Crossover-active regions of the wheat genome are distinguished by DMC1, the chromosome axis, H3K27me3 and signatures of adaptation

Andrew J. Tock[1*,†], Daniel M. Holland[1,†], Wei Jiang[1], Kim Osman[2], Eugenio Sanchez-Moran[2], James D. Higgins[3], Keith J. Edwards[4], Cristobal Uauy[5], F. Chris H. Franklin[2] and Ian R. Henderson[1*]

[1] Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom
[2] School of Biosciences, University of Birmingham, Birmingham B15 2TT, United Kingdom
[3] Department of Genetics and Genome Biology, University of Leicester, Leicester LE1 7RH, United Kingdom
[4] School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, United Kingdom
[5] John Innes Centre, Norwich NR4 7UH, United Kingdom

[†] These authors contributed equally to this work.

* Corresponding authors: irh25@cam.ac.uk and ajt200@cam.ac.uk

CONTENTS:

SUPPLEMENTAL METHODS:

Plant material

Chinese Spring (CS42) wheat plants were grown on F2 compost in controlled-environment growth chambers under the following conditions: 60% humidity, long day settings (16 hours light/8 hours dark) with 400 μmol light intensity, and at 22 °C. For western blot analysis and immunoprecipitation of DMC1 and ASY1, immature pre-emergence spikes were harvested when the flag leaf was fully emerged, approximately 8 weeks after sowing (between Zadoks stages 41 and 43). Leaf tissue from approximately 5-week-old seedlings was harvested for micrococcal nuclease (MNase) digestion and immunoprecipitation of histone modifications. Anthers from cultivar Cadenza were isolated at the required stage of meiosis for immunocytology. Cadenza plants were grown in a controlled environment with long day settings (16 hours light/8 hours dark), at 20 °C, and with 60% relative humidity.

Chromatin immunoprecipitation and sequencing

Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) of DMC1 and ASY1 was performed as reported (Lambing et al. 2020). Approximately 12 grams of immature spikes were ground in liquid nitrogen. Nuclei were extracted in nuclei isolation buffer (1 M sucrose, 60 mM HEPES pH 8.0, 0.6% Triton X-100, 5 mM KCl, 5 mM $MgCl_2$, 5 mM EDTA, 0.4 mM PMSF, 1 mM pepstatin-A, 1×protease inhibitor cocktail), and crosslinked in 1% formaldehyde at room temperature for 25 minutes. The crosslinking reaction was quenched with 125 mM glycine and incubated at room temperature for a further 25 minutes. The nuclei were purified from cellular debris via two rounds of filtration through one layer of Miracloth and centrifuged at 2,500*g* for 25 minutes at 4 °C. The nuclei pellet was resuspended in EB2 buffer (0.25 M sucrose, 1% Triton X-100, 10 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 1 mM EDTA, 5 mM DTT, 0.1 mM PMSF, 1 mM pepstatin-A, 1×protease inhibitor cocktail) and centrifuged at 14,000*g* for 10 minutes at 4 °C.

The nuclei pellet was resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 1% SDS, 10 mM EDTA, 0.1 mM PMSF, 1 mM pepstatin-A) and chromatin was sonicated using a Covaris E220*evolution* with the following settings: power=150V, cycles per burst=200, duty factor=20%, time=180 seconds. Sonicated chromatin was centrifuged at 14,000*g* and the supernatant was extracted and diluted with 1×volume of ChIP dilution buffer (1.1% Triton X-100, 20 mM Tris-HCl pH 8.0, 167 mM NaCl, 1.1 mM EDTA, 1mM pepstatin-A, 1×protease inhibitor cocktail). The chromatin was incubated overnight at 4 °C with either α-DMC1 (rabbit) or α-ASY1 (guinea pig) antibody pre-bound to Protein A or G magnetic beads (Dynabeads, Thermo Fisher) respectively, at a concentration of 1:200, alongside a no-antibody control. We used polyclonal antibodies that were raised against *Arabidopsis* DMC1 in rabbit using a multiple antigenic peptide based on amino acid residues 17–35 of AtDMC1 (Sanchez-Moran et al. 2007), and recombinant wheat ASY1 HORMA domain in guinea pig (Desjardins et al. 2020). The beads were collected on a magnetic rack and washed twice with low-salt wash buffer (150 mM NaCl, 0.1% SDS, 1% Triton X-100, 20 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.4 mM PMSF, 1 mM pepstatin-A, 1×protease inhibitor cocktail) and twice with high-salt wash buffer (500 mM NaCl, 0.1% SDS, 1% Triton X-100, 20 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.4 mM PMSF, 1 mM pepstatin-A, 1×protease inhibitor cocktail). Immunoprecipitated DNA–protein complexes were eluted from the beads with elution buffer (1% SDS, 0.1 M $NaHCO_3$) at 65 °C for 15 minutes. Samples were reverse crosslinked by incubating with 0.24 M NaCl at 65 °C overnight. Proteins were digested with Proteinase K treatment and DNA was purified with phenol:chloroform:isoamyl alcohol (25:24:1) extraction and ethanol precipitation.

Chromatin immunoprecipitation of the histone modifications H3K4me3, H3K9me2 and H3K27me1 was performed using a protocol similar to that described for DMC1 and ASY1, but

using MNase digestion for chromatin fragmentation. MNase digestion was also applied without immunoprecipitation, to generate a library for high-throughput sequencing (MNase-seq). Briefly, nuclei were isolated from 2 grams of leaf tissue and crosslinked in 1% formaldehyde solution. Chromatin was digested with 50 Kunitz units/mL of micrococcal nuclease (MNase, NEB, M0247S) at 37 °C for 10 minutes with occasional vortexing. Digestion was halted by incubating on ice and adding EDTA to a final concentration of 20 mM. An equal amount of nuclei lysis buffer (10 mM Tris pH 8.0, 0.2% SDS, 2% Triton X-100, 0.2% sodium deoxycholate and 1×complete protease inhibitors) was added. Samples were dialyzed by rotating tubes for 2 hours at 4 °C. Dialyzed samples were centrifuged at 14,000*g* for 10 minutes, and the supernatant containing fragmented chromatin was extracted. For MNase-seq, the sample was reverse crosslinked by incubating with 0.25 M NaCl at 65 °C overnight. This was followed by Proteinase K treatment, phenol:chloroform:isoamyl alcohol (25:24:1) extraction and ethanol precipitation. Mononucleosomal fragments were isolated by excising an ~147 bp band from an agarose gel, and 10 ng of gel-extracted DNA was used for MNase-seq library preparation. For immunoprecipitation of histone modifications, 5 µg of antibody for H3K4me3 (ab8580, Abcam), H3K9me2 (ab1220, Abcam) or H3K27me1 (07-448, Sigma-Aldrich) was added to Protein G magnetic beads (Dynabeads, 10003D, Thermo Fisher). Fragmented chromatin was incubated with antibody-bound beads overnight at 4 °C on a rotating wheel. Beads were washed twice in low-salt buffer, once in high-salt buffer, and once in LiCl buffer. Immunoprecipitated DNA–protein complexes were eluted from the beads with elution buffer (1% SDS, 0.1 M $NaHCO_3$) and reverse crosslinked by incubating with 0.25 M NaCl at 65 °C overnight. Proteins were digested with Proteinase K treatment and DNA was purified with phenol:chloroform:isoamyl alcohol (25:24:1) extraction and ethanol precipitation. Mononucleosomal fragments were isolated by excising an ~147-bp band from an agarose gel, and 10 ng of gel-extracted DNA was used for ChIP-seq library preparation. Each ChIP-seq and MNase-seq library was generated using a TruSeq Nano DNA library preparation kit (Illumina) according to the manufacturer's instructions and sequenced for 75-cycles (paired-end) on an Illumina NextSeq 500 instrument.

Western blotting

Total protein extracts and isolated nuclei extracts from immature pre-emergence spikes and seedling leaf tissue were boiled in 2×SDS loading buffer and run on a 4–12% NuPAGE Bis-Tris gel (Invitrogen). Separated proteins were transferred to a PVDF membrane in transfer buffer (25 mM Tris pH 8.3, 192 mM glycine, 20% v/v methanol) overnight at 30 mV at 4 °C. After transfer, the membrane was immersed briefly in TBS-T (20 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.05% Triton X-100) and blocked for 1 hour at room temperature in blocking solution (5% w/v dried milk powder). The blocked membrane was washed 3 times with TBS-T and incubated with the following primary antibodies, diluted in blocking solution, at the following concentrations: α-ASY1 (guinea pig, 1:2000) and α-DMC1 (rabbit, 1:2000). Membranes were incubated with the primary antibody overnight at 4 °C with gentle agitation. Membranes were washed 4 times with TBS-T and then incubated with secondary antibodies, diluted in blocking solution, at the following concentrations: goat α-guinea pig HRP 1:15000 (ab6908, Abcam) and goat α-rabbit HRP 1:15000 (sc-2004, Santa Cruz Biotechnology). The membrane was washed 6 times with TBS-T, detected on film with Amersham ECL Prime Western Blotting Detection Reagent, using an Xograph Compact X4 film processor.

Meiotic immunocytology

For immunolocalization, slides were prepared as described (Armstrong et al. 2002), with the following modifications. Approximately 20 anthers were digested in 20 µL enzyme mix (0.4% cytohelicase, 1.5% sucrose, 1% polyvinylpyrrolidone) in a cavity slide inside a humidified chamber at 37 °C. After 4 minutes anthers were gently crushed to release pollen mother cells (PMCs), anther debris was quickly removed with a needle and digestion continued for a further 3 minutes.

Primary antibodies were used at the following dilutions: α-AtASY1 rabbit or guinea pig, 1:500 (Armstrong et al. 2002); α-AtDMC1 rabbit, 1:200 (Sanchez-Moran et al. 2007); and α-H3K27me3 rabbit, according to the manufacturer's guidelines (C15310069 [CS-069-100], Diagenode). For combined immunofluorescence and fluorescence *in situ* hybridization (FISH) of telomeric repeat sequences, slides were first prepared as for immunolocalization and the primary antibody applied as above. After incubation and washing to remove unbound serum, the primary antibody was blocked with a secondary antibody-biotin conjugate at 1:100 in 1% bovine serum albumen (BSA), in 1×phosphate buffered saline, 0.1% Triton X100 (PBST). Slides were incubated for 30 minutes at 37 °C, washed 3 times with PBST and an *Arabidopsis* telomeric-repeat FISH probe labeled with digoxigenin was applied, as described (Armstrong 2013). Secondary antibodies were FITC (green) and Cy3 or Alexa Fluor 594 (red) conjugates (Sigma; Thermo Fisher). DNA was stained with 5 μg/mL 4´,6´-diamidino-2-phenylindole (DAPI) in Vectashield (Vector Labs). Fluorescence microscopy was carried out using a Nikon Eclipse 90i microscope fitted with a Nikon DS-Qi1Mc camera. NIS Elements software (Nikon) was used to capture images as Z-stacks with a 0.2 μM step and are presented as Z-projections unless otherwise stated. All material was prepared from the wheat cultivar Cadenza.

ChIP-seq and MNase-seq data alignment and processing

Workflows for processing and aligning next-generation sequencing (NGS) reads were developed using Snakemake v3.13.3 (Köster and Rahmann 2012). ChIP-seq and MNase-seq paired-end reads were deduplicated using Dedupe from BBMap v38.22 (Segata et al. 2016), followed by removal of adapter sequences, low-quality bases (Phred+33-scaled quality<20) and reads shorter than 30 bases after trimming using Cutadapt v1.18 (Martin 2011). Deduplicated and trimmed reads were aligned to the RefSeq v1.0 Chinese Spring reference genome assembly (IWGSC 2018) using Bowtie 2 v2.3.4.3 (Langmead and Salzberg 2012), with the following settings: --very-sensitive --no-mixed --no-discordant -k 4. Up to 4 valid alignments were reported for each read pair. Read pairs that aligned equally well to more than one genomic location, without a best-scoring alignment (with Bowtie 2-assigned MAPQ<2), were discarded using Samtools v1.9 (Li et al. 2009). For retained read pairs that aligned to multiple locations, with varying alignment scores, the best-scoring alignment was selected. Alignments with MAPQ≥23 were identified to determine the number of retained alignments that correspond to read pairs that align to only one genomic location. Alignments with more than 6 mismatches or consisting of only one read in a pair were discarded. Read filtering and alignments are summarized in Supplemental Table S1. Using the retained unique and multiple alignments combined in BAM format, read counts per million mapped reads (CPM) in 1-bp and 1-Mb adjacent genomic windows were generated in bigWig and bedGraph formats with the bamCoverage tool from deepTools v3.1.3 (Ramírez et al. 2014). We specified the --extendReads option for read extension to the fragment size defined by the aligned read mates. Reads that aligned to the unassigned chromosomes comprising unanchored scaffolds were excluded from this coverage normalization procedure.

Distal enrichment of DMC1 and ASY1 ChIP-seq coverage was robustly observed when: (i) the alignment MAPQ threshold was varied to retain only unique alignments (MAPQ≥23), or best-scoring multiple alignments combined with unique alignments (MAPQ≥2), or multiple alignments combined with unique alignments (MAPQ≥0; with one alignment randomly selected for read pairs that aligned equally well to more than one location); or (ii) the permitted number of alignment mismatches was varied (Supplemental Figs. S4–S8). Furthermore, coverage for a Chinese Spring chromatin input sequencing library (accession SRR6350669) is more evenly distributed along each chromosome, with less variation between genomic compartments (Supplemental Fig. S9) (IWGSC 2018). Together, this demonstrates that the patterns of DMC1 and ASY1 ChIP-seq enrichment we observe are robust to variations in alignment and filtering parameters.

We note that the proportions of trimmed read pairs with a best-scoring multiple alignment or a unique alignment (with Bowtie 2-assigned MAPQ≥2) that we obtained for DMC1 and ASY1 ChIP-seq are comparable to those observed for published ChIP-seq and input data sets (range=81–88%) (Supplemental Table S1). For example, DMC1 ChIP-seq yielded 185,448,819 trimmed read pairs that aligned with MAPQ≥2 ('Primary alignments') (Supplemental Table S1), constituting 85% of the total trimmed read pairs (218,747,974), and 88% of the trimmed read pairs that aligned with MAPQ≥0 (209,866,855). Similarly, for ASY1 ChIP-seq, 344,590,350 trimmed read pairs aligned with MAPQ≥2, representing 85% of the total trimmed read pairs (405,615,455), and 89% of the trimmed read pairs that aligned with MAPQ≥0 (388,559,032).

In terms of enrichment, we note that the $\log_2$ fold enrichment values we observe for DMC1 and ASY1 ChIP-seq are comparable to those for published histone modification data sets that have well-established distal patterns, such as H3K4me3 and H3K27me3 (Li et al. 2019). The distal enrichment observed for DMC1 and ASY1 is comparable in magnitude to differences in histone modification enrichment along the chromosomes.

