COmapper: High-resolution mapping of meiotic crossovers by long-read sequencing in *Arabidopsis*

Dohwan Byun[1†], Namil Son[1†], Heejin Kim[1], Jaeil Kim[1], Jihye Park[1], Sang-jun Park[1], Hyein Kim[1], Jaebeom Kim[1,2], Juhyun Kim[1,3], Seula Lee[2], Youbong Hyun[2], Piotr A. Ziolkowski[4], Ian R. Henderson[5], and Kyuha Choi[1*]

[1]Department of Life Sciences, Pohang University of Science and Technology, Pohang, Gyeongbuk, Republic of Korea
[2]School of Biological Sciences, Seoul National University, Seoul, Republic of Korea
[3]Plant Quarantine Technology Center, Department of Plant Quarantine, Animal and Plant Quarantine Agency, Gimcheon, Republic of Korea
[4]Laboratory of Genome Biology, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznań, Poland
[5]Department of Plant Sciences, University of Cambridge, Cambridge, UK
*To whom correspondence should be addressed. Email: kyuha@postech.ac.kr

[†]These authors contributed equally to this work.

Abstract

Meiotic crossovers rearrange existing genetic variation between homologous chromosomes, profoundly affecting genomic diversity. Crossovers are typically constrained to 1–3 events per chromosome pair, and their distribution is shaped by chromatin accessibility and DNA polymorphisms. Genome-wide crossover maps can be generated in plants by high-throughput short-read sequencing or linked-read sequencing. Here, we use long-read nanopore sequencing technology to develop a crossover mapping pipeline, COmapper, for high-resolution mapping of genome-wide crossovers from pooled DNA of F_1 hybrid pollen and F_2 recombinant seedlings derived from a cross between *Arabidopsis thaliana* accessions Col and L*er*. We validate the high accuracy of COmapper by applying nanopore long-read sequencing to pooled DNA of *Arabidopsis* F_2 individuals with crossovers mapped by short-read sequencing. Using the COmapper, we constructed high-resolution genomic maps of crossovers using F_1 hybrid pollen and F_2 seedlings in wild type and crossover-elevated *recq4a recq4b* mutant, showing results comparable to short-read sequencing. Crossovers were enriched at gene-proximal promoters in wild type and increased but reshaped by high polymorphism density in *recq4a recq4b*. We propose that COmapper will be widely applicable

for exploring the effects of genetic, epigenetic and environmental changes on the crossover patterns across diverse plant species.

Keywords: Arabidopsis, crossover, meiosis, nanopore long-read sequencing

Introduction

Meiosis consists of a single round of DNA replication and two consecutive rounds of cell division, halving genome ploidy and producing genetically diverse gametes (Villeneuve and Hillers, 2001). During meiosis, homologous chromosomes undergo programmed DNA double-strand breaks (DSBs), pairing, synapsis and homology-directed repair, resulting in crossovers or non-crossovers (Hunter, 2015; Zickler and Kleckner, 2023). Topoisomerase-like SPO11 proteins catalyze the formation of meiotic DSBs and remain covalently attached to the ends of these DSBs (Grelon et al., 2001; Keeney et al., 2014; Lam and Keeney, 2014; Robert et al., 2016). Upon removal of the bound SPO11-oligonucleotides, the 5′ ends of DSBs undergo bidirectional 5′ to 3′ resection, producing 3′ overhangs of single-stranded DNA (ssDNA) (Neale et al., 2005; Lam and Keeney, 2014). This ssDNA invades non-sister chromatids and searches for homology, leading to the formation of recombination intermediates such as displacement loops and double Holliday junctions. Recombination intermediates are eventually resolved into either crossovers or non-crossovers (Mercier et al., 2015; Gray and Cohen, 2016; Wang and Copenhaver, 2018; Zickler and Kleckner, 2023). Although substantial DSBs are initially formed, only a few are repaired into crossovers by class I and class II repair pathways in most species. In plants, the class I pathway is responsible for ~85–90% of all crossovers, which are formed by a group of pro-crossover factors including the E3 ligase HEI10 (Mercier et al., 2015; Wang and Copenhaver, 2018). Class I crossovers are sensitive to interference, a phenomenon that results in multiple crossovers being more widely spaced than would occur randomly. Crossover interference in plants is explained by a coarsening model where the diffusible and aggregating dynamics of HEI10 proteins limit closely spaced crossovers, resulting in 1–3 crossovers per pair of homologous chromosomes (Mercier et al., 2015; Morgan et al., 2021; Girard et al., 2023). In the class II pathway, ~10–15% of all crossovers are non-interfering and mediated by the endonuclease MUS81 (Berchowitz et al., 2007; Zakharyevich et al., 2012). Class II crossovers are inhibited by different pathways involving anti-recombination factors, such as RECQ4A and RECQ4B helicases in plants (Crismani et al., 2012; Séguéla-Arnaud et al., 2015; Mieulet et al., 2018).

In plants, the distribution and frequency of meiotic DSBs and crossovers along chromosomes are non-random and depend on chromatin accessibility (Wijnker et al., 2013; Choi et al., 2013; Shilo et al., 2015; He et al., 2017; Underwood et al., 2018; Choi et al., 2018). Consequently, meiotic crossovers predominantly occur in gene-rich euchromatic regions and are suppressed in heterochromatic pericentromeres and centromeres. At a finer scale, crossovers are highly enriched at gene transcription start sites (TSS) near nucleosome-depleted promoters, but are rare in heterochromatic transposable elements and repeats (Wijnker et al., 2013; Choi et al., 2013; Choi and Henderson, 2015b). In line with this, DNA methylation and histone H3 methylation on lysine 9 (H3K9me2) inhibit crossover formation (Yelina et al., 2015; Underwood et al., 2017; Underwood et al., 2018; Underwood and Choi, 2019; Fernandes et al., 2024). Moreover, a high density of DNA polymorphisms and structural variation suppress crossovers (Choi et al., 2016; Serra et al., 2018a; Dluzewska et al., 2023; Fernandes et al., 2024). However, the frequency of meiotic crossovers correlates with the heterozygosity along chromosomes at a kilobase- (kb) or megabase-pair (Mb) scale in *Arabidopsis* (*Arabidopsis thaliana*), which depends on MSH2, a component of the mismatch repair system (Ziolkowski et al., 2015; Dluzewska et al., 2023). Therefore, mapping of genome-wide crossovers at high resolution is important for understanding the genetic and epigenetic effects on meiotic crossovers and genome evolution in various species (Henderson and Bomblies, 2021; Naish et al., 2021; Kim and Choi, 2022). Short-read sequencing and a probabilistic method for genotyping and crossover prediction have been intensively used to determine crossover numbers and distribution in F_2 individuals between *Arabidopsis* accessions (Wijnker et al., 2013; Rowan et al., 2015; Serra et al., 2018b; Lawrence et al., 2019; Nageswaran et al., 2021; Kim et al., 2022; Kim et al., 2024). Recently, linked-read sequencing has been applied to DNA pooled from hybrid pollen grains or individual pollen nuclei of hybrid plants using microfluidics on a 10x Chromium Controller to map crossovers in *Arabidopsis*, tomato (*Solanum lycopersicum*) and beak sedge (*Rhynchospora breviuscula*) (Sun et al., 2019; Rommel Fuentes et al., 2020b; Castellani et al., 2024).

Here, we develop a crossover mapping pipeline, COmapper (https://github.com/KyuhaChoi-Lab/COmapper), using long-read nanopore sequencing of pooled DNA from F_1 pollen or F_2 recombinant seedlings from *Arabidopsis* hybrids derived from a cross between the accessions Columbia-0 (Col) and Landsberg *erecta*-0 (L*er*). We validate the precision of COmapper for genome-wide crossover analysis through long-read nanopore sequencing of F_2 recombinant individuals whose crossover sites were mapped by short-read sequencing. We demonstrate that COmapper allows high-resolution mapping of crossovers using long-read sequencing of pooled genomic DNA from wild-type and *recq4a recq4b* (*recq4ab*) F_1 hybrid pollen grains or

from their F_2 recombinant seedlings. We propose that COmapper will be useful for determining crossover patterns in diverse plants with different genetic backgrounds and ploidy.

Results

Establishment of COmapper and long-read sequencing to map crossovers

To develop COmapper and enable high-resolution mapping of meiotic crossovers at the single-nucleotide polymorphism (SNP) level using long-read sequencing, we used the two polymorphic *Arabidopsis* accessions, Columbia-0 (Col) and Landsberg *erecta*-0 (L*er*) (**Figure 1A**). We crossed them to generate wild-type (WT) Col × L*er* F_1 hybrid plants, from which we collected recombinant pollen; we also allowed them to self-pollinate to collect F_2 recombinant seeds (**Figure 1A**). We performed genotyping by sequencing (GBS) of individual F_2 recombinant plants using individual indexing, short-read sequencing and the TIGER pipeline, which allows the mapping of crossover sites along chromosomes (**Figure 1B**) (Wijnker et al., 2013; Rowan et al., 2015; Rowan et al., 2019; Nageswaran et al., 2021; Kim et al., 2022; Kim et al., 2024). We then used the GBS data and pooled genomic DNA (gDNA) of F_2 recombinant seedlings as a positive control to establish and validate COmapper. Notably, GBS of F_2 recombinants has been also used to validate the linked-sequencing-based crossover mapping method via microfluidics on the 10x Chromium Controller (**Figure 1C**) (Sun et al., 2019). GBS and linked sequencing both allow the genotyping of individual genomes or large DNA molecules through individual barcoding and multiplexing, amplification, short-read sequencing, demultiplexing and read alignment to determine crossover sites (**Figure 1B** and **C**). By contrast, nanopore long-read sequencing followed by COmapper can directly identify Col-specific or L*er*-specific SNPs within individually sequenced DNA molecules without prior indexing or amplification and can determine the position of crossover sites in individual DNA molecules (**Figure 1D** and **E**). Therefore, the read length obtained by long-read sequencing and the number of SNPs in individually sequenced DNA molecules are critical for generating genome-wide high-resolution maps of crossovers, along with the depth of genome sequence coverage.

To obtain long-read nanopore sequencing reads for the mapping of crossover sites (**Figure 2A**), we isolated mature pollen grains from WT Col × L*er* F_1 hybrid plants and extracted large genomic DNA (gDNA) molecules using our modified CTAB method (see Methods). Our gDNA extraction from pollen grains of Col × L*er* F_1 hybrids resulted in nanopore sequence read lengths of ~11.1 kb for gDNA (mean, 11.1 kb; N50, 14.0 kb), comparable to that of gDNA

extracted from F_2 seedlings using the CTAB method (**Figure 2B** and **Supplemental Figure 1**). To achieve deep sequencing coverage, we used 9 µg gDNA on the PromethION device, which produced reads corresponding to ~100–150 Gb of sequence per flow cell (**Supplemental Figure 1** and **Supplemental Table 1**). We successfully aligned ~90.7% of the nanopore reads to the TAIR10 *Arabidopsis* reference genome using the minimap2 alignment tool (**Figure 2C**). The aligned reads covered the genome uniformly from each chromosome end to end except across centromere-proximal regions, with a mean read coverage frequency of 904.9 (**Figure 2D** and **Supplemental Figure 2**). We used a high-quality SNP dataset distinguishing between Col and L*er* accessions previously used for GBS (Rowan et al., 2015; Serra et al., 2018b; Kim et al., 2024), with a 524-bp mean distance between consecutive SNPs (**Figure 2E** and **F**). We noted that structural variation between Col and L*er* results in decreased coverage of the aligned long reads and fewer SNPs for GBS in the pericentromeric regions near centromeres (**Figure 2D, E** and **Supplemental Figure 3**). Of all aligned long reads, 86.1% contained more than eight SNPs per read, with a mean 19.8 SNPs per read (**Figure 2G**), which were used to determine crossover sites by COmapper.

Development and validation of COmapper

We sought to establish a crossover mapping model for COmapper, which scans for the presence of SNPs from the 5′ to the 3′ end of a long read and identifies simple crossovers (SCOs) where consecutive Col-specific SNPs switch to consecutive L*er*-specific SNPs, or vice versa, by counting the numbers of SNPs (**Figure 3A**). In *Arabidopsis*, most crossovers are SCOs, but a minority are associated with gene conversion (**Figure 3A**) (Wijnker et al., 2013). Therefore, we also considered gene conversion-associated crossovers (GC-COs) when examining the number and positions of consecutive Col-specific or L*er*-specific SNPs to determine crossover sites (**Figure 3B**). Accordingly, we developed and tested six crossover mapping models (**Figure 3B**), which use different numbers of consecutive Col-specific and L*er*-specific SNPs with or without gene conversion within the crossover tract to map crossovers. To validate these models, we performed nanopore sequencing of two independent pooled gDNA samples from 4 or 51 Col × L*er* F_2 recombinant plants in which crossover sites had already been determined by GBS (**Supplemental Figure 4** and **Supplemental Table 2**). We then examined the precision of each crossover mapping model by aligning the long-read molecules to the *Arabidopsis* genome and analyzing their SNPs. This revealed that the 8 SCO and 8 CO models were the most accurate in detecting the true crossover sites mapped by GBS, with precisions of 94.7% (8 SCO) and 94.3% (8 CO),

respectively (Welch′s *t*-test; 8 SCO vs 8 CO $P = 0.657$, 8 CO vs 7 CO $P = 7.05 \times 10^{-10}$) (**Figure 3C, D** and **Supplemental Figure 5**). The 8 CO model, which identifies both SCO and GC-CO types, detected more crossovers than the 8 SCO model, which detected only the SCO type (Welch′s *t*-test, $P = 5.50 \times 10^{-3}$) (**Figure 3D, E, Supplemental Figs 5** and **6**). COmapper with the 8 CO model detected ~7.9% and ~2.1% of the GC-CO type of total crossover sites from the pooled gDNA of 4 and 51 F_2 plants, respectively (**Figure 3E** and **Supplemental Figure 6**). Furthermore, we used COmapper to identify crossover sites in nanopore long reads of pooled hybrid pollen and pooled F_2 seedlings from wild-type (WT) Col × L*er* and *recq4ab* Col × *recq4ab* L*er* plants. Of all crossover sites identified using COmapper with the 8 CO model, ~12.9% and ~8.5% of crossovers were of the GC-CO type in the wild-type (WT) and *recq4ab* Col × *recq4ab* L*er* F_1 hybrid pollen, respectively. Similar to 51 F_2 plants (~2.1%) used for GBS and COmapper, 2.8% and 3.2% of crossovers detected from WT (*n* = 1,000) and *recq4ab* Col × *recq4ab* L*er* F_2 recombinant seedlings (*n* = 1,000) were of the GC-CO type (**Figure 3E**).

