

Methods

COMapper: high-resolution mapping of meiotic crossovers by long-read sequencing in *Arabidopsis*

Dohwan Byun^{1*} , Namil Son^{1*} , Heejin Kim¹ , Jaecil Kim¹ , Jihye Park¹ , Sang-jun Park¹ , Hyein Kim¹ , Jaebeom Kim^{1,2} , Juhyun Kim^{1,3} , Seula Lee² , Youbong Hyun² , Piotr A. Ziolkowski⁴ , Ian R. Henderson⁵  and Kyuha Choi¹ 

¹Department of Life Sciences, Pohang University of Science and Technology, Pohang, 37673, Korea; ²School of Biological Sciences, Seoul National University, Seoul, 08826, Korea; ³Plant Quarantine Technology Center, Department of Plant Quarantine, Animal and Plant Quarantine Agency, Gimcheon, 39660, Korea; ⁴Laboratory of Genome Biology, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznań, 61-614, Poland; ⁵Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK

Author for correspondence:
Kyuha Choi
Email: kyuha@postech.ac.kr

Received: 14 November 2024
Accepted: 2 June 2025

New Phytologist (2025) 247: 1942–1957
doi: 10.1111/nph.70304

Key words: *Arabidopsis thaliana*, crossover, long-read sequencing, meiosis, polymorphism, RECQ4.

Summary

- Meiotic crossovers rearrange existing genetic variation between homologous chromosomes, profoundly affecting genomic diversity. Crossovers are typically constrained to one to three 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 F₁ hybrid pollen and F₂ recombinant seedlings derived from a cross between *Arabidopsis thaliana* accessions Col and Ler. We validate the high accuracy of COMapper by applying nanopore long-read sequencing to the pooled DNA of *Arabidopsis* F₂ individuals with crossovers mapped by short-read sequencing.
- Using the COMapper, we constructed high-resolution genomic maps of crossovers using F₁ hybrid pollen and F₂ seedlings in wild-type (WT) and crossover-elevated *recq4a recq4b* mutants, showing results comparable to short-read sequencing. Crossovers were enriched at gene-proximal promoters in WT 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.

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 & 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 & Kleckner, 2023). Topoisomerase-like SPO11 (sporulation) 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 & 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 & Keeney, 2014). This ssDNA invades nonsister 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 noncrossovers (Mercier *et al.*, 2015; Gray & Cohen, 2016; Wang & Copenhagen, 2018; Zickler & 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 *c.* 85–90% of all crossovers, which are formed by a group of pro-crossover factors including the E3 ligase HUMAN ENHANCER OF CELL INVASION NO. 10 (HEI10) (Mercier *et al.*, 2015; Wang & Copenhagen, 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

*These authors contributed equally to this work.

explained by a coarsening model where the diffusible and aggregating dynamics of HEI10 proteins limit closely spaced crossovers, resulting in one to three crossovers per pair of homologous chromosomes (Mercier *et al.*, 2015; Morgan *et al.*, 2021; Girard *et al.*, 2023). In the class II pathway, *c.* 10–15% of all crossovers are noninterfering and mediated by the endonuclease MMS AND UV SENSITIVE 81 (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 nonrandom and depend on chromatin accessibility (Choi *et al.*, 2013, 2018; Wijnker *et al.*, 2013; Shilo *et al.*, 2015; He *et al.*, 2017; Underwood *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 (Choi *et al.*, 2013; Wijnker *et al.*, 2013; Choi & Henderson, 2015). 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, 2018; Underwood & Choi, 2019; Fernandes *et al.*, 2024). Moreover, a high density of DNA polymorphisms and structural variations suppresses crossovers (Choi *et al.*, 2016; Serra *et al.*, 2018a; Dłuzewska *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 (Mb) scale in *Arabidopsis* (*Arabidopsis thaliana*), which depends on MSH2, a component of the mismatch repair system (Ziolkowski *et al.*, 2015; Dłuzewska *et al.*, 2023). Therefore, the 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 & Bomblies, 2021; Naish *et al.*, 2021; Kim & 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 F₂ 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, 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.*, 2020; 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 F₁ pollen or F₂ recombinant seedlings from *Arabidopsis* hybrids derived from a cross between the accessions Columbia-0 (Col) and Landsberg erecta-0 (Ler). We validate the precision of COMapper for genome-wide crossover analysis through long-read nanopore

sequencing of F₂ 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 (gDNA) from wild-type (WT) and *recq4a recq4b* (*recq4ab*) F₁ hybrid pollen grains or from their F₂ recombinant seedlings. We propose that COMapper will be useful for determining crossover patterns at high resolution in diverse plants with varying genetic backgrounds and ploidy levels.

Materials and Methods

Plant materials

The *Arabidopsis thaliana* (L.) accessions Col and Ler were used as the wild types for Col × Ler hybrid control plants. F₁ and F₂ hybrid plants were cultivated in a growth room under the controlled conditions of 20°C, 50–60% relative humidity, and a 16 h : 8 h, light : dark photoperiod. The *recq4ab* mutants in the Col and Ler backgrounds were used as previously described (Fernandes *et al.*, 2018). *J3pro:J3^{G155R}* plant was used as described (Kim *et al.*, 2024). The *Arabidopsis* accession Di-G (Dijon, NASC stock number: N28205) was obtained from the Arabidopsis Biological Resource Center.

Genotyping by sequencing

Genotyping by sequencing (GBS) libraries were prepared as previously described (Rowan *et al.*, 2015, 2019; Nageswaran *et al.*, 2021; Kim *et al.*, 2024). Briefly, gDNA was extracted from two or three mature leaves per individual plant using the CTAB (1% [w/v] CTAB, 50 mM Tris–HCl pH 8.0, 0.7 M NaCl, 10 mM EDTA) 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 (MM400; Retsch, Haan, Germany). Then, 500 µl of CTAB buffer 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. For each plant, 150 ng of gDNA was used to generate one barcoded sequencing library. The barcoded libraries were then pooled and subjected to paired-end 150-bp sequencing using an Illumina HiSeq-X platform (Macrogen, Seoul, Korea). The TIGER pipeline was used to analyze the sequencing data and identify crossover sites (Rowan *et al.*, 2015).

Preparation of gDNAs for nanopore long-read sequencing

To purify pollen grains from F₁ hybrid plants and extract their gDNA, F₁ plants were planted in soil and grown at 20°C under a 16 h : 8 h, light : dark photoperiod. Whole inflorescences of these hybrid plants were collected and transferred to ice-cold 10% (w/v) sucrose in 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- μ m 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. Proteinase K (P8107S; NEB, Ipswich, MA, USA) and RNase A (20 mg ml⁻¹, 10109142001; Roche, Basel, Switzerland) were added to a final concentration of 80 U ml⁻¹ and 20 μ g ml⁻¹, respectively. To check pollen disruption, a 1- μ l sample was diluted in a 1 : 10 ratio with 10% (w/v) sucrose and observed under a stereomicroscope (M165 FC; Leica, Wetzlar, Germany). Then, 500 μ l 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 (77617; Sigma) 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 μ l of the supernatant was transferred into a new tube, to which a Nanobind disk (Nanobind® Plant Nuclei Kit RT, 102-302-000; PacBio, Menlo Park, CA, USA) was added. The tube containing the Nanobind disk was gently mixed, 400 μ l of 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, 12321D; Invitrogen) 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 μ l EB+ buffer was added onto the disk and gently tapped 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 (Q32853; Thermo Fisher, Waltham, MA, USA) and integrity checked by gel electrophoresis.

For nanopore sequencing of gDNA from 1000 pooled seedlings, freshly harvested F₂ seeds from F₁ hybrid plants were surface-sterilized, germinated, and grown on ½-strength Murashige & Skoog agar medium supplemented with 1% sucrose under controlled conditions (20°C, 16 h : 8 h, light : dark photoperiod). Seedlings were harvested 10 d after germination and immediately frozen in liquid nitrogen. The 10-d-old seedlings were powdered in liquid nitrogen using a pre-chilled mortar and pestle. Genomic DNA was extracted using the CTAB method and quantified with a Qubit dsDNA Broad Range assay kit (Q32853; Thermo Fisher). Using 9 μ g of gDNA from pollen or seedlings, a nanopore long-read sequencing library was prepared with the Ligation Sequencing Kit V14 (SQK-LSK114; Nanopore, Oxford, Oxfordshire, UK). The libraries were sequenced on a PromethION platform (BGI, Tai Po, China).

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 single nucleotide polymorphisms (SNPs) distinguishing the Col and *Ler* accessions, we followed the previous method with modifications (Fernandes *et al.*, 2024). Single nucleotide polymorphisms 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 2160 Col $\times$ *Ler* F₂ plants (Rowan *et al.*, 2019). To retain only high-quality SNPs, the following criteria were applied: (1) genotyping results available for at least 40 F₂ plants, (2) allele frequency between 0.4 and 0.6, and (3) adherence to Mendelian segregation across the 2160 F₂ individuals (chi-square test, *P* > 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 226 864 high-quality Col/*Ler* 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 fewer than 8 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 threefold and exhibiting a high frequency of false positives (FPs) in the negative control provided by the sequencing of Col $\times$ *Ler* F₁ 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 areas and regions where the distance between consecutive SNPs was > 4000 bp were excluded from all analyses to increase the resolution of crossover sites. For the analysis of SNP density around crossover sites, 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).