Bisulfite sequencing data alignment and processing

Published paired-end (2×125 bp) bisulfite sequencing (BS-seq) reads derived from leaves of 2-week-old Chinese Spring plants were retrieved from the European Nucleotide Archive (accession SRR6792678) (IWGSC 2018). Sequencing adapters, low-quality bases (Phred+33-scaled quality<20) and 10 bases from the 5′ end of each read were removed using Trim Galore v0.6.4 (Krueger 2016). Trimmed reads were aligned to the RefSeq v1.0 Chinese Spring assembly (IWGSC 2018) using Bismark v0.20.0 (Krueger and Andrews 2011). Similar to the ChIP-seq and MNase-seq read alignment approach described above, Bismark discards read pairs that align equally well to more than one location, and selects the best alignment for each of the retained read pairs, which represented 62% of trimmed read pairs. Duplicate alignments, representing 19% of the aligned read pairs with a best hit, were discarded. Methylated cytosine calls in CG, CHG and CHH sequence contexts were extracted and context-specific DNA methylation proportions were generated in bedGraph and bigWig formats using the bismark2bedGraph and UCSC bedGraphToBigWig tools.

RNA-seq data alignment and processing

Published paired-end (2×125 bp) RNA-seq libraries derived from Chinese Spring anthers collected at the leptotene–zygotene transition during early meiotic prophase were retrieved from the European Nucleotide Archive (accessions ERR2402972–ERR2402974) (Martín et al. 2018). Ribosomal sequences were removed using BBDuk from BBMap v38.22 (Segata et al. 2016). Adapter sequences, low-quality bases (Phred+33-scaled quality<3 at the beginning and end of each read, and average quality<20 in 4-base sliding windows) and reads shorter than 80 bases after trimming were removed using Trimmomatic v0.38 (Bolger et al. 2014). Trimmed reads were aligned to the RefSeq v1.0 Chinese Spring genome using HISAT2 v2.1.0 (Kim et al. 2015), with the following settings: --no-mixed --no-discordant --no-unal --rna-strandness RF -k 5, and providing splice sites derived from the IWGSC RefSeq v1.1 Chinese Spring annotation of high-confidence protein-coding gene models. Ambiguously aligned read pairs with HISAT2-assigned MAPQ<1 were discarded using Samtools v1.9 (Li et al. 2009). For retained read pairs with multiple valid alignments, the best alignment was selected where available, otherwise an alignment was randomly selected. Uniquely aligned read pairs were extracted by removing alignments with MAPQ<60. Alignments with more than 6 mismatches or consisting of only one read in a pair were discarded. Using combined unique and multiple alignments, per-base transcripts per million (TPM) values were generated in bigWig format with the bamCoverage tool from deepTools v3.1.3 (Ramírez et al. 2014). Reads that aligned to the unassigned chromosome were excluded from this normalization procedure.

Chromosome-scale profiling of chromatin and recombination data

Files containing library-size-normalized ChIP-seq and MNase-seq coverage and DNA methylation proportions in bedGraph format were imported into R statistical computing software v3.5.0 to generate chromosome-scale profiles (1-Mb adjacent genomic windows) to be plotted and for calculating Spearman's rank-order correlation coefficients (r_s). Coverage profiles for a ChIP input sequencing library generated from Chinese Spring (accession SRR6350669) (IWGSC 2018) and for the MNase-seq library generated in the current study were used in conjunction with those for ChIP-seq libraries derived from chromatin fragmented by sonication, or by MNase digestion, respectively, to calculate $\log_2$([ChIP+1]/[control+1]) coverage ratios in 1-Mb adjacent genomic windows. In the context of ChIP-seq, an "input" control sample consists of chromatin that has been crosslinked and sonicated but not immunoprecipitated with an antibody. Chromosome-scale profiles were smoothed by applying a moving average (period=15 1-Mb windows). We used previously defined coordinates delimiting distal (R1 and R3), interstitial (R2a and R2b), proximal (C) and centromeric regions (IWGSC 2018).

Crossover rates (cM/Mb) derived from a Chinese Spring×Renan genetic map, in which each marker has a physical position in the RefSeq v1.0 Chinese Spring assembly, were also analyzed at the chromosome scale. This genetic map was previously constructed using genotyping data for a recombinant inbred population of 430 single-seed descent F_6 individuals at 47,154 filtered polymorphic markers (IWGSC 2018). The genetic map has a mean per-chromosome marker density of 3.22 markers/Mb, and a mean per-chromosome crossover rate of 0.23 cM/Mb. For plotting crossover rates along each chromosome, we used previously calculated mean crossover rates in 10-Mb sliding windows with a 1-Mb step (IWGSC 2018). We also computed mean crossover rates and feature frequencies in 1-Mb adjacent genomic windows to enable calculation of Spearman's correlation coefficients with other data sets profiled in this way. Genes used in all analyses are a representative set of the IWGSC RefSeq v1.1 Chinese Spring annotation of high-confidence protein-coding gene models. We used the NLR-encoding gene annotation (Steuernagel et al. 2020), and genes annotated with the 'defense response' Gene Ontology (GO) term in 'OntologiesForGenes.rds' (available at https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-Gonzalez_etal_2018-06025-Transcriptome-Landscape/data/TablesForExploration/), after retaining only rows containing 'IWGSC+Stress' in the 'ontology' column of this table (Ramírez-González et al. 2018). Elements within transposon superfamilies and subfamilies are those in the IWGSC RefSeq v1.0 Chinese Spring annotation. Correlation coefficients derived from larger numbers of data points are less likely to be exactly zero, the null hypothesis of the attendant significance test. In view of the large number of genomic windows analyzed, *P*-values for r_s correlation coefficients were standardized to represent those based on pairwise values across 100 windows, as described (Good 1982). To illustrate genome-scale trends for each data set, one hundred proportionally scaled windows (100ths) were defined along each chromosome arm orientated from telomere (*TEL*) to centromere (*CEN*), and chromosome profiles in 1-Mb adjacent windows were used to calculate per-scaled-window mean values across all chromosome arms. *TEL–CEN* coverage, cM/Mb and feature frequency profiles were smoothed by applying a moving average (period=3 scaled windows) and plotted.

DMC1 and ASY1 ChIP-seq peak analysis

Sites of local ChIP-seq enrichment were identified by calling peaks using the ranger tool within PeakRanger v1.18 (Feng et al. 2011), providing alignments for a ChIP input sequencing library (accession SRR6350669) (IWGSC 2018) as a control for background. DMC1 and ASY1 peaks were defined by applying *P*-value and false discovery rate (FDR) thresholds of 0.001 and 0.01, with the --pad option enabled and read extension lengths of 200 and 180 nucleotides,

respectively, specified based on the mean fragment length for each library. Peaks in MNase-seq coverage were identified in the same way, but specifying a read extension length of 140 nucleotides. ChIP-seq peaks for histone modifications were defined by applying *P*-value and FDR thresholds of 0.05. Specifically, H3K9ac, H3K27me3, H3K36me3 (IWGSC 2018) H3K4me1 and H3K27ac (Li et al. 2019) ChIP-seq peaks were defined with read extension lengths of 205, 170, 200, 200 and 210 nucleotides, respectively. Alignments for the MNase-seq library were provided as a background control for the definition of peaks in H3K4me3, H3K9me2 and H3K27me1 ChIP-seq libraries, which were derived from chromatin fragmented by MNase digestion, specifying read extension lengths of 135, 135 and 120 nucleotides, respectively. Any overlapping peaks within each data set were merged.

DMC1 and ASY1 ChIP-seq peaks within distal compartments (R1 and R3) and, separately, interstitial and proximal compartments (R2 and C), were evaluated for overlap with other genomic features using permutation tests implemented in the Bioconductor package regioneR v1.6.2 (Gel et al. 2016). Briefly, 10,000 sets of randomly positioned loci of the same number and width distribution, and in the same genomic compartments, as the given DMC1 or ASY1 peak set were defined. For each set, the number of random loci overlapping genomic features within a category of interest was compared with the observed number of peaks overlapping these features. These permuted loci overlap counts were used to calculate the expected number of overlaps (mean number of feature-overlapping random loci from 10,000 sets) and an empirical *P*-value. The corresponding $\log_2$(observed/expected) number of overlaps and significance threshold derived from counts of permuted loci overlaps were plotted as bar graphs. Peak overlap analyses included gene promoter regions up to 1 kb upstream of transcriptional start sites (TSSs), 5′ ends up to 500 bp downstream of TSSs, 3′ ends up to 500 bp upstream of transcriptional termination sites (TTSs), as well as downstream regions up to 1 kb from TTSs. NLR-encoding genes (Steuernagel et al. 2020), genes with putative functions in meiosis (wheat orthologs of known meiotic genes and genes with meiosis-related GO annotations) (Alabdullah et al. 2019), and elements within each transposon superfamily (IWGSC RefSeq v1.0 Chinese Spring annotation) were also analyzed for peak overlap. Compartmentalized DMC1 and ASY1 ChIP-seq peaks in each subgenome were ordered by decreasing $-\log_{10}$(ranger-assigned FDR), and the top 10,000 peaks in each set were evaluated for over-representation of DNA sequence motifs using Weeder v2.0 (Zambelli et al. 2014). Background oligonucleotide frequency files for Weeder v2.0 analyses were generated based on the sequences of randomly positioned loci within the same subgenomic compartments as the peak sets to be analyzed.

Fine-scale profiling of chromatin and recombination data

Profiles around DMC1 and ASY1 ChIP-seq peaks and genes were calculated for ChIP-seq, MNase-seq, BS-seq and RNA-seq data sets by providing the above-described bigWig files to deepTools v3.1.3 (Ramírez et al. 2014) computeMatrix in 'scale-regions' mode. Each feature was divided into non-overlapping, proportionally scaled windows between start and end coordinates, and 2-kb flanking regions were divided into 20-bp windows. Mean values for each data set were calculated within each window, generating a matrix of profiles in which each row represents a feature with 2-kb flanking regions and each column a window. Coverage profiles for a ChIP input sequencing library (accession SRR6350669) (IWGSC 2018), and for the MNase-seq library generated in the current study, were used in conjunction with those for ChIP-seq libraries derived from chromatin fragmented by sonication or by MNase digestion, respectively, to calculate windowed $\log_2$([ChIP+1]/[control+1]) coverage ratios for each feature.

Average frequencies of SNPs and elements within transposon superfamilies were calculated within proportionally scaled windows between feature start and end coordinates, and within 20-bp windows in 2-kb flanking regions, using the normalizeToMatrix function ('w0' method) from the

Bioconductor package EnrichedHeatmap v1.12.0 (Gu et al. 2018). We used published SNPs identified among a panel of 811 geographically diverse bread wheat cultivars and landraces via an exome sequence capture assay (3,039,822 filtered SNPs) (He et al. 2019). SNPs were annotated with their predicted impact on protein function using SnpEff v4.3t (Cingolani et al. 2012), after building a SnpEff database for the IWGSC RefSeq v1.0 Chinese Spring genome assembly with the RefSeq v1.1 gene annotation. For each window within each gene and its 2-kb flanking regions, the binary logarithm of the ratio of SNPs predicted by SnpEff to have a high impact on protein function (potentially deleterious SNPs) to synonymous SNPs ($\log_2$[(dSNP+1)/(sSNP+1)]) was calculated. Equivalent profiles for each data set were calculated around randomly positioned loci of the same number and width distribution in each subgenome.

Meta-profiles were calculated around features grouped by decreasing crossover rate (cM/Mb) derived from the Chinese Spring×Renan genetic map. Specifically, feature boundaries were extended by 1 kb on each side, and mean cM/Mb values were computed within these intervals using previously calculated mean crossover rates in 10-Mb sliding windows with a 1-Mb step (IWGSC 2018). Genes or distally located (R1 and R3) DMC1 or ASY1 ChIP-seq peaks were divided into four groups corresponding to those in the 100th–75th (Quantile 1), 75th–50th (Quantile 2), 50th–25th (Quantile 3) and 25th–0th (Quantile 4) percentiles with regard to their mean cM/Mb values, such that Quantile 1 features have the highest crossover rates. Genes were also divided into four quantiles according to decreasing mean DMC1, ASY1 or H3K27me3 ChIP-seq signal ($\log_2$[(ChIP+1)/(input+1)]), within 1-kb-extended boundaries. Due to a smaller range of cM/Mb values in the crossover-suppressed regions, DMC1 or ASY1 peaks in interstitial and proximal compartments (R2 and C) were divided into three groups corresponding to those in the 100th–66th (Quantile 1), 66th–33rd (Quantile 2) and 33rd–0th (Quantile 3) percentiles for mean cM/Mb. To define control groups for comparison, peaks or genes were randomly allocated to groups, each with the same number of features per chromosome as its correspondingly numbered quantile. Randomly positioned loci were allocated to groups in the same way. Meta-profiles (windowed means and 95% confidence intervals) for each group of peaks, genes or random loci were calculated and plotted using the above-described feature profiles in R version 3.5.0.

Functional annotation composition of gene quantiles

Gene quantiles defined by decreasing crossover rate (cM/Mb), or DMC1, ASY1 or H3K27me3 ChIP-seq signal (across all subgenomes, or within each subgenome) were analyzed for enrichment of 'biological process' Gene Ontology (GO) and high-level GO (GO slim) terms, relative to their representation among all genes in the genome, or the corresponding subgenome, using the Bioconductor package topGO v2.32.0 (Alexa and Rahnenführer 2007) GO and GO slim annotations were obtained from 'OntologiesForGenes.rds' (available at https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-Gonzalez_etal_2018-06025-Transcriptome-Landscape/data/TablesForExploration/), after retaining only rows containing 'IWGSC+Stress' or 'slim IWGSC+Stress', respectively, in the 'ontology' column of this table (Ramírez-González et al. 2018). The default algorithm implemented in topGO ('weight01') was used to identify over-represented GO terms, with Fisher's exact test *P*-values≤0.05.

Gene functional and expression categories were evaluated for over- and under-representation in each gene quantile by applying hypergeometric tests. For analysis of over-representation, a *P*-value derived from a hypergeometric test indicates the probability that the quantile contains more than or equal to the observed number of within-quantile genes in a given category, whereas for under-representation, it denotes the probability that the quantile contains fewer than or equal to this number. We also sampled from the hypergeometric distribution 100,000 times to obtain a probability distribution for plotting the $\log_2$(observed/expected) ratio and significance threshold (α=0.05). Genes previously assigned to different homoeolog expression variation categories,

based on a combined analysis of RNA-seq data from 15 tissue types in Chinese Spring, were retrieved from 'Triads.rds' and 'TriadMovement.rds' (available at https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-Gonzalez_etal_2018-06025-Transcriptome-Landscape/data/TablesForExploration) (Ramírez-González et al. 2018).