To validate the output of COmapper, we used the crossover sites mapped by GBS as true positives from 4 and 51 Col × L*er* F_2 recombinant plants (32 crossovers for 4 F_2 plants, 329 crossovers for 51 F_2 plants). COmapper with the 8 CO model was capable of identifying true positive crossover sites multiple times at each site, while false positive sites, which constituted ~5% of all crossovers, were predominantly mapped only once or twice (**Figure 3F** and **G**). COmapper accurately identified all 32 crossovers that were mapped by GBS in the 4 Col × L*er* F_2 recombinants, with a mean of 18.9 long reads covering each crossover site when reaching a sequencing depth of 40.14 Gb (**Supplemental Figure 4A** and **B**). Of the 329 crossovers mapped by GBS in the 51 Col × L*er* F_2 recombinants, COmapper accurately mapped 271 crossovers, with a mean of 6.07 long reads per crossover site. Notably, COmapper failed to detect the remaining 58 crossovers. COmapper identified the 32 and 271 crossovers from the 4 and 51 Col × L*er* F_2 recombinants, respectively, using long reads with at least eight SNPs per read. The minimum length of the long reads was 0.418 kb and the median length was 3.92 kb (**Figure 3H**). The 58 crossovers mapped by GBS but not by COmapper exhibited a sparser distribution of SNPs, with a greater physical size (median 8.31 kb) encompassing the 8 SNPs required for crossover detection, compared to the median length of 3.92 kb for the 303 crossovers (32 from the 4 F_2 plants and 271 from the 51 F_2 plants) successfully mapped by COmapper (Wilcoxon, $P = 7.78 \times 10^{-8}$) (**Figure 3H**). This provides a rationale for COmapper's failure to detect these 58 crossovers. Specifically, in 41 of 58 crossover sites, the physical length containing 8 SNPs had a median of 14.6 kb, which is larger than genomic average length of 4.72 kb, while we found that the remaining 17 crossovers were not sequenced by chance in the pooled gDNA of 51 F_2 plants. Thus, this validation

analysis of COmapper indicates that successful mapping of meiotic crossovers at SNP resolution by COmapper depends on both the length of long-read sequencing molecules (which need to contain more than eight SNPs) and the depth of sequencing read coverage, as most reads (~99.99%) contain exclusively parental genotypes (**Supplemental Table 1**). Taken together, these results confirm that our COmapper pipeline with an 8 CO model (hereafter referred to simply as 'COmapper') can directly map crossovers in long-read nanopore sequencing molecules at SNP resolution with an overall precision of 94%.

Mapping of crossovers from a pool of Col × L*er* hybrid pollen grains by COmapper

Next, we used the COmapper and long-read nanopore sequencing technology to directly map crossovers from pooled gDNA samples of wild-type (WT) Col × L*er* and *recq4ab* Col × *recq4ab* L*er* F_1 hybrid pollen grains (**Figure 4**). The loss of RECQ4A and RECQ4B function in the *recq4ab* double mutant resulted in a 2.63-fold increase in genetic distance in male meiosis compared to the WT using KASPar technology (**Figure 4A**) (Fernandes et al., 2018). Similarly, COmapper detected a 2.38-fold increase in crossovers per genome (119 Mb) when analyzing long-read sequencing data produced from *recq4ab* Col × *recq4ab* L*er* F_1 hybrid pollen DNA (mean = 3.93 crossovers per genome, ~786.4x genomic coverage) compared to the WT Col × L*er* F_1 hybrid pollen (mean = 1.65 crossovers per genome, 795.8x genomic coverage) (Welch′s *t*-test, $P = 1.22 \times 10^{-12}$) (**Figure 4B**). Importantly, COmapper rarely detected crossovers as false positive crossover events in the negative control sample (leaf genomic DNA from Col × L*er* F_1 hybrids; mean = 0.02 crossovers per genome, 85.6x genomic coverage) compared to the WT Col × L*er* F_1 hybrid pollen (Welch′s *t*-test, $P = 1.10 \times 10^{-10}$) (**Figure 4B**) (**Supplemental Table 1**). These results indicate that COmapper enables the detection of meiotic crossovers in pooled gDNA from Col × L*er* hybrid pollen grains, with an increased number of crossovers in a *recq4ab* background.

Previously, GBS-based crossover maps of *Arabidopsis* Col × L*er* male meiosis using a backcross (BC) population in which the pollen from Col × L*er* F_1 hybrid plants were used to pollinate L*er* stigmas to produce the BCF_1 progeny detected ~4.47 crossovers per genome (Kim et al., 2024). In comparison, COmapper detected ~1.65 crossovers per genome in nanopore long-read sequencing of pooled pollen grains from Col × L*er* hybrid plants, with all detected 1,204 crossovers being non-overlapping. This suggests that a genomic coverage of ~2.7x in nanopore long-read sequencing is required by COmapper to map ~4.47 crossovers per male meiosis. To analyze the genomic landscape of crossovers mapped by COmapper, we normalized crossover frequencies by genomic coverage (1,204 detected

crossovers/795.8x coverage) (**Figure 4C** and **D**). Additionally, using the number of ~4.47 crossovers per genome determined by GBS, we estimated the number of pollen grains from the total number of crossovers detected by COmapper, from which we allowed us to calculate the genetic distance (cM/Mb) across the genome. We then compared the genome-wide crossover landscape obtained by COmapper with the GBS-based crossover map (**Figure 4C** and **D**) (Kim et al., 2024). By plotting the crossovers of male meiosis obtained by both COmapper and GBS along a telomere-to-centromere axis that combined all five chromosomes, we observed that the genomic landscape of crossovers mapped by COmapper correlated with that of GBS (Pearson′s correlation test, R = 0.60, P = 5.74 × 10^{-13}) (**Figure 4C**–**E**). Both showed an enrichment for crossovers along gene-rich chromosome arms, with a further increase in sub-telomeric regions and with crossover suppression in transposable-element-rich pericentromeric regions (**Figure 4C**–**E**). These results are consistent with previous reports that the genomic crossover landscapes correlate with the density of SPO11-1–oligonucleotides, indicative of meiotic DSBs, and are anti-correlated with nucleosome density and DNA methylation (**Figure 4F** and **G**) (Choi et al., 2018; Naish et al., 2021). We also generated a crossover map from *recq4ab* Col × *recq4ab* L*er* F_1 pollen using COmapper. In a *recq4ab* background, we detected greater crossover enrichment at euchromatic chromosome arms than was seen in the WT Col × L*er* F_1 pollen, but a similar crossover suppression at heterochromatic regions as in WT Col × L*er* F_1 pollen (**Figure 4H**); this suggests that heterochromatic factors and structural variation suppress the class II crossovers increased by *recq4ab,* as well as most class I crossovers in pericentromeric regions. We also noted that COmapper detects fewer crossovers at sub-telomeric regions than GBS (**Figure 4C** and **D**). The mean SNP density was lower at sub-telomeric regions than in mid-chromosomal regions (**Supplemental Figure 7A** and **B**). Additionally, the physical distances covered by the eight SNPs required for crossover-site mapping by COmapper were shorter than those identified by GBS in the sub-telomeric, mid-chromosomal and pericentromeric regions (**Supplemental Figure 7C**–**E**). These results indicate that COmapper has limitations in detecting crossovers, as it requires sufficiently long reads to allow genotyping of eight consecutive SNPs within each DNA molecule. Nevertheless, we demonstrate that COmapper can map genome-wide crossovers and reveal the genetic effect of the *recq4ab* double mutant on crossover frequency and distribution by nanopore sequencing of gDNA from pooled hybrid pollen grains.

Mapping of crossovers from pooled F_2 recombinant seedlings by COmapper

In *Arabidopsis*, GBS analysis of F_2 recombinant individuals from self-pollinated Col × L*er* F_1 hybrid plants revealed that each F_2 individual contains ~8 crossovers as a sex-average crossover number (Serra et al., 2018b; Fernandes et al., 2018; Rowan et al., 2019; Nageswaran et al., 2021; Kim et al., 2022; Kim et al., 2024). In agreement with the F_2 GBS results, sex-specific crossover mapping identified ~5 crossovers derived from male meiosis and ~3 crossovers from female meiosis using reciprocal backcrossed mapping populations (Durand et al., 2022; Kim et al., 2024). Using COmapper and nanopore long-read sequencing on the promethION platform that provides ~100–120 Gb of raw data per flow cell, we generated ~584.1x genomic coverage of sequencing data and detected 1,749 crossovers from the pooled gDNA of 51 F_2 recombinant seedlings from self-pollinated Col × L*er* F_1 hybrid, while GBS defined 329 crossovers from the same plant materials. Of these 329 crossovers, 27 separate crossovers (8.2%) were mapped once and 244 crossovers (74.2%) more than once (**Supplemental Figure 4B**). This suggests a possibility that COmapper, with high genome coverage of nanopore long-read sequencing, can detect predominantly non-overlapping crossovers from a pool of F_2 recombinants when more than 51 F_2 seedlings are pooled and sequenced, similar to how it detects all non-overlapping crossovers in pooled pollen grains. We thus hypothesized that COmapper should be able to map ~1,400 non-overlapping crossovers from a gDNA pool from 1,000 F_2 recombinants containing ~8,000 crossovers when sequencing to a depth of ~100 Gb by long-read sequencing. To test this hypothesis, we performed nanopore long-read sequencing of gDNA pooled from 1,000 Col × L*er* F_2 recombinant seedlings (10 days old) derived from WT Col × L*er* and *recq4ab* Col × *recq4ab* L*er* F_1 hybrid seedlings, and mapped their crossovers with COmapper (**Figure 5**). Much as the GBS analysis of Col × L*er* and *recq4ab* Col × *recq4ab* L*er* F_2 individuals (30-day-old plants) detecting a 3.67-fold increase in crossover numbers in the *recq4ab* background compared to WT Col × L*er* F_2 plants (**Figure 5A**) (Serra et al., 2018b; Kim et al., 2024), COmapper detected 3.9-fold more crossovers in the *recq4ab* Col × *recq4ab* L*er* F_2 seedlings compared to WT Col × L*er* F_2 seedlings (**Figure 5B**). COmapper generated a similar crossover landscape from the pooled F_2 seedlings (n = 1,000) as GBS did from adult F_2 individuals, with crossovers being highly enriched along chromosome arms and suppressed over pericentromeric regions in WT and *recq4ab* (Pearson′s correlation test, WT $R = 0.72$, *recq4ab* $R = 0.83$, all $P < 2.20 \times 10^{-16}$) (**Figure 5C–G**). COmapper detected more crossovers along chromosome arms from the pooled *recq4ab* Col × *recq4ab* L*er* F_2 seedlings than were detected in WT Col × L*er* F_2 seedlings, while pericentromeric crossovers remained suppressed in WT and *recq4ab* F_2 seedlings (**Figure 5C** and **E**). Collectively, these results suggest that COmapper can be applied to a pool of hybrid F_2 progeny as well as to F_1 pollen for genome-wide mapping of crossovers and distribution in *Arabidopsis*.

High-resolution mapping of crossovers around genes

In *Arabidopsis*, genome-wide meiotic DSB profiling by sequencing SPO11-1–oligonucleotides (SPO11-1–oligos) revealed that meiotic DSBs occur predominantly in nucleosome-depleted regions, with the highest peak observed over gene-proximal promoters, while the formation of DSBs at the exons of protein-coding genes with high nucleosome density is suppressed (**Figure 6A**) (Underwood et al., 2018; Choi et al., 2018). In agreement with this result, meiotic crossovers in plants are highly enriched around gene TSS (Hellsten et al., 2013; Wijnker et al., 2013; Choi et al., 2013; Choi and Henderson, 2015a; Shilo et al., 2015; Kianian et al., 2018). Because we generated high-resolution crossover maps using COmapper, we examined the position and frequency of crossovers around the TSS and transcription end sites (TES) of all genes in a 4-kb window (2 kb on either side of the TSS or TES). Like meiotic DSBs that were previously determined by sequencing of SPO11-1–oligos (Choi et al., 2018), the crossovers mapped by COmapper in gDNA pooled from WT Col × L*er* F_1 hybrid pollen showed higher enrichment in gene-proximal promoters but lower enrichment in gene bodies, compared to the same number of random sites (**Figure 6A, B** and **Supplemental Figure 8A, B**). The pattern of crossovers determined by our GBS dataset, which had ~5.2x genome coverage for male meiosis (Kim et al., 2024), showed a similar enrichment around gene TSS as that mapped by COmapper in Col × L*er* F_1 hybrid pollen pool (**Figure 6B** and **Supplemental Figure 8B**). Furthermore, COmapper and GBS with ~4.5x genome coverage both detect a higher enrichment of crossovers around gene-proximal promoters and a lower enrichment in gene bodies in WT Col × L*er* F_2 recombinant seedlings compared to random (**Figure 6C** and **Supplemental Figure 8C**), suggesting that chromatin accessibility is a major determinant in shaping crossovers as well as meiotic DSBs around genes in *Arabidopsis*.

