SUPPORTING INFORMATION

for

Genomic features shaping the landscape of meiotic double-strand break hotspots in maize

Yan He, Minghui Wang, Stefanie Dukowic-Schulze, Adele Zhou, Choon-Lin Tiang, Shay Shilo, Gaganpreet K. Sidhu, Steven Eichten, Peter Bradbury, Nathan M. Springer, Edward S. Buckler, Avraham A. Levy, Qi Sun, Jaroslaw Pillardy, Penny M.A. Kianian, Shahryar F. Kianian, Changbin Chen, Wojciech P. Pawlowski[1]

[1] Corresponding author. E-mail: wp45@cornell.edu

This PDF file contains:

Supporting Methods

Tables S1 and S2

Figs. S1 to S13

SUPPORTING METHODS

RAD51 Chromatin Immunoprecipitation (ChIP) to map meiotic double-strand breaks (DSBs) in maize

Anti-RAD51 antibody production

To generate a recombinant RAD51 protein, the full-length CDS of *ZmRAD51A1* was cloned into the pET28a vector using primers ZmRAD51A-L1 and ZmRAD51A-R1 (See SI Appendix, table S1) and expressed in *E. coli* strain BL21 (Promega, Fitchburg, WI, USA). 0.5μg of recombinant protein was injected into a rabbit to raise a polyclonal antibody. The antibody was affinity-purified using the NAB Protein A Plus Spin Kit (Thermo Scientific, Waltham, MA, USA), desalted into 1X PBS buffer using the Slide-A-Lyzer Dialysis Cassette (Thermo Scientific), and concentrated to 1 μg/μl using a protein concentrator (Thermo Scientific) following manufacturer's instructions. The specificity of the antibody was tested using an immunolocalization experiment. Pre-immune rabbit IgG was purified using the same method as used for the anti-RAD51 antibody from serum collected from the same rabbit prior to antigen injection.

ChIP

Male flowers at the zygotene stage of meiotic prophase I were collected from at least twenty maize tassels, fixed in 1% formaldehyde for 10 min, and then quenched in 0.125 mM glycine for 5 min. About 1.5 g of the fixed flower tissue was ground into fine powder in liquid nitrogen and homogenized in Chromatin Extraction Buffer A (10 mM Tris–HCl pH 8.0, 0.4 mM sucrose, 10 mM $MgCl_2$, 1 mM PMSF, 5 mM β-

mercaptoethanol, and 1 tablet of cOmplete Protease Inhibitor Cocktail (Roche Applied Science, Indianapolis, IN, USA) per 50 ml of buffer) for 20 min at 4°C with gentle shaking. The suspension was filtered into a new 50 ml conical tube through 2 layers of Miracloth (EMD Millipore, Billerica, MA, USA), placed in a plastic funnel and centrifuged at 4,000 rpm for 20 min at 4°C. The pellet was resuspended in 1 ml of extraction buffer B (10 mM Tris–HCl pH 8.0, 0.25 M sucrose, 10 mM $MgCl_2$, 1% Triton X-100, 1 mM PMSF, 5 mM β-mercaptoethanol, and 1 tablet of cOmplete Protease Inhibitor Cocktail (Roche Applied Science) per 50 ml of buffer), and centrifuged at 14,000 rpm for 10 min. The pellet was resuspended in 500 µl of Extraction Buffer C (10 mM Tris–HCl pH 8.0, 1.7 M sucrose, 2 mM $MgCl_2$, 0.15% Triton X-100, 1 mM PMSF, 5 mM β-mercaptoethanol, and 1 tablet of cOmplete Protease Inhibitor Cocktail (Roche Applied Science) per 50 ml of buffer), placed on top of 500 µl of Extraction Buffer C cushion, and centrifuged at 14,000 rpm for 1 h at 4°C. The nuclei pellet was resuspended in 500 µl Nuclei Lysis Buffer (50 mM Tris–HCl pH 8.0, 10 mM EDTA, 1% SDS, 1 mM PMSF, and 1 tablet of cOmplete Protease Inhibitor Cocktail (Roche Applied Science) per 50 ml of buffer).

The resulting chromatin was sonicated into fragments of average length of 200 - 400 bp using eight sonicator pulses, five seconds each, and centrifuged at 14,000 rpm for 5 min at 4°C. The supernatant was diluted 10 fold with Dilution Buffer (16.7 mM Tris–HCl pH 8.0, 1.2 mM EDTA, 167 mM NaCl, 1.1% Triton X-100, 1 mM PMSF, and 1 tablet of cOmplete Protease Inhibitor Cocktail (Roche Applied Science) per 50 ml of buffer), and pre-cleared with Dynabeads Protein A (Invitrogen, Carlsbad, CA, USA). 10 µl of the

pre-cleared chromatin was saved as the chromatin input control sample. The remaining chromatin was incubated with antibodies overnight at 4°C, and then captured by Dynabeads Protein A. After incubation, beads were washed 5 times in the following buffers: (i) Low Salt Wash Buffer (20 mM Tris–HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.1% SDS, 1% Triton X-100), (ii). High Salt Wash Buffer (20 mM Tris–HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 0.1% SDS, 1% Triton X-100), (iii) LiCl Buffer (10 mM Tris–HCl pH 8.0, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate), and (iv) TE Buffer, twice (10 mM Tris–HCl pH 8.0, 1 mM EDTA). Chromatin was eluted from beads with 200 μl Elution Buffer (50 mM Tris–HCl pH 8.0, 10 mM EDTA, 200 mM NaCl, 1% SDS), followed by a second round of immunoprecipitation, washes, and elution. Eluted chromatin was treated with RNase for 2 h at 37°C, deproteinized with Proteinase K for 2h at 45°C, and decrosslinked at 65°C for 8 h. DNA was purified using MinElute PCR Purification Columns (Qiagen, Hilden, Germany). DNA concentration was quantified using Quant-iT dsDNA HS Assay Kit (Invitrogen).

ChIP-seq Library Construction and Sequencing

Illumina libraries were constructed using the ChIP-seq DNA Sample Prep Kit (Illumina, San Diego, CA, USA) following manufacturer's protocol, except that DNA size selection was done after the PCR step (1). Libraries were sequenced using Single-end Cluster Generation Kits and 100-cycle Sequencing Kits (Illumina) using the Illumina HiSeq2000 sequencing system following manufacturer's instructions.

Quantitative real time PCR (Q-PCR) was used to confirm enrichment of hotspot DNA by anti-RAD51 ChIP (See SI Appendix, fig. S13). Q-PCR was performed using iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) in a 7500 Real-Time PCR thermocycler (Applied Biosystems, Foster City, CA, USA) according to manufacturer's instructions. Gene copy numbers were calculated using Applied Biosystems SDS Software Version 1.3.1. Hotspots enrichment was calculated first by normalizing the copy number of the target DNA region in the ChIP sample to the input sample, and then by normalizing this ratio to a ratio obtained from the analysis of the *Ubiquitin* gene region using the following equation:

$2^{-[[Ct\ (Hotspot\ region\_ChIP)-Ct\ (Hotspot\ region\_Input)]-[Ct\ (Ubiquitin\ region\_ChIP)-Ct\ (Ubiquitin\ region\_Input)]]}$

All data were averages of three independent experiments. Primer sequences used for q-PCR are listed in SI Appendix, table S1.