Fine-scale analysis of genes encoding nucleotide-binding and leucine-rich repeat (NLR) proteins

Genes encoding NLR immune receptor proteins (Steuernagel et al. 2020), across all subgenomes or within each subgenome, were grouped into four quantiles as described above for all genes, either by decreasing crossover rate (cM/Mb) or by decreasing size of physical clustering. In the latter ranking, Quantile 4 comprises *NLR* singletons only. *NLR* genes were considered clustered where the intervening genomic sequence contained no more than eight genes encoding non-NLRs. Meta-profiles (windowed means and 95% confidence intervals) for each *NLR* gene quantile were calculated and plotted, as described above for all genes. Crossover rate (cM/Mb) quantiles were evaluated for representation of physically clustered and singleton *NLR* genes, using the hypergeometric distribution as described above. Quantiles defined by decreasing crossover rate or cluster size were analyzed for representation of *NLR* genes that were up-regulated 30 minutes after PAMP treatment (with the fungal cell wall component chitin or the bacterial flagellin peptide flg22) (Steuernagel et al. 2020), also based on the hypergeometric distribution. Physical cluster membership of each clustered *NLR* gene was mapped onto the wheat *NLR* phylogeny (Steuernagel et al. 2020). A conservative estimate of monophyletic physical clusters was obtained by identifying clusters whose most recent common ancestor is the parent node of at least one of the *NLR* genes within their respective physical cluster, using the R package phylobase v0.8.10 (Hackathon et al. 2013). *NLR* genes within quantiles defined by decreasing crossover rate were evaluated for representation of those that are members of such monophyletic physical clusters, based on the hypergeometric distribution.

Population genetics analysis

Published exome-sequencing-derived SNPs identified in ten geographically distinct populations of bread wheat cultivars and landraces (sampled from Africa, Middle East, Asia, the Former Soviet Union, Eastern Europe, Western Europe, North America, Central America, South America and Oceania), as reported (He et al. 2019), were analyzed using the R package PopGenome v2.7.5 to compute population genetics statistics for each gene in each regional population (Pfeifer et al. 2014). This included calculation of per-gene R_M, which gives the minimum number of crossover events in the history of a sample (Hudson and Kaplan 1985), SNP frequency, and nucleotide diversity (the mean number of intrapopulation pairwise nucleotide differences per gene, π), each divided by gene width. To analyze genes for evidence of selection within each population, per-gene Tajima's *D* (Tajima 1989) and composite likelihood ratios (CLRs) (Nielsen et al. 2005) were calculated. We divided genes across all subgenomes into quantiles or randomized groups as described above, but with quantiles defined by decreasing intrapopulation width-normalized R_M, DMC1 or ASY1 ChIP-seq signal, or MNase-seq signal. Mean values with 95% confidence intervals for each gene group were calculated and plotted for each of the population genetics statistics described above.

SUPPLEMENTAL REFERENCES:

Alabdullah AK, Borrill P, Martin AC, Ramirez-Gonzalez RH, Hassani-Pak K, Uauy C, Shaw P, Moore G. 2019. A Co-Expression Network in Hexaploid Wheat Reveals Mostly Balanced Expression and Lack of Significant Gene Loss of Homeologous Meiotic Genes Upon Polyploidization. *Front Plant Sci* **10**: 1325.

Alexa A, Rahnenführer J. 2007. Gene set enrichment analysis with topGO. *Bioconductor Improv*.

Armstrong S. 2013. A Time Course for the Analysis of Meiotic Progression in Arabidopsis thaliana. *Methods Mol Biol* **990**: 119–123.

Armstrong SJ, Caryl AP, Jones GH, Franklin FCH. 2002. Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in Arabidopsis and Brassica. *J Cell Sci* **115**: 3645–55.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–20.

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, et al. 2014. Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**: 1249721.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**: 80–92.

Desjardins SD, Ogle DE, Ayoub MA, Heckmann S, Henderson IR, Edwards KJ, Higgins JD. 2020. MutS homologue 4 and MutS homologue 5 Maintain the Obligate Crossover in Wheat Despite Stepwise Gene Loss following Polyploidization. *Plant Physiol* **183**: 1545–1558.

Feng X, Grossman R, Stein L. 2011. PeakRanger: a cloud-enabled peak caller for ChIP-seq data. *BMC Bioinformatics* **12**: 139.

Gel B, Díez-Villanueva A, Serra E, Buschbeck M, Peinado MA, Malinverni R. 2016. regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* **32**: 289–91.

Good IJ. 1982. Standardized tail-area probabilities. *J Stat Comput Simul*.

Gu Z, Eils R, Schlesner M, Ishaque N. 2018. EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. *BMC Genomics* **19**: 234.

Guo X, Su H, Shi Q, Fu S, Wang J, Zhang X, Hu Z, Han F. 2016. De Novo Centromere Formation and Centromeric Sequence Expansion in Wheat and its Wide Hybrids. *PLoS Genet* **12**: e1005997.

Hackathon AR, Cowan P, Vienne D De, Eddelbuettel D, Jombart T, Kembel S, Michon- F, Orme D, Meara BO, Paradis E, et al. 2013. Package “phylobase.” *CRAN*.

He F, Pasam R, Shi F, Kant S, Keeble-Gagnere G, Kay P, Forrest K, Fritz A, Hucl P, Wiebe K, et al. 2019. Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nat Genet* **51**: 896–904.

Hudson RR, Kaplan NL. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**: 147–164.

IWGSC. 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science (80-)* **361**: eaar7191.

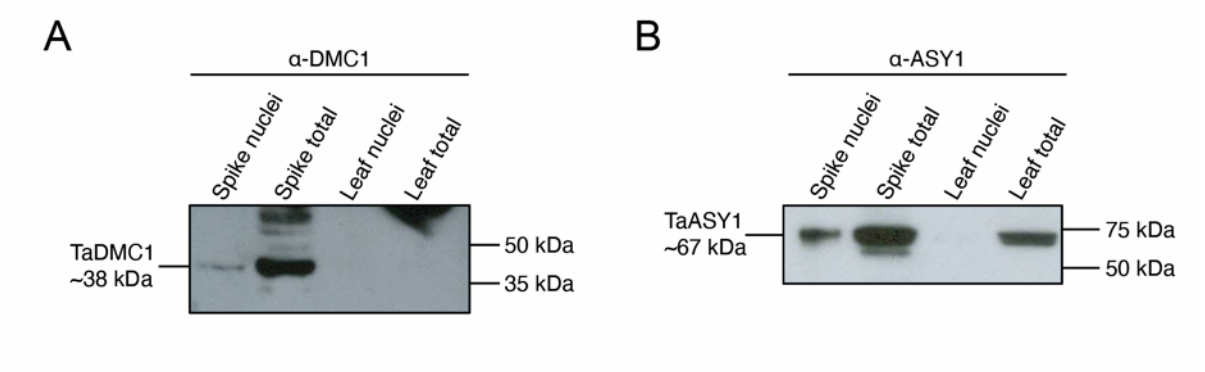
Kim D, Langmead B, Salzberg SL. 2015. HISAT: A fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357–360.

Köster J, Rahmann S. 2012. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* **19**: 2520–2522.

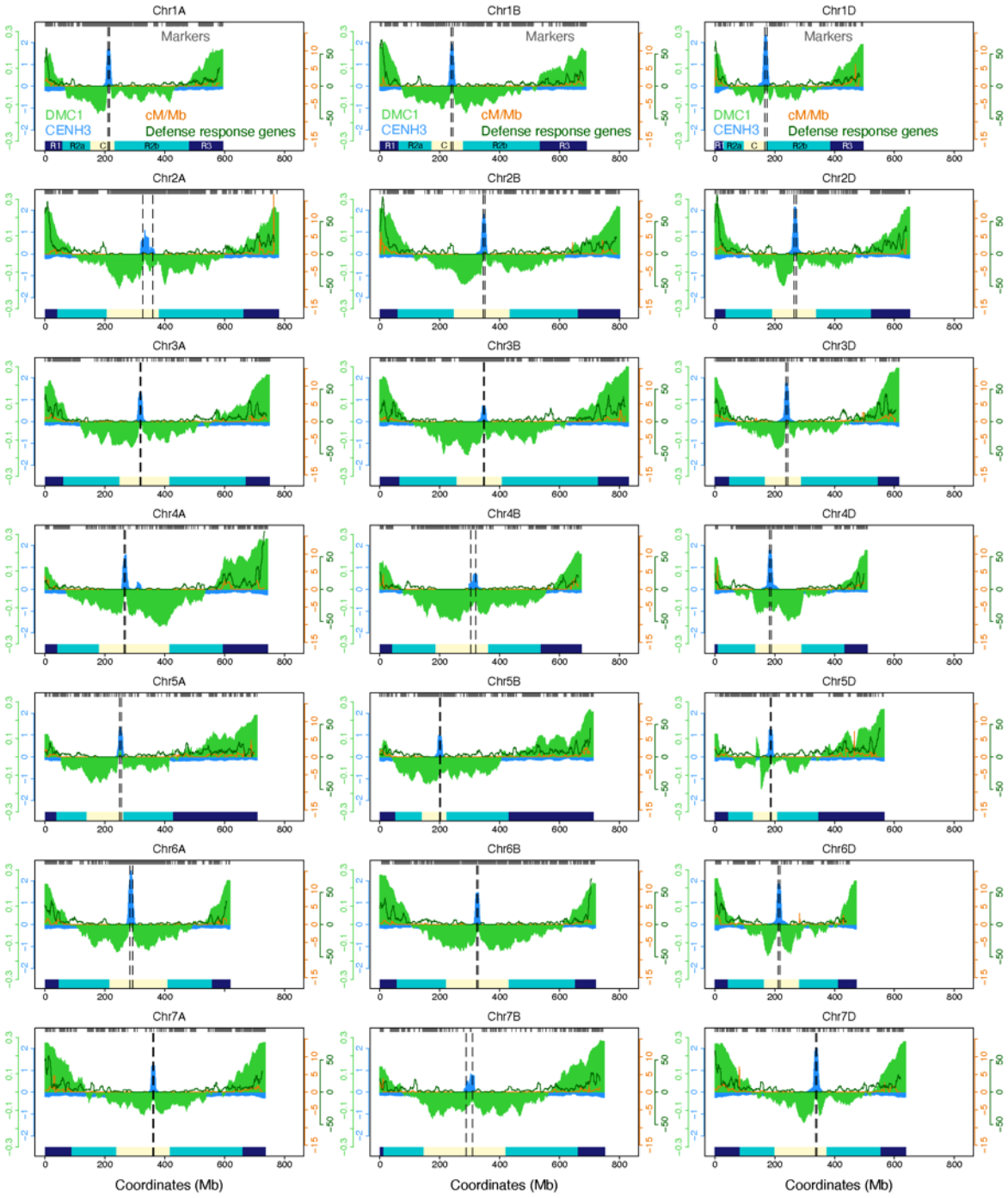
Krueger F. 2016. Trim Galore. *Babraham Bioinforma*.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571–2.
Lambing C, Choi K, Blackwell AR, Henderson IR. 2020. Chromatin Immunoprecipitation of Meiotically Expressed Proteins from Arabidopsis thaliana Flowers. *Methods Mol Biol* **2061**: 219–236.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–9.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–9.
Li Z, Wang M, Lin K, Xie Y, Guo J, Ye L, Zhuang Y, Teng W, Ran X, Tong Y, et al. 2019. The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements. *Genome Biol* **20**: 139.
Martín AC, Borrill P, Higgins J, Alabdullah A, Ramírez-González RH, Swarbreck D, Uauy C, Shaw P, Moore G. 2018. Genome-wide transcription during early wheat meiosis is independent of synapsis, ploidy level, and the Ph1 locus. *Front Plant Sci* **9**: 1791.
Martin M. 2011. Cutadapt. *EMBnet.journal*.
Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. 2005. Genomic scans for selective sweeps using SNP data. *Genome Res* **15**: 1566–1575.
Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. *Mol Biol Evol* **31**: 1929–1936.
Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, Van Ex F, Pasha A, et al. 2018. The transcriptional landscape of polyploid wheat. *Science (80-)* **361**: eaar6089.
Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187–W191.
Sanchez-Moran E, Santos J-L, Jones GH, Franklin FCH. 2007. ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. *Genes Dev* **21**: 2220–33.
Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C, Ormerod KL, Wood DL, Lachner N, et al. 2016. BBMap short-read aligner, and other bioinformatics tools. *Bioinformatics*.
Steuernagel B, Witek K, Krattinger SG, Ramirez-Gonzalez RH, Schoonbeek HJ, Yu G, Baggs E, Witek AI, Yadav I, Krasileva K V., et al. 2020. The NLR-annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiol* **183**: 468–482.
Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
Zambelli F, Pesole G, Pavesi G. 2014. Using Weeder, Pscan, and PscanChIP for the discovery of enriched transcription factor binding Site motifs in nucleotide Sequences. *Curr Protoc Bioinforma* **47**: 2.11.1-31.