The GBS with sparse coverage (~0.1–1x) and TIGER, a GBS pipeline for mapping crossovers are inexpensive and useful for measuring the number of crossovers and interference from many recombinant individuals (Rowan et al., 2015; Rowan et al., 2019). However, when crossovers mapped by GBS and TIGER at different genome coverages (0-1x, 1-2x, and ≥2x) were plotted around TSS and TES for finer scales by analyzing the distance between two SNPs at the crossover site, we found that high-resolution mapping of crossovers by GBS requires a genome coverage of at least 2x, which is comparable to COmapper enabling detection of crossovers at SNP resolution (**Supplemental Figure 9**).

SNP density promotes interfering crossovers and inhibits non-interfering crossovers

We examined the effect of the *recq4ab* double mutant on the pattern of crossovers mapped by COmapper or GBS. The loss of RECQ4A and RECQ4B function resulted in a decreased crossover frequency at the gene TSS and TES, with an increase in crossovers at gene bodies (**Figure 6D** and **Supplemental Figure 8D**). Genetic disruption of *MSH2* was shown to further increase the frequency of class II crossovers in Col × L*er* F_2 progeny in the *recq4ab* background, suggesting that a high density of DNA polymorphisms inhibits the formation of class II crossovers in *recq4ab* (Dluzewska et al., 2023). Consistent with this possibility, we detected peaks for high SNP density between the Col and L*er* accessions at gene TSS and TES, with lower SNP density at distal promoters, gene bodies and random sites (**Figure 6E**). We also observed that the physical distance between two consecutive SNPs at crossover sites is greater in *recq4ab* than in WT based on nanopore long-read sequencing of F_1 pollen and F_2 progeny (**Figure 6F**), suggesting that the increased numbers of class II crossovers in *recq4ab* are sensitive to crossover inhibition mediated by high SNP density and occur frequently at the less dense SNP regions compared to class I crossovers in WT.

To further investigate the effect of SNP density on the distribution of class I and class II crossovers, we plotted the density of SNPs in a 2-kb window, upstream and downstream of the center positions of crossovers mapped by COmapper and GBS, in the WT Col × L*er* and *recq4ab* Col × *recq4ab* L*er* backgrounds (**Figure 6G**). The density of SNPs around the crossovers was higher in WT but lower in the *recq4ab* background compared to that at random sites, with the lowest density seen at the center of crossover positions in both WT and *recq4ab* (**Figure 6G, H**). We also found a similarly higher density of SNPs around crossover sites compared to random sites in the GBS analysis of F_2 individuals from Col expressing the $J3^{G155R}$ allele (*J3pro:*$J3^{G155R}$) × L*er* and Col × Di-G F_1 hybrids (**Supplemental Figure 10**). $J3^{G155R}$, a dominant negative allele of *J3* encoding a HSP40 co-chaperone, increases class I crossovers by enhancing the abundance of HEI10 (Kim et al., 2024). These results suggest that the class I crossovers occur preferentially in the regions with higher SNP density compared to random sites, whereas the increased class II crossovers due to the loss of RECQ4A and RECQ4B function occur mainly in regions with lower SNP density. It is worth noting that the SNP density nearest to the crossover site is lower for both class I and class II crossovers, which is consistent with previous observations that high-density SNPs inhibit crossovers at the crossover hotspot scale (Borts and Haber, 1987; Serra et al., 2018a; Szymanska-Lejman et al., 2023). Taken together, these results indicate that COmapper allows the generation of high-resolution genome-wide crossover maps and reveals the effects of SNP density and genetic mutation of *RECQ4A* and *RECQ4B* on the positioning and frequency of crossovers.

Discussion

Next-generation and linked-read sequencing technologies have contributed to our understanding of the patterning of genome-wide meiotic crossovers in diverse species (Rowan et al., 2015; Serra et al., 2018b; Kianian et al., 2018; Sun et al., 2019; Rommel Fuentes et al., 2020a; Nageswaran et al., 2021; Kim et al., 2022; Durand et al., 2022; Castellani et al., 2024; Kim et al., 2024). Here, using nanopore long-read sequencing (Wang et al., 2021), we developed COmapper that can generate high-resolution maps of meiotic crossovers from pooled pollen and F_2 progeny of F_1 hybrid plants (**Figures 3–6**). We demonstrated that COmapper can map crossovers along chromosomes at high resolution by validating the detection of crossovers in genomic DNA pooled from 4 or 51 recombinant F_2 plants, in which crossovers per each plant were mapped using the GBS approach with individual barcoding.

At a chromosomal scale, COmapper generated genome-wide crossover landscapes comparable to those of GBS in wild-type and *recq4ab* F_1 hybrid pollen and F_2 seedlings from self-pollinated F_1 hybrids, with high enrichment at chromosome arms and suppression at pericentromeres and centromeres (**Figures 4** and **5**). At a finer scale, COmapper revealed a significant enrichment of crossovers at gene promoters and a drop in crossovers within gene bodies, similar to the pattern observed at meiotic DSBs (**Figure 6A–C**). In addition, the analysis of SNP density around crossovers mapped by COmapper in 2-kb windows on either side of crossovers showed that the SNP density was correlated with the number of class I crossovers, but inversely correlated with that of class II crossovers, suggesting opposite responses of the two crossover pathways to SNP density (**Figure 6G** and **H**). These results support the recent observations showing a dual role for MSH2 in promoting class I crossovers and inhibiting class II crossovers in response to polymorphisms (Ziolkowski et al., 2015; Blackwell et al., 2020; Dluzewska et al., 2023; Szymanska-Lejman et al., 2023). Indeed, the *msh2* mutation leads to a further increase in class II crossovers in *recq4ab* while decreasing the number of class I crossovers that were increased in heterozygous regions juxtaposed to homozygous wild-type regions (Dluzewska et al., 2023).

We did observe that COmapper could mis-map a subset of crossovers occurring within large regions with few SNPs, because direct mapping of crossovers by COmapper depends on SNP density and read length, ensuring that each read covers a sufficient number of SNPs in each long read molecule (**Figures 3H, 4D** and **Supplemental Figure 7**). The number of crossovers mapped by COmapper correlated with the sequencing depth of nanopore long-read sequencing. Thus, a deeper nanopore sequencing depth leads to the mapping of more

crossovers for the assembly of genome-wide crossover landscapes, because the majority of aligned sequencing reads represent the parental genotypes (~99.99%), and only ~0.0135% of sequencing reads are crossover recombinant types in Col × L*er* hybrid pollen (**Figure 4** and **Supplemental Table 1**). COmapper does not allow measurement of crossover interference, which requires mapping multiple crossovers along the same chromosome and measuring their physical distance, as can be done in GBS for individual recombinants.

In contrast to GBS, COmapper requires a simple gDNA extraction method and pooled gDNA from F_1 hybrid pollen or F_2 seedlings for nanopore sequencing, instead of gDNA extraction from individual plants and barcoding. Therefore, we propose that COmapper can be widely used to map genome-wide crossovers at high resolution in the following types of studies. First, COmapper should be applicable for performing a rapid and high-throughput screen to detect changes in the number and distribution of meiotic crossovers in crosses between different accessions in *Arabidopsis* and various species. Therefore, COmapper should allow testing for the effects of genetic, epigenetic and environmental changes on crossover patterning. Second, COmapper should be able to detect crossovers in diverse polyploid organisms. Two different autotetraploid *Arabidopsis* accessions, such as natural Wa-1 and Bla-5 or colchicine-doubled tetraploid Col (4x) and L*er* (4x), can be crossed to produce hybrid tetraploid plants and their pollen or progeny can be pooled and sequenced by nanopore sequencing, after which the crossovers can be directly mapped by COmapper. Similarly, COmapper can be useful for analyzing crossover pattern in allopolyploid plants, such as wheat and rapeseed. Third, COmapper should be useful to map genome-wide crossovers in pollen or progeny of outcrossing species, along with calling parental SNPs by nanopore sequencing. Finally, COmapper should allow the measurement of crossover frequency and distribution in meiotic recombination hotspots. Instead of PCR amplification of single ~5–10-kb molecules from a recombination hotspot (Choi et al., 2013), genotyping of pollen or sperm can be performed in pooled samples and combined with nanopore long-read sequencing. Therefore, given the simplicity of sample preparation and the high accuracy in mapping crossovers by long-read nanopore sequencing, we believe that COmapper will be extensively used to understand the mechanisms of meiotic crossover patterning in diverse species beyond plants.

Methods

Plant materials

The *Arabidopsis thaliana* accessions Columbia-0 (Col) and Landsberg *erecta*-0 (L*er*) were used as the wild types for Col × L*er* hybrid control plants. F_1 and F_2 hybrid plants were cultivated in a growth room under the controlled conditions of 20°C, 50–60% relative humidity and 16-h light/8-h dark photoperiod. The *recq4a recq4b* mutants in the Col and L*er* backgrounds were used as previously described (Fernandes et al., 2018). *J3pro:J3*^{G155R} in the Col background was used as described (Kim et al., 2024). The *Arabidopsis thaliana* accession Di-G (Dijon, NASC stock number: N28205) was obtained from the Arabidopsis Biological Resource Center.

Genotyping by sequencing (GBS)

GBS libraries were prepared as previously described (Rowan et al., 2015; Rowan et al., 2019; Nageswaran et al., 2021; Kim et al., 2024). Briefly, genomic DNA (gDNA) was extracted from two or three mature leaves per individual F_2 plant or backcrossed F_1 plant using the CTAB method. Leaves were collected into 2-ml tubes containing four 3-mm stainless steel beads, frozen in liquid nitrogen and ground using a mixer mill (Retsch, MM400). Then, 500 µl of CTAB buffer (1% [w/v] CTAB, 50 mM Tris-HCl pH 8.0, 0.7 M NaCl, 10 mM EDTA) was added to the ground tissue, mixed vigorously and incubated at 65°C for 30 min. Following a chloroform extraction and gDNA precipitation with the addition of isopropanol, the gDNA pellet was washed with 1 ml 70% (v/v) ethanol twice, air-dried and resuspended in nuclease-free water. After removal of RNA using RNase A digestion and precipitation with ethanol, 150 ng of gDNA was used to generate one sequencing library per plant. The gDNA was fragmented with dsDNA Shearase (Zymo Research, E2018-200) and size selected to ~300 bp using AMPure XP magnetic SPRI beads (Beckman-Coulter, A63881). Fragmented gDNA was blunt-ended, A-tailed and ligated with barcoded adapters compatible with Illumina sequencing using a Quick Ligase Kit (NEB, M2200L). Each of the eight barcoded DNA libraries was pooled, size selected and amplified by 12–14 PCR cycles using a KAPA HiFi HotStart ReadyMix kit (Roche, KK2601). The PCR products were then purified using AMPure XP magnetic SPRI beads and quantified using a Bioanalyzer. The barcoded libraries were then pooled and subjected to paired-end 150-bp sequencing using an Illumina HiSeq-X platform (Macrogen, Korea). The TIGER pipeline was used to analyze the sequencing data and identify crossover sites (Rowan et al., 2015).

Preparation of genomic DNAs for nanopore long-read sequencing

To purify pollen grains from F_1 hybrid plants and extract their genomic DNA, F_1 plants were planted in soil and grown at 20°C under a 16-h light/8-h dark photoperiod. Whole inflorescences of these hybrid plants were collected and transferred to ice-cold 10% (w/v)

sucrose in a 50-ml tubes. Flowers were homogenized with a blender (Tefal, BL3051KR), filtered through 80-μm nylon mesh, and centrifuged at 350*g* for 10 min at 4 °C. The supernatant was discarded and the pellet was resuspended in 10% (w/v) sucrose, filtered through a 40-μm cell strainer, and centrifuged at 100*g* for 10 min at 4°C. The purified pellet was frozen in liquid nitrogen and ground with a mortar and pestle. The pellet powder was resuspended in CTAB buffer (1% [w/v] CTAB, 50 mM Tris-HCl pH 8.0, 0.7 M NaCl, 10 mM EDTA). Proteinase K (NEB, P8107S) and RNase A (20 mg/ml, Roche, 10109142001) were added to a final concentration of 80 U/ml and 20 μg/ml, respectively. To check pollen disruption, a 1-μl sample was diluted in a 1:10 ratio with 10% (w/v) sucrose and observed under a stereomicroscope (Leica, M165 FC). Then, 500 μl of the resuspended pellet was transferred to a 1.5-ml tube, to which an equal volume of 25:24:1 (v/v/v) phenol:chloroform:isoamylalcohol (Sigma, 77617) was added, and the mixture was homogenized by gentle manual shaking. The tube was centrifuged at 15,000*g* for 10 min at 4°C, and 400 μl of supernatant was transferred into a new tube, to which a Nanobind disk (Nanobind® Plant Nuclei Kit RT, PacBio, 102-302-000) was added. The tube containing the Nanobind disk was gently mixed, 400 μl isopropanol was added, and the tube was slowly inverted and incubated for 2 h at room temperature on a rotator set at 7 rpm. A magnetic rack (DynaMag™-2, Invitrogen, 12321D) was used to separate the disk and the supernatant, and the supernatant was removed. PW1 buffer in the Nanobind Plant Nuclei Kit was used for washing the disk twice and the supernatant was removed completely by centrifugation. Then, 150 μl EB+ buffer was added onto the disc and gently tabbed and incubated overnight at room temperature. The supernatant was transferred to a new 1.5-ml tube and stored at −20°C. The DNA was quantified using a Qubit dsDNA Broad Range assay kit (Thermo Fisher, Q32853) and integrity checked by gel electrophoresis.