Results

Establishment of COMapper and long-read sequencing to map crossovers

To develop COMapper and enable high-resolution mapping of meiotic crossovers at the SNP level using long-read sequencing, we used the two polymorphic *Arabidopsis* accessions, Col and *Ler* (Fig. 1a). We crossed them to generate WT Col × *Ler* F₁ hybrid plants, from which we collected recombinant pollen; we also allowed them to self-pollinate to collect F₂ recombinant seeds (Fig. 1a). We performed GBS of individual F₂ recombinant plants using individual indexing, short-read sequencing, and the TIGER pipeline, which allows the mapping of crossover sites along chromosomes (Fig. 1b) (Wijnker *et al.*, 2013; Rowan *et al.*, 2015, 2019; Nageswaran *et al.*, 2021; Kim *et al.*, 2022, 2024). We then used the GBS data and pooled gDNA of F₂ recombinant seedlings as a positive control to establish and validate COMapper. Notably, GBS of F₂ recombinants has also been used to validate the linked-sequencing-based crossover mapping method via microfluidics on the 10× Chromium Controller (Fig. 1c) (Sun *et al.*, 2019). Genotyping by sequencing 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 (Fig. 1b,c). By contrast, nanopore long-read sequencing followed by COMapper can directly identify Col-specific or *Ler*-specific SNPs within individually sequenced DNA molecules without prior indexing or amplification and can determine the position of crossover sites in individual DNA molecules (Fig. 1d,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 (Fig. 2a), we isolated mature pollen grains from WT Col × *Ler* F₁ hybrid plants and extracted large gDNA molecules using our modified CTAB method (see the [Materials and Methods](#) section). Our gDNA extraction from pollen grains of Col × *Ler* F₁ hybrids resulted in nanopore sequence read lengths of *c.* 11.1 kb for gDNA (mean, 11.1 kb; N50, 14.0 kb), comparable to that of gDNA extracted from F₂ seedlings using the CTAB method (Fig. 2b, Supporting Information Fig. S1). To achieve deep sequencing coverage, we used *c.* 9 µg gDNA on the PromethION device, which produced reads corresponding to *c.* 100–150 Gb of sequence per flow cell (Fig. S1; Table S1). We successfully aligned *c.* 90.7% of the nanopore reads to the

TAIR10 *Arabidopsis* reference genome using the minimap2 alignment tool (Fig. 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 (Figs 2d, S2). We used a high-quality SNP dataset distinguishing between Col and *Ler* 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 (Fig. 2e,f). We noted that structural variation between Col and *Ler* results in decreased coverage of the aligned long reads and fewer SNPs for GBS in the pericentromeric regions near centromeres (Figs 2d,e, S3). Of all aligned long reads, 86.1% contained more than eight SNPs per read, with a mean of 19.8 SNPs per read (Fig. 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 *Ler*-specific SNPs, or vice versa, by counting the numbers of SNPs (Fig. 3a). In *Arabidopsis*, *c.* 50% of crossovers are associated with gene conversion (Wijnker *et al.*, 2013), which are detected in COMapper data as SCOs characterized by a single switch between parental SNPs. However, a subset of crossovers contains two or more switches between parental SNPs, representing complex crossovers (CCOs) (Schweiger *et al.*, 2024). Therefore, we also considered CCOs when identifying crossover sites by examining the number and positions of consecutive Col-specific or *Ler*-specific SNPs (Fig. 3b). Accordingly, we developed and tested six crossover mapping models (Fig. 3b), which use different numbers of consecutive Col-specific and *Ler*-specific SNPs with or without multiple switches between parental SNPs within the crossover tract to map crossovers. To validate these models, we performed nanopore sequencing of two independent pooled gDNA samples from 4 or 46 Col × *Ler* F₂ recombinant plants in which crossover sites had already been determined by GBS (Fig. S4; Table S2). 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 97.3% (8-SCO) and 97.4% (8-CO), respectively (Welch's *t*-test; 8-SCO vs 8-CO $P = 0.747$, 8-CO vs 7-CO $P = 1.38 \times 10^{-11}$) (Fig. 3c–e). By contrast, the 4-SCO model showed a high number of FPs, likely due to the inherent per-base sequencing error rate (3.98%, *Q* score = 14) of nanopore reads (Fig. S5). The 8-CO model, which identifies both SCO and CCO types, detected more crossovers than the 8-SCO model, which detected only the SCO type (Welch's *t*-test, $P = 2.82 \times 10^{-10}$) (Figs 3c,d, S6, S7). COMapper with the 8-CO model detected *c.* 7.3% and *c.* 1.5% of the CCO type of total crossover sites from the pooled gDNA of 4 and 46 F₂ plants, respectively (Figs 3f, S6, S7). Furthermore, using COMapper with the 8-CO model, we identified crossover

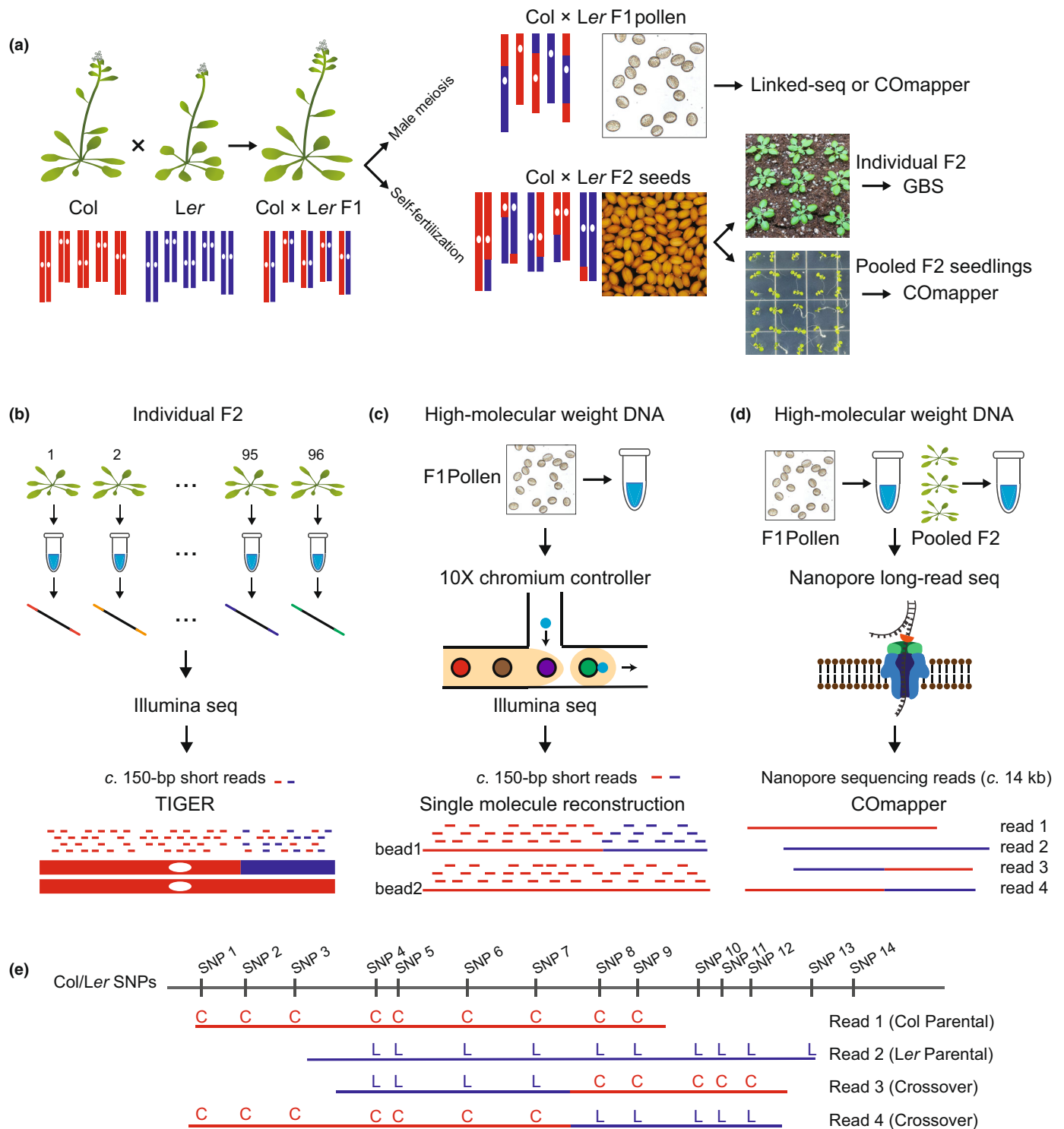


Fig. 1 Genome-wide mapping of crossovers using long-read nanopore sequencing. (a) Diagrams illustrating the preparation of recombinant pollen from F₁ hybrids and F₂ plants to map crossovers in *Arabidopsis thaliana*. (b) Mapping of genome-wide crossovers in F₂ plants by genotyping-by-sequencing (GBS) using multiplex barcoding and the TIGER pipeline. (c) Mapping of crossovers in pollen DNA from F₁ hybrid plants using 10× Chromium Controller and linked-read sequencing. (d) Mapping of crossovers in F₁ pollen and F₂ 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 (Ler) SNP.

