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.

Summary

- 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.

Key words: Crossover, Long-read sequencing, Meiosis, RECQ4, Polymorphism

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 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 & Cohen, 2016; Wang & Copenhaver, 2018; Zickler & Kleckner, 2023). Although

61 substantial DSBs are initially formed, only a few are repaired into crossovers by class
62 I and class II repair pathways in most species. In plants, the class I pathway is
63 responsible for ~85–90% of all crossovers, which are formed by a group of pro-
64 crossover factors including the E3 ligase HEI10 (HUMAN ENHANCER OF CELL
65 INVASION NO.10) (Mercier *et al.*, 2015; Wang & Copenhaver, 2018). Class I crossovers
66 are sensitive to interference, a phenomenon that results in multiple crossovers being
67 more widely spaced than would occur randomly. Crossover interference in plants is
68 explained by a coarsening model where the diffusible and aggregating dynamics of
69 HEI10 proteins limit closely spaced crossovers, resulting in 1–3 crossovers per pair of
70 homologous chromosomes (Mercier *et al.*, 2015; Morgan *et al.*, 2021; Girard *et al.*, 2023).
71 In the class II pathway, ~10–15% of all crossovers are non-interfering and mediated
72 by the endonuclease MUS81 (MMS AND UV SENSITIVE 81) (Berchowitz *et al.*, 2007;
73 Zakharyevich *et al.*, 2012). Class II crossovers are inhibited by different pathways
74 involving anti-recombination factors, such as RECQ4A and RECQ4B helicases in
75 plants (Crismani *et al.*, 2012; Séguéla-Arnaud *et al.*, 2015; Mieulet *et al.*, 2018).

76 In plants, the distribution and frequency of meiotic DSBs and crossovers along
77 chromosomes are non-random and depend on chromatin accessibility (Wijnker *et al.*,
78 2013; Choi *et al.*, 2013, 2018; Shilo *et al.*, 2015; He *et al.*, 2017; Underwood *et al.*, 2018).
79 Consequently, meiotic crossovers predominantly occur in gene-rich euchromatic
80 regions and are suppressed in heterochromatic pericentromeres and centromeres. At
81 a finer scale, crossovers are highly enriched at gene transcription start sites (TSS)
82 near nucleosome-depleted promoters, but are rare in heterochromatic transposable
83 elements and repeats (Wijnker *et al.*, 2013; Choi *et al.*, 2013; Choi & Henderson, 2015). In
84 line with this, DNA methylation and histone H3 methylation on lysine 9 (H3K9me2)
85 inhibit crossover formation (Yelina *et al.*, 2015; Underwood *et al.*, 2017, 2018; Underwood
86 & Choi, 2019; Fernandes *et al.*, 2024). Moreover, a high density of DNA polymorphisms
87 and structural variation suppress crossovers (Choi *et al.*, 2016; Serra *et al.*, 2018a;
88 Dluzewska *et al.*, 2023; Fernandes *et al.*, 2024). However, the frequency of meiotic
89 crossovers correlates with the heterozygosity along chromosomes at a kilobase- (kb)
90 or megabase-pair (Mb) scale in *Arabidopsis* (*Arabidopsis thaliana*), which depends on
91 MSH2, a component of the mismatch repair system (Ziolkowski *et al.*, 2015; Dluzewska
92 *et al.*, 2023). Therefore, mapping of genome-wide crossovers at high resolution is

93 important for understanding the genetic and epigenetic effects on meiotic crossovers
94 and genome evolution in various species (Henderson & Bomblies, 2021; Naish *et al.*, 2021;
95 Kim & Choi, 2022). Short-read sequencing and a probabilistic method for genotyping
96 and crossover prediction have been intensively used to determine crossover numbers
97 and distribution in F_2 individuals between *Arabidopsis* accessions (Wijnker *et al.*, 2013;
98 Rowan *et al.*, 2015; Serra *et al.*, 2018b; Lawrence *et al.*, 2019; Nageswaran *et al.*, 2021; Kim
99 *et al.*, 2022, 2024). Recently, linked-read sequencing has been applied to DNA pooled
100 from hybrid pollen grains or individual pollen nuclei of hybrid plants using microfluidics
101 on a 10x Chromium Controller to map crossovers in *Arabidopsis*, tomato (*Solanum*
102 *lycopersicum*) and beak sedge (*Rhynchospora breviuscula*) (Sun *et al.*, 2019; Rommel
103 Fuentes *et al.*, 2020; Castellani *et al.*, 2024).

104 Here, we develop a crossover mapping pipeline, COmapper
105 (https://github.com/KyuhaChoi-Lab/COmapper), using long-read nanopore
106 sequencing of pooled DNA from F_1 pollen or F_2 recombinant seedlings from
107 *Arabidopsis* hybrids derived from a cross between the accessions Columbia-0 (Col)
108 and Landsberg *erecta*-0 (L*er*). We validate the precision of COmapper for genome-
109 wide crossover analysis through long-read nanopore sequencing of F_2 recombinant
110 individuals whose crossover sites were mapped by short-read sequencing. We
111 demonstrate that COmapper allows high-resolution mapping of crossovers using long-
112 read sequencing of pooled genomic DNA from wild-type and *recq4a recq4b* (*recq4ab*)
113 F_1 hybrid pollen grains or from their F_2 recombinant seedlings. We propose that
114 COmapper will be useful for determining crossover patterns at high resolution in
115 diverse plants with different genetic backgrounds and ploidy.

116

117 **Materials and Methods**

118 **Plant materials**

119 The *Arabidopsis* accessions Col and L*er* were used as the wild types for Col × L*er*
120 hybrid control plants. F_1 and F_2 hybrid plants were cultivated in a growth room under
121 the controlled conditions of 20°C, 50–60% relative humidity and 16-h light/8-h dark
122 photoperiod. The *recq4ab* mutants in the Col and L*er* backgrounds were used as

123 previously described (Fernandes *et al.,* 2018). *J3pro:J3*^{G155R} plant was used as
124 described (Kim *et al.,* 2024). The *Arabidopsis* accession Di-G (Dijon, NASC stock
125 number: N28205) was obtained from the Arabidopsis Biological Resource Center.

126

127 **Genotyping by sequencing (GBS)**

128 GBS libraries were prepared as previously described (Rowan *et al.,* 2015, 2019;
129 Nageswaran *et al.,* 2021; Kim *et al.,* 2024). Briefly, genomic DNA (gDNA) was extracted
130 from two or three mature leaves per individual plant using the CTAB (1% [w/v] CTAB,
131 50 mM Tris-HCl pH 8.0, 0.7 M NaCl, 10 mM EDTA) method. Leaves were collected
132 into 2-ml tubes containing four 3-mm stainless steel beads, frozen in liquid nitrogen
133 and ground using a mixer mill (Retsch, MM400). Then, 500 µl of CTAB buffer was
134 added to the ground tissue, mixed vigorously and incubated at 65°C for 30 min.
135 Following a chloroform extraction and gDNA precipitation with the addition of
136 isopropanol, the gDNA pellet was washed with 1 ml 70% (v/v) ethanol twice, air-dried
137 and resuspended in nuclease-free water. 150 ng of gDNA was used to generate one
138 barcoded sequencing library per plant. The barcoded libraries were then pooled and
139 subjected to paired-end 150-bp sequencing using an Illumina HiSeq-X platform
140 (Macrogen, Korea). The TIGER pipeline was used to analyze the sequencing data and
141 identify crossover sites (Rowan *et al.,* 2015).