For nanopore sequencing of gDNA from 1,000 pooled seedlings, 10-day-old seedlings were ground in liquid nitrogen using a mortar and pestle. The ground tissue was resuspended in four volumes of CTAB buffer (1% [w/v] CTAB, 50 mM Tris-HCl pH 8.0, 0.7 M NaCl, 10 mM EDTA) and incubated at 65°C for 30 min. Following chloroform extraction, isopropanol precipitation and removal of RNAs as above, the gDNA pellet was resuspended in 150 μl TE (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA) buffer and gDNA was quantified using a Qubit dsDNA Broad Range assay kit (Thermo Fisher, Q32853). Nine micrograms of gDNA from pollen or seedlings was used to construct a nanopore long-read sequencing library using a Ligation Sequencing Kit V14 (Nanopore, SQK-LSK114). The libraries were sequenced using a PromethION platform (BGI, Hong Kong).

Alignment of nanopore sequencing reads to the reference genome

The FASTQ files containing the nanopore sequencing reads were aligned to the TAIR10 reference genome sequence using minimap2 with the following settings "-ax map-ont -cs"(Li, 2018). To obtain uniquely aligned long reads, sambamba was employed with the following settings: "-F "not unmapped and not duplicate and mapping_quality >= 30 and sequence_length >= 100""(Tarasov et al., 2015). Bases with a quality (QUAL) score <14 were replaced with 'N' using an awk script, thereby addressing genotyping errors that were a consequence of sequencing errors.

COmapper pipeline

To obtain high-quality SNPs distinguishing the Col and L*er* accessions, we followed the previous method with modifications (Fernandes et al., 2024). SNPs within syntenic regions were identified by comparing the reference genomes of Col and Ler using SyRI (Goel et al., 2019). These SNPs were validated using sequencing data from a total of 2,160 Col × L*er* F_2 plants were used (Rowan et al., 2019). To retain only high-quality SNPs, the following criteria were applied: (1) genotyping results available for at least 40 F_2 plants, (2) allele frequency between 0.4 and 0.6, and (3) adherence to Mendelian segregation across the 2,160 F_2 individuals (chi-square test, *P*-value > 0.1). These sites were then integrated with an independent set of high-quality SNP calls from the previous report (Yelina et al., 2015).

The genotypes for the SNPs in masked alignments were converted to string format using a reference list of 226,684 high-quality Col/L*er* SNPs. Genotype strings with two successive identical genotypes at either end (CC-CC or LL-LL) were identified as parental reads. Genotype strings that satisfied the following conditions were identified as crossover reads: The two successive SNP genotypes at one end differed from those at the other end (CC-LL or LL-CC), and the four successive SNP genotypes before and after the crossover site matched the genotype at the respective end of the genotype string (CCCC-LLLL or LLLL-CCCC). Genotype strings with less than eight characters or those not identified as either crossover or parental reads were excluded from the calculation of crossover frequency. Regions exhibiting a depth of coverage exceeding the average genome read depth by three-fold and exhibiting a high frequency of false positives in the negative control provided by the sequencing of Col × L*er* F_1 leaf gDNA were excluded from the analysis. The standard output provided the input alignment information, the genotyping summary, the total number of crossover and parental reads, the length of the crossover and parental reads and the information regarding crossover reads in a CSV file.

Bioinformatic analysis of crossovers and genomic features.

To analyze the genome-wide pattern of crossovers as well as genes, transposable elements, DNA methylation, SPO11-1–oligos, and MNase-seq (micrococcal nuclease digestion with deep sequencing) data, the TAIR10 genome was tiled into 100-kb bins. DNA methylation levels in each bin were used for plotting. The coverage of SPO11-1–oligo and MNase-seq data was normalized by sequencing reads of randomly digested gDNA before plotting (Choi et al., 2018). The number of overlaps for genes, transposable elements, and crossovers within each bin was calculated and plotted. The coverage of SPO11-1–oligo and MNase-seq data and the frequency of crossovers around genes were calculated using deepTools computeMatrix (v. 3.5.4) to generate gene metaplots. Regions within the pericentromeric regions and where the distance between consecutive SNPs was >4,000 bp were excluded from all analyses to increase resolution of crossover sites. For the analysis of SNP density around crossover site, the SNP frequency in 10-bp bins within a 4-kb window centered on each crossover site was calculated using deepTools computeMatrix (v. 3.5.4).

Data and code availability

The nanopore sequencing data generated in this study have been deposited in the ArrayExpress database at EMBL-EBL (http://www.ebi.ac.uk/arrayexpress). Detailed information on the nanopore long-read sequencing and other data for GBS and genomic features used in this study is included in **Supplemental Table 2**. Source data are provided with this paper. All custom code used in this study, including the COmapper pipeline, is available at GitHub (https://github.com/KyuhaChoi-Lab/COmapper).

Funding

This work was funded by a Samsung Science and Technology Foundation grant (SSTF-BA2202-09) and the National Research Foundation of Korea (RS-2024-00335818, RS-2024-00407469).

Author contributions

D.B., N.S. and K.C. designed the project. D.B., N.S., and J.K. developed COmapper pipeline. D.B., N.S., H.K., J.K., J.P., S.P., H.K., J.K., S.L, Y.H., P.A.Z., I.R.H, and K.C. performed

experiments and analysis. D.B., N.S. and K.C. wrote the original draft of manuscript. All authors revised the manuscript.

Acknowledgements

We thank Raphael Mercier (MPI for Plant Breeding Research) for providing *recq4ab* seeds. The authors declare no competing interests.

References

Berchowitz, L. E., Francis, K. E., Bey, A. L., and Copenhaver, G. P. (2007). The Role of AtMUS81 in Interference-Insensitive Crossovers in A. thaliana. *PLoS Genet.* **3**:e132.

Blackwell, A. R., Dluzewska, J., Szymanska- Lejman, M., Desjardins, S., Tock, A. J., Kbiri, N., Lambing, C., Lawrence, E. J., Bieluszewski, T., Rowan, B., et al. (2020). MSH 2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis . *EMBO J.* **39**:1–22.

Borts, R. H., and Haber, J. E. (1987). Meiotic recombination in yeast: alteration by multiple heterozygosities. *Science* **237**:1459–65.

Castellani, M., Zhang, M., Thangavel, G., Mata-Sucre, Y., Lux, T., Campoy, J. A., Marek, M., Huettel, B., Sun, H., Mayer, K. F. X., et al. (2024). Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning. *Nat. Plants* **10**:423–438.

Choi, K., and Henderson, I. R. (2015a). Meiotic recombination hotspots - A comparative view. *Plant J.* **83**:52–61.

Choi, K., and Henderson, I. R. (2015b). Meiotic recombination hotspots - a comparative view. *Plant J.* **83**:52–61.

Choi, K., Zhao, X., Kelly, K. A., Venn, O., Higgins, J. D., Yelina, N. E., Hardcastle, T. J., Ziolkowski, P. A., Copenhaver, G. P., Franklin, F. C. H., et al. (2013). Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* **45**:1327–1336.

Choi, K., Reinhard, C., Serra, H. H., Ziolkowski, P. A. P. A. P., Underwood, C. J. C. J., Zhao, X., Hardcastle, T. J. T. J. T., Yelina, N. E. N. N. E., Griffin, C., Jackson, M., et al. (2016). Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes. *PLoS Genet.* **12**:e1006179.

Choi, K., Zhao, X., Tock, A. J. A. J., Lambing, C., Underwood, C. J. C. J., Hardcastle, T. J. T. J., Serra, H., Kim, J. J., Cho, H. S. H. S., Kim, J. J., et al. (2018). Nucleosomes and DNA methylation shape meiotic DSB frequency in Arabidopsis thaliana transposons and gene regulatory regions. *Genome Res.* **28**:532–546.

Crismani, W., Girard, C., Froger, N., Pradillo, M., Santos, J. L., Chelysheva, L., Copenhaver, G. P., Horlow, C., and Mercier, R. (2012). FANCM Limits Meiotic Crossovers. *Science* **336**:1588–1590.

Dluzewska, J., Dziegielewski, W., Szymanska-Lejman, M., Gazecka, M., Henderson, I. R., Higgins, J. D., and Ziolkowski, P. A. (2023). MSH2 stimulates interfering and inhibits non-interfering crossovers in response to genetic polymorphism. *Nat. Commun.* **14**.

Durand, S., Lian, Q., Jing, J., Ernst, M., Grelon, M., Zwicker, D., and Mercier, R. (2022). Joint control of meiotic crossover patterning by the synaptonemal complex and HEI10 dosage. *Nat. Commun.* **13**.

Fernandes, J. B., Séguéla-Arnaud, M., Larchevêque, C., Lloyd, A. H., and Mercier, R. (2018). Unleashing meiotic crossovers in hybrid plants. *Proc. Natl. Acad. Sci. U. S. A.* **115**:2431–2436.

Fernandes, J. B., Naish, M., Lian, Q., Burns, R., Tock, A. J., Rabanal, F. A., Wlodzimierz, P., Habring, A., Nicholas, R. E., Weigel, D., et al. (2024). Structural variation and DNA methylation shape the centromere-proximal meiotic crossover landscape in Arabidopsis. *Genome Biol.* **25**:1–31.

Girard, C., Zwicker, D., and Mercier, R. (2023). The regulation of meiotic crossover distribution: a coarse solution to a century-old mystery? *Biochem. Soc. Trans.* **51**:1179–1190.

Goel, M., Sun, H., Jiao, W. B., and Schneeberger, K. (2019). SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**:1–13.

Gray, S., and Cohen, P. E. (2016). Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation. *Annu. Rev. Genet.* **50**:175–210.

Grelon, M., Vezon, D., Gendrot, G., and Pelletier, G. (2001). AtSPO11-1 is necessary for ef®cient meiotic recombination in plants. *EMBO J.* **20**:589–600.

He, Y., Wang, M., Dukowic-Schulze, S., Zhou, A., Tiang, C.-L., Shilo, S., Sidhu, G. K., Eichten, S., Bradbury, P., Springer, N. M., et al. (2017). Genomic features shaping the landscape of meiotic double-strand-break hotspots in maize. *Proc. Natl. Acad. Sci. U. S. A.* **114**:12231–12236.

Hellsten, U., Wright, K. M., Jenkins, J., Shu, S., Yuan, Y., Wessler, S. R., Schmutz, J., Willis, J. H., and Rokhsar, D. S. (2013). Fine-scale variation in meiotic recombination in Mimulus inferred from population shotgun sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **110**:19478–82.

Henderson, I. R., and Bomblies, K. (2021). Evolution and Plasticity of Genome-Wide Meiotic Recombination Rates Advance Access published 2021.

Hunter, N. (2015). Meiotic recombination: The essence of heredity. *Cold Spring Harb. Perspect. Biol.* **7**:1–36.

Keeney, S., Lange, J., and Mohibullah, N. (2014). Self-organization of meiotic recombination initiation: general principles and molecular pathways. *Annu. Rev. Genet.* **48**:187–214.

Kianian, P. M. A., Wang, M., Simons, K., Ghavami, F., He, Y., Dukowic-Schulze, S., Sundararajan, A., Sun, Q., Pillardy, J., Mudge, J., et al. (2018). High-resolution crossover mapping reveals similarities and differences of male and female recombination in maize. *Nat. Commun.* **9**:2370.

Kim, H., and Choi, K. (2022). Fast and Precise : How to Measure Meiotic Crossovers in Arabidopsis. *Mol. Cells* **45**:273–283.

Kim, J., Park, J., Kim, H., Son, N., Kim, E., Kim, J. J., Byun, D., Lee, Y., Park, Y. M., Nageswaran, D. C., et al. (2022). Arabidopsis HEAT SHOCK FACTOR BINDING PROTEIN is required to limit meiotic crossovers and HEI10 transcription. *EMBO J.* **41**:e109958.

Kim, H., Kim, J., Son, N., Kuo, P., Morgan, C., Chambon, A., Byun, D., Park, J., Lee, Y., Park, Y. M., et al. (2024). Control of meiotic crossover interference by a proteolytic chaperone network. *Nat. Plants* **10**:453–468.

Lam, I., and Keeney, S. (2014). Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harb. Perspect. Biol.* **7**:a016634.

Lawrence, E. J., Gao, H., Tock, A. J., Lambing, C., Blackwell, A. R., Feng, X., and Henderson, I. R. (2019). Natural Variation in TBP-ASSOCIATED FACTOR 4b Controls Meiotic Crossover and Germline Transcription in Arabidopsis. *Curr. Biol.* **29**:2676-2686.e3.

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**:3094–3100.

Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N., and Grelon, M. (2015). The Molecular Biology of Meiosis in Plants. *Annu. Rev. Plant Biol.* **66**:297–327.

Mieulet, D., Aubert, G., Bres, C., Klein, A., Droc, G., Vieille, E., Rond-Coissieux, C., Sanchez, M., Dalmais, M., Mauxion, J.-P., et al. (2018). Unleashing meiotic crossovers in crops. *Nat. Plants* **4**:1010–1016.

Morgan, C., Fozard, J. A., Hartley, M., Henderson, I. R., Bomblies, K., and Howard, M. (2021). Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in Arabidopsis. *Nat. Commun.* **12**:4674.