Nucleosome occupancy mapping

Nuclei pellet was resuspended in 500 µl Digestion Buffer (50 mM Tris-HCl pH8.0, 5 mM $CaCl_2$, 0.1 mM PMSF, and 1 tablet of cOmplete Protease Inhibitor Cocktail (Roche Applied Science) per 50 ml of buffer). Samples were sonicated for 5 second and treated with 1 U/µl of micrococcal nuclease (NEB, Ipswich, MA, USA) for 10 min at room temperature. The reaction was stopped with 10 mM EDTA and DNA was prepared as in the ChIP protocol. The purified DNA was separated in a 2% agarose gel and fragments ~150 bp in size were recovered. Approximately 100 ng of mononucleosome DNA was used for Illumina library construction. As a control, randomly fragmented chromatin was prepared by sonication to produce 200 - 500 bp fragments.

Immunolocalization and Immuno-Fluorescence *in situ* hybridization (FiSH)

Immunolocalization experiments were performed as previously described (2). Staged anthers were fixed for 45 min in buffer containing 4% paraformaldehyde to best preserve chromatin structure. Meiocytes were extruded from anthers and embedded in 5% polyacrylamide to maintain the three-dimensional structure of the nuclei. Primary antibodies, rabbit anti-ZmRAD51, mouse anti-HsH3K4me3 (Abcam, Cambridge, MA), and mouse anti-HsH3K9me2 (Zymo Research, Irvine, CA), were diluted 1:500. Secondary antibodies were diluted 1:50.

Immuno-FiSH experiments were also conducted as previously described (3), using a probe specific to the maize centromere sequence (4) and the anti-ZmRAD51 antibody diluted 1:500.

Computational analyses

Processing and mapping Illumina reads to the maize genome scaffold

The following pipeline was used to process Illumina ChIP-seq reads.

(i) Base calling and initial data processing were performed using the standard Illumina protocol. Reads that passed quality control were aligned to the maize B73 reference genome sequence (release AGPv2) using the Burrows-Wheeler Aligner (BWA) (5).

(ii) The first ambiguous base was removed from sequencing reads.

(iii) The reads were progressively trimmed at the 3' termini, 1 bp at a time, until they could be mapped to the maize genome scaffold with no more than 2 mismatches. Only reads that were longer than 40 bp after trimming were aligned.

(iv) Reads that mapped to multiple genome locations (43 – 57% of reads, depending on the experiment) were filtered out. SAMtools (6) was used to convert the uniquely mapping reads to BAM files.

To compare the two independent RAD51 ChIP experiments, Illumina reads uniquely mapping to the maize genome scaffold were counted using a 2 kb sliding window. The tag counts were normalized to the total number of reads aligned to the genome in each experiment. We conducted two independent RAD51 ChIP experiments and found that they produced highly similar results (Pearson Correlation Coefficient $R^2 = 0.86$; See SI Appendix, table S2). Thus, the two datasets were combined in further analyses.

Peak calling

To identify peaks in ChIP-seq datasets, we used MACS (version 2.0.10) (7), which applies a dynamic Poisson distribution mode to differentiate ChIP-enriched regions from non-enriched regions. For the RAD51 ChIP, peak calling was performed with the following parameters: bandwidth = 800 bp, shift size = 400 bp, MACS mode off, and q-value cutoff = 0.01. Parameters for peak calling in the H3K4me3 ChIP experiments were set at: bandwidth = 300 bp, shift size = 150 bp, MACS mode off, and q-value cutoff = 0.01. In both RAD51 and H3K4me3 ChIP experiments, input chromatin was used to identify regions of enrichment. To call peaks in the RAD51 ChIP, we also used as controls (i) ChIP conducted using pre-immune IgG on meiotic chromatin and (ii) ChIP conducted using the anti-RAD51 antibody on leaf tissue chromatin (See SI Appendix, fig. S2).

Generating and analyzing DSB hotspot maps

The Integrative Genomics Viewer (IGV) (8) was used to visualize DSB hotspot maps. WIG files were generated using $\log_2$-transformed ratios of normalized counts of reads in hotspot regions. Peak annotation was performed using PeakAnalyzer (9). Hotspot mid-points were used to assign hotspot locations to annotated genome features, such as 5' UTRs, 3' UTRs, introns, exons, etc. For genes with multiple gene models, only one, randomly selected model was used. For each annotation category, percentage distribution was calculated and plotted using the R software.

Identification of repeats

To characterize DSB hotspots located in repetitive regions of the maize genome we used RepeatMasker (http://www.repeatmasker.org). To determine whether hotspot regions were enriched in specific repeat classes, a random set of Illumina reads from the chromatin input dataset was used as a control. First, we generated three sets of 10,000 random reads from the chromatin input dataset. We then used RMBlast to compare DSB hotspot sequences and the random sequences to RepBase (10). The three random sequence sets produced highly similar results, and one of them was selected at random to compute enrichment using a one-sided binomial test.

MHS identification

To identify hotspot-associated DNA motifs, a *de novo* motif identification program rGADEM (11) was used with a P-value cut-off of 0.0002, eValue of -5, and numGeneration of 500. After removing peaks overlapping repetitive elements, DNA

sequences corresponding to non-repetitive region hotspots were extracted from the B73 reference genome sequence and used as seeds for the GADEM analysis. The motif search was performed with oligonucleotide sizes ranging from 6 to 20. JASPAR CORE database (12) was used to compare the *de novo* motifs to a list of reference motifs.

Finding MHS in CML228 and in Mo17

To identify the number of MHS copies in *CML228 and in* Mo17, as assembled genome scaffolds are not available for these inbreds, we used Illumina whole-genome shotgun sequences. K-mer sequences 31 bp in length were generated from the Illumina sequencing reads using the Jellyfish software (13). Genome wide k-mer coverage depth was calculated using the following formula: $M = N * (L – K + 1) / L$ where $N = C * L / G$ and M = k-mer coverage, N = base coverage; L = read length, K = k-mer size, C= read count, and G = genome size. K-mers that were only in single copies were removed as they were considered to represent sequencing errors. The remaining k-mers were scanned using a position weight matrix (PWM) to identify those with 95% or more identity to the B73 MHS. The numbers of MHS copies were estimated as numbers of k-mers that match the motif sequence normalized to the genome wide k-mer coverage depth.

CO motif analysis

We analyzed a collection of maize COs that included two previously published datasets (14, 15) as well as a set of 7221 CO events generated by NRGene (Ness-Ziona, Israel) by genotyping-by-sequencing (GBS) (16) 200 F_2 plants from a cross of B73 and a

commercial maize variety from Israel. Altogether, these datasets contained 104 COs that were mapped to within 2 kb. These 104 COs were compared to a random sample of similarly sized fragments originating from the whole genome and screened for presence of sequence motifs using HOMER and MEME (17, 18). In MEME, we used two types of background models, randomized sequences with base composition similar to the dataset and random sampling from the whole genome. Both tools in all conditions identified an A-rich motif (19) and the CC-rich repeat motif as the most significant motifs in maize CO events. TOMTOM, a tool in MEME, was used for motif comparison.