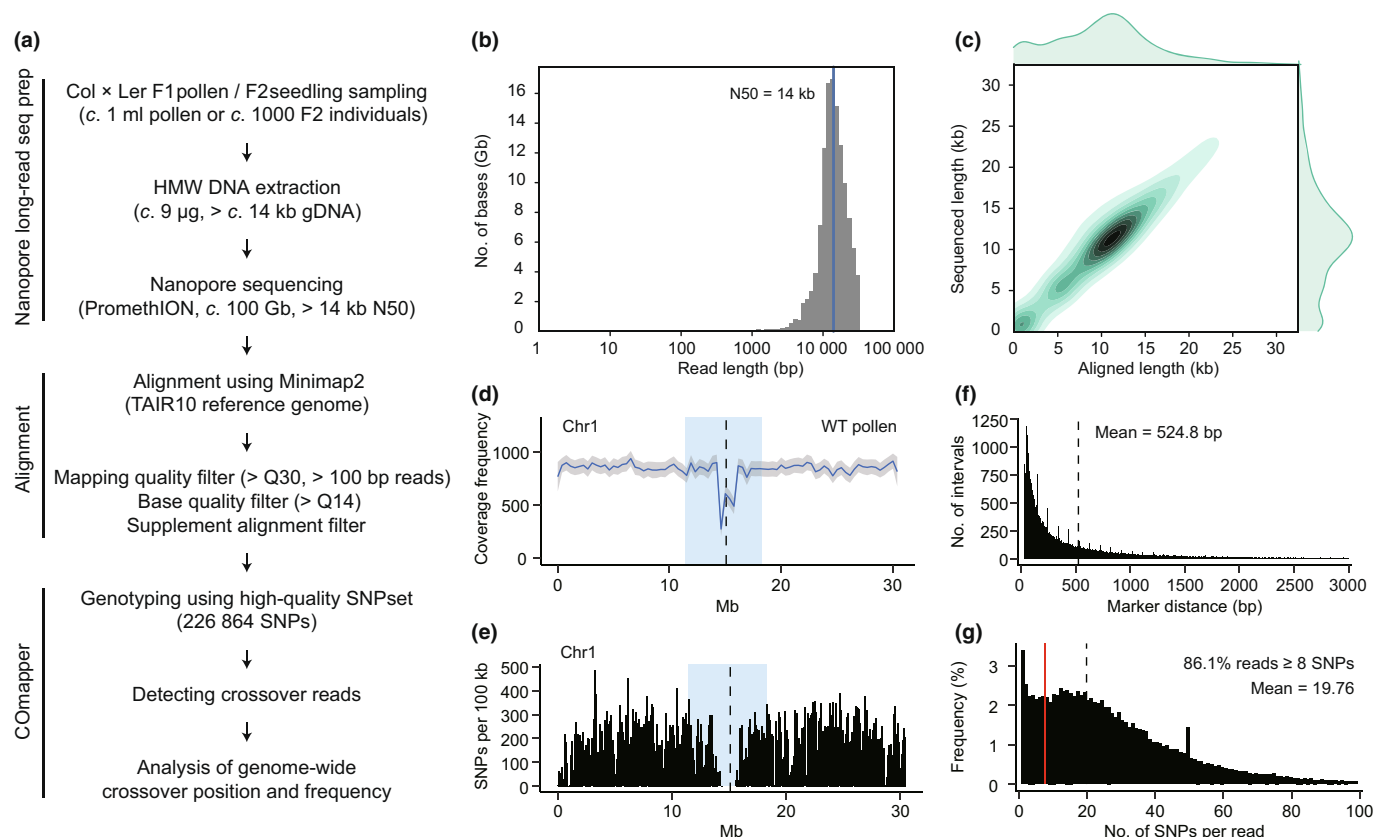


Fig. 2 Development of COMapper to map crossovers using long-read sequencing in *Arabidopsis thaliana*. (a) Procedure for crossover mapping using nanopore long-read sequencing and the COMapper pipeline. (b) Distribution of read length and number of pollen DNAs sequenced by nanopore long-read sequencing. (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 × Ler F₁ 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 single nucleotide polymorphisms (SNPs) between the Col and Ler accessions, in 100-kb sliding windows. (f) Distribution of inter-SNP distances between consecutive SNPs differentiating the Col and Ler 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.

sites in nanopore reads of pooled hybrid pollen and pooled F₂ seedlings of WT Col × Ler and *recq4ab* Col × *recq4ab* Ler. Among detected crossovers, c. 8.9% (WT) and c. 7.8% (*recq4ab*) were CCOs in F₁ hybrid pollen. In F₂ seedlings ($n = 1000$), CCOs accounted for 2.1% (WT) and 2.9% (*recq4ab*), similar to the 2.1% observed in 46 F₂ plants analyzed by GBS and COMapper (Fig. 3f).

To validate the output of COMapper, we used the crossover sites mapped by GBS and TIGER with manual curation as true positives (TPs) from 4 and 46 Col × Ler F₂ recombinant plants (33 crossovers for 4 F₂ plants, 342 crossovers for 46 F₂ plants) (Dataset S1, S2). COMapper with the 8-CO model was capable of identifying TP crossover sites multiple times per site, whereas c. 3% FPs were detected only once or twice (Fig. 3g,h). Analysis of these FPs revealed four sources: chimeric reads with error clusters (16 in 4 F₂, 16 in 46 F₂), substitution errors (5 in 4 F₂, 12 in 46 F₂), chimeric reads with large indels (3 in 46 F₂), and mitotic recombination (10 in 46 F₂), reflecting limitations of nanopore sequencing (Fig. S8; Dataset S3).

COMapper accurately identified all 33 crossovers that were mapped by GBS in the 4 F₂ recombinants, with a mean of 18.9 long reads covering each crossover site when reaching a sequencing depth of 40.14 Gb (Fig. S4a,b). Of the 342 crossovers mapped by GBS in the 46 F₂ recombinants, COMapper accurately mapped 299 crossovers, with a mean of 6.07 long reads per crossover site. Notably, COMapper failed to detect the remaining 43 crossovers. To analyze the false negatives, we categorized the detected (299) and undetected crossovers (43) from the 46 F₂ pool according to the 8-SNP length around the crossover site (Fig. S9a,b). While crossovers were more concentrated in the SNP-rich regions, the proportion of undetected crossovers relative to the total crossover (detected and undetected crossovers) increases as the 8-SNP length increases (Fig. S9a,b). It indicates the dependence of COMapper's crossover mapping sensitivity on SNP density. However, the undetected crossovers (false negatives) were not concentrated at certain regions of the genome but randomly distributed at a genome-wide scale (Fig. S9c; Dataset S4).

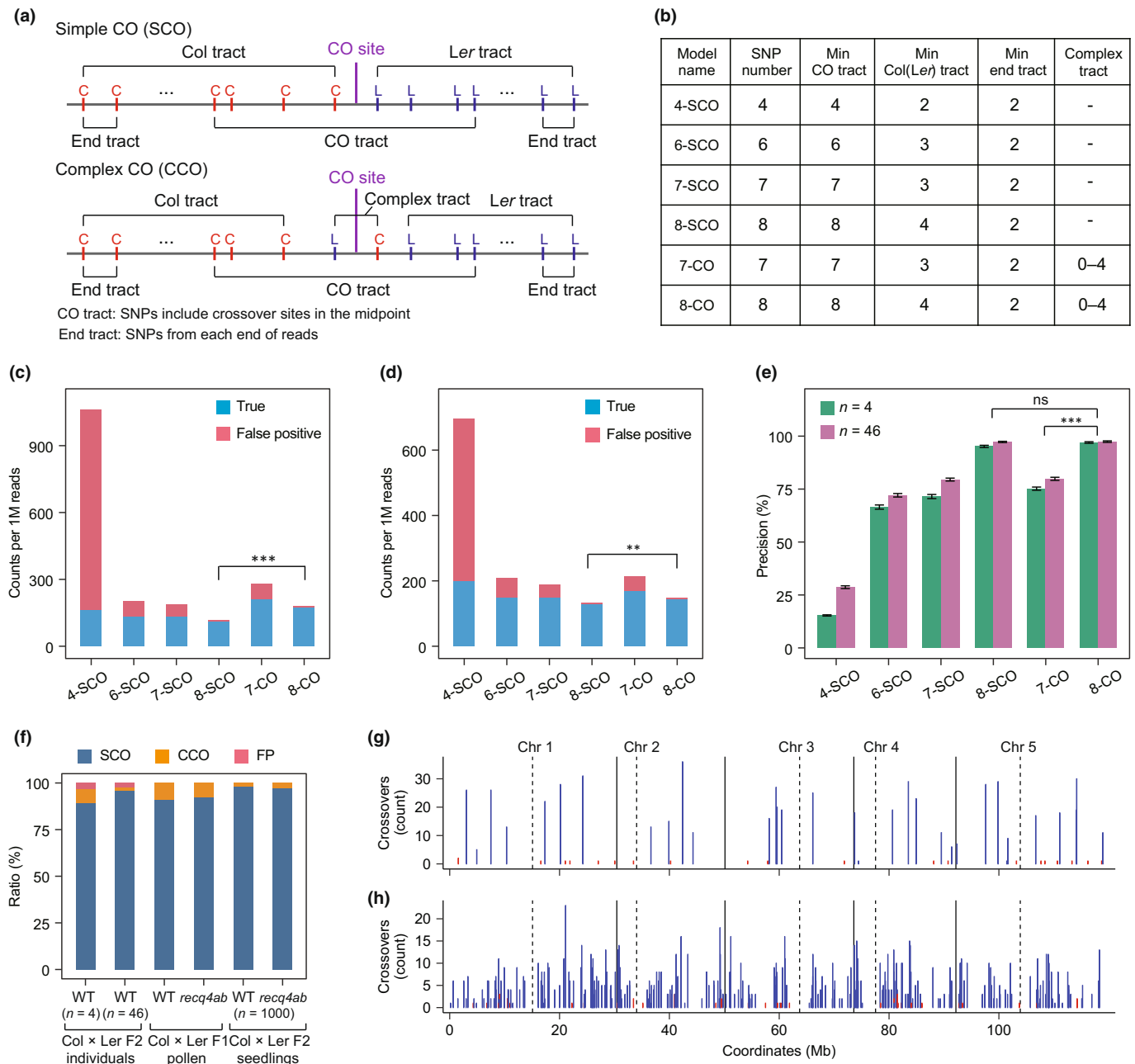


Fig. 3 Validation of COMapper using *Arabidopsis thaliana* Col $\times$ Ler F₂ recombinant plants. (a) Diagram showing simple crossover (SCO) and complex crossover (CCO) types. (b) Summary table of the tested COMapper models with different numbers of single nucleotide polymorphisms (SNPs) and minimum crossover (Min CO) and complex tracts. (c) Number of crossovers detected per million reads in different COMapper models in 4 Col $\times$ Ler F₂ plants (Kim *et al.*, 2024). True crossovers (blue) were detected by COMapper models among the crossovers mapped by genotyping by sequencing (GBS) in 4 F₂ plants, and false crossovers (red) indicate the crossovers that differ from the crossovers defined by GBS. (d) As in (c), but showing the number of crossovers detected per million reads in 46 Col $\times$ Ler F₂ plants. (e) Precision (True positive/(True positive + False positive)) of crossover detection in the different COMapper models shown in (c). Error bar indicates SD. In (c), (d), and (e), asterisks indicate significant differences between models (**, $P < 0.01$; ***, $P < 0.001$; ns (not significant) $P > 0.05$; Welch's t -test). (f) Relative proportion of crossover types in 4 F₂ plants (genotyped by GBS), 46 F₂ plants (genotyped by GBS), Col-0 $\times$ Ler F₁ hybrid pollen DNA, *recq4ab* Col $\times$ *recq4ab* Ler hybrid pollen DNA, 1000 Col-0 $\times$ Ler F₂ seedlings, and 1000 *recq4ab* Col $\times$ *recq4ab* Ler F₂ 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 46 F₂ plants compared to the crossovers identified by GBS. (g) Frequency of crossovers mapped by COMapper along the five *Arabidopsis* chromosomes on a continuous x-axis in 4 Col $\times$ Ler F₂ recombinants, where crossovers in each F₂ plant were mapped by GBS. Blue bars represent true positive crossovers mapped by GBS; red bars represent false positive crossovers. (h) As in (g), but from 46 Col $\times$ Ler F₂ recombinants.