Nageswaran, D. C., Kim, J., Lambing, C., Kim, J., Park, J., Kim, E.-J., Cho, H. S., Kim, H., Byun, D., Park, Y. M., et al. (2021). HIGH CROSSOVER RATE1 encodes PROTEIN PHOSPHATASE X1 and restricts meiotic crossovers in Arabidopsis. *Nat. Plants* **7**:452–467.

Naish, M., Alonge, M., Wlodzimierz, P., Tock, A. J., Abramson, B. W., Schmücker, A., Mandáková, T., Jamge, B., Lambing, C., Kuo, P., et al. (2021). The genetic and epigenetic landscape of the Arabidopsis centromeres. *Science* **374**:eabi7489.

Neale, M. J., Pan, J., and Keeney, S. (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**:1053–1057.

Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Guichard, V., Bourbon, H.-M., and de Massy, B. (2016). The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* **351**:943–949.

Rommel Fuentes, R., Hesselink, T., Nieuwenhuis, R., Bakker, L., Schijlen, E., van Dooijeweert, W., Diaz Trivino, S., de Haan, J. R., Sanchez Perez, G., Zhang, X., et al. (2020a). Meiotic recombination profiling of interspecific hybrid F1 tomato pollen by linked read sequencing. *Plant J.* **102**:480–492.

Rommel Fuentes, R., Hesselink, T., Nieuwenhuis, R., Bakker, L., Schijlen, E., van Dooijeweert, W., Diaz Trivino, S., de Haan, J. R., Sanchez Perez, G., Zhang, X., et al. (2020b). Meiotic recombination profiling of interspecific hybrid F1 tomato pollen by linked read sequencing. *Plant J.* **102**:480–492.

Rowan, B. A., Patel, V., Weigel, D., and Schneeberger, K. (2015). Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping. *G3 Genes, Genomes, Genet.* **5**:385–398.

Rowan, B. A., Heavens, D., Feuerborn, T. R., Tock, A. J., Henderson, I. R., and Weigel, D. (2019). An Ultra High-Density Arabidopsis thaliana Crossover Map That Refines the Influences of Structural Variation and Epigenetic Features. *Genetics* **213**:771–787.

Séguéla-Arnaud, M., Crismani, W., Larchevêque, C., Mazel, J., Froger, N., Choinard, S., Lemhemdi, A., Macaisne, N., Van Leene, J., Gevaert, K., et al. (2015). Multiple mechanisms limit meiotic crossovers: TOP3α and two BLM homologs antagonize crossovers in parallel to FANCM. *Proc. Natl. Acad. Sci. U. S. A.* **112**:4713–4718.

Serra, H., Choi, K., Zhao, X., Blackwell, A. R., Kim, J., and Henderson, I. R. (2018a). Interhomolog polymorphism shapes meiotic crossover within the Arabidopsis RAC1 and RPP13 disease resistance genes. *PLoS Genet.* **14**:1–28.

Serra, H., Lambing, C., Griffin, C. H., Topp, S. D., Nageswaran, D. C., Underwood, C. J., Ziolkowski, P. A., Séguéla-Arnaud, M., Fernandes, J. B., Mercier, R., et al. (2018b). Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. *Proc. Natl. Acad. Sci. U. S. A.* **115**:2437–2442.

Shilo, S., Melamed-Bessudo, C., Dorone, Y., Barkai, N., and Levy, A. A. (2015). DNA Crossover Motifs Associated with Epigenetic Modifications Delineate Open Chromatin Regions in Arabidopsis. *Plant Cell* **27**:2427–36.

Sun, H., Rowan, B. A., Flood, P. J., Brandt, R., Fuss, J., Hancock, A. M., Michelmore, R. W., Huettel, B., and Schneeberger, K. (2019). Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nat. Commun.* **10**:4310.

Szymanska-Lejman, M., Dziegielewski, W., Dluzewska, J., Kbiri, N., Bieluszewska, A., Poethig, R. S., and Ziolkowski, P. A. (2023). The effect of DNA polymorphisms and natural variation on crossover hotspot activity in Arabidopsis hybrids. *Nat. Commun.* **14**:1–16.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., and Prins, P. (2015). Sambamba: Fast processing of NGS alignment formats. *Bioinformatics* **31**:2032–2034.

Underwood, C. J., and Choi, K. (2019). Heterogeneous transposable elements as silencers, enhancers and targets of meiotic recombination. *Chromosoma* **128**.

Underwood, C. J., Henderson, I. R., and Martienssen, R. A. (2017). Genetic and epigenetic variation of transposable elements in Arabidopsis. *Curr. Opin. Plant Biol.* **36**:135–141.

Underwood, C. J., Choi, K., Lambing, C., Zhao, X., Serra, H., Borges, F., Simorowski, J., Ernst, E., Jacob, Y., Henderson, I. R. I. R. I. R., et al. (2018). Epigenetic activation of meiotic recombination near Arabidopsis thaliana centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Res.* **28**:519–531.

Villeneuve, A. M., and Hillers, K. J. (2001). Whence meiosis? *Cell* **106**:647–50.

Wang, Y., and Copenhaver, G. P. (2018). Meiotic Recombination: Mixing It Up in Plants. *Annu. Rev. Plant Biol.* **69**:577–609.

Wang, Y., Zhao, Y., Bollas, A., Wang, Y., and Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* **39**:1348–1365.

Wijnker, E., Velikkakam James, G., Ding, J., Becker, F., Klasen, J. R., Rawat, V., Rowan, B. A., de Jong, D. F., de Snoo, C. B., Zapata, L., et al. (2013). The genomic landscape of meiotic crossovers and gene conversions in Arabidopsis thaliana. *Elife* **2**:e01426.

Yelina, N. E., Lambing, C., Hardcastle, T. J., Zhao, X., Santos, B., and Henderson, I. R. (2015). DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. *Genes Dev.* **29**:2183–2202.

Zakharyevich, K., Tang, S., Ma, Y., and Hunter, N. (2012). Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* **149**:334–347.

Zickler, D., and Kleckner, N. (2023). Meiosis: Dances Between Homologs. *Annu. Rev. Genet.* **57**:1–63.

Ziolkowski, P. A., Berchowitz, L. E., Lambing, C., Yelina, N. E., Zhao, X., Kelly, K. A., Choi, K., Ziolkowska, L., June, V., Sanchez-Moran, E., et al. (2015). Juxtaposition of Heterozygous and Homozygous Regions Causes Reciprocal Crossover Remodelling via Interference During Arabidopsis Meiosis. *Elife* **4**:e03708.

Figure legends

Figure 1. Genome-wide mapping of crossovers using long-read nanopore sequencing.

(A) Diagrams illustrating the preparation of recombinant pollen from F_1 hybrids and F_2 plants to map crossovers. **(B)** Mapping of genome-wide crossovers in F_2 plants by genotyping-by-sequencing (GBS) using multiplex barcoding and the TIGER pipeline. **(C)** Mapping of crossovers in pollen DNA from F_1 hybrid plants using 10x Chromium Controller and linked-read sequencing. **(D)** Mapping of crossovers in F_1 pollen and F_2 plants using nanopore long-read sequencing and COmapper. **(E)** High-resolution mapping of crossover sites by analyzing single-nucleotide polymorphisms (SNPs) within long-read DNA molecules. Red C indicates a Columbia-0 (Col)-specific SNP; purple L indicates Landsberg *erecta*-0 (L*er*) SNP.

Figure 2. Development of COmapper to map crossovers using long-read sequencing.

(A) Procedure for crossover mapping using nanopore long-read sequencing and the COmapper pipeline. **(B)** Distribution of read length and number of pollen DNAs sequenced by nanopore long-read sequencing. **(C)** Density distribution of the lengths of DNA molecules sequenced by nanopore long-read sequencing and aligned to the *Arabidopsis* reference genome using Minimap2. **(D)** Coverage frequency of long reads from Col × L*er* F_1 pollen DNA along *Arabidopsis* chromosome 1. The vertical dashed line indicates the centromere. The pericentromeric region is highlighted in blue. **(E)** As in (**D**), but showing the density of SNPs between the Col and L*er* accessions, in 100-kb sliding windows. **(F)** Distribution of inter-SNP distances between consecutive SNPs differentiating the Col and L*er* accessions. The vertical dashed line indicates the mean distance between consecutive SNPs. **(G)** Distribution of SNP number and frequency of SNPs per nanopore sequencing read. The vertical dashed line indicates the mean number of SNPs per read, and the vertical red line indicates the position corresponding to eight SNPs in a read.

Figure 3. Validation of COmapper using Col × L*er* F_2 recombinant plants.

(A) Diagram showing simple crossover (SCO) and gene-conversion-associated crossover (GC-CO) types. **(B)** Summary table of the tested COmapper models with different numbers of SNPs and minimum crossover (Min CO) and gene conversion (GC) tracts. **(C)** Number of crossovers detected per million reads in different COmapper models. True crossovers (blue) were detected by COmapper models among the crossovers mapped by GBS in 51 F_2 plants, and false crossovers (red) indicate the crossovers that differ from the crossovers defined by

GBS. **(D)** Precision (True positive/(True positive + False positive)) of crossover detection in the different COmapper models shown in (**C**). In (**C**) and (**D**), asterisks indicate significant differences between models (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; Welch's *t*-test). **(E)** Relative proportion of crossover types in 4 F_2 plants (genotyped by GBS), 51 F_2 plants (genotyped by GBS), Col-0 × L*er* F_1 hybrid pollen DNA, *recq4ab* Col × *recq4ab* L*er* hybrid pollen DNA, 1,000 Col-0 × L*er* F_2 seedlings and 1,000 *recq4ab* Col × *recq4ab* L*er* F_2 seedlings detected by the established model (8 CO) of COmapper. FP (red) indicates the proportion of false positive crossovers identified by COmapper in 4 and 51 F_2 plants compared to the crossovers identified by GBS. **(F)** Frequency of crossovers mapped by COmapper along the five *Arabidopsis* chromosomes on a continuous *x*-axis in 4 Col × L*er* F_2 recombinants, where crossovers in each F_2 plant were mapped by GBS. Blue bars represent true positive crossovers mapped by GBS; red bars represent false positive crossovers. **(G)** As in (**F**), but from 51 Col × L*er* F_2 recombinants. **(H)** Length density of nanopore sequencing reads containing eight SNPs for detected or non-detected crossovers in 4 and 51 Col × L*er* F_2 recombinant plants by COmapper. Vertical dashed lines indicate the median length (in bp) of eight SNP strings, 3.92 kb for detected crossovers (blue) and 8.31 kb (red) for non-detected crossovers, respectively.

Figure 4. Mapping of crossovers from pooled F_1 hybrid pollen DNA by COmapper.
(A) Number of crossovers per genome of Col × L*er* and *recq4ab* Col × *recq4ab* L*er* male meiosis in backcrossing populations (BCF_1)(Fernandes et al., 2018). Data are presented as means ± SD. *n* indicates the number of backcrossed F_1 plants. **(B)** Number of crossovers per genomic coverage of nanopore long-read sequencing data of gDNA from Col × L*er* F_1 leaves, Col × L*er* F_1 pollen and *recq4ab* Col × *recq4ab* L*er* F_1 pollen. Each data point indicates crossovers per genome from a randomly selected 100 reads. Data are presented as means ± SD from 10 replicates. **(C)** Normalized frequency of crossovers mapped by GBS (blue, n = 190; 852 crossovers) (cM/Mb) and COmapper (red, n = ~277; 1,246 crossovers) (crossovers/genome) along the five *Arabidopsis* chromosomes on a continuous *x*-axis in Col × L*er* male meiosis. **(D)** As in (**C**), but showing normalized crossover frequency (cM/Mb, crossovers/genome) along the chromosome telomere (TEL)-to-centromere (CEN) axis. **(E)** Correlation between crossover frequencies in 1-Mb bins of Col × L*er* BCF_1 individuals (measured by GBS) and Col × L*er* F_1 pollen (measured by COmapper) across the genome, analyzed using Pearson's correlation test. **(F)** As in (**C**), but showing the density of genes and transposable elements. **(G)** As in (**C**), but showing the density of SPO11–oligonucleotides (SPO11-1–oligo) and MNase sequencing data (MNase-seq). **(H)** As in (**C**), but showing the

density of crossovers mapped by COmapper in Col × L*er* and *recq4ab* Col × *recq4ab* L*er* F_1 pollen (*n* = ~159; 2,009 crossovers).