Nucleosome occupancy analysis

Illumina sequence reads from micrococcal nuclease mapping experiments were mapped to DSB hotspots regions and counted using 500 bp sliding windows with a step size of 1bp. The counts were normalized by the total number of aligned reads.

DNA methylation analyses

To analyze methylation levels of DSB and CO sites, DSB hotspots, we used data deposited at http://genomaize.org/cgi-bin/hgTables, which were obtained by immunoprecipitation of methylated DNA from seedling tissue with an anti-5-methylcytosine antibody, followed by DNA hybridization to a high-density tiling microarray (20). We compared methylation levels at sites of the 3,126 DSB hotspots, the 104 COs mapped to within 2 kb (see CO motif analysis), and fragments of similar size selected from the entire maize genome using the U-test (21).

To examine the methylation status of MHS, whole genome bisulfite sequencing (MethylC-seq) was performed (22). Briefly, 173 million 100 bp paired-end bisulfite sequencing reads from B73 seedling tissue were mapped to the maize reference genome (B73 RefGen v2) with the Bismark aligner (23) (v0.7.2; flags: -n 2, -l 50). Methylated cytosines were extracted using the Bismark methylation extractor under standard parameters. Positions of CG, CHG, and CHH methylation sites were mapped relative to MHS positions across the genome and average CG, CHG, and CHH levels were calculated across 100 bins encompassing MHS and 100 bp regions on each side of the motif.

CO distribution

To examine CO hotspots, we used the maize Nested Association Mapping (NAM) population (24). This recombinant inbred line population was generated by crossing a set of 25 diverse maize inbreds to B73. Even though recombination patterns in NAM are likely affected by all parental genotypes, the strongest hotspots in the combined population are expected to be those of B73, as B73 is the common parent in the 25 crosses. The NAM population was genotyped using the genotyping-by-sequencing approach (16, 25). Positions of COs were inferred from the genotyping data by identifying breakpoints between parental SNP alleles as described by Esch *et al*. (26).

REFERENCES

1. Smagulova F*, et al.* (2011) Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* 472:375-378.

2. 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.

3. Armstrong SJ, Sanchez-Moran E, & Franklin FC (2009) Cytological analysis of *Arabidopsis thaliana* meiotic chromosomes. *Methods Mol. Biol.* 558:131-145.

4. Pawlowski WP*, et al.* (2004) Coordination of meiotic recombination, pairing, and synapsis by PHS1. *Science* 303:89-92.

5. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760.

6. Li H*, et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078-2079.

7. Zhang Y*, et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137.

8. Thorvaldsdottir H, Robinson JT, & Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14:178-192.

9. Salmon-Divon M, Dvinge H, Tammoja K, & Bertone P (2010) PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics* 11:415.

10. Jurka J*, et al.* (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* 110:462-467.

11. Li L (2009) GADEM: a genetic algorithm guided formation of spaced dyads coupled with an EM algorithm for motif discovery. *J. Comput. Biol.* 16:317-329.

12. Bryne JC*, et al.* (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.* 36:D102-106.

13. Marcais G & Kingsford C (2011) A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27:764-770.

14. Rodgers-Melnick E*, et al.* (2015) Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl. Acad. Sci. USA* 112:3823-3828.

15. Li X, Li L, & Yan J (2015) Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nat. Commun.* 6:6648.

16. Elshire RJ*, et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379.

17. Heinz S*, et al.* (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38:576-589.

18. Bailey TL*, et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37:W202-208.

19. Shilo S, Melamed-Bessudo C, Dorone Y, Barkai N, & Levy AA (2015) DNA crossover motifs associated with epigenetic modifications delineate open chromatin regions in Arabidopsis. *Plant Cell* 27:2427-2436.

20. Eichten SR*, et al.* (2011) Heritable epigenetic variation among maize inbreds. *PLoS Genet* 7:e1002372.

21. Mann HB & Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* 18:50-60.

22. Eichten SR*, et al.* (2013) Epigenetic and genetic influences on DNA methylation variation in maize populations. *Plant Cell* 25:2783-2797.

23. Krueger F & Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27:1571-1572.

24. McMullen MD*, et al.* (2009) Genetic properties of the maize nested association mapping population. *Science* 325:737-740.

25. Glaubitz JC*, et al.* (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346.

26. Esch E, Szymaniak JM, Yates H, Pawlowski WP, & Buckler ES (2007) Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency. *Genetics* 177:1851-1858.

Table S1. List of PCR primers used in the study.

Primer name	Sequence (5'-3')	Purpose
ZmRAD51A-1L	GGATCCATGTCGTCGGCGGCGCAGCA	Clone full-length CDS sequence of ZmRAD51A into pET28a
ZmRAD51A-1R	CTCGAGTCAATCCTTAACATCTGCAA	
Chr1-L	ACGCCGCACAGATTATGCTGG	ChIP-qPCR of a region on chromosome 1 (33808250 - 133808700)
Chr1-R	ACGGCCAGTACGTGTGCCCTCAG	
Chr5-L	AGGGCGGATCTCCACCCGAG	ChIP-qPCR of a region on chromosome 5 (192481900 - 192482400)
Chr5-R	CGGGCGAGCGCATCTGCGCCT	
Chr9-L	CCAGACTTCTGCTCGGTGTAC	ChIP-qPCR of a region on chromosome 9 (11778246 - 11778537)
Chr9-R	AGTCCGGCATGGACATGTGGCT	
UBQ10-L	CCTCTGAGGCGGAGAACAAGGTGA	ChIP-qPCR of a region in the promoter of *Ubiquitin10* (chr5: 82444264 - 82444445)
UBQ10-R	ATTGACTGGCAAGACCATCACCT	

Table S2. Pearson Correlation Coefficients R^2 between RAD51 ChIP-seq data and three controls. Anti-RAD51 = ChIP conducted using the anti-RAD51 antibody on meiotic chromatin. Input = Illumina-sequenced meiotic input chromatin that was not subjected to ChIP. IgG = ChIP conducted using pre-immune IgG instead of the anti-RAD51 antibody on meiotic chromatin. Anti-RAD51 on leaf = ChIP conducted using the anti-RAD51 antibody on leaf tissue chromatin.