142

143 **Preparation of genomic DNAs for nanopore long-read sequencing**

144 To purify pollen grains from F_1 hybrid plants and extract their genomic DNA, F_1 plants
145 were planted in soil and grown at 20°C under a 16-h light/8-h dark photoperiod. Whole
146 inflorescences of these hybrid plants were collected and transferred to ice-cold 10%
147 (w/v) sucrose in a 50-ml tubes. Flowers were homogenized with a blender (Tefal,
148 BL3051KR), filtered through 80-µm nylon mesh, and centrifuged at 350*g* for 10 min at
149 4 °C. The supernatant was discarded and the pellet was resuspended in 10% (w/v)
150 sucrose, filtered through a 40-µm cell strainer, and centrifuged at 100*g* for 10 min at
151 4°C. The purified pellet was frozen in liquid nitrogen and ground with a mortar and
152 pestle. The pellet powder was resuspended in CTAB buffer. Proteinase K (NEB,

153 P8107S) and RNase A (20 mg/ml, Roche, 10109142001) were added to a final
154 concentration of 80 U/ml and 20 μg/ml, respectively. To check pollen disruption, a 1-
155 μl sample was diluted in a 1:10 ratio with 10% (w/v) sucrose and observed under a
156 stereomicroscope (Leica, M165 FC). Then, 500 μl of the resuspended pellet was
157 transferred to a 1.5-ml tube, to which an equal volume of 25:24:1 (v/v/v)
158 phenol:chloroform:isoamylalcohol (Sigma, 77617) was added, and the mixture was
159 homogenized by gentle manual shaking. The tube was centrifuged at 15,000*g* for 10
160 min at 4°C, and 400 μl of supernatant was transferred into a new tube, to which a
161 Nanobind disk (Nanobind® Plant Nuclei Kit RT, PacBio, 102-302-000) was added. The
162 tube containing the Nanobind disk was gently mixed, 400 μl isopropanol was added,
163 and the tube was slowly inverted and incubated for 2 h at room temperature on a
164 rotator set at 7 rpm. A magnetic rack (DynaMag™-2, Invitrogen, 12321D) was used to
165 separate the disk and the supernatant, and the supernatant was removed. PW1 buffer
166 in the Nanobind Plant Nuclei Kit was used for washing the disk twice and the
167 supernatant was removed completely by centrifugation. Then, 150 μl EB+ buffer was
168 added onto the disc and gently tabbed and incubated overnight at room temperature.
169 The supernatant was transferred to a new 1.5-ml tube and stored at −20°C. The DNA
170 was quantified using a Qubit dsDNA Broad Range assay kit (Thermo Fisher, Q32853)
171 and integrity checked by gel electrophoresis.

172 For nanopore sequencing of gDNA from 1,000 pooled seedlings, freshly harvested F_2
173 seeds from F_1 hybrid plants were surface-sterilized, germinated, and grown on 1/2 MS
174 agar medium supplemented with 1% sucrose under controlled conditions (20°C, 16-
175 hour light/8-hour dark photoperiod). Seedlings were harvested 10 days after
176 germination and immediately frozen in liquid nitrogen. The 10-day-old seedlings were
177 powdered in liquid nitrogen using a pre-chilled mortar and pestle. gDNA was extracted
178 using the CTAB method and quantified using a Qubit dsDNA Broad Range assay kit
179 (Thermo Fisher, Q32853). Nine micrograms of gDNA from pollen or seedlings was
180 used to construct a nanopore long-read sequencing library using a Ligation
181 Sequencing Kit V14 (Nanopore, SQK-LSK114). The libraries were sequenced on a
182 PromethION platform (BGI, Hong Kong).

183

184 **Alignment of nanopore sequencing reads to the reference genome**

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

192

193 **COmapper pipeline**

194 To obtain high-quality SNPs distinguishing the Col and L*er* accessions, we followed
195 the previous method with modifications (Fernandes *et al.,* 2024). SNPs within syntenic
196 regions were identified by comparing the reference genomes of Col and L*er* using
197 SyRI (Goel *et al.,* 2019). These SNPs were validated using sequencing data from a total
198 of 2,160 Col × L*er* F_2 plants were used (Rowan *et al.,* 2019). To retain only high-quality
199 SNPs, the following criteria were applied: (1) genotyping results available for at least
200 40 F_2 plants, (2) allele frequency between 0.4 and 0.6, and (3) adherence to Mendelian
201 segregation across the 2,160 F_2 individuals (chi-square test, *P*-value > 0.1). These
202 sites were then integrated with an independent set of high-quality SNP calls from the
203 previous report (Yelina *et al.,* 2015).

204 The genotypes for the SNPs in masked alignments were converted to string format
205 using a reference list of 226,864 high-quality Col/L*er* SNPs. Genotype strings with two
206 successive identical genotypes at either end (CC-CC or LL-LL) were identified as
207 parental reads. Genotype strings that satisfied the following conditions were identified
208 as crossover reads: The two successive SNP genotypes at one end differed from
209 those at the other end (CC-LL or LL-CC), and the four successive SNP genotypes
210 before and after the crossover site matched the genotype at the respective end of the
211 genotype string (CCCC-LLLL or LLLL-CCCC). Genotype strings with less than eight
212 characters or those not identified as either crossover or parental reads were excluded
213 from the calculation of crossover frequency. Regions exhibiting a depth of coverage

214 exceeding the average genome read depth by three-fold and exhibiting a high
215 frequency of false positives in the negative control provided by the sequencing of Col
216 × L*er* F_1 leaf gDNA were excluded from the analysis. The standard output provided
217 the input alignment information, the genotyping summary, the total number of
218 crossover and parental reads, the length of the crossover and parental reads and the
219 information regarding crossover reads in a CSV file.

220

221 **Bioinformatic analysis of crossovers and genomic features.**

222 To analyze the genome-wide pattern of crossovers as well as genes, transposable
223 elements, DNA methylation, SPO11-1–oligos, and MNase-seq (micrococcal nuclease
224 digestion with deep sequencing) data, the TAIR10 genome was tiled into 100-kb bins.
225 DNA methylation levels in each bin were used for plotting. The coverage of SPO11-
226 1–oligo and MNase-seq data was normalized by sequencing reads of randomly
227 digested gDNA before plotting (Choi *et al.,* 2018). The number of overlaps for genes,
228 transposable elements, and crossovers within each bin was calculated and plotted.
229 The coverage of SPO11-1–oligo and MNase-seq data and the frequency of crossovers
230 around genes were calculated using deepTools computeMatrix (v. 3.5.4) to generate
231 gene metaplots. Regions within the pericentromeric regions and where the distance
232 between consecutive SNPs was >4,000 bp were excluded from all analyses to
233 increase resolution of crossover sites. For the analysis of SNP density around
234 crossover site, the SNP frequency in 10-bp bins within a 4-kb window centered on
235 each crossover site was calculated using deepTools computeMatrix (v. 3.5.4).

236

237 **Accession numbers**

238 Arabidopsis Genome Initiative locus identifiers for the genes mentioned in this article
239 are as follows: AT3G13170 for SPO11-1, AT1G10930 for RECQ4A, AT1G60930 for
240 RECQ4B, AT1G53490 for HEI10, AT4G30870 for MUS81, AT3G18524 for MSH2,
241 and AT3G44110 for J3.

242

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*) (Fig. **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 (Fig. **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 (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 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 (Fig. **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 (Fig. **1b,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 (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 × 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

274 kb), comparable to that of gDNA extracted from F_2 seedlings using the CTAB method
275 (Figs **2b, S1**). To achieve deep sequencing coverage, we used ~9 µg gDNA on the
276 PromethION device, which produced reads corresponding to ~100–150 Gb of
277 sequence per flow cell (Fig. **S1**; Table **S1**). We successfully aligned ~90.7% of the
278 nanopore reads to the TAIR10 *Arabidopsis* reference genome using the minimap2
279 alignment tool (Fig. **2c**). The aligned reads covered the genome uniformly from each
280 chromosome end to end except across centromere-proximal regions, with a mean
281 read coverage frequency of 904.9 (Figs **2d, S2**). We used a high-quality SNP dataset
282 distinguishing between Col and L*er* accessions previously used for GBS (Rowan *et al.,*
283 2015; Serra *et al.,* 2018b; Kim *et al.,* 2024), with a 524-bp mean distance between
284 consecutive SNPs (Fig. **2e,f**). We noted that structural variation between Col and L*er*
285 results in decreased coverage of the aligned long reads and fewer SNPs for GBS in
286 the pericentromeric regions near centromeres (Figs **2d,e, S3**). Of all aligned long
287 reads, 86.1% contained more than eight SNPs per read, with a mean 19.8 SNPs per
288 read (Fig. **2g**), which were used to determine crossover sites by COmapper.

