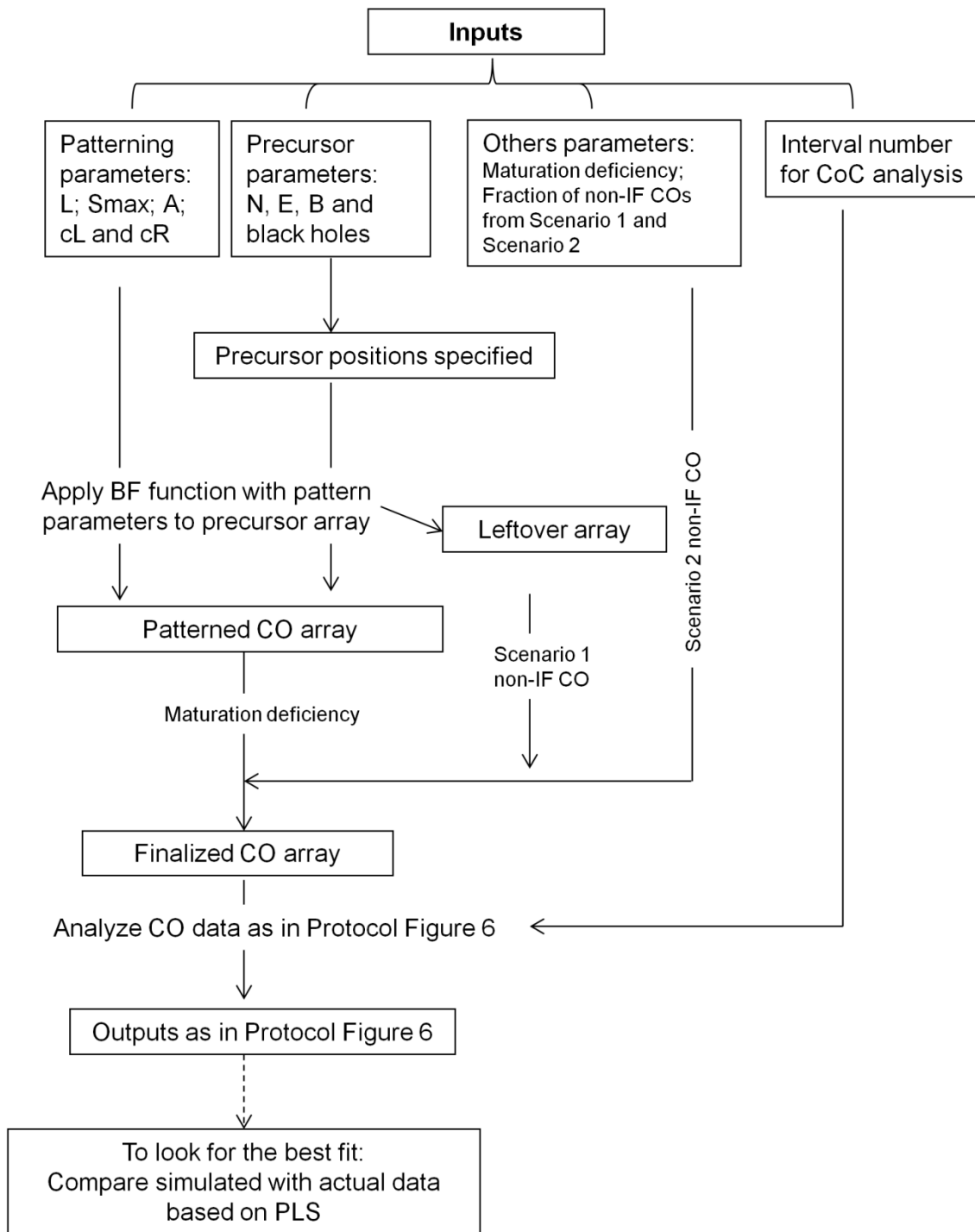


**Protocol Figure 4. Effects of physical chromosome length.** If different chromosomes are of the different physical lengths, the number of precursors is proportional to physical length as shown by experimental analysis. Thus, in BF simulations of such cases, (N) must be increased in proportion to chromosome length. Additionally, assuming that interference spreads over the same physical distance in both cases, as shown by experimental analysis for several examples, and given that the BF model normalizes every chromosome length to "1", the input value of (L) must be adjusted inversely to chromosome length such that it corresponds to the same physical distance in all cases. Simulations were performed in which (N) and (L) are varied coordinately as required to accommodate different physical chromosome lengths. Such simulations show that the number of COs varies in proportion to the number of precursors (Panel A) whereas CoC relationships are identical in all cases (Panel B).



**Protocol Figure 5. Program work flow chart for Option A: Analysis of CO data.**





**Protocol Figure 6. Program work flow chart for Option B: A Single BF Simulation.**