As COMapper depends on the depth of sequencing, we conducted a comprehensive analysis of how sequencing coverage influences COMapper's performance in terms of TP and FP crossover read detection, precision ($TP/[TP + FP]$), and false negatives (crossovers detected by GBS but not by COMapper) (Fig. S10). We found that both TP and FP reads, including those from the same crossover sites, increase proportionally with sequencing depth, keeping precision stable across coverage levels (Fig. S10a–f). As coverage per F_2 increases, false negatives decrease. At the same time, the discovery rate of new unique crossovers declines due to repeated detection of previously identified sites (Fig. S10g,h). Additionally, we found that pooling more individuals increases the likelihood of capturing nonoverlapping crossover sites, thereby reducing redundancy and maximizing the number of distinct crossovers (Fig. S10i). This validation shows that SNP-resolution crossover mapping by COMapper requires long reads with at least eight SNPs and sufficient sequencing depth (Table S1). 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 97%.

To further assess COMapper's performance in the *recq4ab* background, we re-sequenced gDNA from 39 *recq4ab* Col $\times$ Ler F_2 individuals—previously analyzed by GBS—using nanopore sequencing at a total coverage of 410 $\times$ (*c.* 10.5 $\times$ per individual) (Fig. S11). COMapper achieved *c.* 96% precision, comparable to the 97% observed in the WT dataset, demonstrating reliable crossover detection in both genotypes. In addition, we confirmed that COMapper is highly robust for crossover detection using the Col-CEN (Telomere-to-Telomere) reference genome as well as TAIR10, based on 46 F_2 individuals (Fig. S12).

Mapping of crossovers from a pool of Col $\times$ Ler 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 WT Col $\times$ Ler and *recq4ab* Col $\times$ *recq4ab* Ler F_1 hybrid pollen grains (Fig. 4). Our GBS analysis showed a 3.12-fold increase in crossovers during male meiosis in the *recq4ab* mutant compared to the WT, consistent with the previous report (Fig. 4a) (Fernandes *et al.*, 2018). Similarly, COMapper detected a 2.41-fold increase in crossovers per genome (119 Mb) when analyzing long-read sequencing data produced from *recq4ab* F_1 hybrid pollen DNA (mean = 3.92 crossovers per genome, *c.* 763.8 $\times$ coverage) compared to the WT (mean = 1.62 crossovers per genome, 1072.3 $\times$ coverage) (Welch's *t*-test, $P = 1.22 \times 10^{-12}$) (Fig. 4b). Importantly, COMapper rarely detected crossovers as FP crossover events in the negative control sample (leaf gDNA from Col $\times$ Ler F_1 hybrids; mean = 0.03 crossovers per genome, 82.8 $\times$ coverage) compared to the WT Col $\times$ Ler F_1 hybrid pollen (Welch's *t*-test, $P = 5.51 \times 10^{-10}$) (Fig. 4b; Table S1). These results indicate that COMapper enables the detection of meiotic crossovers in pooled gDNA from Col $\times$ Ler hybrid pollen grains, with an increased number of crossovers in a *recq4ab* background.

Previously, GBS-based crossover maps of *Arabidopsis* Col $\times$ Ler male meiosis using a backcross (BC) population detected *c.* 4.47 crossovers per genome (Kim *et al.*, 2024). In comparison, COMapper detected *c.* 1.65 crossovers per genome in nanopore long-read sequencing of pooled pollen grains from Col $\times$ Ler hybrid plants, with all detected 1204 crossovers being nonoverlapping. This suggests an inherent feature of the COMapper method, which samples crossover molecules directly from long reads, rather than reconstructing full meiotic products, as is done in GBS. To further validate COMapper's performance, we performed an independent COMapper experiment using pollen from WT Col $\times$ Ler F_1 plants, confirming the robustness of its crossover detection (Figs 4b, S13). Despite detecting fewer absolute events, our validation demonstrates that COMapper provides a consistent and representative subset of crossovers relative to sequencing coverage.

To compare genome-wide crossover patterns detected by COMapper and GBS using the Col-CEN complete genome (Kim *et al.*, 2024), we calculated relative crossover rates as the proportion of crossovers in each 100-kb genomic bin, normalized to the total number of crossovers (Fig. 4c–f). We used Fisher's exact test to assess crossover frequency differences and found no significant regional differences between COMapper and GBS, in either WT or *recq4ab* datasets (Fig. 4c,e). We also observed that the genomic landscape of crossovers mapped by COMapper correlated with that of GBS (Fig. 4d,f), indicating that COMapper and GBS produce highly comparable genome-wide crossover landscapes. Notably, we identified a *c.* 3.2 Mb inversion (Chr4: 14 564 462–17 818 793) in *recq4ab* via nanopore long-read analysis, which likely underlies crossover suppression observed in both GBS and COMapper (Fig. 4e), consistent with the previous report (Serra *et al.*, 2018b). Inversions interfere with homolog pairing and synapsis, and single crossovers within them can produce dicentric or acentric chromosomes, leading to suppression or rare double crossovers (Rönsplies *et al.*, 2022). Although not statistically significant, COMapper detected slightly more crossovers within the *recq4ab* inversion than GBS, possibly due to selection against nonviable recombinants in the GBS BCF₁ population (Fig. 4e).

Despite the comparable crossover landscapes generated by both approaches, we noted that COMapper detects fewer crossovers at sub-telomeric regions than GBS by plotting the crossovers of male meiosis along a telomere-to-centromere axis that combined all five chromosomes (Fig. S14a). The mean SNP density was lower at sub-telomeric regions than in mid-chromosomal regions (Fig. S14b,c). 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 (Fig. S14d–f).

To further assess the effect of SNP density on the crossover detection sensitivity of COMapper, we divided the genome into four categories based on the SNP density, defined by the 8-SNP span (Fig. S15a,b), and compared the distribution of COMapper-detected crossovers across these regions with those of GBS-detected crossovers (Fig. S15c). In the WT male GBS