Treatment		**Anti-RAD51**		**Input**		**IgG**
	Replicate	1	2	1	2	
Anti-RAD51	1	-				
	2	0.86	-			
Input	1	0.48	0.58	-		
	2	0.48	0.56	0.92	-	
IgG		0.66	0.76	0.85	0.81	-
Anti-RAD51 on leaf		0.07	0.07	0.19	0.24	0.18

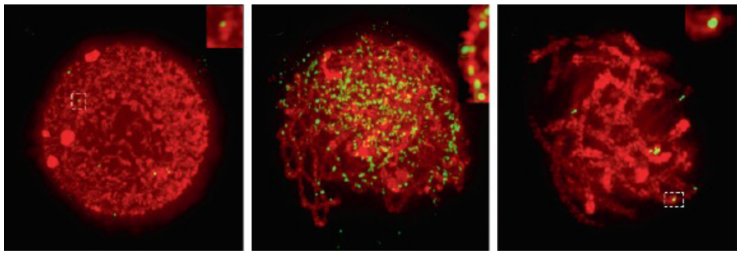

Fig. S1. Distribution of RAD51 foci in maize meiocytes at different stages of early prophase I. Red = DAPI-stained chromatin; green = RAD51. Small segments of chromosomes are shown at a higher magnification in insets in upper right corners of images.

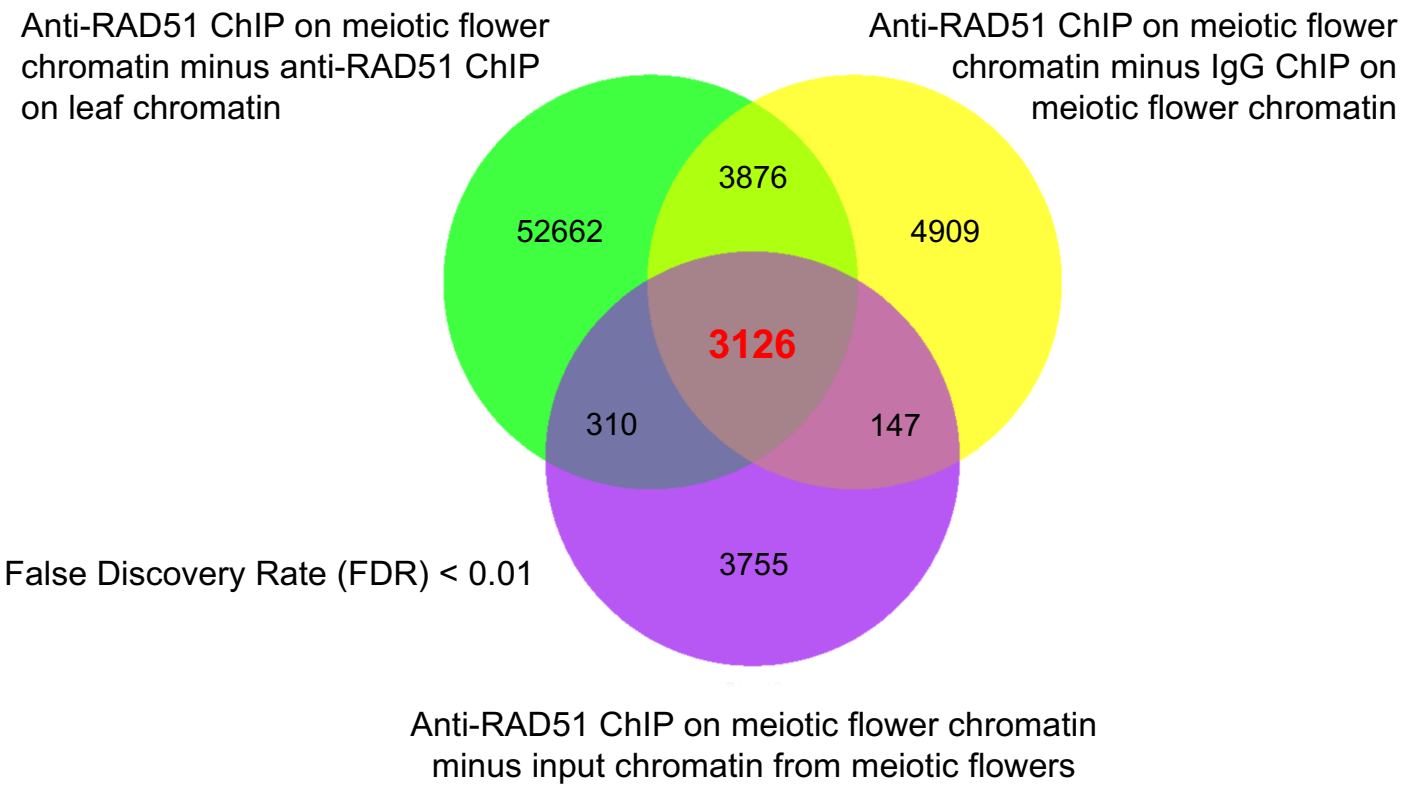

Fig. S2. A Venn diagram depicting the process of identification of DSB hotspots in maize using RAD51 ChIP seq data compared to three controls: (i) Illumina-sequenced input chromatin, (ii) ChIP products generated by using unspecific rabbit immunoglobulin instead of the anti-RAD51 antibody, and (iii) ChIP products generated using the anti-RAD51 antibody on leaf tissue chromatin. DSB hotspots are the 3,126 regions showing enrichments in anti-RAD51 antibody ChIP compared to the three controls.