289

290 **Development and validation of COmapper**

291 We sought to establish a crossover mapping model for COmapper, which scans for
292 the presence of SNPs from the 5′ to the 3′ end of a long read and identifies simple
293 crossovers (SCOs) where consecutive Col-specific SNPs switch to consecutive L*er*-
294 specific SNPs, or vice versa, by counting the numbers of SNPs (Fig. **3a**). In
295 *Arabidopsis*, approximately 50% of crossovers are associated with gene conversion
296 (Wijnker *et al.,* 2013), which are detected in COmapper data as SCOs characterized by
297 a single switch between parental SNPs. However, a subset of crossovers contain two
298 or more switches between parental SNPs, representing complex crossover (CCOs)
299 (Schweiger *et al.,* 2024). Therefore, we also considered CCOs when identifying
300 crossover sites by examining the number and positions of consecutive Col-specific or
301 L*er*-specific SNPs (Fig. **3b**). Accordingly, we developed and tested six crossover
302 mapping models (Fig. **3b**), which use different numbers of consecutive Col-specific
303 and L*er*-specific SNPs with or without multiple switches between parental SNPs within
304 the crossover tract to map crossovers. To validate these models, we performed

305 nanopore sequencing of two independent pooled gDNA samples from 4 or 46 Col ×
306 L*er* F_2 recombinant plants in which crossover sites had already been determined by
307 GBS (Fig. **S4**; Table **S2**). We then examined the precision of each crossover mapping
308 model by aligning the long-read molecules to the *Arabidopsis* genome and analyzing
309 their SNPs. This revealed that the 8-SCO and 8-CO models were the most accurate
310 in detecting the true crossover sites mapped by GBS, with precisions of 97.3% (8-
311 SCO) and 97.4% (8-CO), respectively (Welch′s *t*-test; 8-SCO vs 8-CO P = 0.747, 8-
312 CO vs 7-CO $P = 1.38\times 10^{-11}$) (Figs **3c–e**). In contrast, 4-SCO model showed a high
313 number of false positives, likely due to the inherent per-base sequencing error rate
314 (3.98%, Q score = 14) of nanopore reads (Fig. **S5**). The 8-CO model, which identifies
315 both SCO and CCO types, detected more crossovers than the 8-SCO model, which
316 detected only the SCO type (Welch′s *t*-test, $P = 2.82 \times 10^{-10}$) (Figs **3c,d, S6,7**).
317 COmapper with the 8-CO model detected ~7.3% and ~1.5% of the CCO type of total
318 crossover sites from the pooled gDNA of 4 and 46 F_2 plants, respectively (Figs **3f,**
319 **S6,7**). Furthermore, using COmapper with the 8-CO model, we identified crossover
320 sites in nanopore reads of pooled hybrid pollen and pooled F_2 seedlings of wild-type
321 (WT) Col × L*er* and *recq4ab* Col × *recq4ab* L*er*. Among detected crossovers, ~8.9%
322 (WT) and ~7.8% (*recq4ab*) were CCOs in F_1 hybrid pollen. In F_2 seedlings (n = 1,000),
323 CCOs accounted for 2.1% (WT) and 2.9% (*recq4ab*) similar to the 2.1% observed to
324 46 F_2 plants analyzed by GBS and COmapper (Fig. **3f**).

325 To validate the output of COmapper, we used the crossover sites mapped by GBS
326 and TIGER with manual curation as true positives from 4 and 46 Col × L*er* F_2
327 recombinant plants (33 crossovers for 4 F_2 plants, 342 crossovers for 46 F_2 plants)
328 (Dataset **S1** and **S2**). COmapper with the 8-CO model was capable of identifying true
329 positive crossover sites multiple times per site, whereas ~3% false positives were
330 detected only once or twice (Fig. **3g,h**). Analysis of these false positives revealed four
331 sources: chimeric reads with error clusters (16 in 4 F_2, 16 in 46 F_2), substitution errors
332 (5 in 4 F_2, 12 in 46 F_2), chimeric reads with large indels (3 in 46 F_2), and mitotic
333 recombination (10 in 46 F_2), reflecting limitations of nanopore sequencing (Fig. **S8**;
334 Dataset **S3**).

335 COmapper accurately identified all 33 crossovers that were mapped by GBS in the 4
336 F_2 recombinants, with a mean of 18.9 long reads covering each crossover site when

337 reaching a sequencing depth of 40.14 Gb (Fig. **S4a**,**b**). Of the 342 crossovers mapped
338 by GBS in the 46 F_2 recombinants, COmapper accurately mapped 299 crossovers,
339 with a mean of 6.07 long reads per crossover site. Notably, COmapper failed to detect
340 the remaining 43 crossovers. To analyze the false negatives, we categorized the
341 detected (299) and undetected crossovers (43) from the 46 F_2 pool according to the
342 8-SNP length around the crossover site (Fig. **S9a,b**). While crossovers were more
343 concentrated in the SNP-rich regions, the proportion of undetected crossovers relative
344 to the total crossover (detected and undetected crossovers) increases as the 8-SNP
345 length increases (Fig. **S9a,b**). It indicates the dependance of COmapper's crossover
346 mapping sensitivity on SNP density. However, the undetected crossovers (false
347 negatives) were not concentrated at certain region of genome but randomly distributed
348 at genome-wide scale (Fig. **S9c**; Dataset **S4**).

349 As COmapper depends on the depth of sequencing, we conducted a comprehensive
350 analysis of how sequencing coverage influences COmapper's performance in terms
351 of true positive (TP) and false positive (FP) crossover read detection, precision (TP /
352 [TP + FP]), and false negatives (crossovers detected by GBS but not by COmapper)
353 (Fig. **S10**). We found that both TP and FP reads, including those from the same
354 crossover sites, increase proportionally with sequencing depth, keeping precision
355 stable across coverage levels (Fig. **S10a–f**). As coverage per F_2 increases, false
356 negatives decrease. At the same time, the discovery rate of new unique crossovers
357 declines due to repeated detection of previously identified sites (Fig. **S10g,h**).
358 Additionally, we found that pooling more individuals increases the likelihood of
359 capturing non-overlapping crossover sites, thereby reducing redundancy and
360 maximizing the number of distinct crossovers (Fig. **S10i**). This validation shows that
361 SNP-resolution crossover mapping by COmapper requires long reads with at least
362 eight SNPs and sufficient sequencing depth (Table **S1**). Taken together, these results
363 confirm that our COmapper pipeline with an 8-CO model (hereafter referred to simply
364 as 'COmapper') can directly map crossovers in long-read nanopore sequencing
365 molecules at SNP resolution with an overall precision of 97%.

366 To further assess COmapper's performance in the *recq4ab* background, we re-
367 sequenced genomic DNA from 39 *recq4ab* Col × L*er* F_2 individuals—previously
368 analyzed by GBS—using nanopore sequencing at a total coverage of 410× (~10.5×

369 per individual) (Fig. **S11**). COmapper achieved ~96% precision, comparable to the 97%
370 observed in the WT dataset, demonstrating reliable crossover detection in both
371 genotypes. In addition, we confirmed that COmapper is highly robust to Col-CEN
372 (Telomere-to-Telomere) reference genome as well as TAIR10 in crossover detection
373 using 46 F_2 individuals (Fig. **S12**).

374

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

377 Next, we used the COmapper and long-read nanopore sequencing technology to
378 directly map crossovers from pooled gDNA samples of WT Col × L*er* and *recq4ab* Col
379 × *recq4ab* L*er* F_1 hybrid pollen grains (Fig. **4**). Our GBS analysis showed a 3.12-fold
380 increase in crossovers during male meiosis in the *recq4ab* mutant compared to the
381 WT, consistent with the previous report (Fig. **4a**) (Fernandes *et al.,* 2018). Similarly,
382 COmapper detected a 2.41-fold increase in crossovers per genome (119 Mb) when
383 analyzing long-read sequencing data produced from *recq4ab* F_1 hybrid pollen DNA
384 (mean = 3.92 crossovers per genome, ~763.8x coverage) compared to the WT (mean
385 = 1.62 crossovers per genome, 1,072.3x coverage) (Welch′s *t*-test, $P = 1.22 \times 10^{-12}$)
386 (Fig. **4b**). Importantly, COmapper rarely detected crossovers as false positive
387 crossover events in the negative control sample (leaf genomic DNA from Col × L*er* F_1
388 hybrids; mean = 0.03 crossovers per genome, 82.8x coverage) compared to the WT
389 Col × L*er* F_1 hybrid pollen (Welch′s *t*-test, $P = 5.51 \times 10^{-10}$) (Fig. **4b**; Table **S1**). These
390 results indicate that COmapper enables the detection of meiotic crossovers in pooled
391 gDNA from Col × L*er* hybrid pollen grains, with an increased number of crossovers in
392 a *recq4ab* background.

393 Previously, GBS-based crossover maps of *Arabidopsis* Col × L*er* male meiosis using
394 a backcross (BC) population detected ~4.47 crossovers per genome (Kim *et al.,* 2024).
395 In comparison, COmapper detected ~1.65 crossovers per genome in nanopore long-
396 read sequencing of pooled pollen grains from Col × L*er* hybrid plants, with all detected
397 1,204 crossovers being non-overlapping. This suggests an inherent feature of
398 COmapper method, which samples crossover molecules directly from long reads,
399 rather than reconstructing full meiotic products, as is done in GBS. To further validate

400 COmapper's performance, we performed an independent COmapper experiment
401 using pollen from WT Col × L*er* F_1 plants, confirming the robustness of its crossover
402 detection (Figs **4b, S13**). Despite detecting fewer absolute events, our validation
403 demonstrates that COmapper provide a consistent and representative subset of
404 crossovers relative to sequencing coverage.