SUPPLEMENTAL FIGURES:

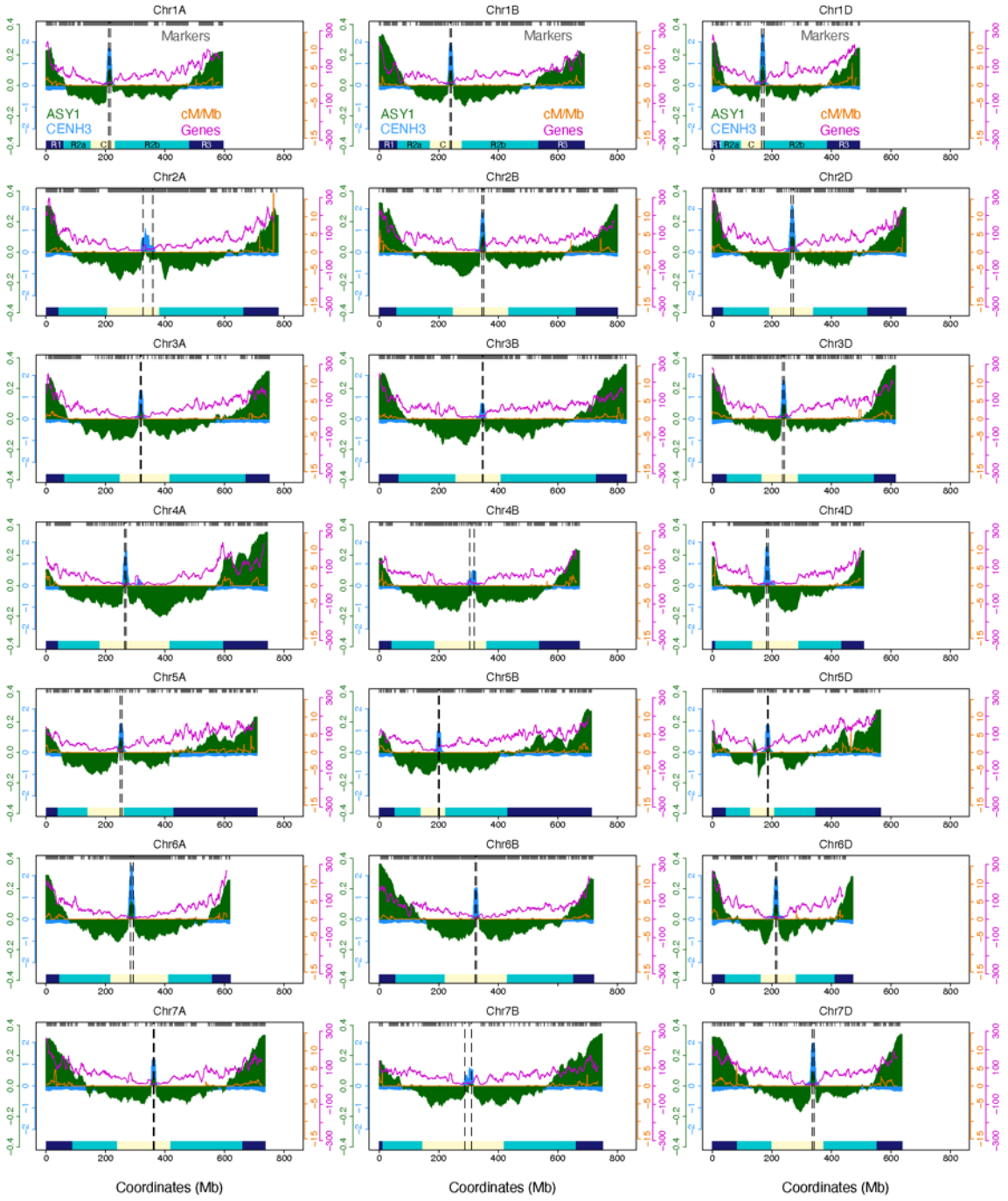


Supplemental Figure S1. Validation of α-DMC1 and α-ASY1 antibodies via western blotting. **A.** Western blot analysis using α-DMC1 (rabbit) (Sanchez-Moran et al. 2007) and protein extracts from immature pre-emergence spikes from Chinese Spring wheat. **B.** Western blot analysis using α-ASY1 (guinea pig) (Desjardins et al. 2020) and protein extracts from immature pre-emergence spikes from Chinese Spring wheat.



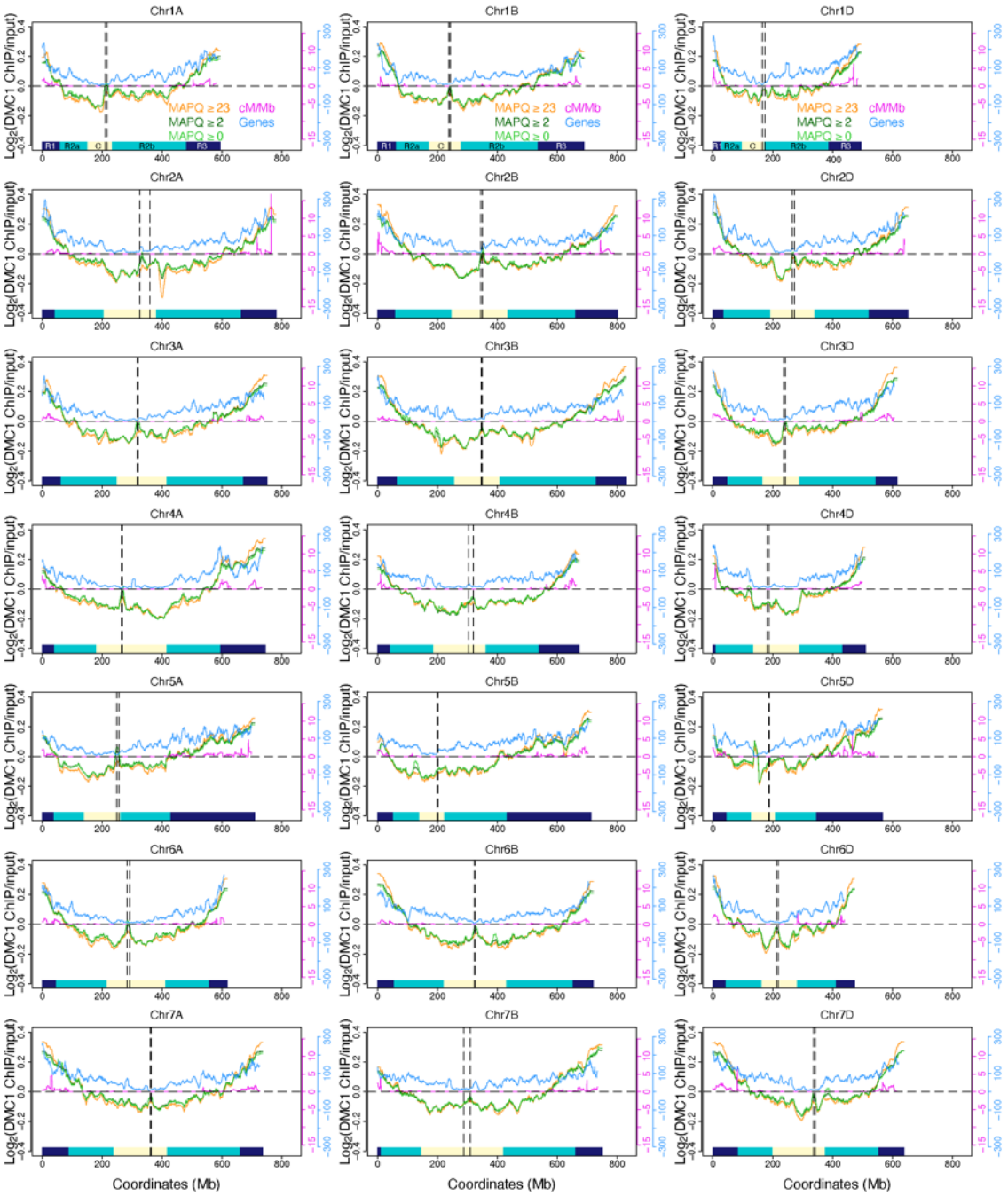
Supplemental Figure S2. Genome-wide landscapes of the meiotic recombinase DMC1, CENH3, crossover rate and defense-response genes. Chromosome profiles of DMC1 (green, $\log_2$[ChIP/input]) and CENH3 (light blue, $\log_2$[ChIP/input]) (Guo et al. 2016) ChIP-seq in 1-Mb adjacent windows, smoothed by applying a moving average (period=15 1-Mb windows). ChIP-seq chromosome profiles (shading) are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018)) and gene frequencies in 10-Mb sliding windows with a 1-Mb step (lines). Gene frequency profiles correspond to genes annotated with the 'defense response' Gene Ontology (GO) term (Ramírez-González et al. 2018). Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise

boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018). Gray ticks denote genomic coordinates of genetic markers used to construct the Chinese Spring×Renan genetic map.



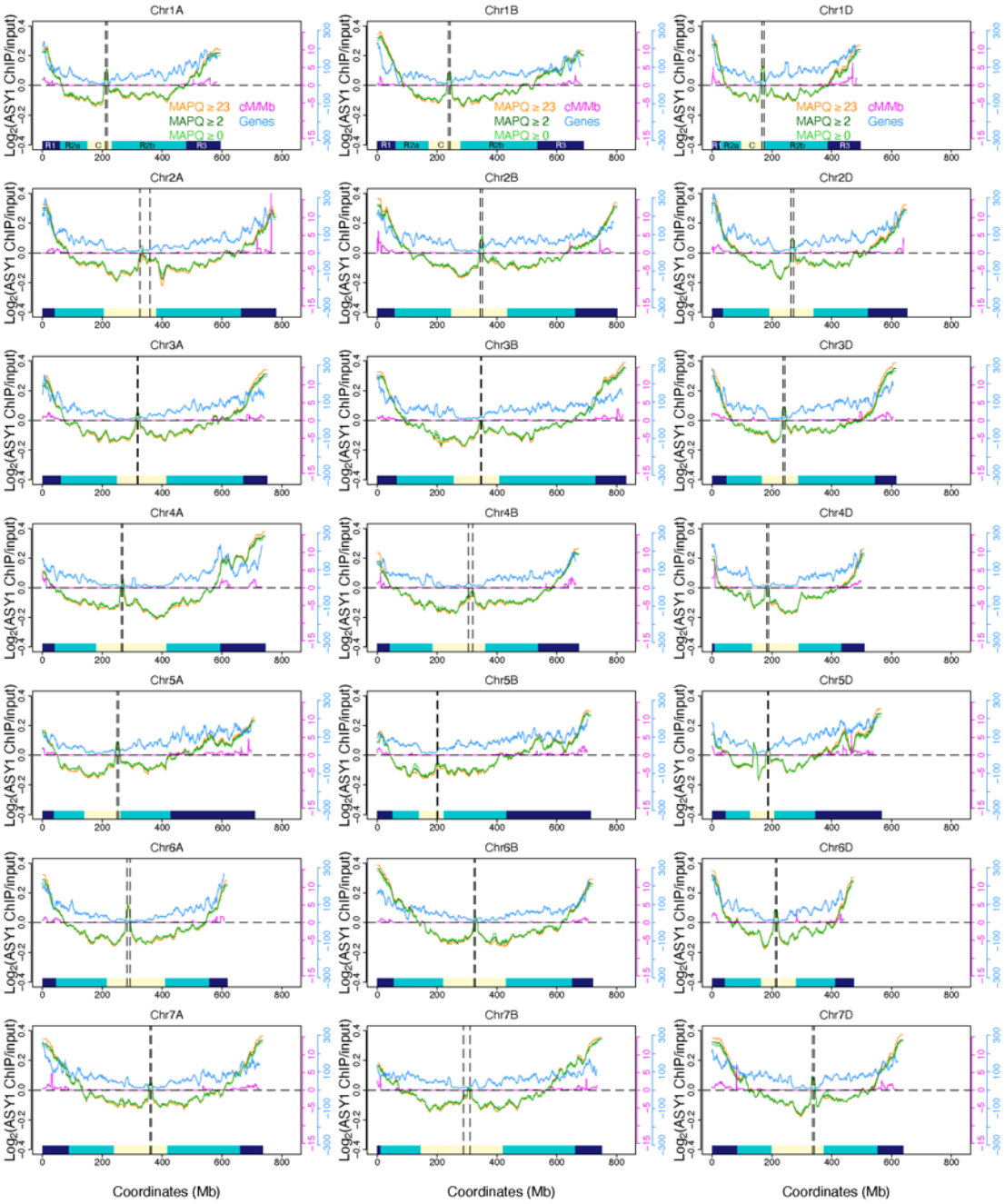
Supplemental Figure S3. Genome-wide landscapes of the meiotic axis protein ASY1, CENH3, crossover rate and genes. Chromosome profiles of ASY1 (dark green, $\log_2$[ChIP/input]) and CENH3 (light blue, $\log_2$[ChIP/input]) (Guo et al. 2016) ChIP-seq in 1-Mb adjacent windows, smoothed by applying a moving average (period=15 1-Mb windows). ChIP-seq chromosome profiles (shading) are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018)) and gene frequencies in 10-Mb sliding windows with a 1-Mb step (lines). Gene frequency profiles correspond to a representative set of the IWGSC RefSeq v1.1 Chinese Spring annotation of high-confidence protein-coding gene models (IWGSC 2018). Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines)

regions are indicated along the x-axis (IWGSC 2018). Gray ticks denote genomic coordinates of genetic markers used to construct the Chinese Spring×Renan genetic map.



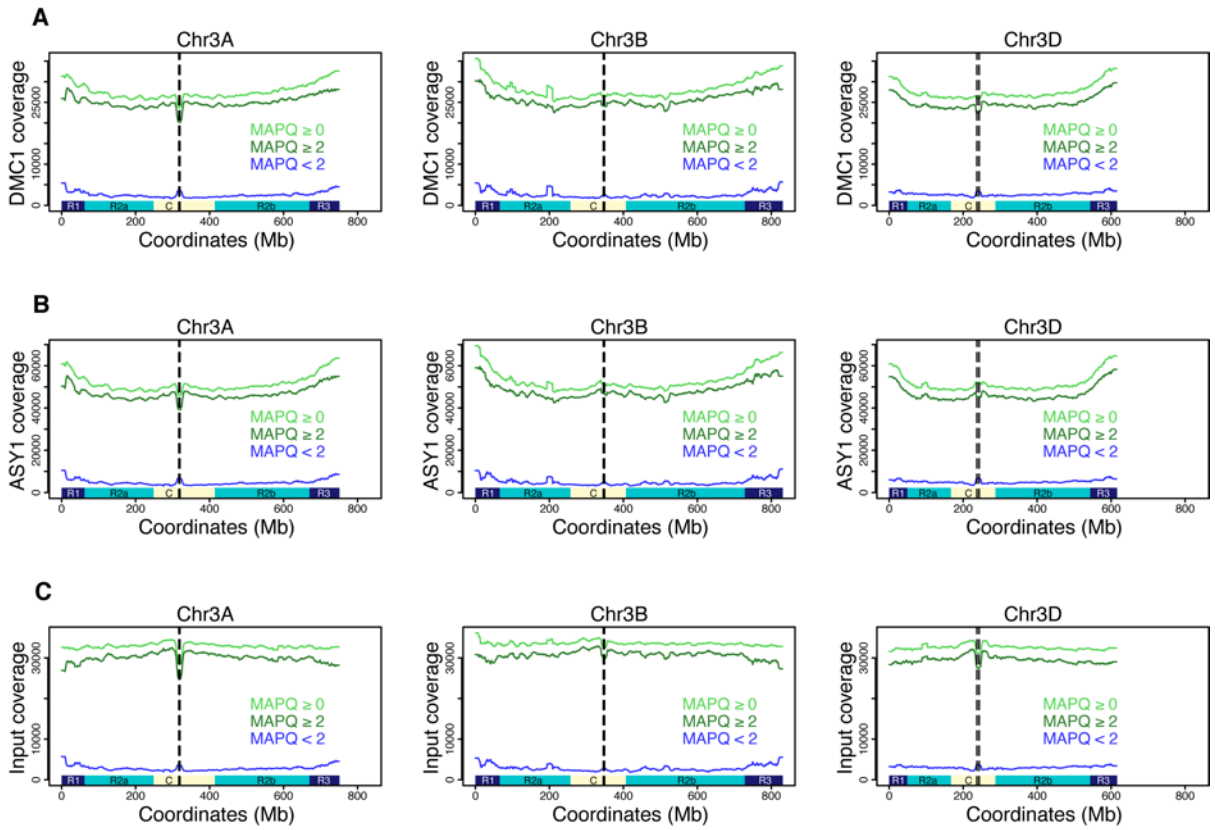
Supplemental Figure S4. Genome-wide landscape of DMC1 ChIP-seq enrichment after varying the alignment MAPQ score threshold. Chromosome profiles of DMC1 ChIP-seq in 1-Mb adjacent windows (log_2[ChIP/input], using a Chinese Spring chromatin input sequencing library with accession SRR6350669 (IWGSC 2018)), smoothed by applying a moving average (period=15 1-Mb windows). Bowtie 2-assigned MAPQ≥23 (orange) corresponds to unique alignments, MAPQ≥2 (dark green) corresponds to best-scoring multiple alignments combined with unique alignments, and MAPQ≥0 (light green) corresponds to multiple alignments (with one alignment randomly selected for read pairs that aligned equally well to more than one location) combined with unique alignments. Alignments containing more than 6 mismatches were discarded. ChIP-seq chromosome profiles are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018), pink) and gene frequencies (light blue) in

10-Mb sliding windows with a 1-Mb step. Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018).