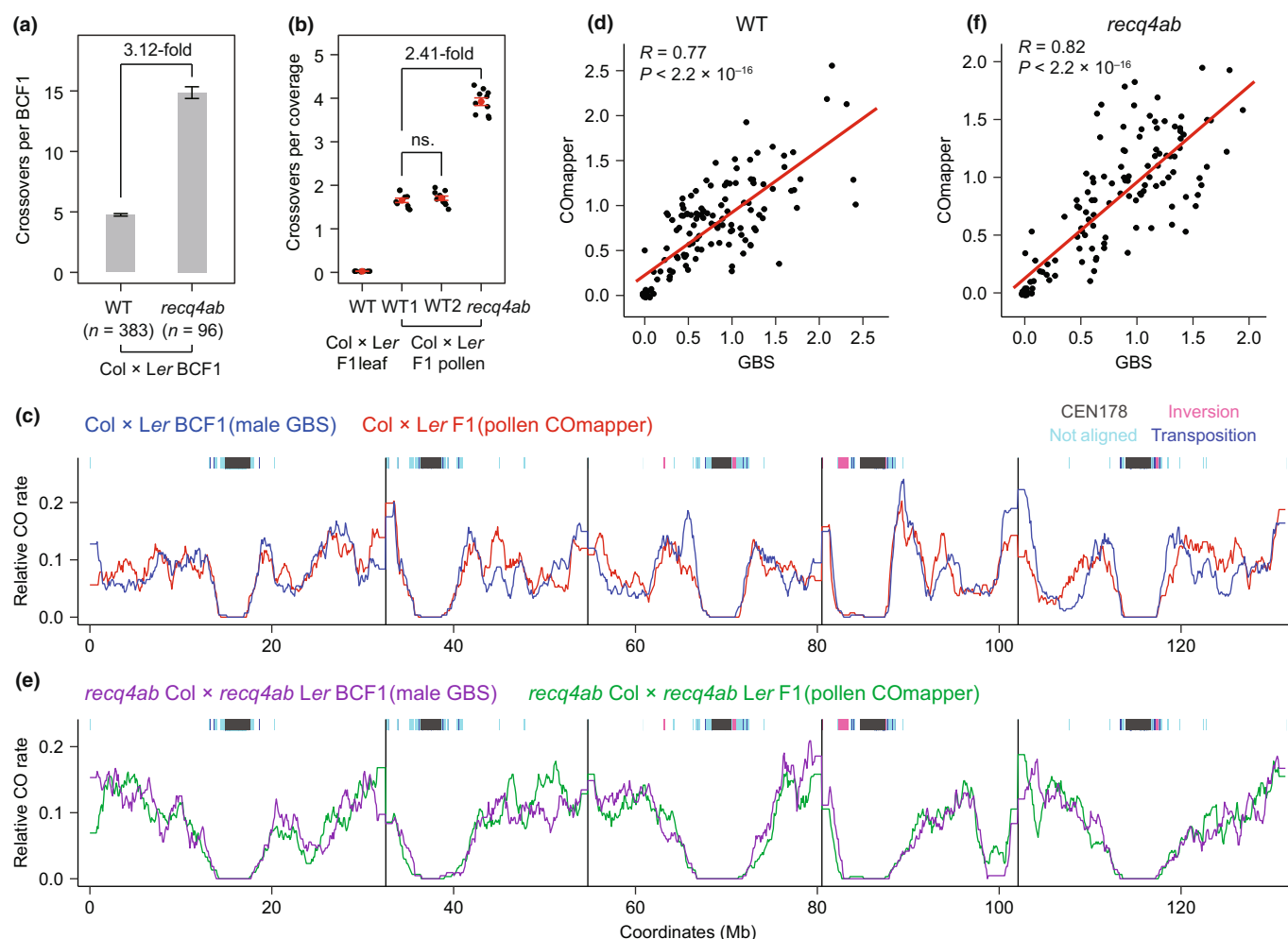


Fig. 4 Mapping of crossovers from pooled F₁ hybrid pollen DNA by COMapper in *Arabidopsis thaliana*. (a) Number of male crossovers per backcross F₁ (BCF₁) individual calculated from genotyping by sequencing (GBS) data of wild-type (WT) Col × Ler BCF₁ ($n = 383$) (Kim *et al.*, 2024) and *recq4ab* Col × *recq4ab* Ler BCF₁ ($n = 96$) plants. n indicates the number of BCF₁ plants generated by backcrossing pollen from Col × Ler F₁ hybrids to Ler ovules in both WT and *recq4ab* backgrounds. Data are presented as mean ± SD. (b) Number of crossovers per genomic coverage of nanopore sequencing data of genomic DNA (gDNA) from WT Col × Ler F₁ leaves, WT Col × Ler F₁ pollen (WT set1, WT set2), and *recq4ab* Col × *recq4ab* Ler F₁ pollen. Each data point indicates crossovers per genome from a randomly selected million reads. Data are presented as mean ± SD from 10 replicates (not significant (ns) $P > 0.05$; Welch's t -test). (c) Relative frequency of crossovers mapped by Col × Ler male GBS (blue, $n = 1824$ crossovers) and Col × Ler F₁ pollen COMapper (red, $n = 1679$ crossovers) along the five *Arabidopsis* chromosomes on a continuous x-axis in Col × Ler male meiosis using Col-CEN complete genome. Fisher's exact test was applied to each 100-kb bin in every 1 Mb window to assess differences in relative crossover rates. (d) Correlation between crossover frequencies in 1-Mb bins of Col × Ler BCF₁ individuals (measured by GBS) and Col × Ler F₁ pollen (measured by COMapper) across the genome, assessed using Pearson's correlation test. (e) As in (c), but showing the relative crossover frequency of crossovers mapped by *recq4ab* Col × *recq4ab* Ler male GBS (blue, $n = 1428$ crossovers) and *recq4ab* Col × *recq4ab* Ler F₁ pollen COMapper (red, $n = 1963$ crossovers). (f) As in (d), but showing the correlation between crossover frequencies of crossovers mapped by *recq4ab* Col × *recq4ab* Ler BCF₁ individuals and *recq4ab* Col × *recq4ab* Ler F₁ pollen.

dataset, *c.* 14.6% of crossovers were found in pericentromeric and subtelomeric regions with 8-SNP spans ≥ 10 kb; by contrast, only 1.8% of crossovers in the WT pollen COMapper dataset were detected in these SNP-poor regions (Fig. S15c). Since GBS-based crossover detection is not affected by SNP density, this result indicates that COMapper has reduced sensitivity in SNP-poor regions and preferentially detects crossovers in SNP-rich areas. However, because crossovers in SNP-poor pericentromeric regions account for only a small fraction of total crossovers, and a subset of SNP-poor regions are interspersed among SNP-rich regions (Fig. S15a), the overall effect of SNP density on crossover rate profiles along chromosomes was minimal when analyzed at

the megabase scale. Nonetheless, comparisons of crossover rates between COMapper and GBS should be interpreted cautiously in SNP-poor regions, such as subtelomeric and pericentromeric regions, particularly at high resolution (e.g. kilobase scale), where detection sensitivity may differ significantly. COMapper requires long reads (≥ 10 –15 kb) containing at least 8 SNPs for reliable crossover detection, underscoring its limitation in regions with sparse SNP density. Despite this, we demonstrate that COMapper can effectively map genome-wide crossovers and capture the genetic effect of the *recq4ab* mutant on crossover frequency and distribution by nanopore sequencing of pooled hybrid pollen DNA.

Mapping of crossovers from pooled F₂ recombinant seedlings by COMapper

Genotyping by sequencing analysis of Col × *Ler* F₂ individuals revealed *c.* 8 crossovers per F₂ (Fernandes *et al.*, 2018; Serra *et al.*, 2018b; Rowan *et al.*, 2019; Nageswaran *et al.*, 2021; Kim *et al.*, 2022, 2024). From the COMapper analysis of 40 Gb nanopore sequencing data from the 4 F₂ pool and 100 Gb data from the 46 F₂ pool, the proportions of nonoverlapping crossovers relative to the total detected crossovers were 5.23% (33/630) and 18.02% (299/1659), respectively (Fig. S10). Thus, we sought to increase the proportion of detected nonoverlapping crossovers by increasing the number of F₂ individuals included in the pool. To determine pool size, we conducted subsampling analysis using a 4 F₂ pool and fitted a generalized linear model (GLM) ($y \sim -0.196 \times \log_e(x) + 0.873$) (Fig. S10i). Based on the COMapper results from the 46 WT F₂ pool, *c.* 1600 crossovers are expected to be detected from 100 Gb of sequencing data. According to our GLM model, sequencing 1000 F₂ individuals at a total depth of 100 Gb—equivalent to *c.* 0.82× coverage per F₂—would yield 90.2% nonoverlapping crossovers, corresponding to 1472 unique crossover events.

To test this hypothesis, we performed nanopore long-read sequencing of gDNA pooled from 1000 Col × *Ler* F₂ recombinant seedlings (10-d-old) derived from WT Col × *Ler* and *recq4ab* Col × *recq4ab* *Ler* F₁ hybrid seedlings, and mapped their crossovers with COMapper (Fig. 5). Much as the GBS analysis of Col × *Ler* and *recq4ab* Col × *recq4ab* *Ler* F₂ individuals (30-d-old plants) detected a 3.67-fold increase in crossover numbers in the *recq4ab* background compared to WT Col × *Ler* F₂ plants (Fig. 5a) (Serra *et al.*, 2018b; Kim *et al.*, 2024), COMapper detected 3.86-fold more crossovers in the *recq4ab* Col × *recq4ab* *Ler* F₂ seedlings compared to WT Col × *Ler* F₂ seedlings (Fig. 5b). COMapper generated crossover landscapes from the pooled F₂ seedlings (*n* = 1000) similar to those obtained by GBS from adult F₂ individuals in both WT and *recq4ab* backgrounds, with crossovers highly enriched along chromosome arms and suppressed in pericentromeric and centromeric regions (Fig. 5c–f). Notably, COMapper detected lower crossover frequencies in specific subtelomeric regions of chromosomes 1 and 3 in *recq4ab* than in WT (Fig. 5e). Collectively, these results suggest that COMapper can be applied to a pool of hybrid F₂ progeny as well as to F₁ pollen for genome-wide mapping of crossovers and distribution in *Arabidopsis*.

High-resolution mapping of meiotic 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 (Fig. 6a) (Choi *et al.*, 2018; Underwood *et al.*, 2018). In agreement with this result, meiotic crossovers in plants are

highly enriched around gene TSS (Choi *et al.*, 2013; Hellsten *et al.*, 2013; Wijnker *et al.*, 2013; Choi & Henderson, 2015; 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 × *Ler* F₁ hybrid pollen showed higher enrichment in gene-proximal promoters but lower enrichment in gene bodies, compared to the same number of random sites (Figs 6a,b, S16a,b). The pattern of crossovers determined by our GBS dataset, which had *c.* 5.2× genome coverage for male meiosis (Kim *et al.*, 2024), showed a similar enrichment around gene TSS as that mapped by COMapper in Col × *Ler* F₁ hybrid pollen pool (Figs 6b, S16b). Furthermore, COMapper and GBS with *c.* 4.5× genome coverage both detect a higher enrichment of crossovers around gene-proximal promoters and a lower enrichment in gene bodies in WT Col × *Ler* F₂ recombinant seedlings compared to random (Figs 6c, S16c), 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 (*c.* 0.1–1×) 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, 2019). However, when crossovers mapped by GBS and TIGER at different genome coverages (0–1×, 1–2×, and ≥ 2×) were plotted around TSS and TES at finer scales—based on the distance between the two SNPs flanking each crossover site—we found that high-resolution crossover mapping by GBS requires a genome coverage of at least 3×, which is comparable to COMapper enabling detection of crossovers at SNP resolution (Fig. S17).

SNP density promotes interfering crossovers and inhibits noninterfering crossovers

We examined the effect of the *recq4ab* mutant on the pattern of crossovers mapped by COMapper or GBS. The loss of RECQ4A/B function resulted in a decreased crossover frequency at the gene TSS and TES, with an increase in crossovers at gene bodies (Figs 6d, S16d). Genetic disruption of *MSH2* was shown to further increase the frequency of class II crossovers in Col × *Ler* F₂ 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 *Ler* accessions at gene TSS and TES, with lower SNP density at distal promoters, gene bodies, and random sites (Fig. 6e). We also found that the physical distance between consecutive SNPs at crossover sites is greater in *recq4ab* than in WT based on nanopore long-read sequencing of F₁ pollen and F₂ progeny (Fig. 6f), suggesting that the increased class II crossovers in *recq4ab* are more sensitive to SNP-mediated crossover