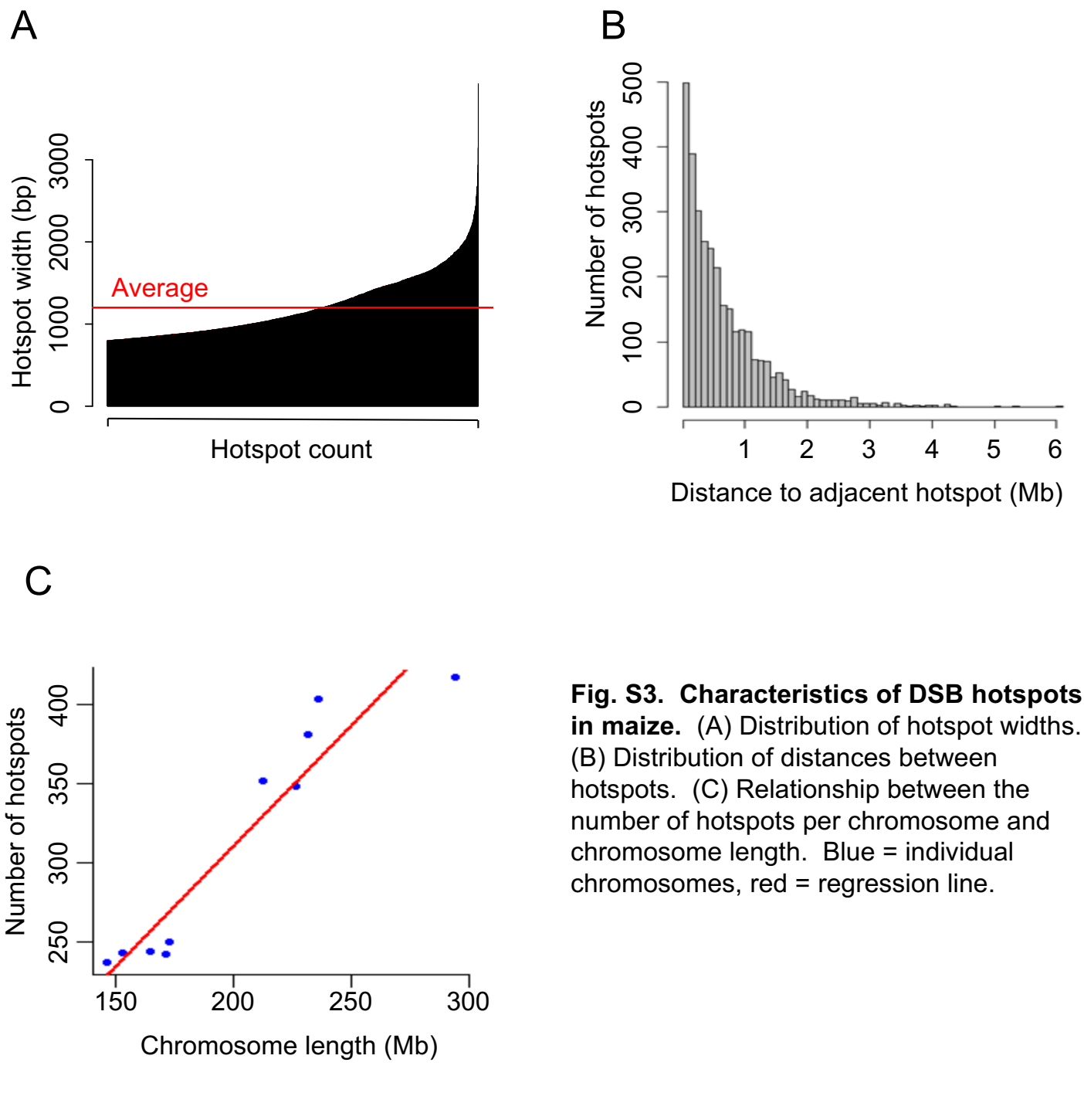

Fig. S3. Characteristics of DSB hotspots in maize. (A) Distribution of hotspot widths. (B) Distribution of distances between hotspots. (C) Relationship between the number of hotspots per chromosome and chromosome length. Blue = individual chromosomes, red = regression line.

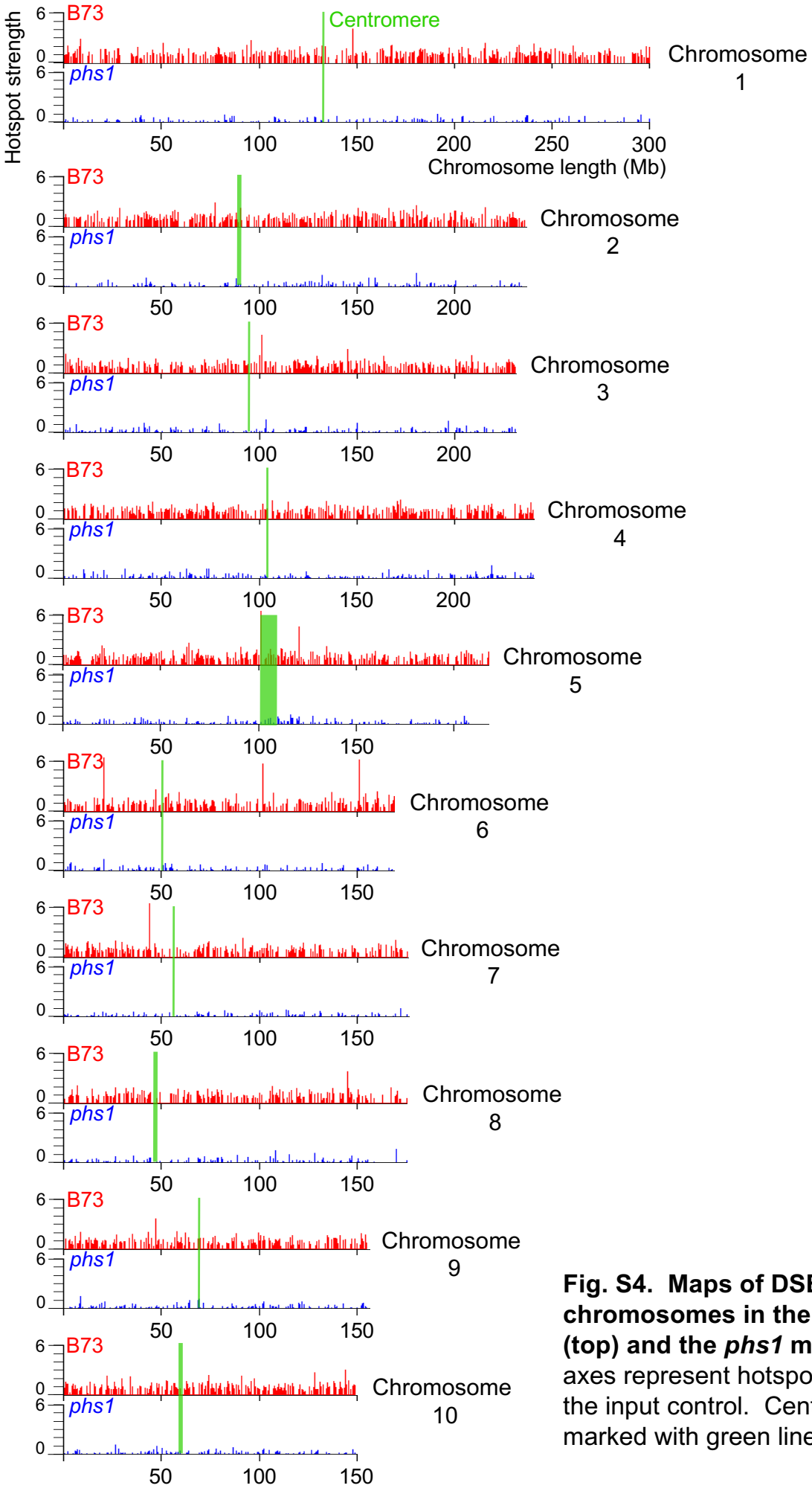

Fig. S4. Maps of DSB hotspots on maize chromosomes in the B73 inbred of maize (top) and the *phs1* mutant (bottom). The Y axes represent hotspot strength compared to the input control. Centromeres positions are marked with green lines.

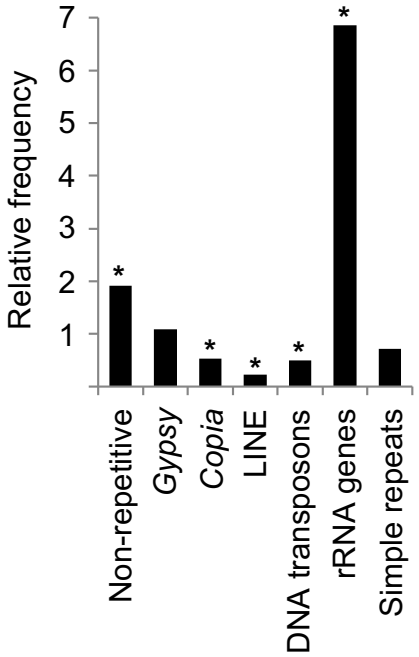

Fig. S5. Distribution of DSB hotspots in various genome components relative to random distribution generated using computer simulation (n = 10,000) that takes into account the relative content of each element in Illumina reads that align unambiguously to the assembled maize reference genome sequence. Note that the relative frequency of hotspots in rRNA genes is likely an overestimate as the number of rRNA gene repeats is underestimated in the reference genome.