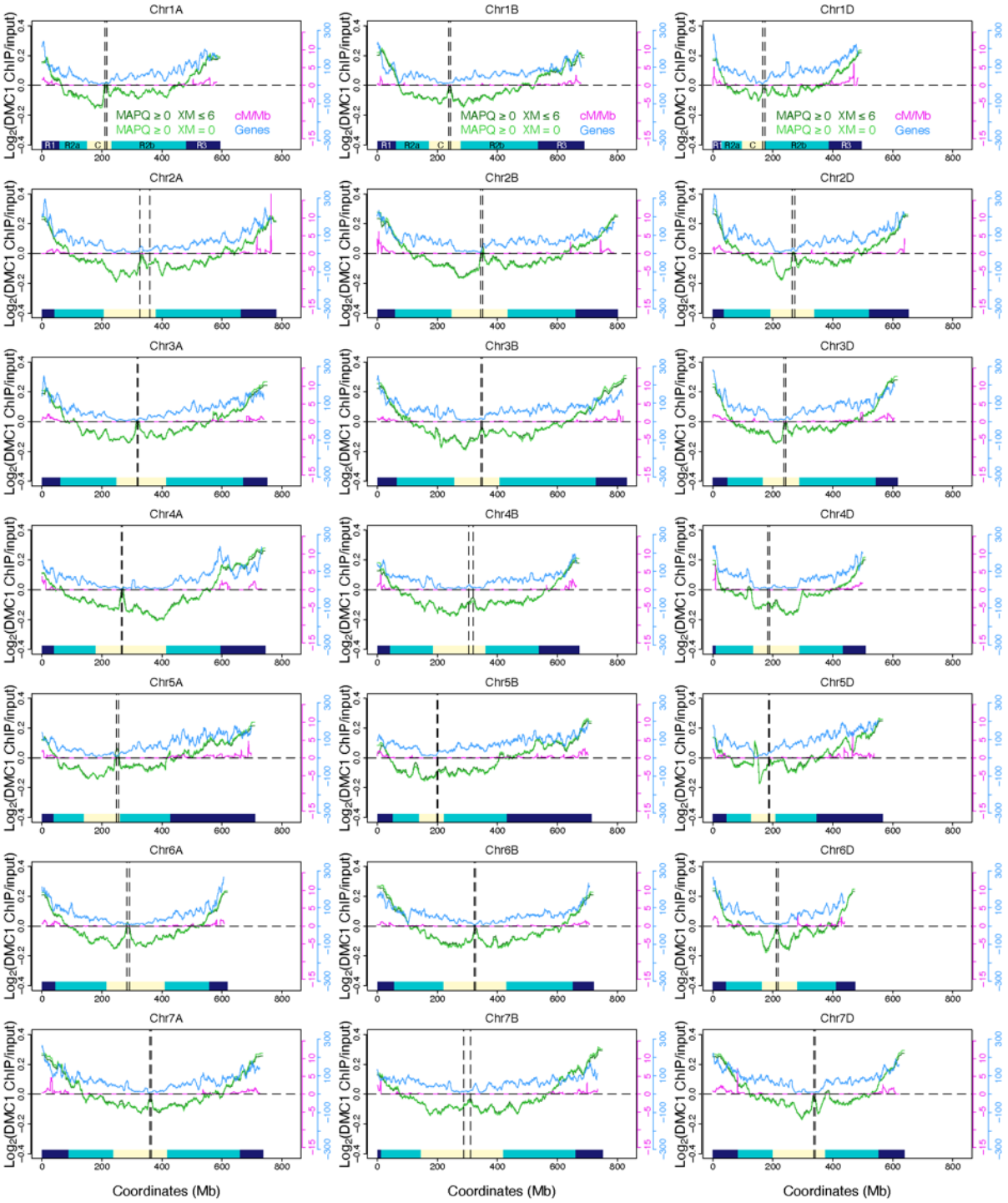


Supplemental Figure S5. Genome-wide landscape of ASY1 ChIP-seq enrichment after varying the alignment MAPQ score threshold. Chromosome profiles of ASY1 ChIP-seq in 1-Mb adjacent windows (log_2[ChIP/input], using a Chinese Spring chromatin input sequencing library with accession SRR6350669 (IWGSC 2018)), smoothed by applying a moving average (period=15 1-Mb windows). Bowtie 2-assigned MAPQ≥23 (orange) corresponds to unique alignments, MAPQ≥2 (dark green) corresponds to best-scoring multiple alignments combined with unique alignments, and MAPQ≥0 (light green) corresponds to multiple alignments (with one alignment randomly selected for read pairs that aligned equally well to more than one location) combined with unique alignments. Alignments containing more than 6 mismatches were discarded. ChIP-seq chromosome profiles are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018), pink), and gene frequencies (light blue) in

10-Mb sliding windows with a 1-Mb step. Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018).

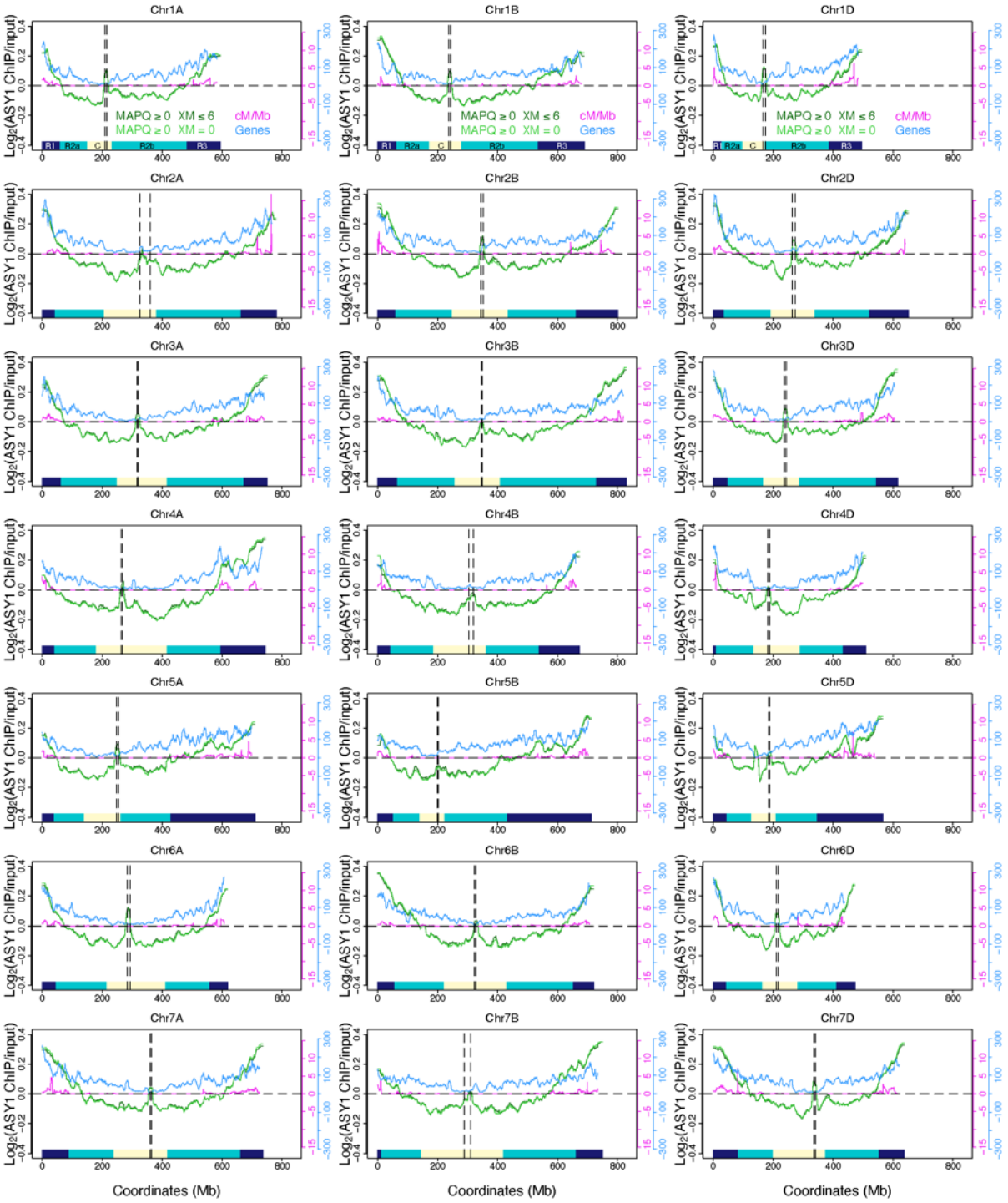


Supplemental Figure S6. Genomic landscapes of DMC1 and ASY1 ChIP and chromatin input sequencing libraries after varying the alignment MAPQ score threshold. A. Raw coverage profiles along a representative set of homoeologous chromosomes (3A, 3B and 3D) for DMC1 ChIP-seq in 1-Mb adjacent windows, smoothed by applying a moving average (period=15 1-Mb windows). Bowtie 2-assigned MAPQ≥0 (light green) corresponds to multiple alignments (with one alignment randomly selected for read pairs that aligned equally well to more than one location) combined with unique alignments, MAPQ≥2 (dark green) corresponds to best-scoring multiple alignments combined with unique alignments, and MAPQ<2 (blue) corresponds to read pairs that aligned equally well to multiple locations, without a best-scoring alignment (with one alignment randomly selected for each read pair). Alignments containing more than 6 mismatches were discarded. Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018). **B.** As in A, but for ASY1 ChIP-seq. **C.** As in A, but for a Chinese Spring chromatin input sequencing library with accession SRR6350669 (IWGSC 2018).



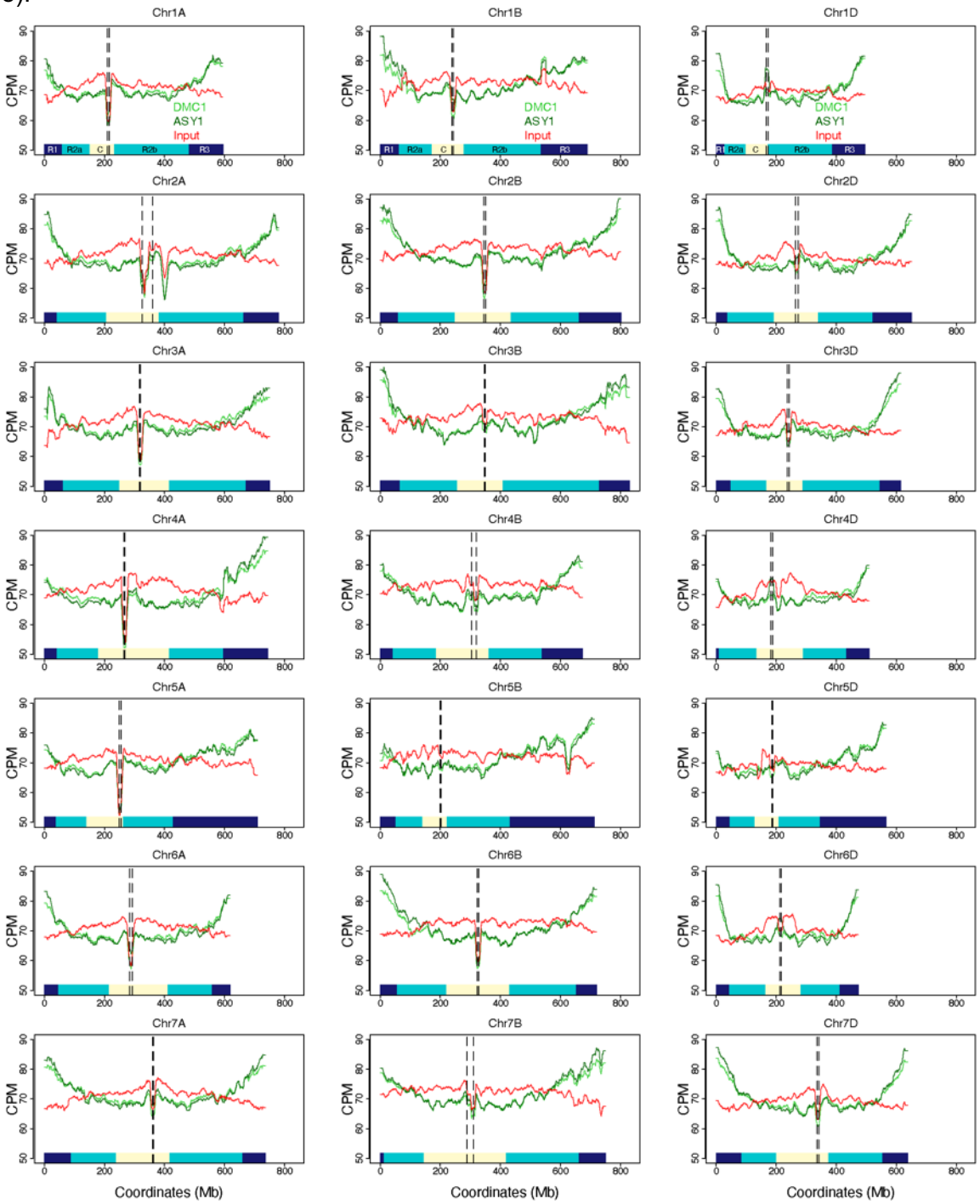
Supplemental Figure S7. Genome-wide landscape of DMC1 ChIP-seq enrichment after varying the alignment mismatch threshold. Chromosome profiles of DMC1 ChIP-seq in 1-Mb adjacent windows ($\log_2$[ChIP/input], using a Chinese Spring chromatin input sequencing library with accession SRR6350669 (IWGSC 2018)), smoothed by applying a moving average (period=15 1-Mb windows). ChIP-seq chromosome profiles are shown after discarding alignments with more than 6 mismatches (MAPQ≥0 XM≤6, dark green), or with any mismatches (MAPQ≥0 XM=0, light green). ChIP-seq chromosome profiles are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018), pink), and gene frequencies

(light blue) in 10-Mb sliding windows with a 1-Mb step. Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018).