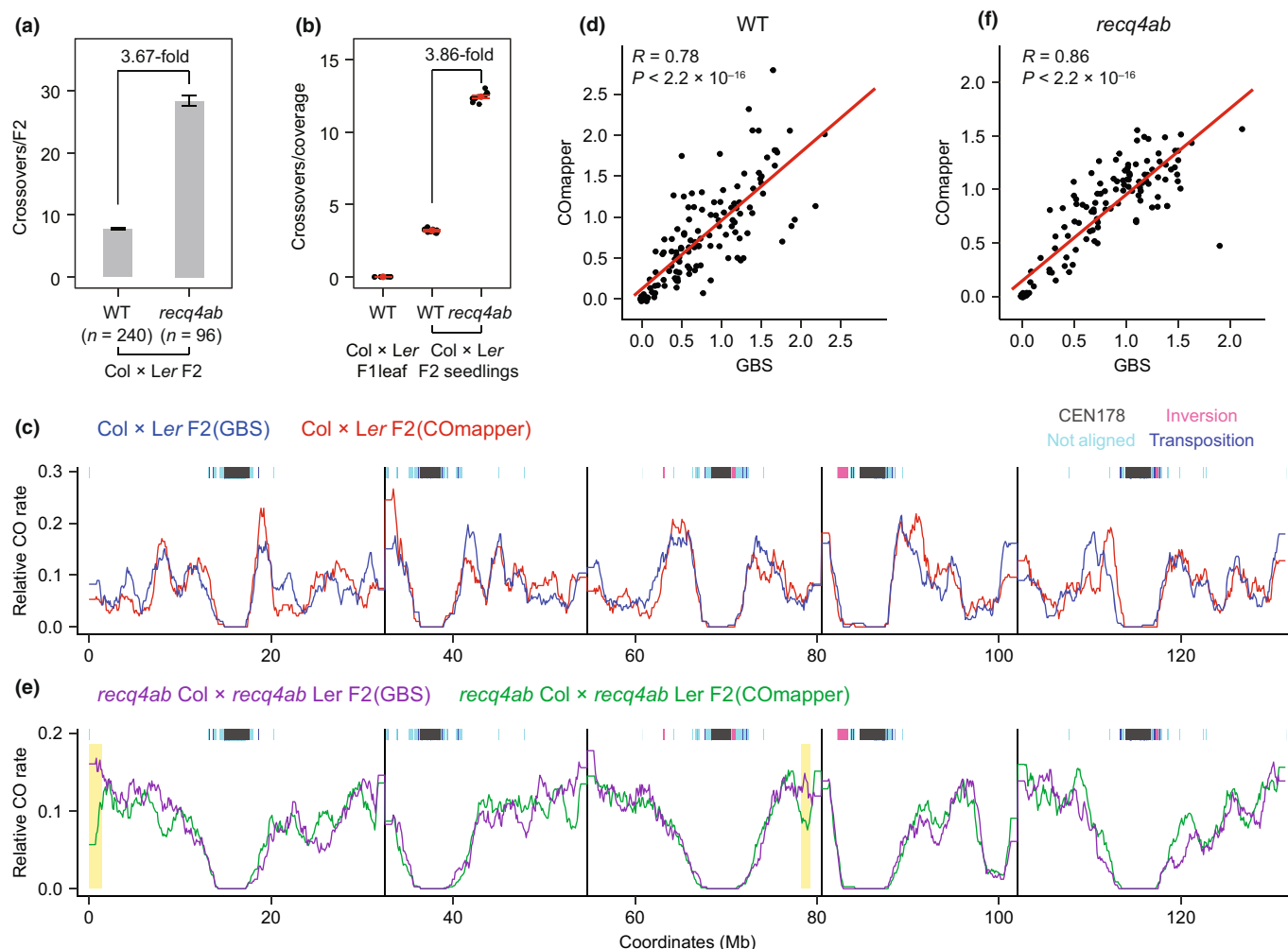


Fig. 5 Mapping of crossovers from pooled F_2 recombinant seedlings by COmapper in *Arabidopsis thaliana*. (a) Number of crossovers per F_2 individual calculated from genotyping by sequencing (GBS) data of wild-type (WT) Col $\times$ Ler F_2 ($n = 240$) (Kim *et al.*, 2024) and *recq4ab* Col $\times$ *recq4ab* Ler F_2 ($n = 96$) plants. n indicates the number of F_2 individual plants. Data are presented as mean $\pm$ SD. (b) Number of crossovers per genome of nanopore sequencing data from gDNA of Col $\times$ Ler F_1 leaves, Col $\times$ Ler F_2 , and *recq4ab* Col $\times$ *recq4ab* Ler F_2 seedlings ($n = 1000$). Each data point indicates the number of crossovers per genome from a randomly selected 1 million reads. Data are presented as mean $\pm$ SD from 10 replicates. (c) Relative frequency of crossovers mapped by GBS (blue, $n = 1852$ crossovers) and COmapper (red, $n = 1390$ crossovers) along the five *Arabidopsis* chromosomes on a continuous x-axis in WT Col $\times$ Ler F_2 plants (Col-CEN). Fisher's exact test was applied to each 100-kb bin to assess differences in relative crossover rates. Bins with false discovery rate (FDR) below 0.05 are highlighted in yellow. (d) Correlation between crossover frequencies in 1-Mb bins of Col $\times$ Ler F_2 individuals (measured by GBS) and Col $\times$ Ler F_2 seedlings (measured by COmapper) across the genome, assessed using Pearson's correlation test. (e) As in (c), but showing the relative frequency of crossovers mapped by GBS in *recq4ab* Col $\times$ *recq4ab* Ler F_2 plants (purple, $n = 2731$ crossovers) or by COmapper in *recq4ab* Col $\times$ *recq4ab* Ler F_2 seedlings (green, $n = 5760$ crossovers). (f) As in (d), but showing the correlation between crossover frequencies of crossovers mapped by *recq4ab* Col $\times$ *recq4ab* Ler F_2 individuals and *recq4ab* Col $\times$ *recq4ab* Ler F_2 seedlings.

inhibition and tend to occur in regions with lower SNP density 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 $\times$ Ler and *recq4ab* Col $\times$ *recq4ab* Ler backgrounds (Fig. 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* (Fig. 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 ($J3^{pro}:J3^{G155R}$) $\times$ Ler and Col $\times$ Di-G F_1 hybrids (Fig. S18). $J3^{G155R}$, a dominant negative allele of $J3$ encoding an 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/B 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,

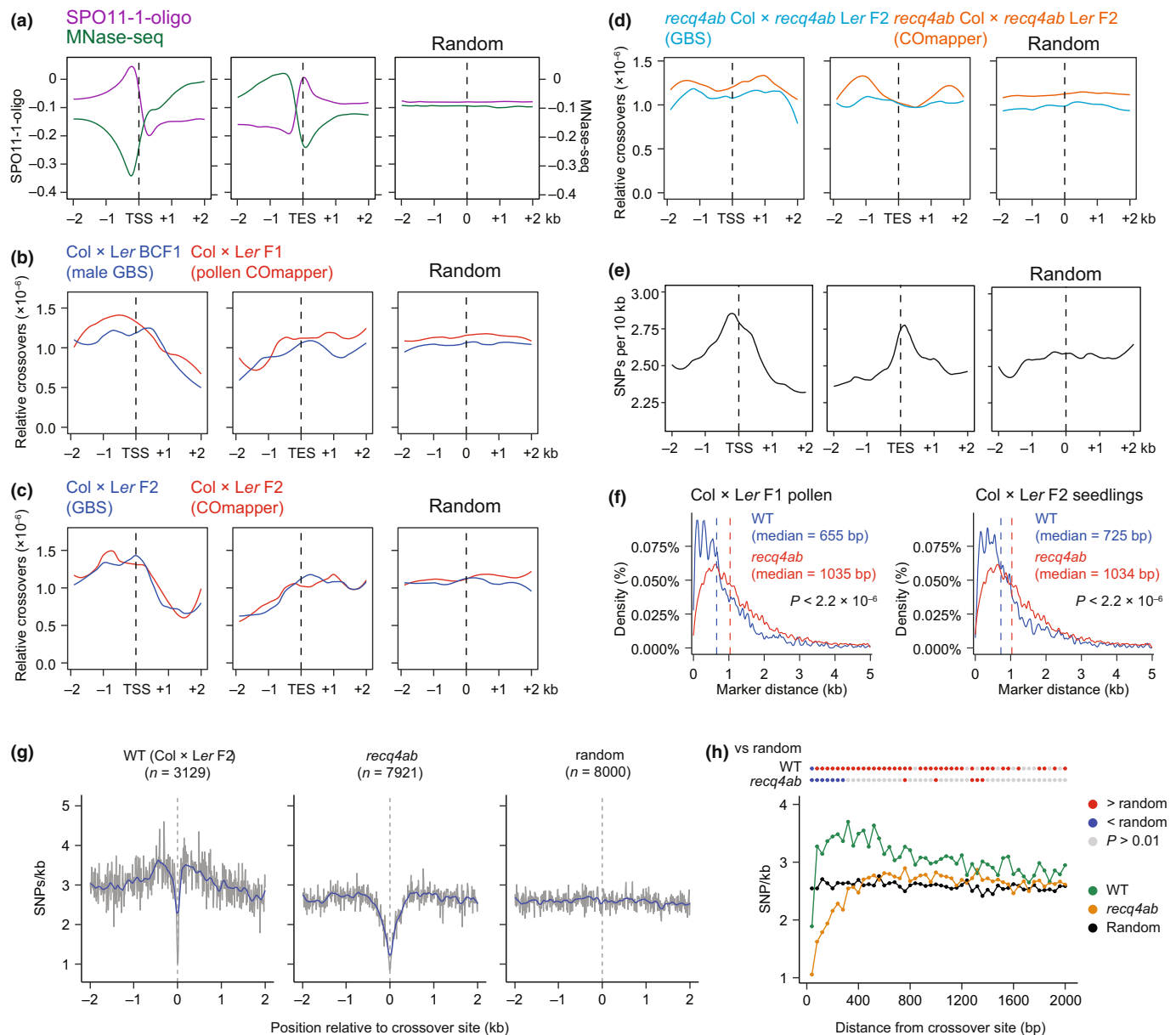


Fig. 6 Analysis of crossovers and single nucleotide polymorphisms (SNPs) around genes in *Arabidopsis thaliana*. (a) Metaplots showing the density of SPO11-1-oligonucleotides (SPO11-1-oligo) and nucleosomes from MNase sequencing (MNase-seq) around gene transcription start sites (TSS) and transcription end sites (TES) in a 2-kb window (Choi *et al.*, 2018). (b) As in (a), but showing the relative rate of crossovers mapped by genotyping by sequencing (GBS; blue, $n = 1824$ crossovers) described previously (Rowan *et al.*, 2019; Kim *et al.*, 2024) and COMapper (red, $n = 1679$ crossovers) in *Arabidopsis* Col × Ler male meiosis. (c) As in (b), but showing the relative rate of crossovers mapped by GBS (blue, 1852 crossovers) and COMapper (red, $n = 1391$ crossovers) in *Arabidopsis* Col × Ler F₂ plants. (d) As in (c), but showing the relative rate of crossovers mapped by GBS (skyblue, $n = 2731$ crossovers) and COMapper (orange, $n = 5760$ crossovers) in *Arabidopsis* recq4ab Col × recq4ab Ler F₂ plants. (e) As for (a), but showing the density of SNPs between Col and Ler accessions. (f) Density and physical distance between two consecutive SNPs at crossover sites in wild-type (WT) Col × Ler and recq4ab Col × recq4ab Ler for F₁ pollen and F₂ progeny. The vertical dashed line indicates the 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 × Ler F₂ and recq4ab Col × recq4ab Ler F₂ 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 nonsignificant differences.