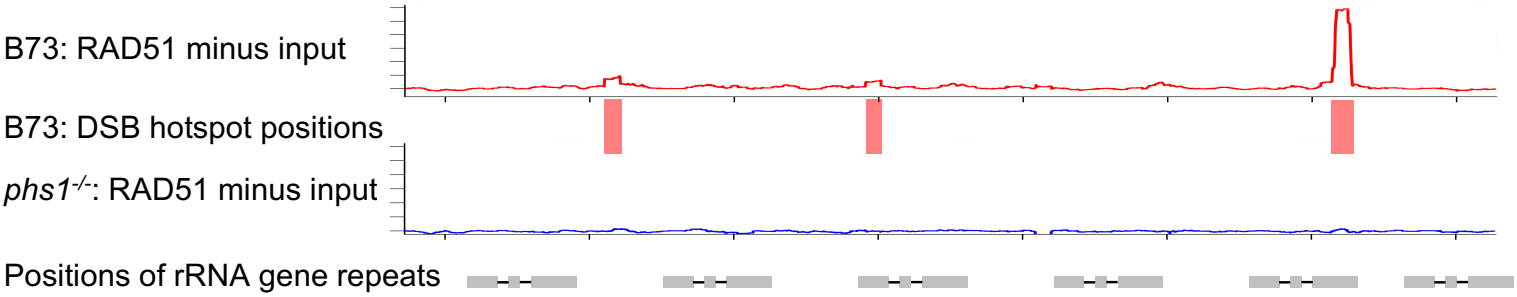

Fig. S6. A diagram of the 5S rRNA gene repeat region on maize chromosome 2 showing sites of DSB formation. RAD51 ChIP-seq reads from the B73 inbred and the *phs1* mutant are presented after subtracting chromatin input controls using a 1 kb sliding window. The Y axes represent hotspot strength.

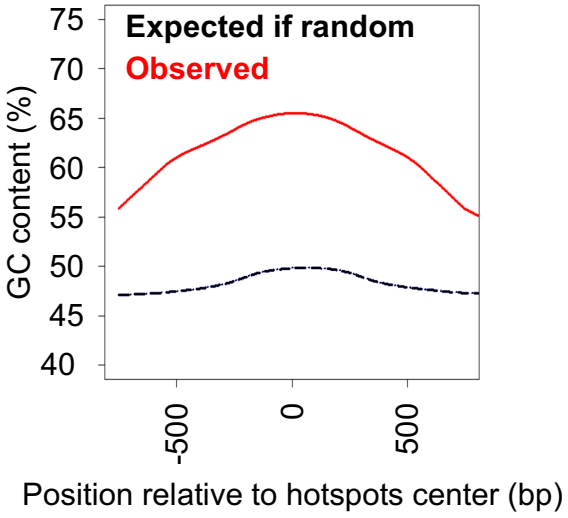

Fig. S7. GC content of the sites of DSB hotspots.

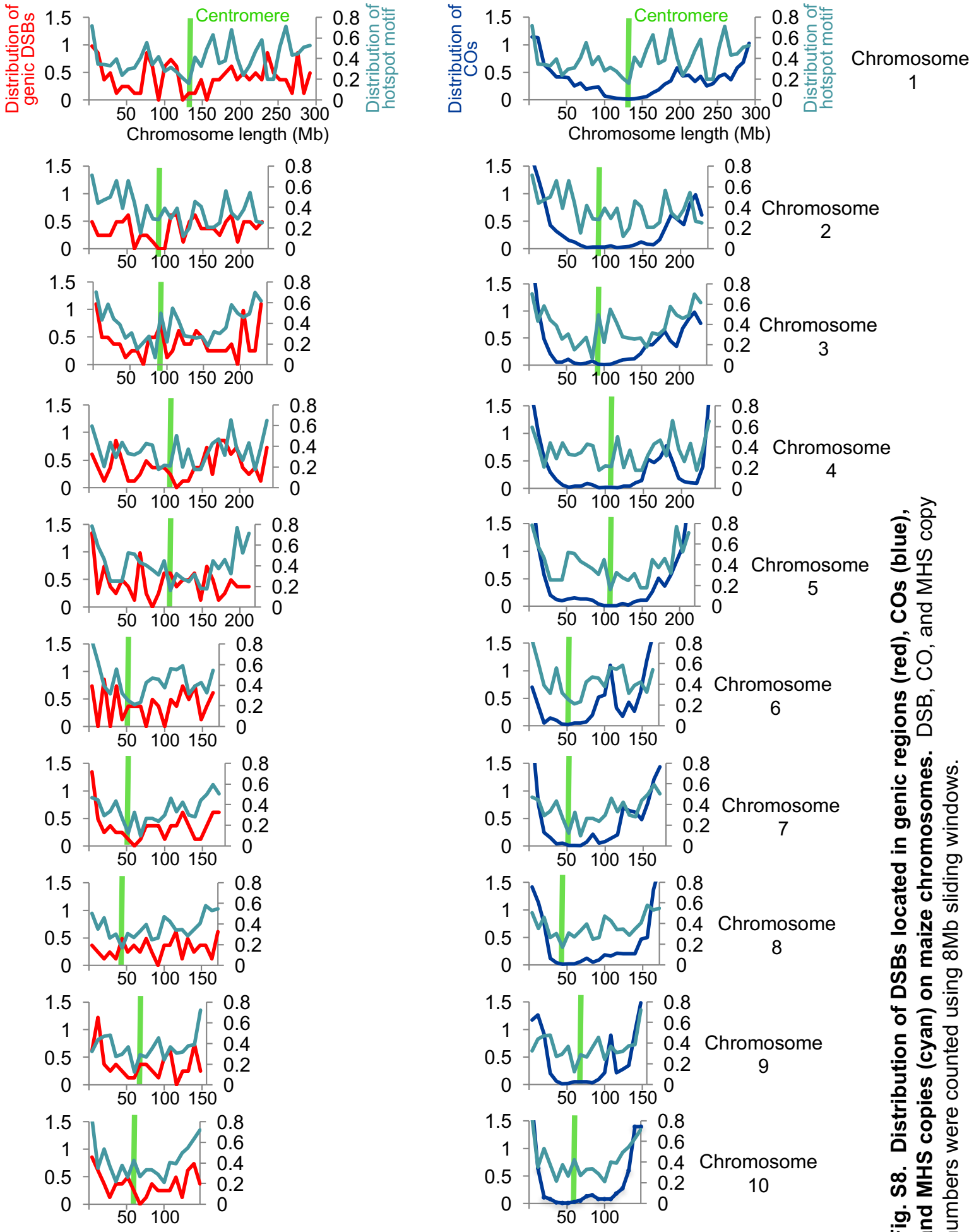

Fig. S8. Distribution of DSBs located in genic regions (red), COs (blue), and MHS copies (cyan) on maize chromosomes. DSB, CO, and MHS copy numbers were counted using 8Mb sliding windows.