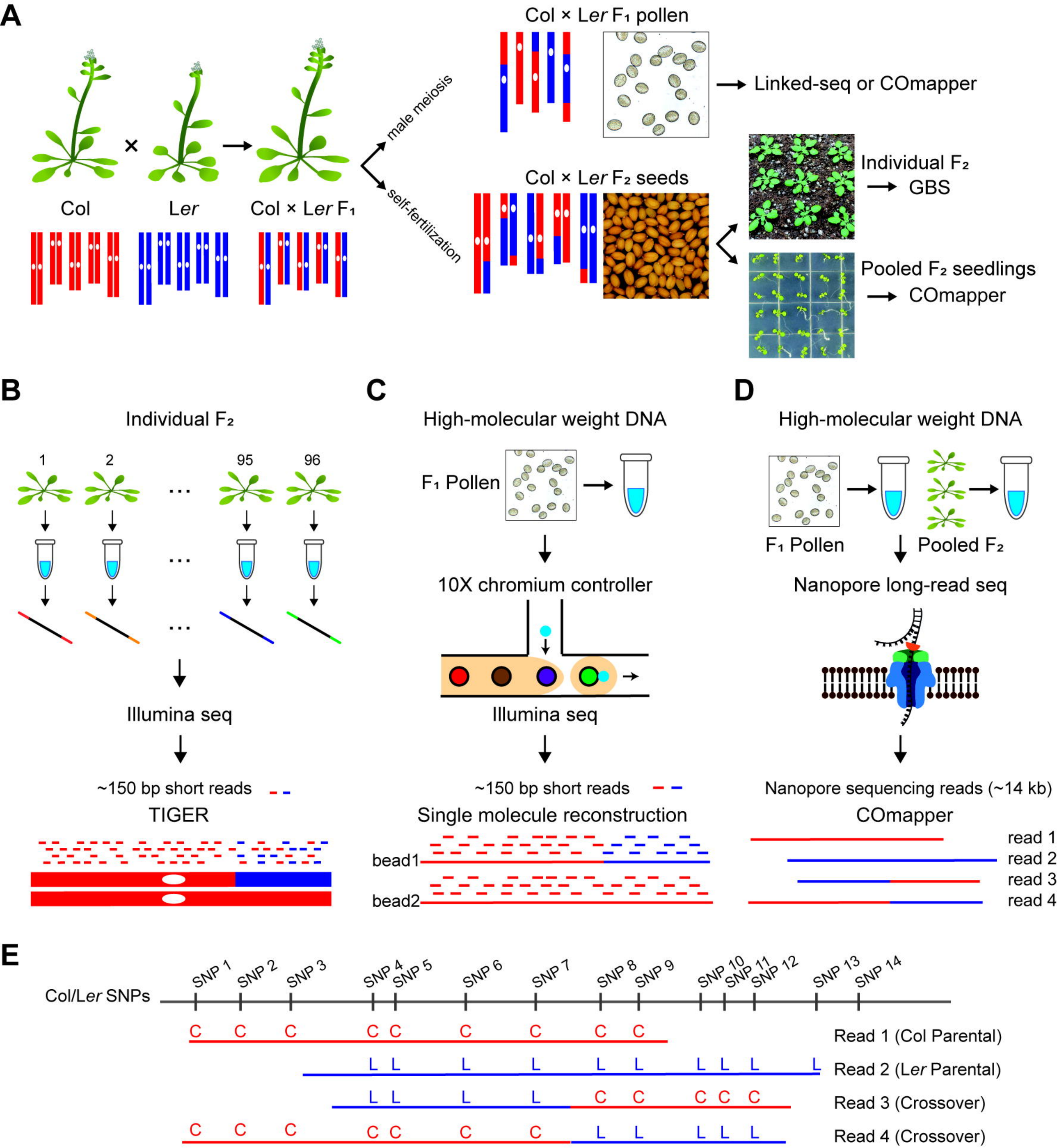
Figure 5. Mapping of crossovers from pooled F_2 recombinant seedlings by COmapper.

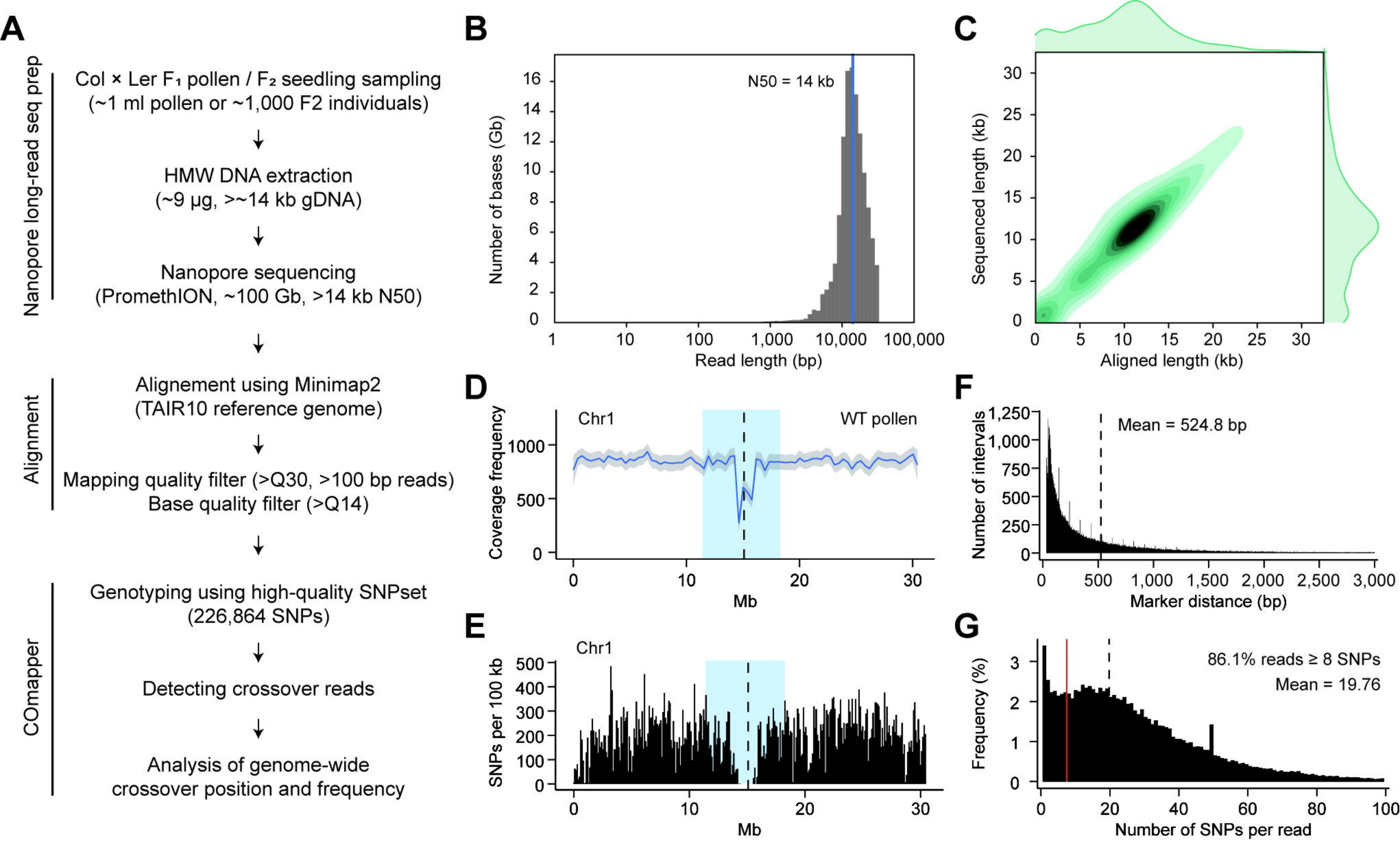
(A) Number of crossovers per F_2 individual calculated from GBS data of wild-type Col × L*er* F_2 (*n* = 240) and *recq4ab* Col × *recq4ab* L*er* F_2 (*n* = 96) plants. *n* indicates the number of F_2 individual plants. Data are presented as means ± SD. **(B)** Number of crossovers per genome of nanopore long-read sequencing data from gDNA of Col × L*er* F_1 leaves, Col × L*er* F_2 and *recq4ab* Col × *recq4ab* L*er* F_2 seedlings (*n* = 1,000). Each data point indicates the number of crossovers per genome from a randomly selected 1 million reads. Data are presented as means ± SD from 10 replicates. **(C)** Normalized frequency of crossovers mapped by GBS (blue, *n* = 240; 1,852 crossovers) (cM/Mb) and COmapper (red, *n* = ~179; 1,433 crossovers) (crossovers/genome) along the five *Arabidopsis* chromosomes on a continuous *x*-axis in WT Col × L*er* F_2 plants. *n* indicates the number of F_2 individual plants. **(D)** As in (**C**), but showing normalized crossover frequency (cM/Mb, crossovers/genome) along the chromosome telomere (TEL) to centromere (CEN) axis. **(E)** As in (**C**), but showing the crossovers mapped by GBS in *recq4ab* Col × *recq4ab* L*er* F_2 plants (*n* = 96; 2,731 crossovers) or by COmapper in *recq4ab* Col × *recq4ab* L*er* F_2 seedlings (*n* = ~192; 5,757 crossovers). **(F, G)** Correlation between crossover frequencies in 1-Mb bins of Col × L*er* F_2 seedlings (**F**) and *recq4ab* Col x *recq4ab* L*er* F_2 seedlings (**G**) measured by GBS and COmapper across the genome, analyzed using Pearson′s correlation test.

Figure 6. Analysis of crossovers and SNPs around genes.

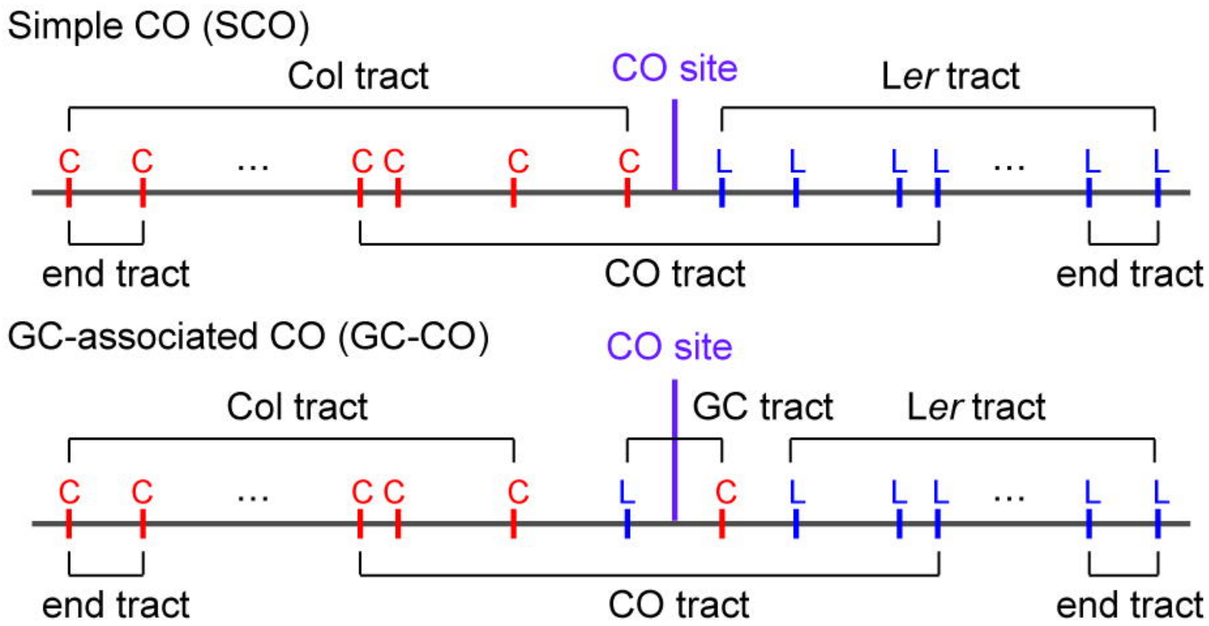
(A) Metaplots showing the density of SPO11-1–oligonucleotides (SPO11-1–oligo) and nucleosomes from MNase sequencing (MNase-seq) around gene transcription start sites (TSS) and transcription end sites (TES) in a 2-kb window. (**B)** As in (**A**), but showing the density of crossovers mapped by GBS (blue, 852 crossovers) and COmapper (red, 1,246 crossovers) in *Arabidopsis* Col × L*er* male meiosis. **(C)** As in (**B**), but showing the density of crossovers mapped by GBS (blue, 1,852 crossovers) and COmapper (red, 1,433 crossovers) in *Arabidopsis* Col × L*er* F_2 plants. **(D)** As in (**C**), but showing the density of crossovers mapped by GBS (skyblue, 2,731 crossovers) and COmapper (orange, 5,757 crossovers) in *Arabidopsis recq4ab* Col × *recq4ab* L*er* F_2 plants. **(E)** As for (**A**), but showing the density of SNPs between Col and L*er* accessions. **(F)** Density and physical distance between two

consecutive SNPs at crossover sites in WT Col × L*er* and *recq4ab* Col × *recq4ab* L*er* for F_1 pollen and F_2 progeny. The vertical dashed line indicates median distance (bp) between two consecutive SNPs at the crossover site. Significance between genotypes was examined using a Wilcoxon test. **(G)** Metaplots showing SNP density around crossovers mapped by GBS and COmapper in WT Col × L*er* F_2 and *recq4ab* Col × *recq4ab* L*er* F_2 progeny compared to random positions. **(H)** As in (**G**), but showing the significance of changes in SNP density according to distance from the crossover site. Each dot represents the SNP density per 40-bp bin. Significance between genotype and a random position per interval was examined using a Wilcoxon test. Red and blue dots indicate higher and lower SNP density than random, respectively; gray dots indicate non-significant differences.