405 To compare genome-wide crossover patterns detected by COmapper and GBS using
406 the Col-CEN complete genome (Kim *et al.,* 2024), we calculated relative crossover rates
407 as the proportion of crossovers in each 100-kb genomic bin, normalized to the total
408 number of crossovers (Fig. **4c–f**). We used Fisher's exact test to assess crossover
409 frequency differences and found no significant regional differences between
410 COmapper and GBS, in either WT or *recq4ab* datasets (Fig. **4c,e**). We also observed
411 that the genomic landscape of crossovers mapped by COmapper correlated with that
412 of GBS (Fig. **4d,f**), indicating that COmapper and GBS produce highly comparable
413 genome-wide crossover landscapes. Notably, we identified a ~3.2 Mb inversion (Chr4:
414 14,564,462–17,818,793) in *recq4ab* via nanopore long-read analysis, which likely
415 underlies crossover suppression observed in both GBS and COmapper, consistent
416 with the previous report (Serra *et al.,* 2018b). Inversions interfere with homolog pairing
417 and synapsis, and single crossovers within them can produce dicentric or acentric
418 chromosomes, leading to suppression or rare double crossovers (Rönspies *et al.,* 2022).
419 Although not statistically significant, COmapper detected slightly more crossovers
420 within the *recq4ab* inversion (Fig. **4e**) than GBS, possibly due to selection against non-
421 viable recombinants in the F_2 population.

422 Despite the comparable crossover landscapes generated by both approaches, we
423 noted that COmapper detects fewer crossovers at sub-telomeric regions than GBS by
424 plotting the crossovers of male meiosis along a telomere-to-centromere axis that
425 combined all five chromosomes (Fig. **S14a**). The mean SNP density was lower at sub-
426 telomeric regions than in mid-chromosomal regions (Fig. **S14b,c**). Additionally, the
427 physical distances covered by the eight SNPs required for crossover-site mapping by
428 COmapper were shorter than those identified by GBS in the sub-telomeric, mid-
429 chromosomal and pericentromeric regions (Fig. **S14d–f**).

430 To further assess the effect of SNP density on the crossover detection sensitivity of
431 COmapper, we divided the genome into four categories based on the SNP density,
432 defined by the 8-SNP span (Fig. **S15a,b**), and compared the distribution of COmapper-
433 detected crossovers across these regions with those of GBS-detected crossovers (Fig.
434 **S15c**). In the WT male GBS dataset, ~14.6% of crossovers were found in
435 pericentromeric and subtelomeric regions with 8-SNP spans ≥10 kb, in contrast, only
436 1.8% of crossovers in WT pollen COmapper dataset were detected in these SNP-poor
437 regions (Fig. **S15c**). Since GBS-based crossover detection is not affected by SNP
438 density, this result indicates that COmapper has reduced sensitivity in SNP-poor
439 regions and preferentially detects crossovers in SNP-rich areas. However, because
440 crossovers in SNP-poor pericentromeric regions account for only a small fraction of
441 total crossovers, and a subset of SNP-poor regions are interspersed among SNP-rich
442 regions. (Fig. **S15a**), the overall effect of SNP density on crossover rate profiles along
443 chromosomes was minimal when analyzed at the megabase scale. Nonetheless,
444 comparisons of crossover rates between COmapper and GBS should be interpreted
445 cautiously in SNP-poor regions, such as subtelomeric and pericentromeric regions,
446 particularly at high resolution (e.g., kilobase scale), where detection sensitivity may
447 differ signficantly. COmapper requires long reads (≥10–15 kb) containing at least eight
448 SNPs for reliable crossover detection, underscoring its limitation in regions with spare
449 SNP density. Despite this, we demonstrate that COmapper can effectively map
450 genome-wide crossovers and capture the genetic effect of the *recq4ab* mutant on
451 crossover frequency and distribution by nanopore sequencing of pooled hybrid pollen
452 DNA.

453

454 **Mapping of crossovers from pooled F_2 recombinant seedlings by COmapper**

455 GBS analysis of Col × L*er* F_2 individuals revealed ~8 crossovers per F_2 (Serra *et al.*,
456 2018b; Fernandes *et al.*, 2018; Rowan *et al.*, 2019; Nageswaran *et al.*, 2021; Kim *et al.*, 2022,
457 2024). From the COmapper analysis of 40 Gb nanopore sequencing data from the 4
458 F_2 pool and 100 Gb data from the 46 F_2 pool, the proportions of non-overlapping
459 crossovers relative to the total detected crossovers were 5.23% (33/630) and 18.02%
460 (299/1659), respectively (Fig. **S10**). Thus, we sought to increase the proportion of

461 detected non-overlapping crossovers by increasing the number of F_2 individuals
462 included in the pool. To determine pool size, we conducted subsampling analysis
463 using 4 F_2 pool and fitted generalized linear model (GLM) ($y \sim -0.196 \times ln(x) + 0.873$)
464 (Fig. **S10i**). Based on the COmapper results from the 46 WT F_2 pool, approximately
465 1,600 crossovers are expected to be detected from 100 Gb of sequencing data.
466 According to our GLM model, sequencing 1,000 F_2 individuals at a total depth of 100
467 Gb—equivalent to ~0.82× coverage per F_2—would yield 90.2% non-overlapping
468 crossovers, corresponding to 1,472 unique crossover events.

469 To test this hypothesis, we performed nanopore long-read sequencing of gDNA pooled
470 from 1,000 Col × L*er* F_2 recombinant seedlings (10 days old) derived from WT Col ×
471 L*er* and *recq4ab* Col × *recq4ab* L*er* F_1 hybrid seedlings, and mapped their crossovers
472 with COmapper (Fig. **5**). Much as the GBS analysis of Col × L*er* and *recq4ab* Col ×
473 *recq4ab* L*er* F_2 individuals (30-day-old plants) detecting a 3.67-fold increase in
474 crossover numbers in the *recq4ab* background compared to WT Col × L*er* F_2 plants
475 (Fig. **5a**) (Serra *et al.,* 2018b; Kim *et al.,* 2024), COmapper detected 3.86-fold more
476 crossovers in the *recq4ab* Col × *recq4ab* L*er* F_2 seedlings compared to WT Col × L*er*
477 F_2 seedlings (Fig. **5b**). COmapper generated crossover landscapes from the pooled
478 F_2 seedlings (n = 1,000) similar to those obtained by GBS from adult F_2 individuals in
479 both WT and *recq4ab* backgrounds, with crossovers highly enriched along
480 chromosome arms and suppressed in pericentromeric and centromeric regions (Fig.
481 **5c**–**f**). Notably, COmapper detected lower crossover frequencies in specific
482 subtelomeric regions of chromosome 1 and 3 in *recq4ab* compared to WT (Fig. **5e**).
483 Collectively, these results suggest that COmapper can be applied to a pool of hybrid
484 F_2 progeny as well as to F_1 pollen for genome-wide mapping of crossovers and
485 distribution in *Arabidopsis*.

486

487 **High-resolution mapping of meiotic crossovers around genes**

488 In *Arabidopsis*, genome-wide meiotic DSB profiling by sequencing SPO11-1–
489 oligonucleotides (SPO11-1–oligos) revealed that meiotic DSBs occur predominantly
490 in nucleosome-depleted regions, with the highest peak observed over gene-proximal
491 promoters, while the formation of DSBs at the exons of protein-coding genes with high

492 nucleosome density is suppressed (Fig. **6a**) (Underwood *et al.*, 2018; Choi *et al.*, 2018).
493 In agreement with this result, meiotic crossovers in plants are highly enriched around
494 gene TSS (Hellsten *et al.*, 2013; Wijnker *et al.*, 2013; Choi *et al.*, 2013; Choi & Henderson,
495 2015; Shilo *et al.*, 2015; Kianian *et al.*, 2018). Because we generated high-resolution
496 crossover maps using COmapper, we examined the position and frequency of
497 crossovers around the TSS and transcription end sites (TES) of all genes in a 4-kb
498 window (2 kb on either side of the TSS or TES). Like meiotic DSBs that were previously
499 determined by sequencing of SPO11-1–oligos (Choi *et al.*, 2018), the crossovers
500 mapped by COmapper in gDNA pooled from WT Col × L*er* F_1 hybrid pollen showed
501 higher enrichment in gene-proximal promoters but lower enrichment in gene bodies,
502 compared to the same number of random sites (Figs **6a,b, S16a,b**). The pattern of
503 crossovers determined by our GBS dataset, which had ~5.2x genome coverage for
504 male meiosis (Kim *et al.*, 2024), showed a similar enrichment around gene TSS as that
505 mapped by COmapper in Col × L*er* F_1 hybrid pollen pool (Figs **6b, S16b**). Furthermore,
506 COmapper and GBS with ~4.5x genome coverage both detect a higher enrichment of
507 crossovers around gene-proximal promoters and a lower enrichment in gene bodies
508 in WT Col × L*er* F_2 recombinant seedlings compared to random (Figs **6c, S16c**),
509 suggesting that chromatin accessibility is a major determinant in shaping crossovers
510 as well as meiotic DSBs around genes in *Arabidopsis*.