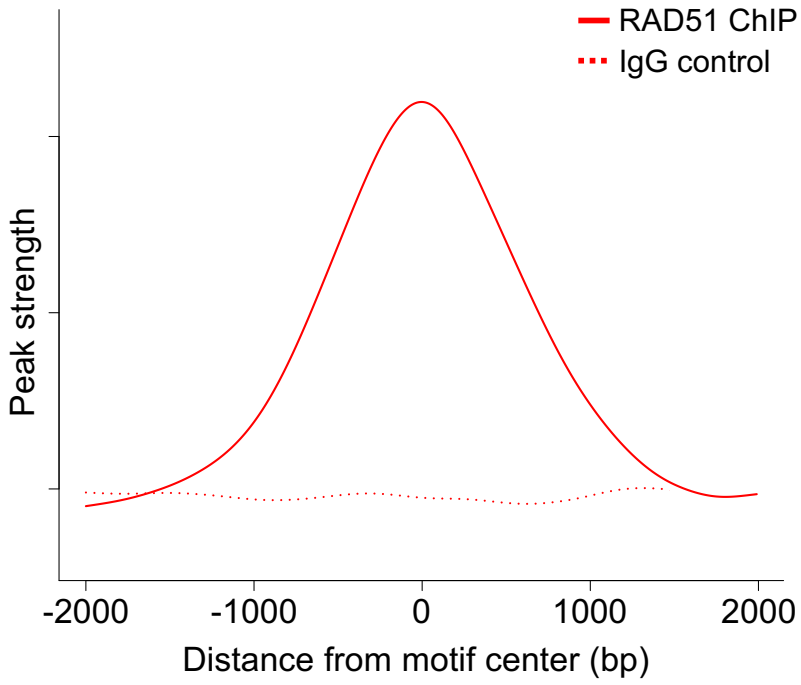

Fig. S9. Distribution of RAD51 ChIP and IgG control signal around MHS copies that are sites of DSB hotspots.

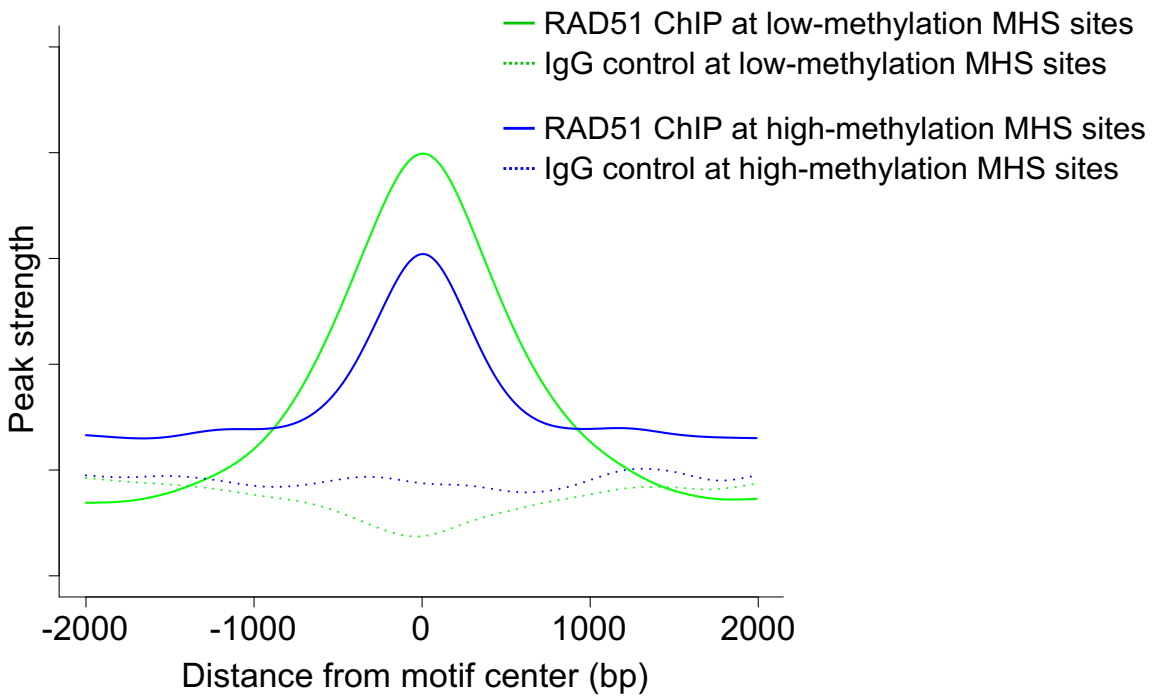

Fig. S10. Distribution of RAD51 ChIP and IgG control signal around MHS copies exhibiting varying levels of C methylation. Seventy-five % of methylated C residues was used as the threshold between high and low methylation.