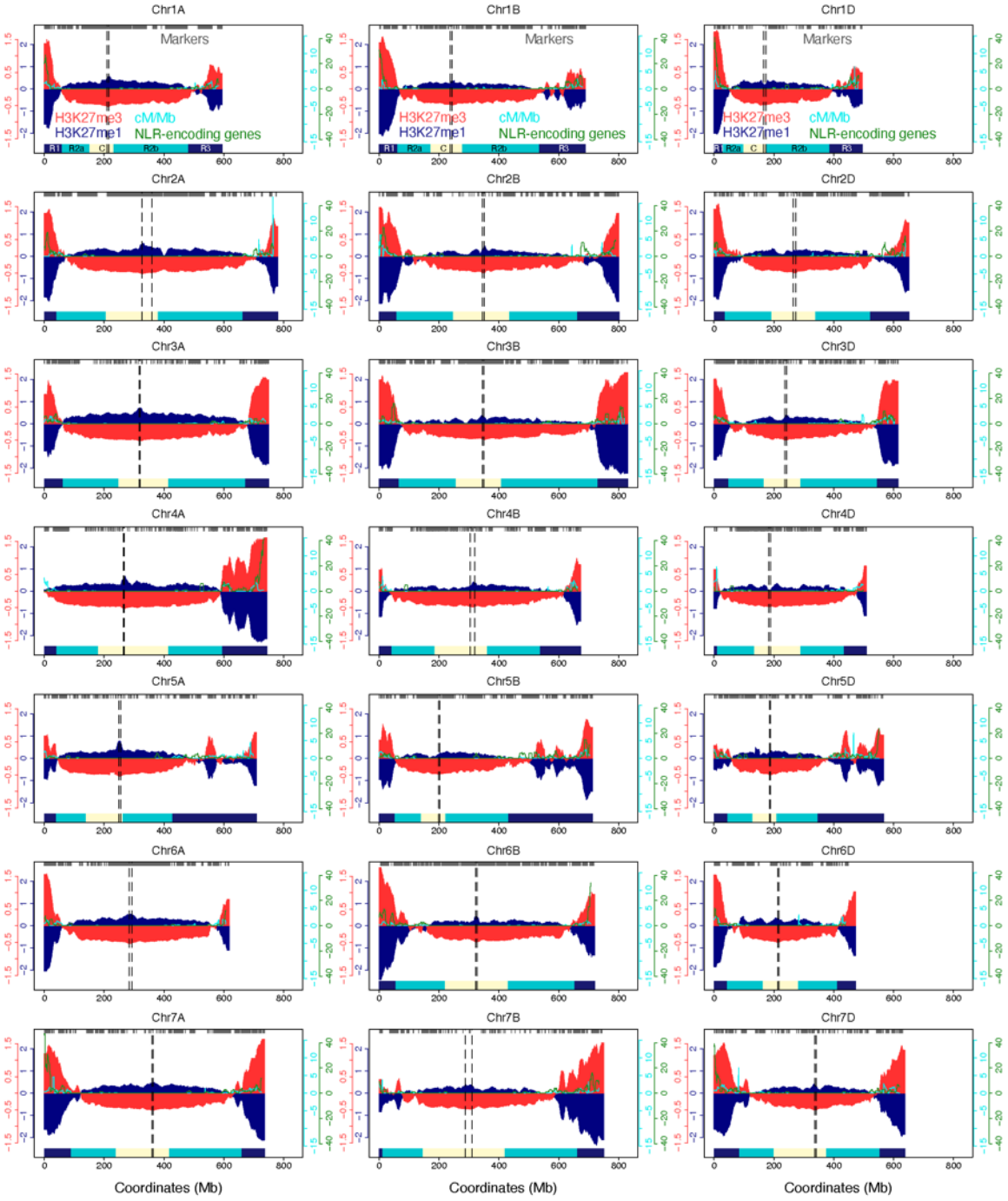
Supplemental Figure S8. Genome-wide landscape of ASY1 ChIP-seq enrichment after varying the alignment mismatch threshold. Chromosome profiles of ASY1 ChIP-seq in 1-Mb adjacent windows (log_2[ChIP/input], using a Chinese Spring chromatin input sequencing library with accession SRR6350669 (IWGSC 2018)), smoothed by applying a moving average (period=15 1-Mb windows). ChIP-seq chromosome profiles are shown after discarding alignments with more than 6 mismatches (MAPQ≥0 XM≤6, dark green), or with any mismatches (MAPQ≥0 XM=0, light green). ChIP-seq chromosome profiles are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018), pink), and gene frequencies

(light blue) in 10-Mb sliding windows with a 1-Mb step. Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018).



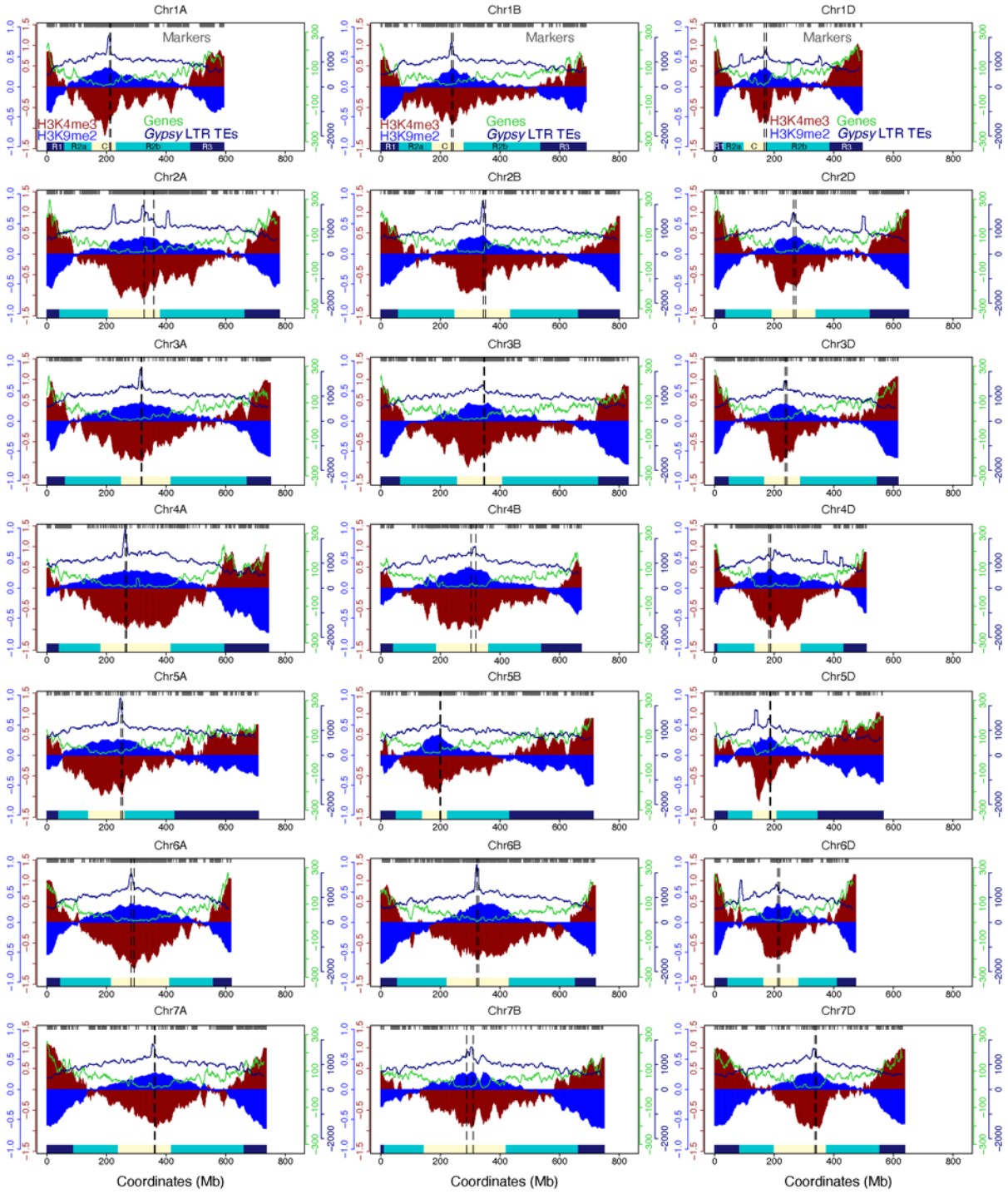
Supplemental Figure S9. Genome-wide landscapes of DMC1 and ASY1 ChIP and chromatin input sequencing libraries. Chromosome profiles of coverage (counts per million mapped reads, CPM) for DMC1 (light green) and ASY1 (dark green) ChIP-seq in 1-Mb adjacent windows, compared to those for a Chinese Spring chromatin input sequencing library (red) with accession SRR6350669 (IWGSC 2018), smoothed by applying a moving average (period=15 1-Mb

windows). In each case, best-scoring multiple alignments were combined with unique alignments (MAPQ≥2), and alignments containing more than 6 mismatches were discarded. Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018).



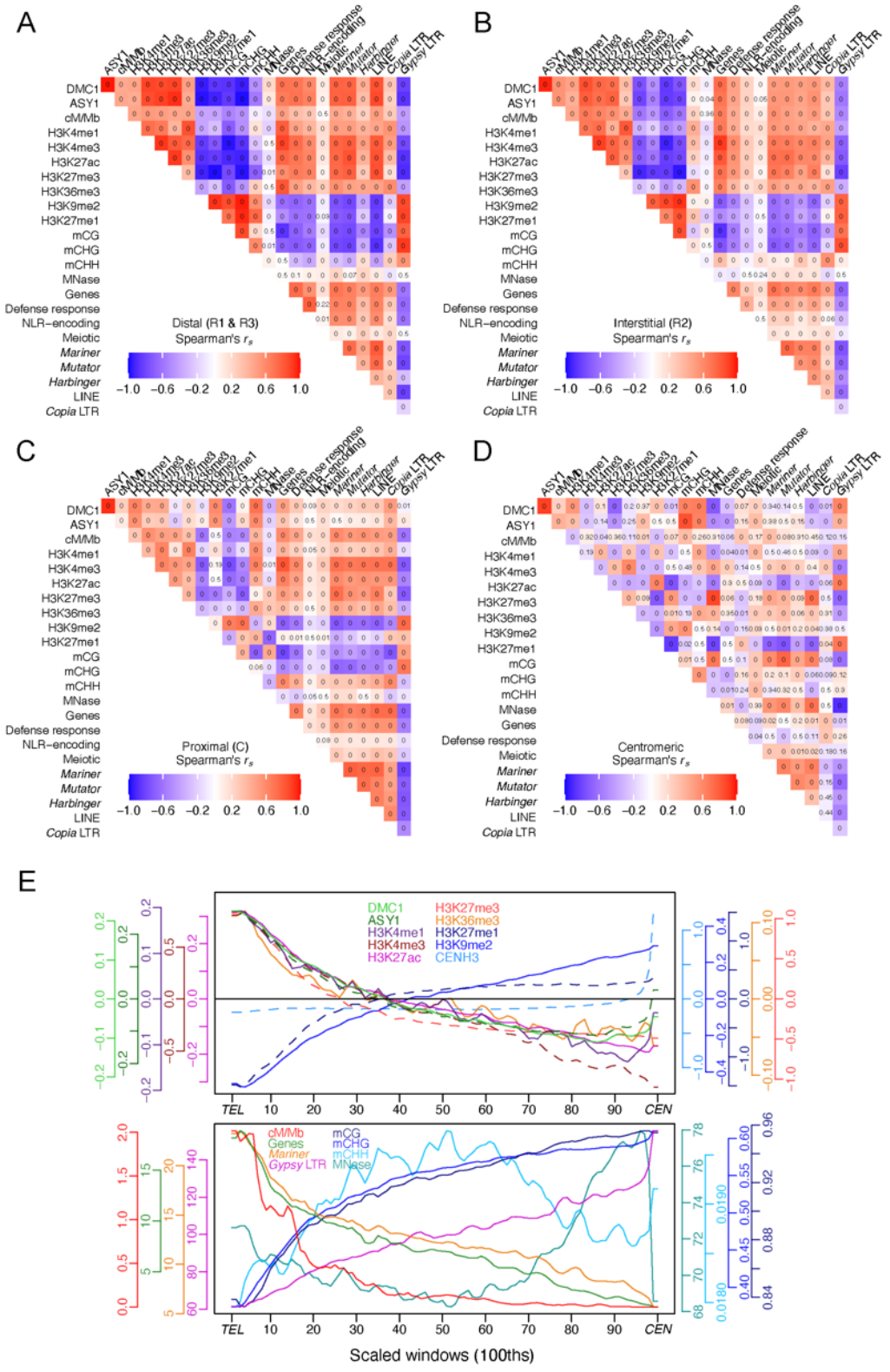
Supplemental Figure S10. Genome-wide landscapes of H3K27me1, H3K27me3, crossover rate and *NLR* genes. Chromosome profiles of H3K27me1 (navy, $\log_2$[ChIP/MNase]) and H3K27me3 (red, $\log_2$[ChIP/input]) (IWGSC 2018) ChIP-seq in 1-Mb adjacent windows, smoothed by applying a moving average (period=15 1-Mb windows). ChIP-seq chromosome profiles (shading) are overlaid with crossover rates (cM/Mb, derived from a Chinese Spring×Renan genetic map (IWGSC 2018)) and gene frequencies in 10-Mb sliding windows with a 1-Mb step (lines). Gene frequency profiles correspond to genes encoding nucleotide-binding and leucine-rich repeat (NLR) proteins (Steuernagel et al. 2020). Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC

2018). Gray ticks denote genomic coordinates of genetic markers used to construct the Chinese Spring×Renan genetic map.