A
Col × Ler F₁ pollen
Linked-seq or COmapper
male meiosis
Col
Ler
Col × Ler F₁
self-fertilization
Col × Ler F₂ seeds
Individual F₂
GBS
Pooled F₂ seedlings
COmapper
B
Individual F₂
1
2
95
96
Illumina seq
~150 bp short reads
TIGER
C
High-molecular weight DNA
F₁ Pollen
10X chromium controller
Illumina seq
~150 bp short reads
Single molecule reconstruction
bead1
bead2
D
High-molecular weight DNA
F₁ Pollen
Pooled F₂
Nanopore long-read seq
Nanopore sequencing reads (~14 kb)
COmapper
read 1
read 2
read 3
read 4
E
Col/Ler SNPs
SNP 1
SNP 2
SNP 3
SNP 4
SNP 5
SNP 6
SNP 7
SNP 8
SNP 9
SNP 10
SNP 11
SNP 12
SNP 13
SNP 14
Read 1 (Col Parental)
Read 2 (Ler Parental)
Read 3 (Crossover)
Read 4 (Crossover)


A
Nanopore long-read seq prep
Col × Ler F₁ pollen / F₂ seedling sampling
(~1 ml pollen or ~1,000 F2 individuals)
HMW DNA extraction
(~9 µg, >~14 kb gDNA)
Nanopore sequencing
(PromethION, ~100 Gb, >14 kb N50)
Alignment
Alignement using Minimap2
(TAIR10 reference genome)
Mapping quality filter (>Q30, >100 bp reads)
Base quality filter (>Q14)
COmapper
Genotyping using high-quality SNPset
(226,864 SNPs)
Detecting crossover reads
Analysis of genome-wide
crossover position and frequency
B
N50 = 14 kb
Number of bases (Gb)
Read length (bp)
C
Sequenced length (kb)
Aligned length (kb)
D
Chr1
WT pollen
Coverage frequency
Mb
E
Chr1
SNPs per 100 kb
Mb
F
Mean = 524.8 bp
Number of intervals
Marker distance (bp)
G
86.1% reads ≥ 8 SNPs
Mean = 19.76
Frequency (%)
Number of SNPs per read

A

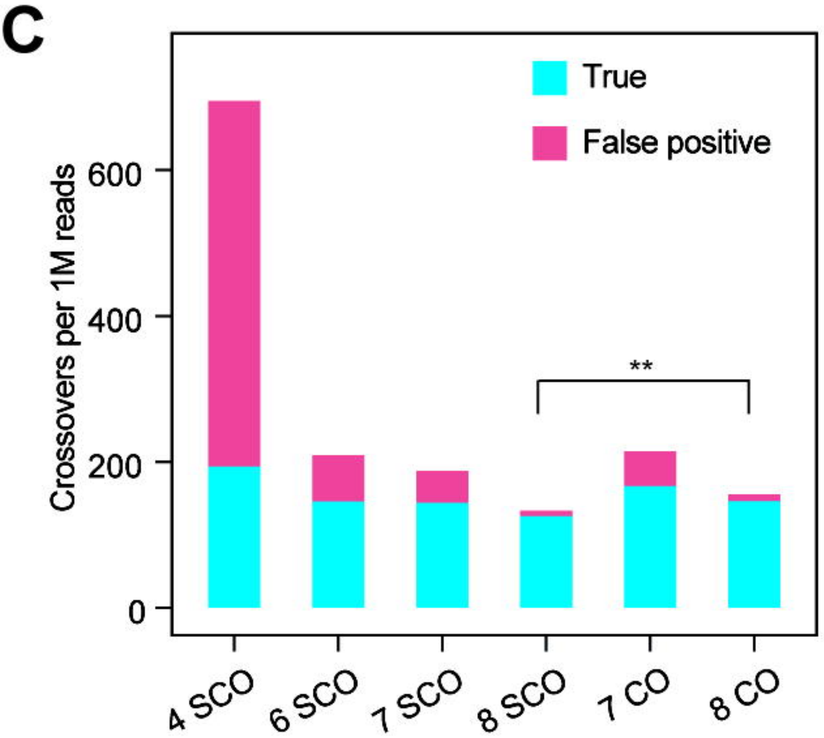


CO tract: SNPs include crossover sites in the midpoint
end tract: SNPs from each end of reads

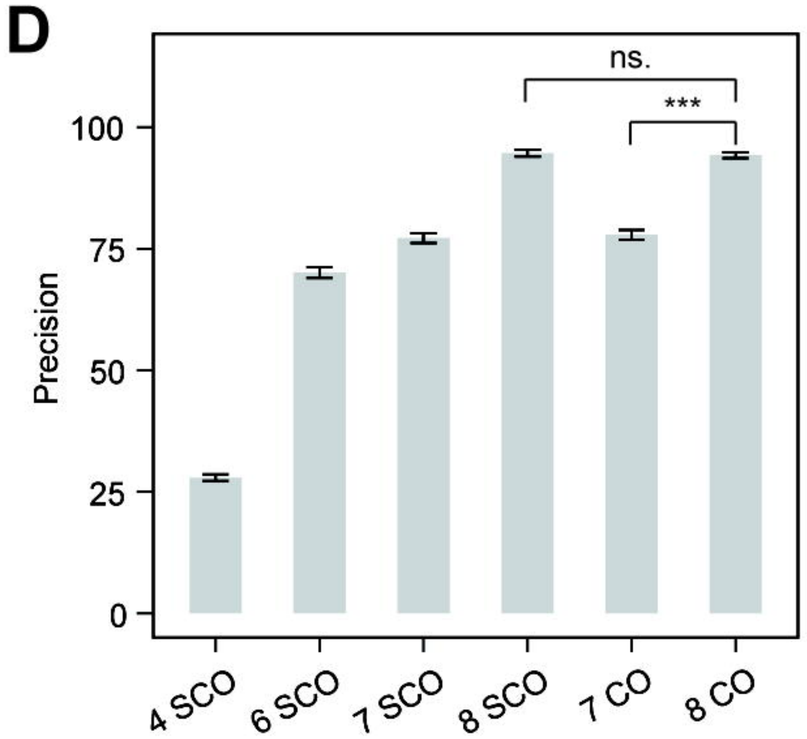
B

Model name	SNP number	Min CO tract	Min Col(Ler) tract	Min end tract	GC tract
4 SCO	4	4	2	2	-
6 SCO	6	6	3	2	-
7 SCO	7	7	3	2	-
8 SCO	8	8	4	2	-
7 CO	7	7	3	2	0~4
8 CO	8	8	4	2	0~4