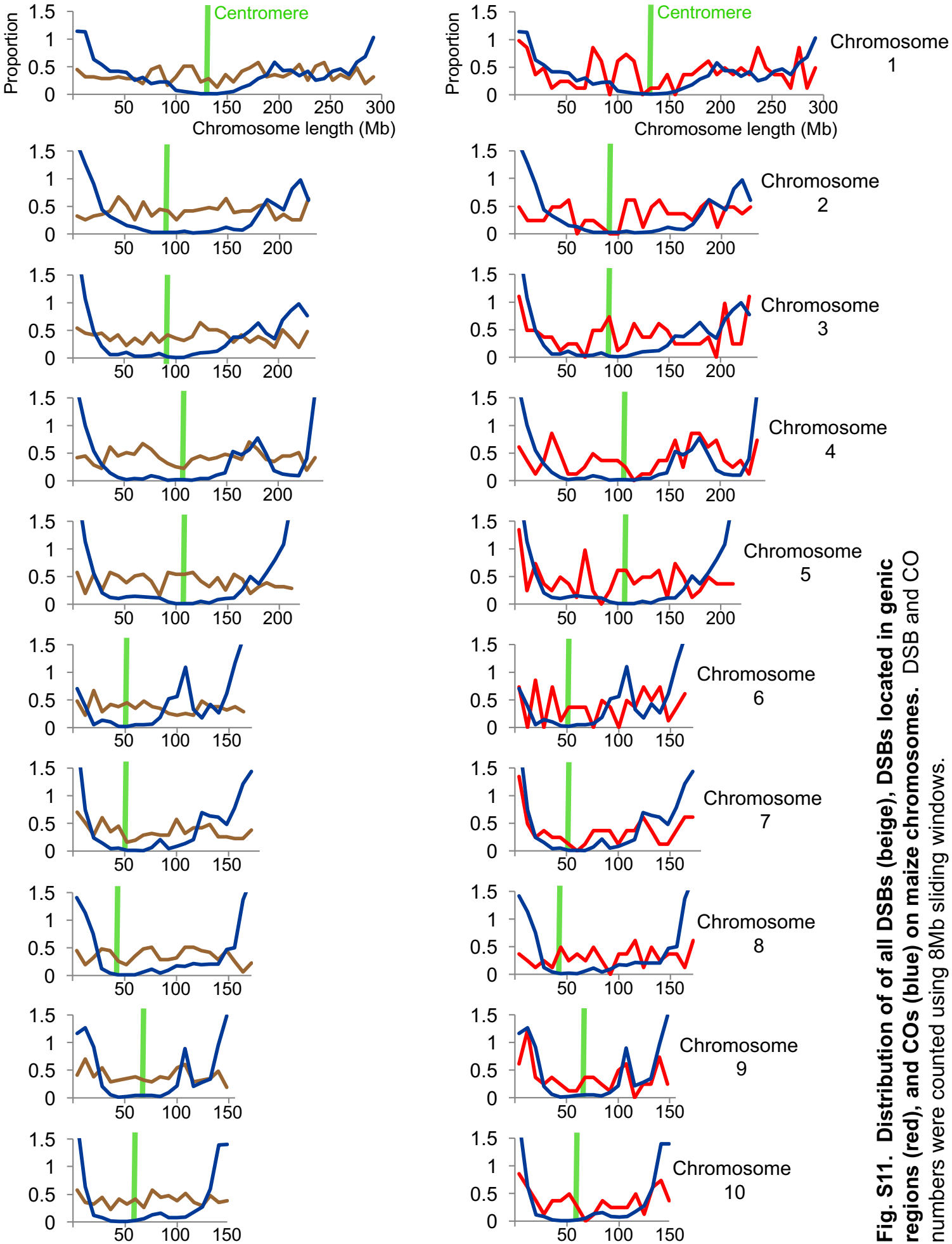

Fig. S11. Distribution of of all DSBs (beige), DSBs located in genic regions (red), and COs (blue) on maize chromosomes. DSB and CO numbers were counted using 8Mb sliding windows.

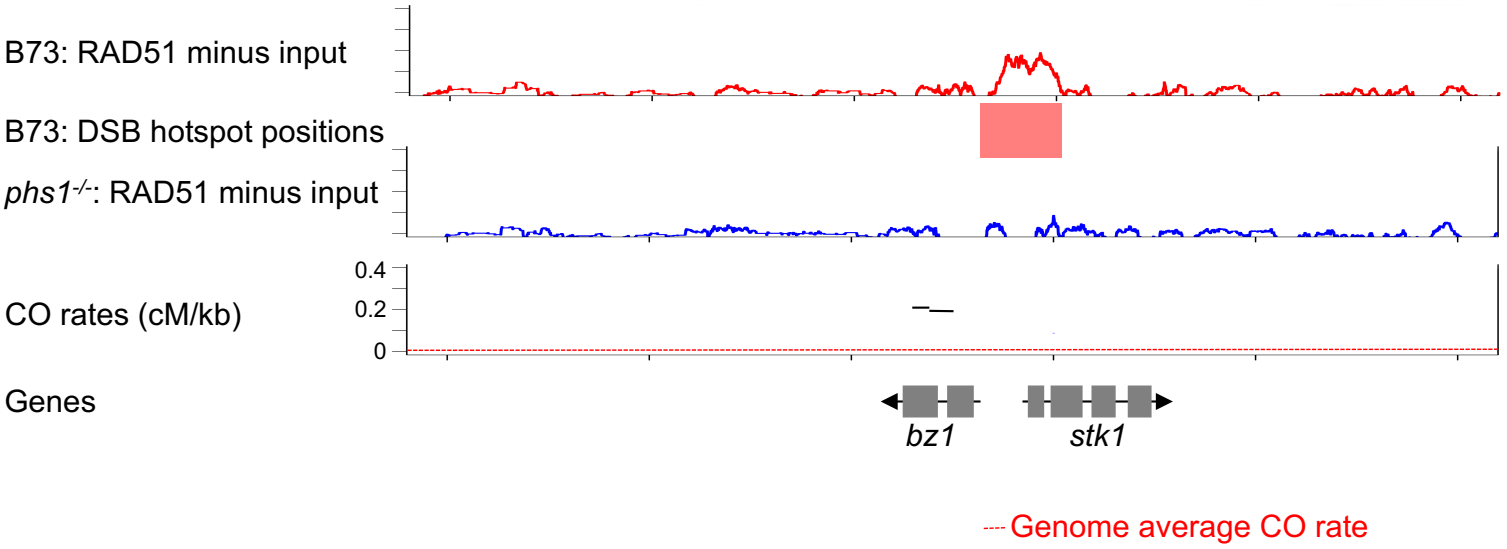

Fig. S12. DSB formation in the *bronze* region on maize chromosome 9. RAD51 ChIP seq reads from the B73 inbred and the *phs1* mutant are presented relative to chromatin input controls using a 1 kb sliding window. The Y axes represent hotspot strength. CO rates in the *bz1* gene (1, 2) are plotted in comparison to the genome average.

1. Dooner HK, *et al.* (1985) A molecular genetic analysis of insertion mutations in the *bronze* locus in maize. *Mol Gen Genet* 20):240-246.
2. Dooner HK & Martinez-Ferez IM (1997) Recombination occurs uniformly within the *bronze* gene, a meiotic recombination hotspot in the maize genome. *Plant Cell* 9:1633-1646.

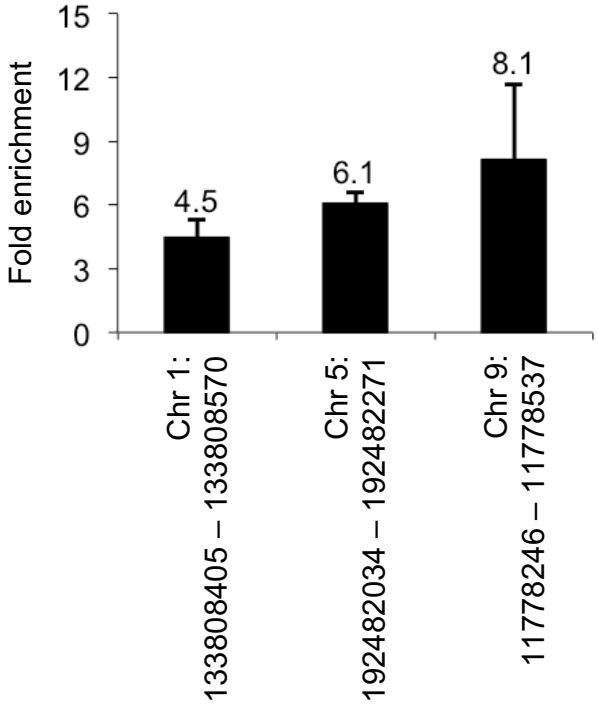

Fig. S13. Quantitative real-time PCR confirmation of enrichment of hotspot DNA by RAD51 ChIP at three hotspots. Enrichment values are ratios of immunoprecipitated to input DNA, were normalized using a region in the promoter of the *Ubiquitin10* gene, and are means of three independent experiments.