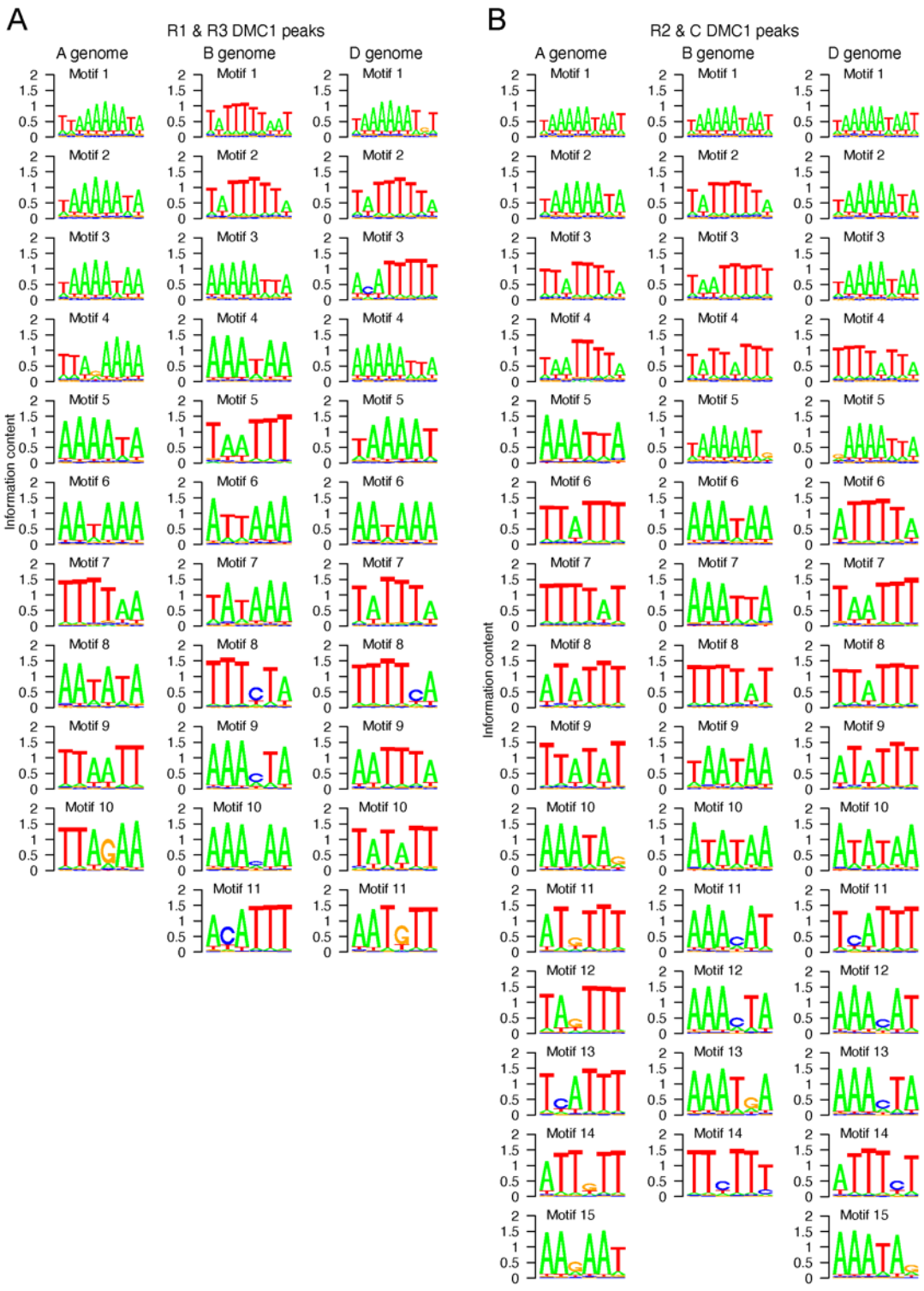
Supplemental Figure S11. Genome-wide landscapes of H3K4me3, H3K9me2, genes and *Gypsy* LTR transposons. Chromosome profiles of H3K4me3 (dark red, $\log_2$[ChIP/MNase]) and H3K9me2 (blue, $\log_2$[ChIP/MNase]) ChIP-seq in 1-Mb adjacent windows, smoothed by applying a moving average (period=15 1-Mb windows). ChIP-seq chromosome profiles (shading) are overlaid with gene and TE frequencies in 10-Mb sliding windows with a 1-Mb step (lines). Gene frequency profiles correspond to a representative set of the IWGSC RefSeq v1.1 Chinese Spring annotation of high-confidence protein-coding gene models (IWGSC 2018). *Gypsy* LTR transposable element (TE) frequency profiles correspond to elements in the IWGSC RefSeq v1.0

Chinese Spring annotation (IWGSC 2018). Previously defined coordinates delimiting distal (R1 and R3, navy boxes), interstitial (R2a and R2b, turquoise boxes), proximal (C, cream boxes) and centromeric (vertical dashed lines) regions are indicated along the x-axis (IWGSC 2018).

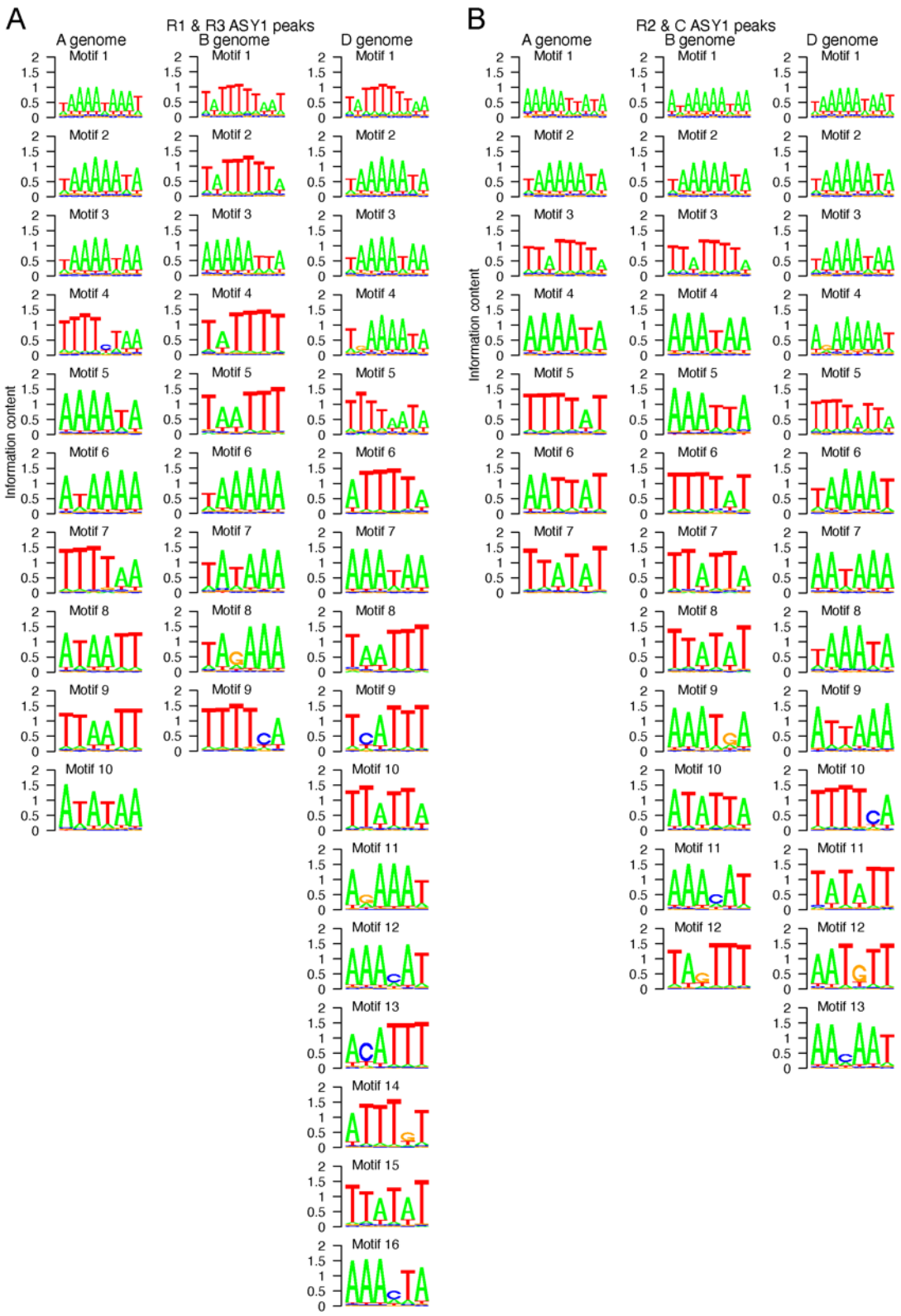


Supplemental Figure S12. Correlation analyses of chromatin, recombination and gene and transposon annotations. A. Spearman's rank-order correlation coefficients (r_s) for the indicated parameter pairs profiled in 1-Mb adjacent windows within distal compartments (R1 and R3), with cells coded according to the color-gradient correlation scale. *P*-values for r_s correlation

coefficients were standardized to represent those based on pairwise values across 100 windows and are indicated within each cell. Included published data sets are H3K4me1 and H3K27ac ChIP-seq (log_2[ChIP/input]) (Li et al. 2019), H3K27me3 and H3K36me3 ChIP-seq (log_2[ChIP/input]) (IWGSC 2018), and whole-genome bisulfite sequencing-derived DNA methylation (mCG, mCHH and mCHG proportions) (IWGSC 2018). **B.** As in A, but for interstitial compartments (R2). **C.** As in A, but for proximal compartments. **D.** As in A, but for centromeric regions. **E.** Mean log_2(ChIP/control), crossover rate (cM/Mb), genomic feature frequency, DNA methylation proportion and MNase-seq profiles along proportionally scaled chromosome arms (100 scaled windows per chromosome arm), oriented from telomere (*TEL*) to centromere (*CEN*), smoothed by applying a moving average (period=3 scaled windows).

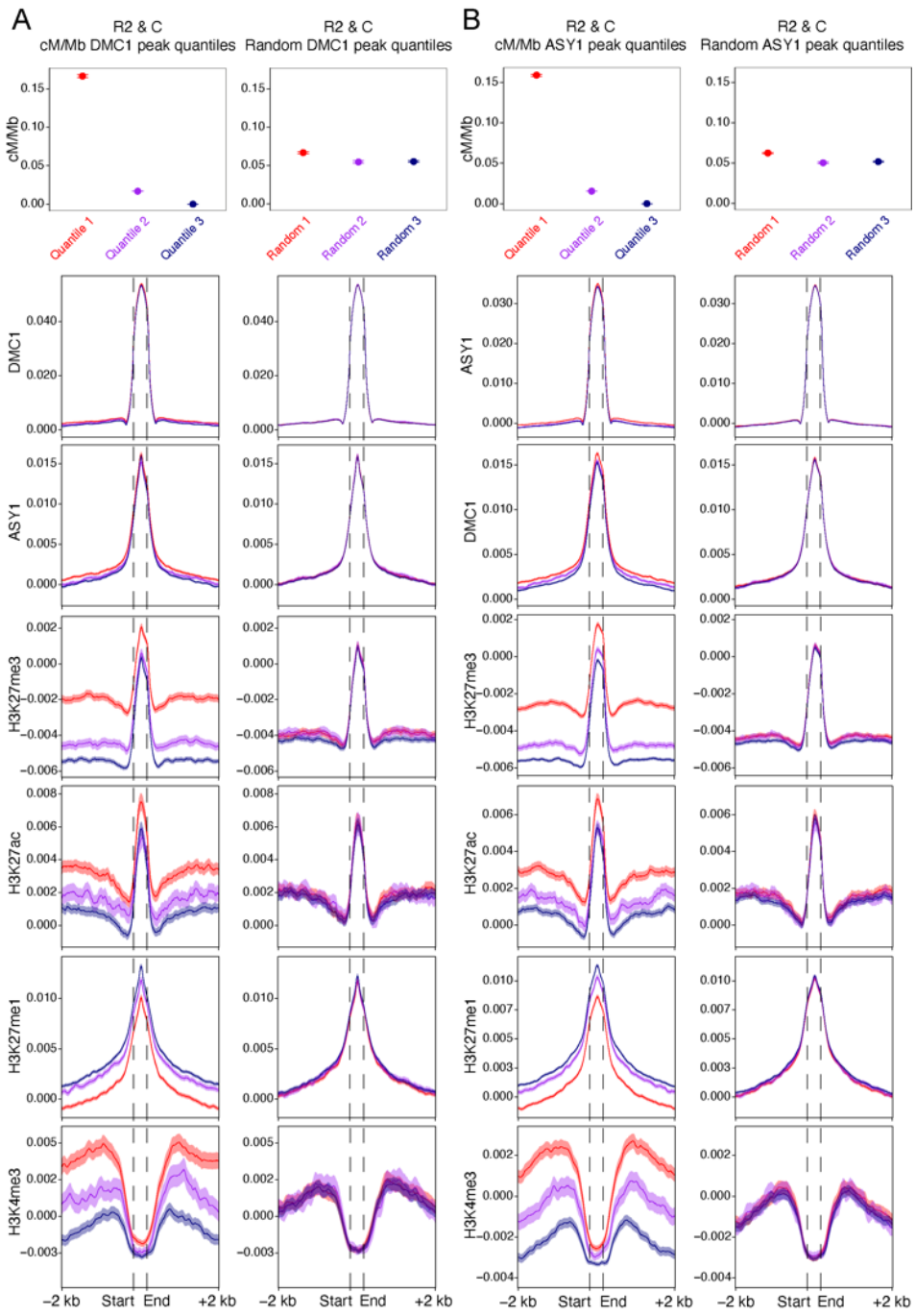


Supplemental Figure S13. DNA sequence motif enrichment within DMC1 ChIP-seq peaks. **A.** Over-represented DNA sequence motifs identified by Weeder v2.0 among the top 10,000 DMC1 ChIP-seq peaks within the distal (R1 and R3) compartments of each subgenome, with peaks ranked by decreasing $-\log_{10}$(ranger-assigned FDR). **B.** As in A, but for DMC1 ChIP-seq peaks within the interstitial and proximal (R2 and C) compartments of each subgenome.