511 The GBS with sparse coverage (~0.1–1x) and TIGER, a GBS pipeline for mapping
512 crossovers are inexpensive and useful for measuring the number of crossovers and
513 interference from many recombinant individuals (Rowan *et al.*, 2015, 2019). However,
514 when crossovers mapped by GBS and TIGER at different genome coverages (0–1×,
515 1–2×, and ≥2×) were plotted around TSS and TES at finer scales—based on the
516 distance between the two SNPs flanking each crossover site—we found that high-
517 resolution crossover mapping by GBS requires a genome coverage of at least 3×,
518 which is comparable to COmapper enabling detection of crossovers at SNP resolution
519 (Fig. **S17**).

520

521 **SNP density promotes interfering crossovers and inhibits non-interfering**
522 **crossovers**

523 We examined the effect of the *recq4ab* mutant on the pattern of crossovers mapped
524 by COmapper or GBS. The loss of RECQ4A/B function resulted in a decreased
525 crossover frequency at the gene TSS and TES, with an increase in crossovers at gene
526 bodies (Figs **6d, S16d**). Genetic disruption of *MSH2* was shown to further increase the
527 frequency of class II crossovers in Col × L*er* F_2 progeny in the *recq4ab* background,
528 suggesting that a high density of DNA polymorphisms inhibits the formation of class II
529 crossovers in *recq4ab* (Dluzewska *et al.,* 2023). Consistent with this possibility, we
530 detected peaks for high SNP density between the Col and L*er* accessions at gene
531 TSS and TES, with lower SNP density at distal promoters, gene bodies and random
532 sites (Fig. **6e**). We also found that the physical distance between consecutive SNPs
533 at crossover sites is greater in *recq4ab* than in WT based on nanopore long-read
534 sequencing of F_1 pollen and F_2 progeny (Fig. **6f**), suggesting that the increased class
535 II crossovers in *recq4ab* are more sensitive to SNP-mediated crossover inhibition and
536 tend to occur in regions with lower SNP density compared to class I crossovers in WT.

537 To further investigate the effect of SNP density on the distribution of class I and class
538 II crossovers, we plotted the density of SNPs in a 2-kb window, upstream and
539 downstream of the center positions of crossovers mapped by COmapper and GBS, in
540 the WT Col × L*er* and *recq4ab* Col × *recq4ab* L*er* backgrounds (Fig. **6g**). The density
541 of SNPs around the crossovers was higher in WT but lower in the *recq4ab* background
542 compared to that at random sites, with the lowest density seen at the center of
543 crossover positions in both WT and *recq4ab* (Fig. **6g,h**). We also found a similarly
544 higher density of SNPs around crossover sites compared to random sites in the GBS
545 analysis of F_2 individuals from Col expressing the $J3^{G155R}$ allele (*J3pro:*$J3^{G155R}$) × L*er*
546 and Col × Di-G F_1 hybrids (Fig. **S18**). $J3^{G155R}$, a dominant negative allele of *J3*
547 encoding a HSP40 co-chaperone, increases class I crossovers by enhancing the
548 abundance of HEI10 (Kim *et al.,* 2024). These results suggest that the class I crossovers
549 occur preferentially in the regions with higher SNP density compared to random sites,
550 whereas the increased class II crossovers due to the loss of RECQ4A/B function occur
551 mainly in regions with lower SNP density. It is worth noting that the SNP density
552 nearest to the crossover site is lower for both class I and class II crossovers, which is
553 consistent with previous observations that high-density SNPs inhibit crossovers at the
554 crossover hotspot scale (Borts & Haber, 1987; Serra *et al.,* 2018a; Szymanska-Lejman *et al.,*

555 2023). Taken together, these results indicate that COmapper allows the generation of
556 high-resolution genome-wide crossover maps and reveals the effects of SNP density
557 and genetic mutation of *RECQ4A/B* on the positioning and frequency of crossovers.

558

559 **Discussion**

560 Next-generation and linked-read sequencing technologies have contributed to our
561 understanding of the patterning of genome-wide meiotic crossovers in diverse species
562 (Rowan *et al.*, 2015; Serra *et al.*, 2018b; Kianian *et al.*, 2018; Sun *et al.*, 2019; Nageswaran *et*
563 *al.*, 2021; Kim *et al.*, 2022, 2024; Durand *et al.*, 2022; Castellani *et al.*, 2024). Here, using
564 nanopore long-read sequencing (Wang *et al.*, 2021), we developed COmapper that can
565 generate high-resolution maps of meiotic crossovers from pooled pollen and F_2
566 progeny of F_1 hybrid plants (Figs. **3–6**). We demonstrated that COmapper can map
567 crossovers along chromosomes at high resolution by validating the detection of
568 crossovers in genomic DNA pooled from 4 or 46 recombinant F_2 plants, in which
569 crossovers per each plant were mapped using the GBS approach with individual
570 barcoding.

571 At a chromosomal scale, COmapper generated genome-wide crossover landscapes
572 comparable to those of GBS in WT and *recq4ab* F_1 hybrid pollen and F_2 seedlings from
573 self-pollinated F_1 hybrids, with high enrichment at chromosome arms and suppression
574 at pericentromeres and centromeres (Figs. **4,5**). At a finer scale, COmapper revealed
575 a strong enrichment of crossovers at gene promoters and a reduction within gene
576 bodies, resembling the distribution of meiotic DSBs (Fig. **6a–c**). Analysis of SNP
577 density around crossovers showed distinct patterns: it was positively correlated with
578 class I crossovers but negatively with class II (Fig. **6g,h**), demonstrating differential
579 sensitivity of the two crossover pathways to sequence polymorphisms (Ziolkowski *et al.*,
580 2015; Blackwell *et al.*, 2020; Dluzewska *et al.*, 2023; Szymanska-Lejman *et al.*, 2023).

581 We observed that COmapper can mis-map some crossovers occurring within SNP-
582 poor genomic regions, because accurate crossover mapping depends on both SNP
583 density and read length—each long read must cover a sufficient number of SNPs (Figs.
584 **S14**). In genomes with low SNP density, alternative methods like GBS, linked-read

585 sequencing or Hi-C may offer more reliable crossover detection. Like COmapper, Hi-
586 C and linked-read approaches can be applied to pooled samples, but typically provide
587 lower crossover resolution. GBS requires at least 3x coverage to achieve crossover
588 detection resolution comparable to COmapper. In pooled samples, the number of
589 crossovers detected by COmapper increased proportionally with nanopore
590 sequencing depth (Table **S3**), indicating that deeper coverage enables more
591 comprehensive reconstruction of genome-wide crossover landscapes (Figs **4, 5**; Table
592 **S1**). However, COmapper does not retain individual-level information, such as the
593 number and position of crossovers per sample, which is essential for quantitative trait
594 locus mapping. Additionally, it cannot be used to assess crossover interference, which
595 requires measuring the physical distances between crossovers along the same
596 chromosome, as can be done with GBS in individually genotyped recombinants (Table
597 **S3**).