which is consistent with previous observations that high-density SNPs inhibit crossovers at the crossover hotspot scale (Borts & 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/B* 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; Kianian *et al.*, 2018; Serra *et al.*, 2018b; Sun *et al.*, 2019; Nageswaran *et al.*, 2021; Durand *et al.*, 2022; Kim *et al.*, 2022, 2024; Castellani *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 F₂ progeny of F₁ hybrid plants (Figs 3–6). We demonstrated that COMapper can map crossovers along chromosomes at high resolution by validating the detection of crossovers in gDNA pooled from 4 or 46 recombinant F₂ plants, in which crossovers per 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 WT and *recq4ab* F₁ hybrid pollen and F₂ seedlings from self-pollinated F₁ hybrids, with high enrichment at chromosome arms and suppression at pericentromeres and centromeres (Figs 4, 5). At a finer scale, COMapper revealed a strong enrichment of crossovers at gene promoters and a reduction within gene bodies, resembling the distribution of meiotic DSBs (Fig. 6a–c). Analysis of SNP density around crossovers showed distinct patterns: It was positively correlated with class I crossovers but negatively with class II (Fig. 6g,h), demonstrating differential sensitivity of the two crossover pathways to sequence polymorphisms (Ziolkowski *et al.*, 2015; Blackwell *et al.*, 2020; Dlużewska *et al.*, 2023; Szymanska-Lejman *et al.*, 2023).

We observed that COMapper can mis-map some crossovers occurring within SNP-poor genomic regions because accurate crossover mapping depends on both SNP density and read length—each long read must cover a sufficient number of SNPs (Fig. S14). In genomes with low SNP density, alternative methods like GBS, linked-read sequencing, or Hi-C may offer more reliable crossover detection. Like COMapper, Hi-C and linked-read approaches can be applied to pooled samples, but typically provide lower crossover resolution. GBS requires at least 3× coverage to achieve crossover detection resolution comparable to COMapper. In pooled samples, the number of crossovers detected by COMapper increased proportionally with nanopore sequencing depth (Table S3), indicating that deeper coverage enables more comprehensive reconstruction of genome-wide crossover landscapes (Figs 4, 5; Table S1). However, COMapper does not retain individual-level information, such as the number and position of crossovers per sample, which is essential for quantitative trait locus mapping. Additionally, it cannot be used to assess crossover interference, which requires measuring the physical distances between crossovers along the same chromosome, as can be done with GBS in individually genotyped recombinants (Table S3).

In contrast to GBS, COMapper requires a simple gDNA extraction method and pooled gDNA from F₁ hybrid pollen or F₂ 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 diploid and polyploid organisms. However, its successful application in complex genomes depends on the SNP density in crossover regions. Therefore, assessing the density and distribution of SNPs across parental genomes is crucial before applying COMapper. For example, we analyzed SNP distribution between two tomato species (*Solanum lycopersicum* and *S. pimpinellifolium*), demonstrating the scalability of COMapper (Fig. S19). Ongoing improvements in nanopore sequencing technology, such as longer read lengths and lower sequencing error rates, will further enhance its utility for large and complex genomes. For polyploids, COMapper can be applied to hybrid autotetraploid *Arabidopsis* plants derived from natural accessions (e.g. Wa-1 × Bla-5) or colchicine-doubled tetraploid (e.g. Col (4×) × Ler (4×)). Likewise, COMapper could facilitate crossover analysis 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 *c.* 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.

Acknowledgements

We thank Raphael Mercier (MPI for Plant Breeding Research) for providing *recq4ab* seeds. 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).

Competing interests

None declared.

Author contributions

DB, NS and KC designed the project. DB, NS and Jaebeom K developed the COMapper pipeline. DB, NS, HK, Jaehil K, JP, S-jP, HK, Juhyun K, SL, YH, PAZ, IRH and KC performed

experiments and analysis. DB, NS and KC wrote the original draft of the manuscript. All authors revised the manuscript. DB and NS contributed equally to this work.

ORCID

Dohwan Byun  <https://orcid.org/0000-0002-1980-2252>
 Kyuha Choi  <https://orcid.org/0000-0002-4072-3807>
 Ian R. Henderson  <https://orcid.org/0000-0001-5066-1489>
 Youbong Hyun  <https://orcid.org/0000-0002-9421-7061>
 Heejin Kim  <https://orcid.org/0000-0002-6144-775X>
 Hyein Kim  <https://orcid.org/0009-0006-9903-6530>
 Jaebeom Kim  <https://orcid.org/0009-0000-9961-8836>
 Jaeil Kim  <https://orcid.org/0000-0001-8322-5624>
 Juhyun Kim  <https://orcid.org/0000-0001-5462-3391>
 Seula Lee  <https://orcid.org/0009-0005-1822-2466>
 Jihye Park  <https://orcid.org/0000-0002-9619-6575>
 Sang-jun Park  <https://orcid.org/0009-0005-9799-6821>
 Namil Son  <https://orcid.org/0000-0003-4352-5627>
 Piotr A. Ziolkowski  <https://orcid.org/0000-0001-7673-6565>

Data availability

Arabidopsis Genome Initiative locus identifiers for the genes mentioned in this article are as follows: AT3G13170 for SPO11-1, AT1G10930 for RECQ4A, AT1G60930 for RECQ4B, AT1G53490 for HEI10, AT4G30870 for MUS81, AT3G18524 for MSH2, and AT3G44110 for J3. The nanopore sequencing data generated in this study have been deposited in the Array Express database at EMBL-EBL (<http://www.ebi.ac.uk/arrayexpress>): E-MTAB-14367 (<https://www.ebi.ac.uk/biostudies/ArrayExpress/studies/E-MTAB-14367>); E-MTAB-14368 (<https://www.ebi.ac.uk/biostudies/ArrayExpress/studies/E-MTAB-14368>); E-MTAB-14369 (<https://www.ebi.ac.uk/biostudies/ArrayExpress/studies/E-MTAB-14369>); E-MTAB-14381 (<https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-14381>); E-MTAB-15082 (<https://www.ebi.ac.uk/biostudies/ArrayExpress/studies/E-MTAB-15082>); E-MTAB-15083 (<https://www.ebi.ac.uk/biostudies/ArrayExpress/studies/E-MTAB-15083>). Detailed information on the nanopore long-read sequencing and other data for GBS and genomic features used in this study is included in Table S2. 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>).

References

Berchowitz LE, Francis KE, Bey AL, Copenhaver GP. 2007. The role of AtMUS81 in interference-insensitive crossovers in *A. thaliana*. *PLoS Genetics* 3: e132.
 Blackwell AR, Dłuzewska J, Szymanska-Lejman M, Desjardins S, Tock AJ, Kbir N, Lambing C, Lawrence EJ, Bieluszewski T, Rowan B *et al.* 2020. MSH2 shapes the meiotic crossover landscape in relation to interhomolog polymorphism in Arabidopsis. *EMBO Journal* 39: 1–22.
 Borts RH, Haber JE. 1987. Meiotic recombination in yeast: alteration by multiple heterozygosities. *Science* 237: 1459–1465.

Castellani M, Zhang M, Thangavel G, Mata-Sucre Y, Lux T, Campoy JA, Marek M, Huettel B, Sun H, Mayer KFX *et al.* 2024. Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning. *Nature Plants* 10: 423–438.
 Choi K, Henderson IR. 2015. Meiotic recombination hotspots – a comparative view. *The Plant Journal* 83: 52–61.
 Choi K, Reinhard C, Serra H, Ziolkowski PA, Underwood CJ, Zhao X, Hardcastle TJ, Yelina NE, Griffin C, Jackson M *et al.* 2016. Recombination rate heterogeneity within arabidopsis disease resistance genes. *PLoS Genetics* 12: e1006179.
 Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA, Copenhaver GP, Franklin FCH *et al.* 2013. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nature Genetics* 45: 1327–1336.
 Choi K, Zhao X, Tock AJ, Lambing C, Underwood CJ, Hardcastle TJ, Serra H, Kim J, Cho HS, Kim J *et al.* 2018. Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Research* 28: 532–546.
 Crismani W, Girard C, Froger N, Pradillo M, Santos JL, Chelysheva L, Copenhaver GP, Horlow C, Mercier R. 2012. FANCM limits meiotic crossovers. *Science* 336: 1588–1590.
 Dłuzewska J, Dziegielewska W, Szymanska-Lejman M, Gazecka M, Henderson IR, Higgins JD, Ziolkowski PA. 2023. MSH2 stimulates interfering and inhibits non-interfering crossovers in response to genetic polymorphism. *Nature Communications* 14: 6716.
 Durand S, Lian Q, Jing J, Ernst M, Grelon M, Zwicker D, Mercier R. 2022. Joint control of meiotic crossover patterning by the synaptonemal complex and HEI10 dosage. *Nature Communications* 13: 5999.
 Fernandes JB, Naish M, Lian Q, Burns R, Tock AJ, Rabanal FA, Włodzimierz P, Habring A, Nicholas RE, Weigel D *et al.* 2024. Structural variation and DNA methylation shape the centromere-proximal meiotic crossover landscape in Arabidopsis. *Genome Biology* 25: 1–31.
 Fernandes JB, Séguéla-Arnaud M, Larchevêque C, Lloyd AH, Mercier R. 2018. Unleashing meiotic crossovers in hybrid plants. *Proceedings of the National Academy of Sciences, USA* 115: 2431–2436.
 Girard C, Zwicker D, Mercier R. 2023. The regulation of meiotic crossover distribution: a coarse solution to a century-old mystery? *Biochemical Society Transactions* 51: 1179–1190.
 Goel M, Sun H, Jiao WB, Schneeberger K. 2019. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biology* 20: 1–13.
 Gray S, Cohen PE. 2016. Control of meiotic crossovers: from double-strand break formation to designation. *Annual Review of Genetics* 50: 175–210.
 Grelon M, Vezon D, Gendrot G, Pelletier G. 2001. AtSPO11-1 is necessary for efficient meiotic recombination in plants. *EMBO Journal* 20: 589–600.
 He Y, Wang M, Dukowicz-Schulze S, Zhou A, Tiang C-L, Shilo S, Sidhu GK, Eichten S, Bradbury P, Springer NM *et al.* 2017. Genomic features shaping the landscape of meiotic double-strand-break hotspots in maize. *Proceedings of the National Academy of Sciences, USA* 114: 12231–12236.
 Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, Wessler SR, Schmutz J, Willis JH, Rokhsar DS. 2013. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. *Proceedings of the National Academy of Sciences, USA* 110: 19478–19482.
 Henderson IR, Bomblies K. 2021. Evolution and plasticity of genome-wide meiotic recombination rates. *Annual Review of Genetics* 55: 1–21.
 Hunter N. 2015. Meiotic recombination: the essence of heredity. *Cold Spring Harbor Perspectives in Biology* 7: 1–36.
 Keeney S, Lange J, Mohibullah N. 2014. Self-organization of meiotic recombination initiation: general principles and molecular pathways. *Annual Review of Genetics* 48: 187–214.
 Kianian PMA, Wang M, Simons K, Ghavami F, He Y, Dukowicz-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. *Nature Communications* 9: 2370.
 Kim H, Choi K. 2022. Fast and precise: how to measure meiotic crossovers in Arabidopsis. *Molecules and Cells* 45: 273–283.