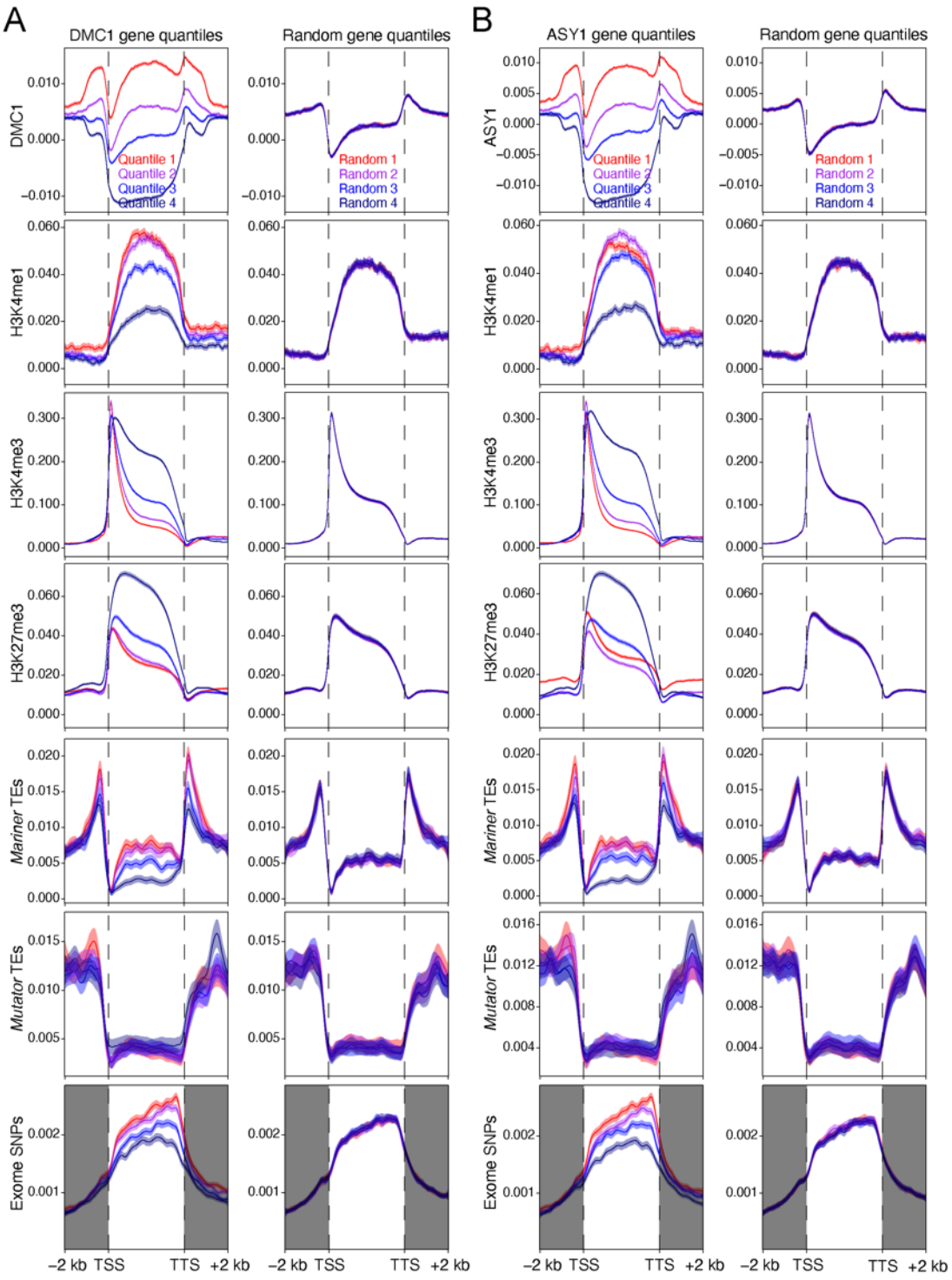
Supplemental Figure S14. DNA sequence motif enrichment within ASY1 ChIP-seq peaks. **A.** Over-represented DNA sequence motifs identified by Weeder v2.0 among the top 10,000 ASY1 ChIP-seq peaks within the distal (R1 and R3) compartments of each subgenome, with

peaks ranked by decreasing $-\log_{10}$(ranger-assigned FDR). **B.** As in A, but for ASY1 ChIP-seq peaks within the interstitial and proximal (R2 and C) compartments of each subgenome.



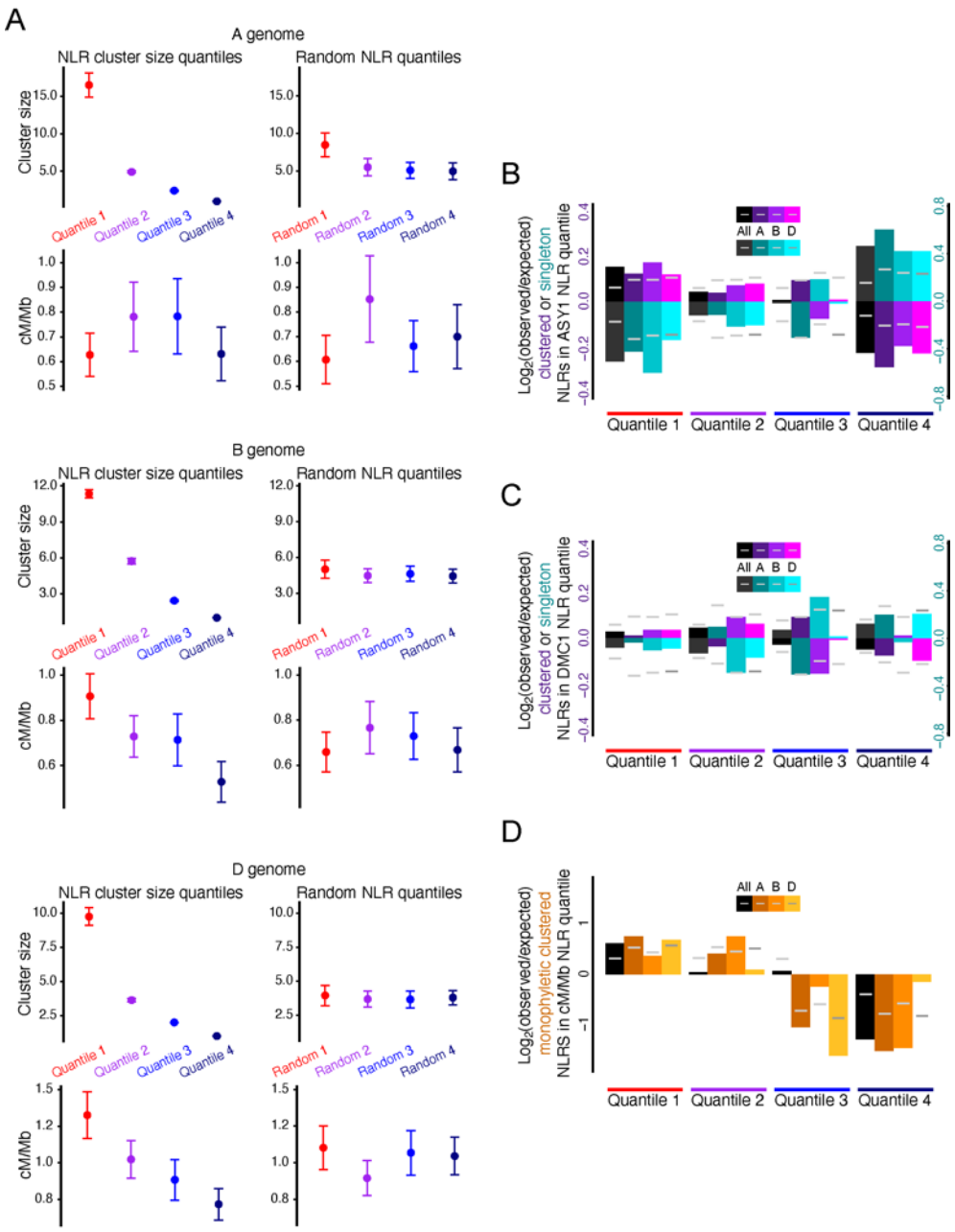
Supplemental Figure S15. Chromatin profiles around interstitial and proximal DMC1 and ASY1 ChIP-seq peaks grouped by decreasing crossover rate. A. Interstitial and proximal (R2 and C) DMC1 ChIP-seq peaks were divided into three groups corresponding to those in the 100th—66th (Quantile 1), 66th—33rd (Quantile 2) and 33rd—0th (Quantile 3) percentiles with regard to their mean crossover rate (cM/Mb, derived from the Chinese Spring×Renan genetic map) between 1 kb upstream of peak start coordinates and 1 kb downstream of peak end coordinates (left), or into three randomized groups (right). Solid circles denote the mean crossover rate for each group of peaks (error bars=95% confidence intervals for mean cM/Mb values). Meta-profiles show windowed mean values (solid lines) for each group of peaks and 2-kb flanking regions (transparent ribbons=95% confidence intervals). ChIP-seq coverage meta-profiles of

DMC1, ASY1, H3K27me3 (IWGSC 2018) and H3K27ac (Li et al. 2019) are derived from $\log_2$(ChIP/input) profiles, and those for H3K27me1 and H3K4me3 from $\log_2$(ChIP/MNase) profiles. **B.** As in A, but for interstitial and proximal ASY1 ChIP-seq peaks.



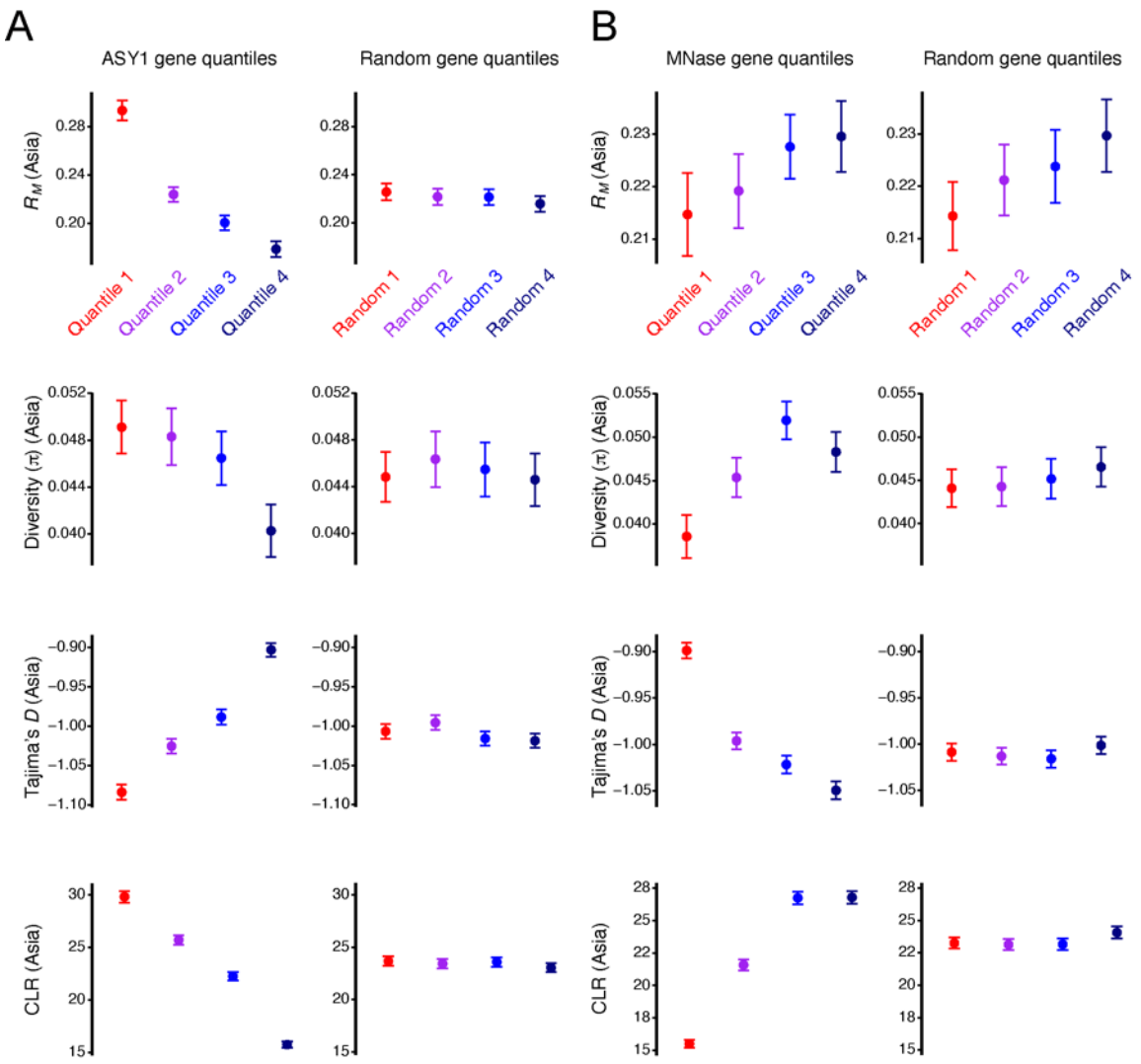
Supplemental Figure S16. Chromatin profiles, transposon density and SNPs around genes grouped by decreasing DMC1 or ASY1 ChIP-seq coverage located in interstitial and proximal compartments. A. Genes were divided into four groups corresponding to those in the 100th—75th (Quantile 1), 75th—50th (Quantile 2), 50th—25th (Quantile 3) and 25th—0th (Quantile 4) percentiles with regard to their mean $\log_2$(DMC1 ChIP/input) coverage values between 1 kb upstream of transcriptional start sites (TSSs) and 1 kb downstream of transcriptional termination sites (TTSs) (left), or into four randomized groups (right). Meta-profiles show

windowed mean values (solid lines) for each group of genes and 2-kb flanking regions (transparent ribbons=95% confidence intervals). ChIP-seq coverage meta-profiles of DMC1, ASY1, H3K4me1 (Li et al. 2019) and H3K27me3 (IWGSC 2018) are derived from $\log_2$(ChIP/input) profiles, and those for H3K4me3 from $\log_2$(ChIP/MNase) profiles. Meta-profiles of elements within the *Mariner* or *Mutator* transposon superfamilies or exome-sequencing-derived SNPs (He et al. 2019) are based on windowed transposon or SNP frequencies. SNPs outside of transcribed regions were retained for meta-profile calculation (gray shading). **B.** As in A, but with genes divided into four groups according to their mean $\log_2$(ASY1 ChIP/input) coverage values (left), or into four randomized groups (right).



Supplemental Figure S17. Genes encoding nucleotide-binding and leucine-rich repeat (NLR) proteins grouped by decreasing physical clustering, ASY1, DMC1 or crossover rate. **A.** *NLR* genes (Steuernagel et al. 2020) in each subgenome were divided into four groups corresponding to those in the 100th–75th (Quantile 1), 75th–50th (Quantile 2), 50th–25th (Quantile 3) and 25th–0th (Quantile 4) percentiles with regard to physical cluster size (left), or into four randomized groups (right). Solid circles denote the mean cluster size or crossover rate for each group of genes (error bars=95% confidence intervals for mean cluster size or cM/Mb values). **B.** *NLR* gene quantiles defined by decreasing mean $\log_2$(ASY1 ChIP/input) coverage were evaluated for representation of physically clustered or singleton *NLR* genes. Log_2(observed/expected) ratios (bar graphs) and significance thresholds (horizontal gray lines, α=0.05) were calculated by sampling from the hypergeometric distribution 100,000 times. Analyses were performed across all subgenomes and within each subgenome (All, A, B and D).

C. As in B, but analyzing *NLR* gene quantiles defined by decreasing mean $\log_2$(DMC1 ChIP/input) coverage. **D.** As in B, but analyzing *NLR* gene quantiles defined by decreasing mean crossover rate for representation of physically clustered *NLR* genes that are monophyletic.



Supplemental Figure S18. Historical recombination, nucleotide diversity and signatures of selection at genes grouped by decreasing ASY1 or nucleosome occupancy. A. Genes were divided into four groups corresponding to those in the 100th–75th (Quantile 1), 75th–50th (Quantile 2), 50th–25th (Quantile 3) and 25th–0th (Quantile 4) percentiles with regard to their mean $\log_2$(ASY1 ChIP/input) coverage (left), or into four randomized groups (right). Solid circles denote the mean value of the indicated parameter for each group of genes (error bars=95% confidence intervals for mean values). Nucleotide diversity (the mean number of intrapopulation pairwise nucleotide differences per gene, π) was divided by gene width. Per-gene Tajima's *D* and composite likelihood ratios (CLRs) were also calculated. **B.** As in A, but with genes divided into four groups according to decreasing mean MNase-seq coverage (left), or into four randomized groups (right).