598 In contrast to GBS, COmapper requires a simple gDNA extraction method and pooled
599 gDNA from F_1 hybrid pollen or F_2 seedlings for nanopore sequencing, instead of gDNA
600 extraction from individual plants and barcoding. Therefore, we propose that
601 COmapper can be widely used to map genome-wide crossovers at high resolution in
602 the following types of studies. First, COmapper should be applicable for performing a
603 rapid and high-throughput screen to detect changes in the number and distribution of
604 meiotic crossovers in crosses between different accessions in *Arabidopsis* and various
605 species. Therefore, COmapper should allow testing for the effects of genetic,
606 epigenetic and environmental changes on crossover patterning. Second, COmapper
607 should be able to detect crossovers in diverse diploid and polyploid organisms.
608 However, its successful application in complex genomes depends on the SNP density
609 in crossover regions. Therefore, assessing the density and distribution of SNPs across
610 parental genomes is crucial before applying COmapper. For example, we analyzed
611 SNP distribution between two tomato cultivars (*Solanum lycopersicum* and *S.*
612 *pimpinellifolium*), demonstrating the scalability of COmapper (Fig. **S19**). Ongoing
613 improvements in nanopore sequencing technology such as longer read lengths and
614 lower sequencing error rates will further enhance its utility for large and complex
615 genomes. For polyploids, COmapper can be applied to hybrid autotetraploid
616 *Arabidopsis* plants derived from natural accessions (e.g., Wa-1 × Bla-5) or colchicine-

617 doubled tetraploid (e.g., Col (4x) × L*er* (4x)). Likewise, COmapper could facilitate
618 crossover analysis in allopolyploid plants such as wheat and rapeseed. Third,
619 COmapper should be useful to map genome-wide crossovers in pollen or progeny of
620 outcrossing species, along with calling parental SNPs by nanopore sequencing.
621 Finally, COmapper should allow the measurement of crossover frequency and
622 distribution in meiotic recombination hotspots. Instead of PCR amplification of single
623 ~5–10-kb molecules from a recombination hotspot (Choi *et al.,* 2013), genotyping of
624 pollen or sperm can be performed in pooled samples and combined with nanopore
625 long-read sequencing. Therefore, given the simplicity of sample preparation and the
626 high accuracy in mapping crossovers by long-read nanopore sequencing, we believe
627 that COmapper will be extensively used to understand the mechanisms of meiotic
628 crossover patterning in diverse species beyond plants.

629

630 **Acknowledgements**

631 We thank Raphael Mercier (MPI for Plant Breeding Research) for providing *recq4ab*
632 seeds. This work was funded by a Samsung Science and Technology Foundation
633 grant (SSTF-BA2202-09) and the National Research Foundation of Korea (RS-2024-
634 00335818, RS-2024-00407469).

635

636 **Competing interests**

637 None declared.

638 **Author contributions**

639 D.B., N.S. and K.C. designed the project. D.B., N.S., and J.K. developed COmapper
640 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.
641 performed experiments and analysis. D.B., N.S. and K.C. wrote the original draft of
642 manuscript. All authors revised the manuscript.

643

644 **Data availability**

645 The nanopore sequencing data generated in this study have been deposited in the
646 ArrayExpress database at EMBL-EBL (http://www.ebi.ac.uk/arrayexpress). Detailed
647 information on the nanopore long-read sequencing and other data for GBS and
648 genomic features used in this study is included in **Supplementary Table S2**. Source
649 data are provided with this paper. All custom code used in this study, including the
650 COmapper pipeline, is available at GitHub (https://github.com/KyuhaChoi-
651 Lab/COmapper).

652

653 ## References

654 **Berchowitz LE, Francis KE, Bey AL, Copenhaver GP**. **2007**. The Role of AtMUS81 in Interference-
655 Insensitive Crossovers in A. thaliana. *PLoS Genetics* **3**: e132.

656 **Blackwell AR, Dluzewska J, Szymanska-Lejman M, Desjardins S, Tock AJ, Kbiri N, Lambing C,**
657 **Lawrence EJ, Bieluszewski T, Rowan B, *et al.* 2020**. MSH 2 shapes the meiotic crossover landscape in
658 relation to interhomolog polymorphism in Arabidopsis . *The EMBO Journal* **39**: 1–22.

659 **Borts RH, Haber JE**. **1987**. Meiotic recombination in yeast: alteration by multiple heterozygosities.
660 *Science (New York, N.Y.)* **237**: 1459–65.

661 **Castellani M, Zhang M, Thangavel G, Mata-Sucre Y, Lux T, Campoy JA, Marek M, Huettel B, Sun H,**
662 **Mayer KFX, *et al.* 2024**. Meiotic recombination dynamics in plants with repeat-based
663 holocentromeres shed light on the primary drivers of crossover patterning. *Nature Plants* **10**: 423–
664 438.

665 **Choi K, Henderson IR**. **2015**. Meiotic recombination hotspots - A comparative view. *Plant Journal* **83**:
666 52–61.

667 **Choi K, Reinhard C, Serra HH, Ziolkowski PAPAP, Underwood CJCJ, Zhao X, Hardcastle TJTJT, Yelina**
668 **NENNE, Griffin C, Jackson M, *et al.* 2016**. Recombination Rate Heterogeneity within Arabidopsis
669 Disease Resistance Genes (R Michelmore, Ed.). *PLoS Genetics* **12**: e1006179.

670 **Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA, Copenhaver**
671 **GP, Franklin FCH, *et al.* 2013**. Arabidopsis meiotic crossover hot spots overlap with H2A.Z
672 nucleosomes at gene promoters. *Nature Genetics* **45**: 1327–1336.

673 **Choi K, Zhao X, Tock AJAJ, Lambing C, Underwood CJCJ, Hardcastle TJTJ, Serra H, Kim JJ, Cho HSHS,**
674 **Kim JJ, *et al.* 2018**. Nucleosomes and DNA methylation shape meiotic DSB frequency in Arabidopsis
675 thaliana transposons and gene regulatory regions. *Genome Research* **28**: 532–546.

676 **Crismani W, Girard C, Froger N, Pradillo M, Santos JL, Chelysheva L, Copenhaver GP, Horlow C,**
677 **Mercier R**. **2012**. FANCM Limits Meiotic Crossovers. *Science (New York, N.Y.)* **336**: 1588–1590.

Dluzewska J, Dziegielewski 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**.

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**.

Fernandes JB, Naish M, Lian Q, Burns R, Tock AJ, Rabanal FA, Wlodzimierz 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 of the United States of America* **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 ef®cient meiotic recombination in plants. *EMBO Journal* **20**: 589–600.

He Y, Wang M, Dukowic-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 of the United States of America* **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 of the United States of America* **110**: 19478–82.

Henderson IR, Bomblies K. **2021**. Evolution and Plasticity of Genome-Wide Meiotic Recombination Rates. : 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.

711 **Kianian PMA, Wang M, Simons K, Ghavami F, He Y, Dukowic-Schulze S, Sundararajan A, Sun Q,**
712 **Pillardy J, Mudge J, *et al.* 2018**. High-resolution crossover mapping reveals similarities and
713 differences of male and female recombination in maize. *Nature Communications* **9**: 2370.

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

716 **Kim H, Kim J, Son N, Kuo P, Morgan C, Chambon A, Byun D, Park J, Lee Y, Park YM, *et al.* 2024**.
717 Control of meiotic crossover interference by a proteolytic chaperone network. *Nature Plants* **10**:
718 453–468.

719 **Kim J, Park J, Kim H, Son N, Kim E, Kim JJ, Byun D, Lee Y, Park YM, Nageswaran DC, *et al.* 2022**.
720 Arabidopsis HEAT SHOCK FACTOR BINDING PROTEIN is required to limit meiotic crossovers and
721 HEI10 transcription. *The EMBO Journal* **41**: e109958.

722 **Lam I, Keeney S**. **2014**. Mechanism and regulation of meiotic recombination initiation. *Cold Spring*
723 *Harbor perspectives in biology* **7**: a016634.

724 **Lawrence EJ, Gao H, Tock AJ, Lambing C, Blackwell AR, Feng X, Henderson IR**. **2019**. Natural
725 Variation in TBP-ASSOCIATED FACTOR 4b Controls Meiotic Crossover and Germline Transcription in
726 Arabidopsis. *Current Biology* **29**: 2676-2686.e3.

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

728 **Mercier R, Mézard C, Jenczewski E, Macaisne N, Grelon M**. **2015**. The Molecular Biology of Meiosis
729 in Plants. *Annual Review of Plant Biology* **66**: 297–327.

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

732 **Morgan C, Fozard JA, Hartley M, Henderson IR, Bomblies K, Howard M**. **2021**. Diffusion-mediated
733 HEI10 coarsening can explain meiotic crossover positioning in Arabidopsis. *Nature Communications*
734 **12**: 4674.

735 **Nageswaran DC, Kim J, Lambing C, Kim J, Park J, Kim E-J, Cho HS, Kim H, Byun D, Park YM, *et al.***
736 **2021**. HIGH CROSSOVER RATE1 encodes PROTEIN PHOSPHATASE X1 and restricts meiotic crossovers
737 in Arabidopsis. *Nature Plants* **7**: 452–467.