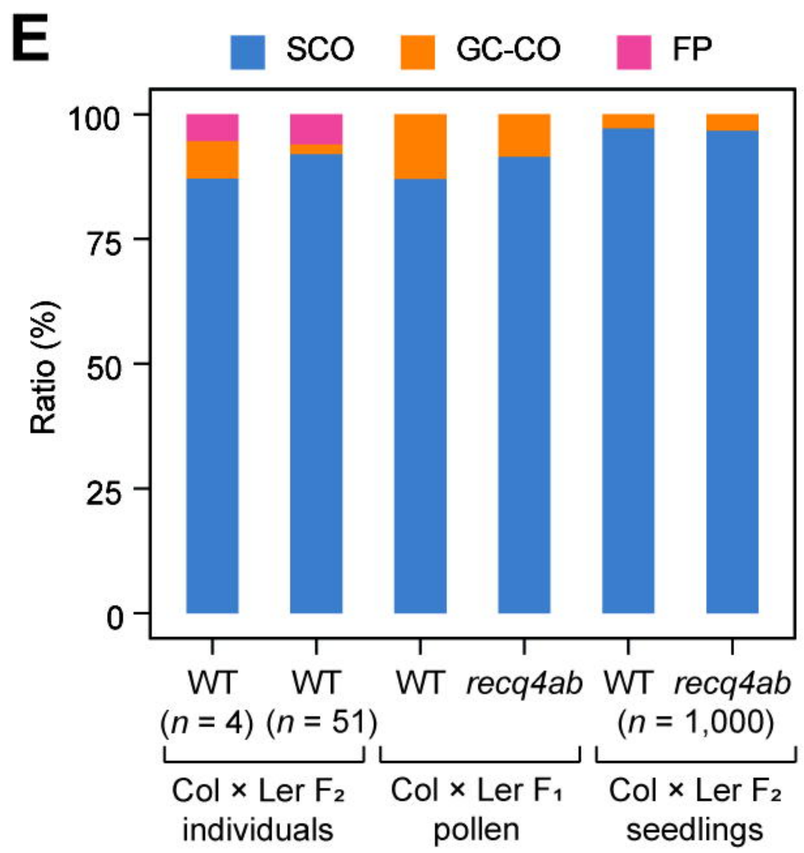
C



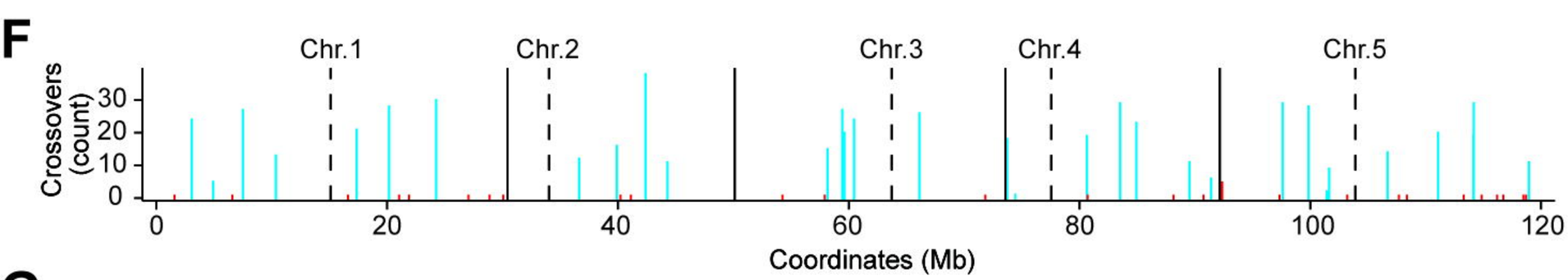
D



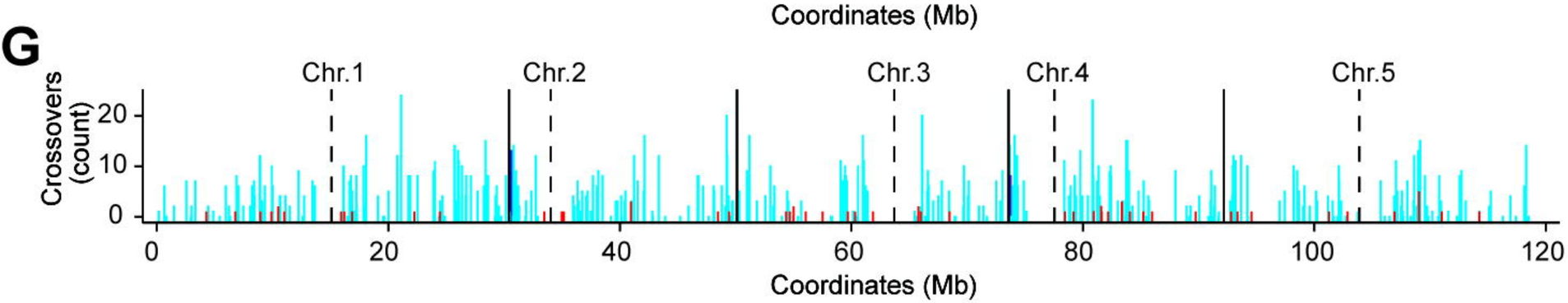
E



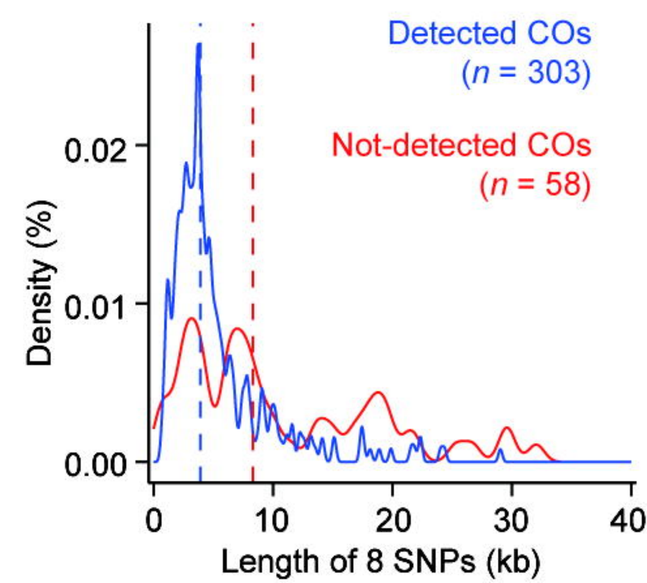
F

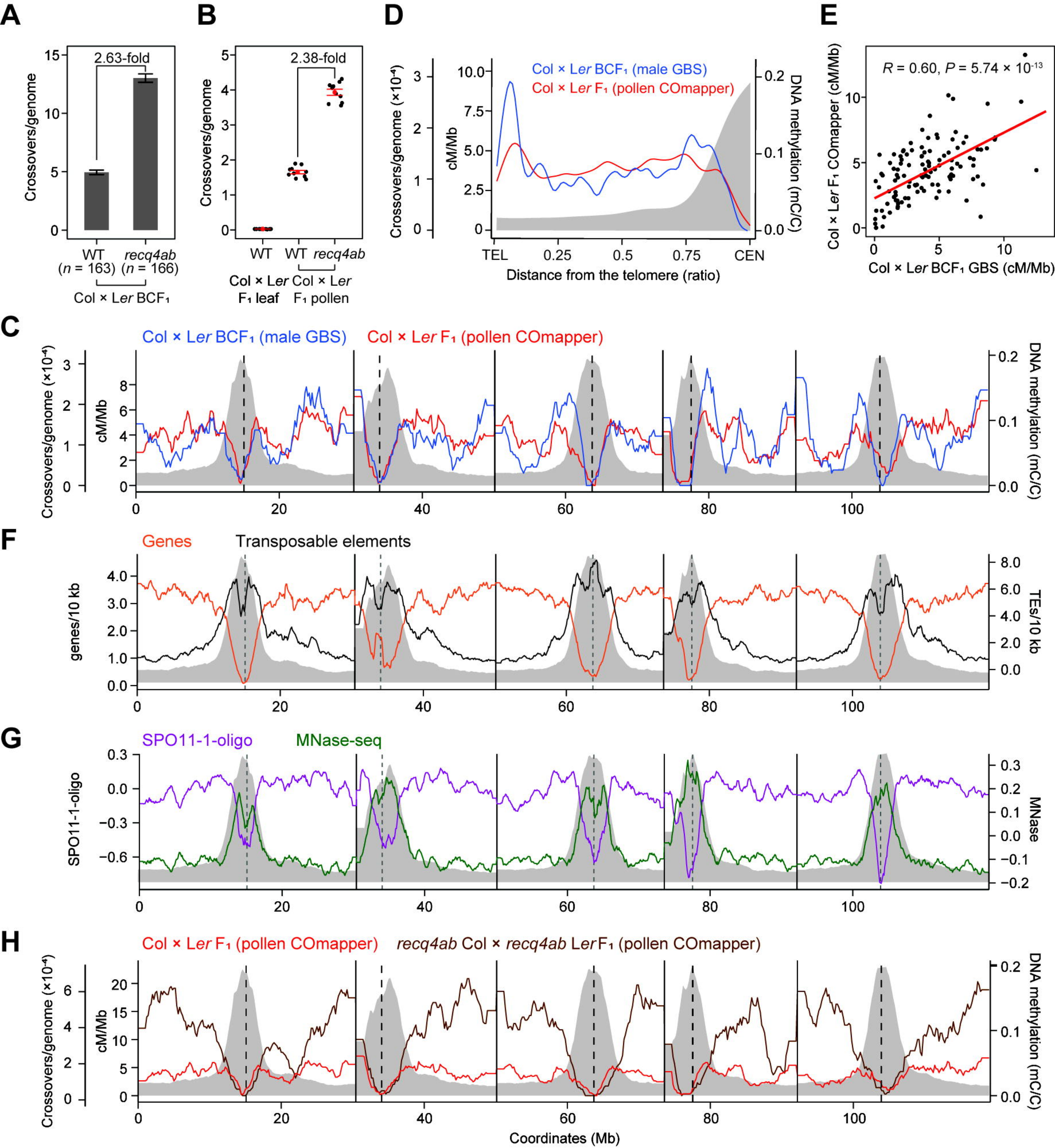


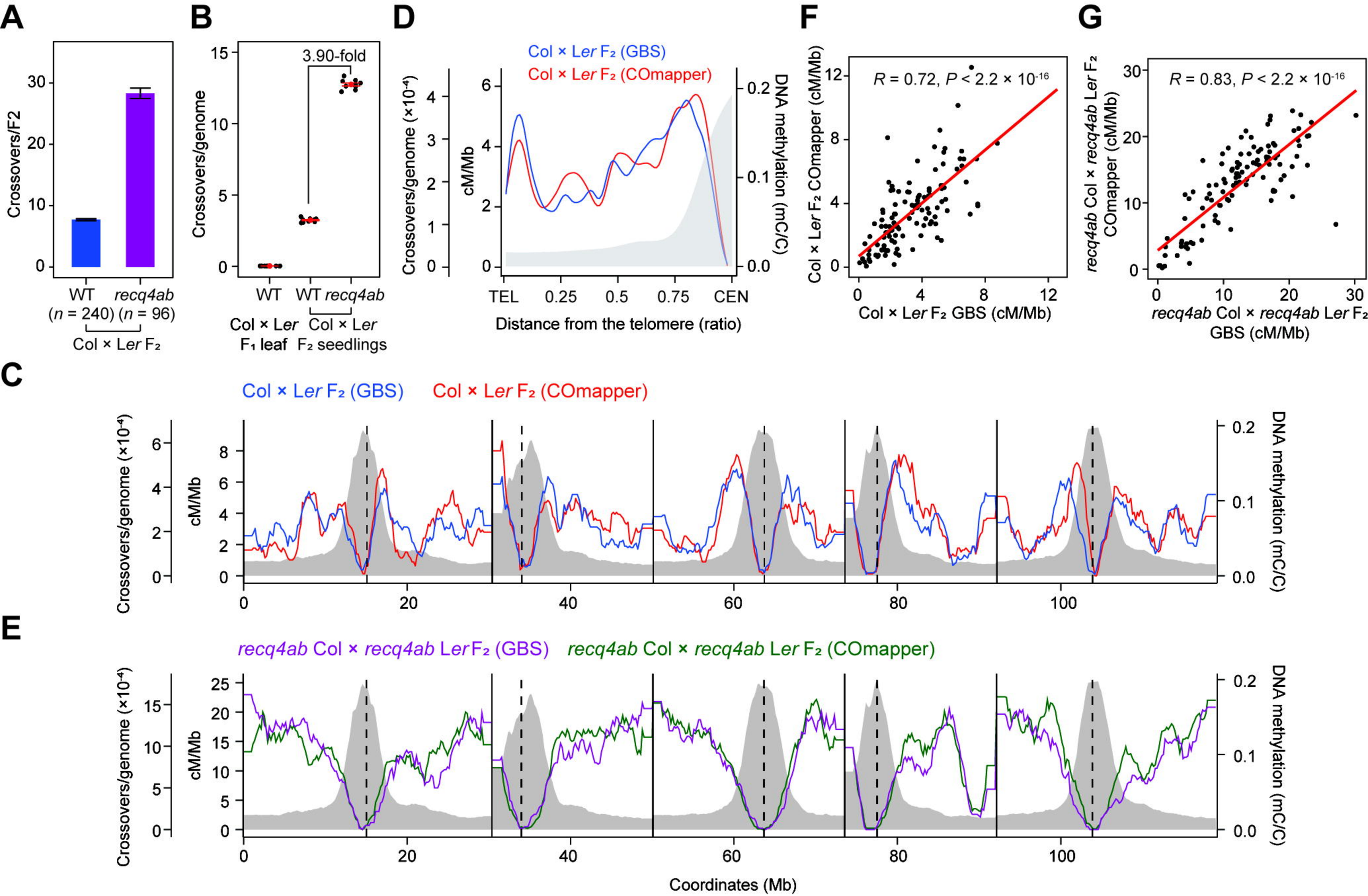
G

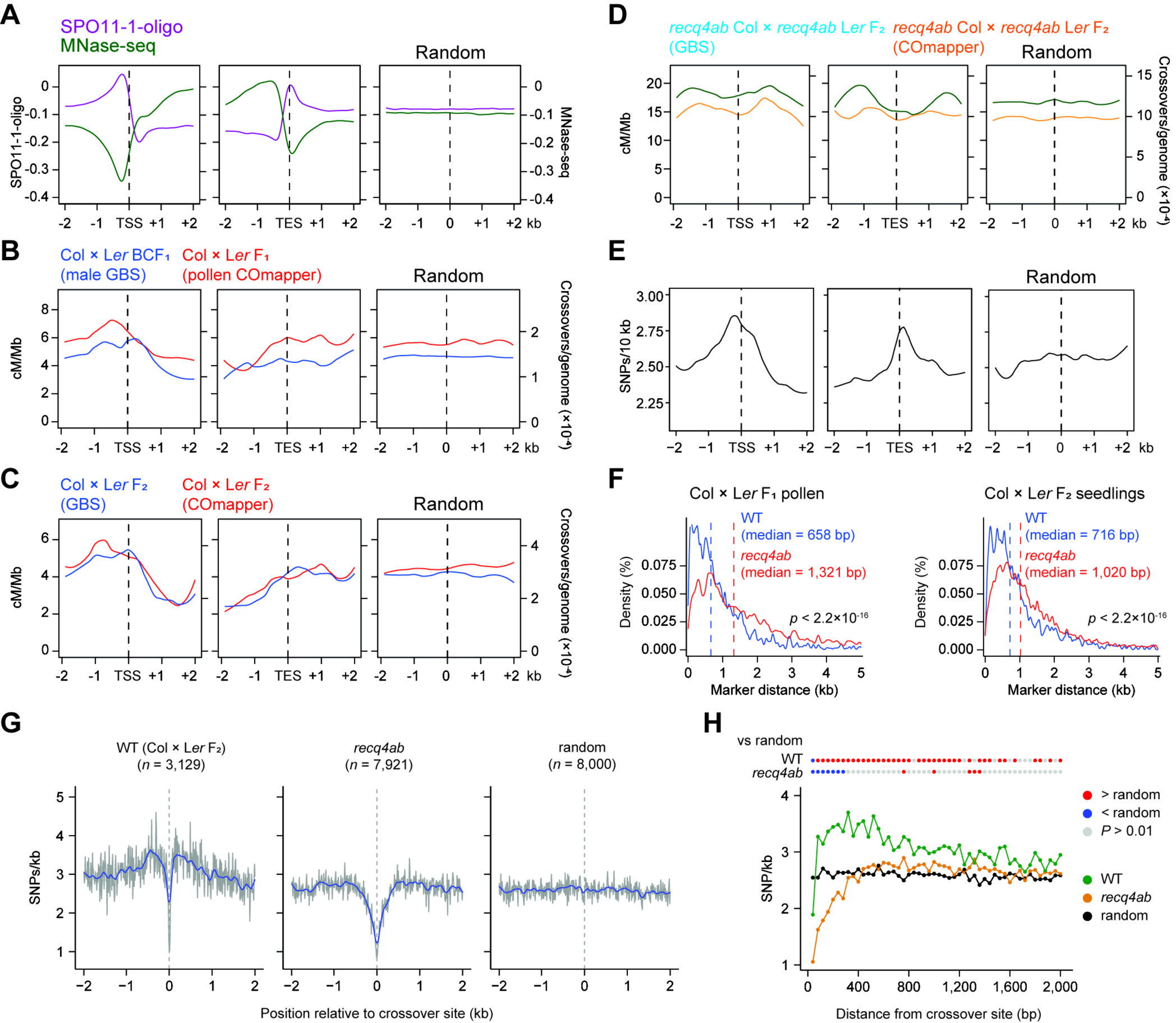


H




A
2.63-fold
Crossovers/genome
WT (n = 163)
recq4ab (n = 166)
Col × Ler BCF1
B
2.38-fold
Crossovers/genome
WT
WT
recq4ab
Col × Ler F1 leaf
Col × Ler F1 pollen
D
Col × Ler BCF1 (male GBS)
Col × Ler F1 (pollen COmapper)
Crossovers/genome (×10-4)
cM/Mb
DNA methylation (mC/C)
TEL
0.25
0.5
0.75
CEN
Distance from the telomere (ratio)
E
R = 0.60, P = 5.74 × 10-13
Col × Ler F1 COmapper (cM/Mb)
Col × Ler BCF1 GBS (cM/Mb)
C
Col × Ler BCF1 (male GBS)
Col × Ler F1 (pollen COmapper)
Crossovers/genome (×10-4)
cM/Mb
DNA methylation (mC/C)
F
Genes
Transposable elements
genes/10 kb
TEs/10 kb
G
SPO11-1-oligo
MNase-seq
SPO11-1-oligo
MNase
H
Col × Ler F1 (pollen COmapper)
recq4ab Col × recq4ab LerF1 (pollen COmapper)
Crossovers/genome (×10-4)
cM/Mb
DNA methylation (mC/C)
Coordinates (Mb)


A
Crossovers/F2
WT
recq4ab
(n = 240) (n = 96)
Col × Ler F₂
B
3.90-fold
Crossovers/genome
WT
WT recq4ab
Col × Ler
Col × Ler
F₁ leaf
F₂ seedlings
D
Col × Ler F₂ (GBS)
Col × Ler F₂ (COmapper)
Crossovers/genome (×10⁻⁴)
cM/Mb
DNA methylation (mC/C)
TEL
0.25
0.5
0.75
CEN
Distance from the telomere (ratio)
F
R = 0.72, P < 2.2 × 10⁻¹⁶
Col × Ler F₂ COmapper (cM/Mb)
Col × Ler F₂ GBS (cM/Mb)
G
R = 0.83, P < 2.2 × 10⁻¹⁶
recq4ab Col × recq4ab Ler F₂ COmapper (cM/Mb)
recq4ab Col × recq4ab Ler F₂ GBS (cM/Mb)
C
Col × Ler F₂ (GBS)
Col × Ler F₂ (COmapper)
E
recq4ab Col × recq4ab Ler F₂ (GBS)
recq4ab Col × recq4ab Ler F₂ (COmapper)
Coordinates (Mb)


A
SPO11-1-oligo
MNase-seq
Random
SPO11-1-oligo
MNase-seq
-2 -1 TSS +1 +2
-2 -1 TES +1 +2
-2 -1 0 +1 +2 kb
B
Col × Ler BCF1 (male GBS)
Col × Ler F1 (pollen COmapper)
Random
cM/Mb
Crossovers/genome (×10-4)
C
Col × Ler F2 (GBS)
Col × Ler F2 (COmapper)
Random
cM/Mb
Crossovers/genome (×10-4)
D
recq4ab Col × recq4ab Ler F2 (GBS)
recq4ab Col × recq4ab Ler F2 (COmapper)
Random
cM/Mb
Crossovers/genome (×10-4)
E
Random
SNPs/10 kb
F
Col × Ler F1 pollen
WT (median = 658 bp)
recq4ab (median = 1,321 bp)
p < 2.2×10-16
Density (%)
Marker distance (kb)
Col × Ler F2 seedlings
WT (median = 716 bp)
recq4ab (median = 1,020 bp)
p < 2.2×10-16
Density (%)
Marker distance (kb)
G
WT (Col × Ler F2) (n = 3,129)
recq4ab (n = 7,921)
random (n = 8,000)
SNPs/kb
Position relative to crossover site (kb)
H
vs random
WT
recq4ab
> random
< random
P > 0.01
WT
recq4ab
random
SNP/kb
Distance from crossover site (bp)