- Kim H, Kim J, Son N, Kuo P, Morgan C, Chambon A, Byun D, Park J, Lee Y, Park YM *et al.* 2024. Control of meiotic crossover interference by a proteolytic chaperone network. *Nature Plants* 10: 453–468.
- Kim J, Park J, Kim H, Son N, Kim E, Kim J, Byun D, Lee Y, Park YM, Nageswaran DC *et al.* 2022. Arabidopsis HEAT SHOCK FACTOR BINDING PROTEIN is required to limit meiotic crossovers and HEI10 transcription. *EMBO Journal* 41: e109958.
- Lam I, Keeney S. 2014. Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harbor Perspectives in Biology* 7: a016634.
- Lawrence EJ, Gao H, Tock AJ, Lambing C, Blackwell AR, Feng X, Henderson IR. 2019. Natural variation in TBP-ASSOCIATED FACTOR 4b controls meiotic crossover and germline transcription in Arabidopsis. *Current Biology* 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, Grelon M. 2015. The molecular biology of meiosis in plants. *Annual Review of Plant Biology* 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. *Nature Plants* 4: 1010–1016.
- Morgan C, Fozard JA, Hartley M, Henderson IR, Bomblies K, Howard M. 2021. Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in Arabidopsis. *Nature Communications* 12: 4674.
- Nageswaran DC, Kim J, Lambing C, Kim J, Park J, Kim E-J, Cho HS, Kim H, Byun D, Park YM *et al.* 2021. HIGH CROSSOVER RATE1 encodes PROTEIN PHOSPHATASE X1 and restricts meiotic crossovers in Arabidopsis. *Nature Plants* 7: 452–467.
- Naish M, Alonge M, Wlodzimierz P, Tock AJ, Abramson BW, 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 MJ, Pan J, 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, 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 Doijeweert W, Diaz Trivino S, de Haan JR, Sanchez Perez G, Zhang X *et al.* 2020. Meiotic recombination profiling of interspecific hybrid F1 tomato pollen by linked read sequencing. *The Plant Journal* 102: 480–492.
- Rönspies M, Schmidt C, Schindele P, Lieberman-Lazarovich M, Houben A, Puchta H. 2022. Massive crossover suppression by CRISPR–Cas-mediated plant chromosome engineering. *Nature Plants* 8: 1153–1159.
- Rowan BA, Heavens D, Feuerborn TR, Tock AJ, Henderson IR, 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.
- Rowan BA, Patel V, Weigel D, Schneeberger K. 2015. Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping. *G3: Genes, Genomes, Genetics* 5: 385–398.
- Schweiger R, Lee S, Zhou C, Yang T-P, Smith K, Li S, Sanghvi R, Neville M, Mitchell E, Nessa A *et al.* 2024. Insights into non-crossover recombination from long-read sperm sequencing. *bioRxiv*. doi: [10.1101/2024.07.05.602249](https://doi.org/10.1101/2024.07.05.602249).
- 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: TOP3 α and two BLM homologs antagonize crossovers in parallel to FANCM. *Proceedings of the National Academy of Sciences, USA* 112: 4713–4718.
- Serra H, Choi K, Zhao X, Blackwell AR, Kim J, Henderson IR. 2018a. Interhomolog polymorphism shapes meiotic crossover within the Arabidopsis RAC1 and RPP13 disease resistance genes. *PLoS Genetics* 14: 1–28.
- Serra H, Lambing C, Griffin CH, Topp SD, Nageswaran DC, Underwood CJ, Ziolkowski PA, Séguéla-Arnaud M, Fernandes JB, Mercier R *et al.* 2018b. Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. *Proceedings of the National Academy of Sciences, USA* 115: 2437–2442.
- Shilo S, Melamed-Bessudo C, Dorone Y, Barkai N, Levy AA. 2015. DNA crossover motifs associated with epigenetic modifications delineate open chromatin regions in Arabidopsis. *Plant Cell* 27: 2427–2436.
- Sun H, Rowan BA, Flood PJ, Brandt R, Fuss J, Hancock AM, Michelmore RW, Huettig B, Schneeberger K. 2019. Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nature Communications* 10: 4310.
- Szymanska-Lejman M, Dziegielewska W, Dłuzewska J, Kbir N, Bieluszewska A, Poethig RS, Ziolkowski PA. 2023. The effect of DNA polymorphisms and natural variation on crossover hotspot activity in Arabidopsis hybrids. *Nature Communications* 14: 1–16.
- Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 31: 2032–2034.
- Underwood CJ, Choi K. 2019. Heterogeneous transposable elements as silencers, enhancers and targets of meiotic recombination. *Chromosoma* 128: 279–296.
- Underwood CJ, Choi K, Lambing C, Zhao X, Serra H, Borges F, Simorowski J, Ernst E, Jacob Y, Henderson IR *et al.* 2018. Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Research* 28: 519–531.
- Underwood CJ, Henderson IR, Martienssen RA. 2017. Genetic and epigenetic variation of transposable elements in Arabidopsis. *Current Opinion in Plant Biology* 36: 135–141.
- Villeneuve AM, Hillers KJ. 2001. Whence meiosis? *Cell* 106: 647–650.
- Wang Y, Copenhaver GP. 2018. Meiotic recombination: mixing it up in plants. *Annual Review of Plant Biology* 69: 577–609.
- Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. 2021. Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology* 39: 1348–1365.
- Wijnker E, Velikkakam James G, Ding J, Becker F, Klasen JR, Rawat V, Rowan BA, de Jong DF, de Snoo CB, Zapata L *et al.* 2013. The genomic landscape of meiotic crossovers and gene conversions in *Arabidopsis thaliana*. *eLife* 2: e01426.
- Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR. 2015. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. *Genes & Development* 29: 2183–2202.
- Zakharyevich K, Tang S, Ma Y, Hunter N. 2012. Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* 149: 334–347.
- Zickler D, Kleckner N. 2023. Meiosis: dances between homologs. *Annual Review of Genetics* 57: 1–63.
- Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, 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.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 Allele frequency maps of 46 and 4 F2 individuals.

Dataset S2 Information of crossover sites in 46 and 4 F2 individuals.

Dataset S3 Screenshots of false positives.

Dataset S4 Fisher's exact test for Fig. S9c.

Fig. S1 Distribution of read length from long-read nanopore sequencing.

Fig. S2 Nanopore sequencing read coverage.

Fig. S3 Comparison of genome structure and single nucleotide polymorphisms between the Col and *Ler* accessions.

Fig. S4 Validation of COMapper for mapping crossovers using Col $\times$ *Ler* F2 individuals.

Fig. S5 Estimating false positive rates attributable to nanopore sequencing errors.

Fig. S6 Representative simple crossover and complex crossover sites identified by COMapper in nanopore sequencing reads.

Fig. S7 Representative complex crossover sites identified by both COMapper and GBS allele frequency analysis.

Fig. S8 Sources of false positives in COMapper results from pooled 4 or 46 Col $\times$ *Ler* F2 individuals.

Fig. S9 Analysis of crossovers not detected by COMapper from 46 Col $\times$ *Ler* F2 individuals.

Fig. S10 Impact of sequencing coverage on COMapper performance.

Fig. S11 Validation of COMapper using *recq4ab* Col $\times$ *recq4ab* *Ler* F2 recombinant plants.

Fig. S12 Validation of compatibility between TAIR10 and Col-CEN genomes.

Fig. S13 Comparison of genome-wide crossover landscapes from WT Col $\times$ *Ler* F1 pollen COMapper replicates.

Fig. S14 Comparison of SNP distances at crossover sites mapped by GBS and COMapper.

Fig. S15 Genomic region classification based on 8-SNP length and corresponding crossover distributions.

Fig. S16 Crossovers occur more frequently at gene promoters and less frequently at gene bodies than a random distribution.

Fig. S17 Resolution of crossovers mapped by GBS and COMapper.

Fig. S18 Class I crossovers are associated with higher SNP density.

Fig. S19 Analysis of SNP density between *Solanum lycopersicum* and *Solanum pimpinellifolium*.

Table S1 Summary of long-read nanopore sequencing libraries used in this study.

Table S2 Detailed information for the sequencing libraries used in this study.

Table S3 Comparison of genome-wide crossover mapping techniques.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Disclaimer: The New Phytologist Foundation remains neutral with regard to jurisdictional claims in maps and in any institutional affiliations.