738 **Naish M, Alonge M, Wlodzimierz P, Tock AJ, Abramson BW, Schmücker A, Mandáková T, Jamge B,**
739 **Lambing C, Kuo P, *et al.* 2021**. The genetic and epigenetic landscape of the Arabidopsis centromeres.
740 *Science (New York, N.Y.)* **374**: eabi7489.

741 **Neale MJ, Pan J, Keeney S**. **2005**. Endonucleolytic processing of covalent protein-linked DNA double-
742 strand breaks. *Nature* **436**: 1053–1057.

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

Rommel Fuentes R, Hesselink T, Nieuwenhuis R, Bakker L, Schijlen E, van Dooijeweert 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. *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*: 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: TOP3α and two BLM homologs antagonize crossovers in parallel to FANCM. *Proceedings of the National Academy of Sciences of the United States of America* **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 of the United States of America* **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. *The Plant cell* **27**: 2427–36.

Sun H, Rowan BA, Flood PJ, Brandt R, Fuss J, Hancock AM, Michelmore RW, Huettel B, Schneeberger K. **2019**. Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nature Communications* **10**: 4310.

Szymanska-Lejman M, Dziegielewski W, Dluzewska J, Kbiri 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.

783 **Underwood CJ, Choi K**. **2019**. Heterogeneous transposable elements as silencers, enhancers and
784 targets of meiotic recombination. *Chromosoma* **128**.

785 **Underwood CJ, Choi K, Lambing C, Zhao X, Serra H, Borges F, Simorowski J, Ernst E, Jacob Y,**
786 **Henderson IRIRIR, *et al.* 2018**. Epigenetic activation of meiotic recombination near Arabidopsis
787 thaliana centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Research* **28**: 519–
788 531.

789 **Underwood CJ, Henderson IR, Martienssen RA**. **2017**. Genetic and epigenetic variation of
790 transposable elements in Arabidopsis. *Current Opinion in Plant Biology* **36**: 135–141.

791 **Villeneuve AM, Hillers KJ**. **2001**. Whence meiosis? *Cell* **106**: 647–50.

792 **Wang Y, Copenhaver GP**. **2018**. Meiotic Recombination: Mixing It Up in Plants. *Annual Review of*
793 *Plant Biology* **69**: 577–609.

794 **Wang Y, Zhao Y, Bollas A, Wang Y, Au KF**. **2021**. Nanopore sequencing technology, bioinformatics
795 and applications. *Nature Biotechnology* **39**: 1348–1365.

796 **Wijnker E, Velikkakam James G, Ding J, Becker F, Klasen JR, Rawat V, Rowan BA, de Jong DF, de**
797 **Snoo CB, Zapata L, *et al.* 2013**. The genomic landscape of meiotic crossovers and gene conversions
798 in Arabidopsis thaliana. *eLife* **2**: e01426.

799 **Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR**. **2015**. DNA methylation
800 epigenetically silences crossover hot spots and controls chromosomal domains of meiotic
801 recombination in Arabidopsis. *Genes & Development* **29**: 2183–2202.

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

804 **Zickler D, Kleckner N**. **2023**. Meiosis: Dances Between Homologs. *Annual Review of Genetics* **57**: 1–
805 63.

806 **Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, Choi K, Ziolkowska L, June V,**
807 **Sanchez-Moran E, *et al.* 2015**. Juxtaposition of Heterozygous and Homozygous Regions Causes
808 Reciprocal Crossover Remodelling via Interference During Arabidopsis Meiosis. *eLife* **4**: e03708.

809

810

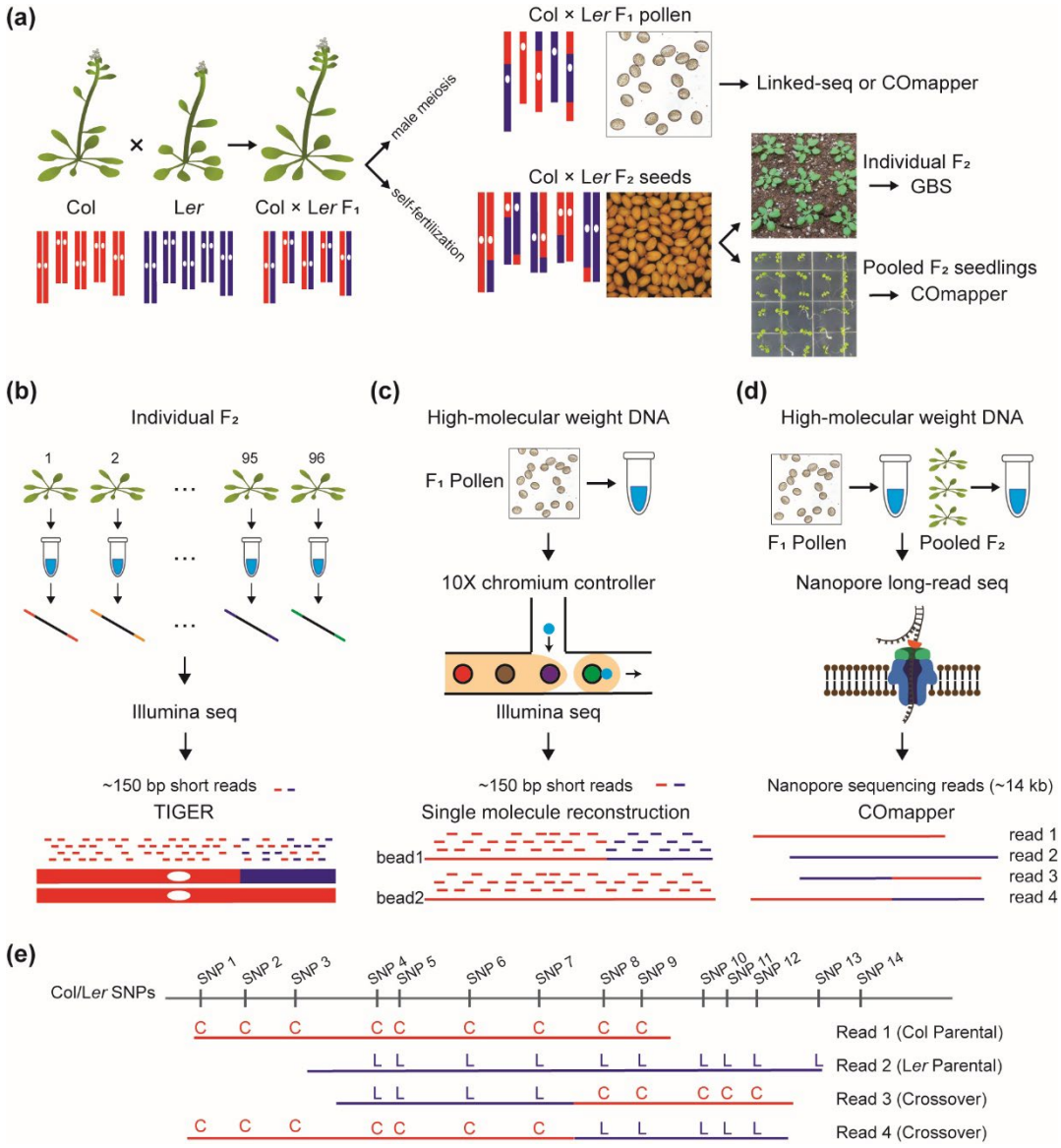


Fig. 1 Genome-wide mapping of crossovers using long-read nanopore sequencing. (**a**) iagrams illustrating the preparation of recombinant pollen from F_1 hybrids and F_2 plants to map crossovers. (**b**) Mapping of genome-wide crossovers in F_2 plants by genotyping-by-sequencing (GBS) using multiplex 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.

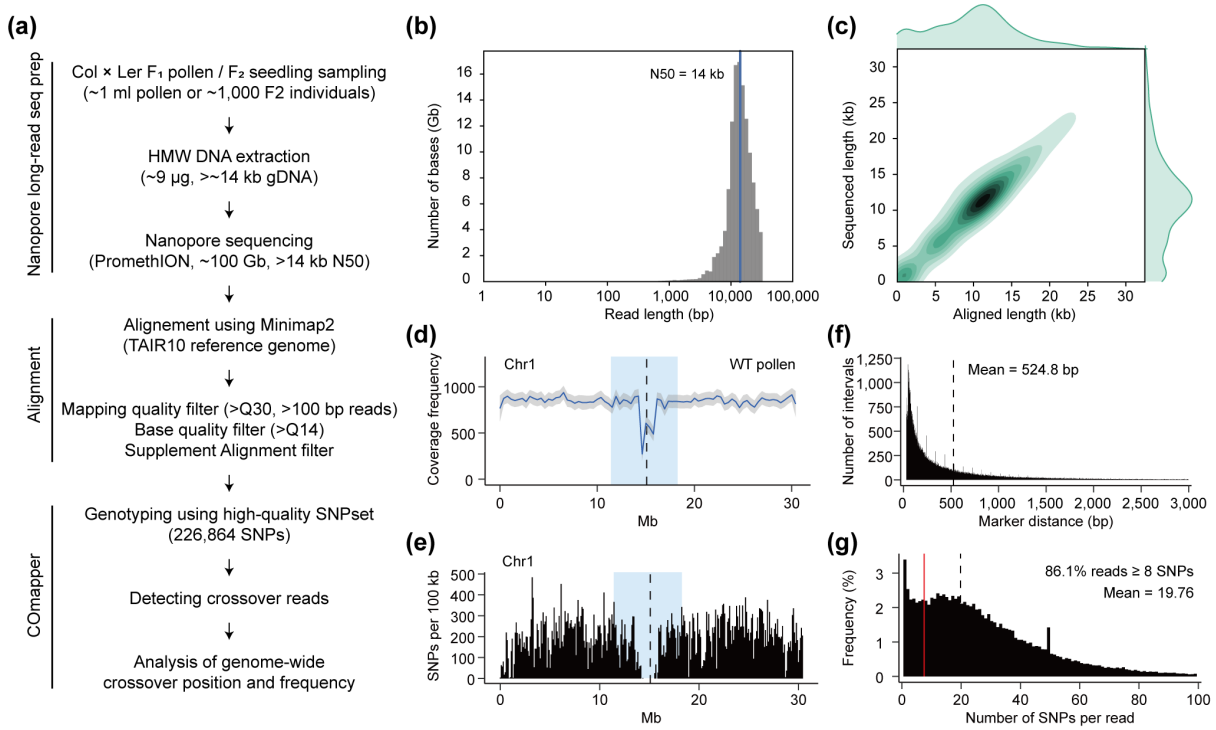
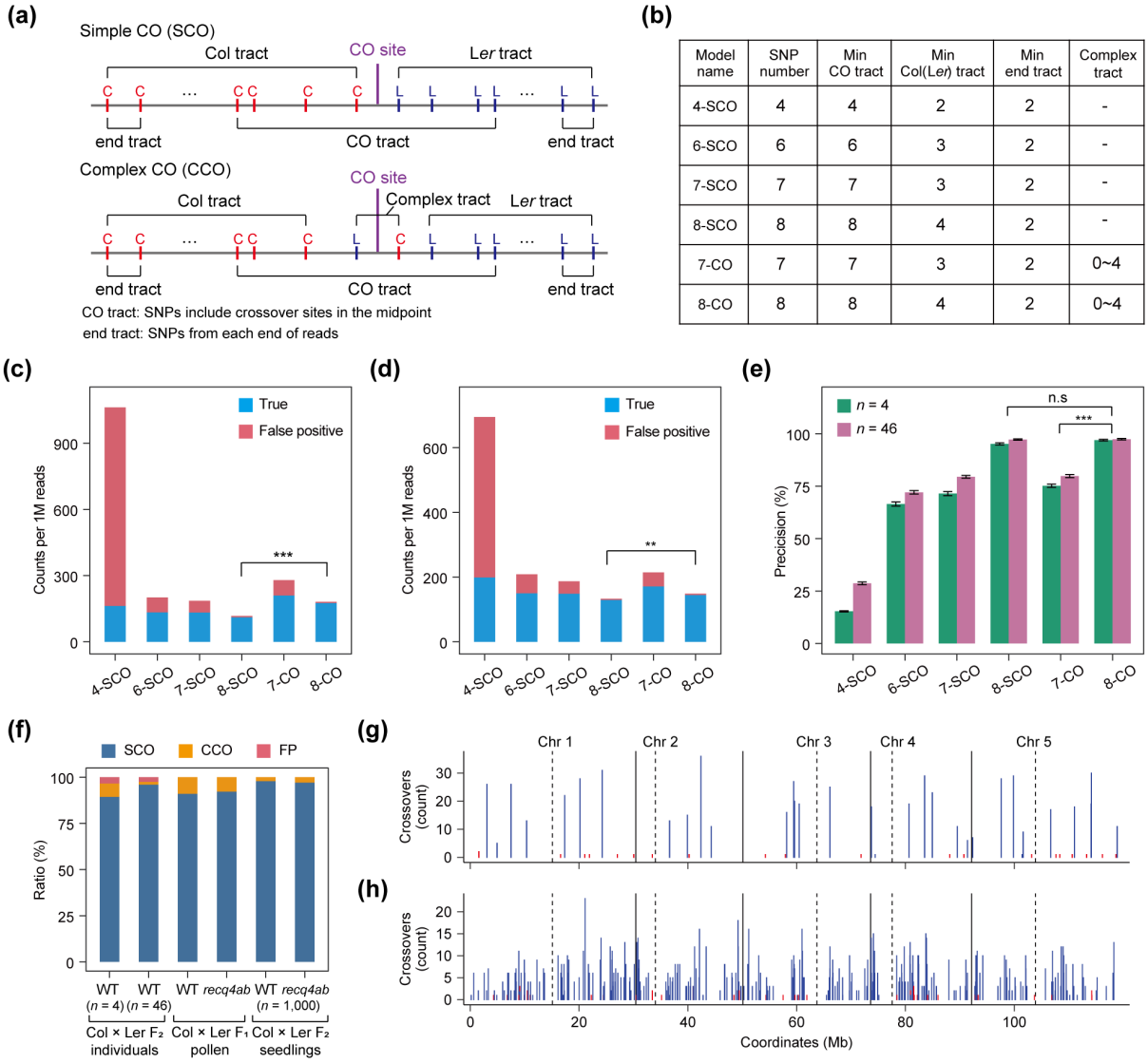


Fig. 2 Development of COmapper to map crossovers using long-read sequencing. (**a**) Procedure for crossover mapping using nanopore long-read sequencing and the COmapper pipeline. (**b**) Distribution of read length and number of pollen DNAs sequenced by nanopore long-read sequencing. (**c**) Density 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.



Model name	SNP number	Min CO tract	Min Col(Ler) tract	Min end tract	Complex 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

Fig. 3

841
842 **Fig. 3 Validation of COmapper using Col × L*er* F_2 recombinant plants.** (**a**) Diagram
843 showing simple crossover (SCO) and complex (CCO) types. (**b**) Summary table of the
844 tested COmapper models with different numbers of SNPs and minimum crossover
845 (Min CO) and complex tracts. (**c**) Number of crossovers detected per million reads in
846 different COmapper models in 4 Col × L*er* F_2 plants (Kim *et al.,* 2024). True crossovers
847 (blue) were detected by COmapper models among the crossovers mapped by GBS in
848 4 F_2 plants, and false crossovers (red) indicate the crossovers that differ from the
849 crossovers defined by GBS. (**d**) as in (**c**), but showing number of crossovers detected
850 per million reads in 46 Col × L*er* F_2 plants. (**e**) Precision (True positive/(True positive
851 + False positive)) of crossover detection in the different COmapper models shown in
852 (C). In (C) and (D), asterisks indicate significant differences between models (*P <
853 0.05, **P < 0.01, ***P < 0.001; Welch′s *t*-test). (**f**) Relative proportion of crossover
854 types in 4 F_2 plants (genotyped by GBS), 46 F_2 plants (genotyped by GBS), Col-0 ×
855 L*er* F_1 hybrid pollen DNA, *recq4ab* Col × *recq4ab* L*er* hybrid pollen DNA, 1,000 Col-0
856 × L*er* F_2 seedlings and 1,000 *recq4ab* Col × *recq4ab* L*er* F_2 seedlings detected by the
857 established model (8-CO) of COmapper. FP (red) indicates the proportion of false
858 positive crossovers identified by COmapper in 4 and 46 F_2 plants compared to the
859 crossovers identified by GBS. (**g**) Frequency of crossovers mapped by COmapper
860 along the five *Arabidopsis* chromosomes on a continuous *x*-axis in 4 Col × L*er* F_2
861 recombinants, where crossovers in each F_2 plant were mapped by GBS. Blue bars
862 represent true positive crossovers mapped by GBS; red bars represent false positive
863 crossovers. (**h**) As in (**g**), but from 46 Col × L*er* F_2 recombinants.
864

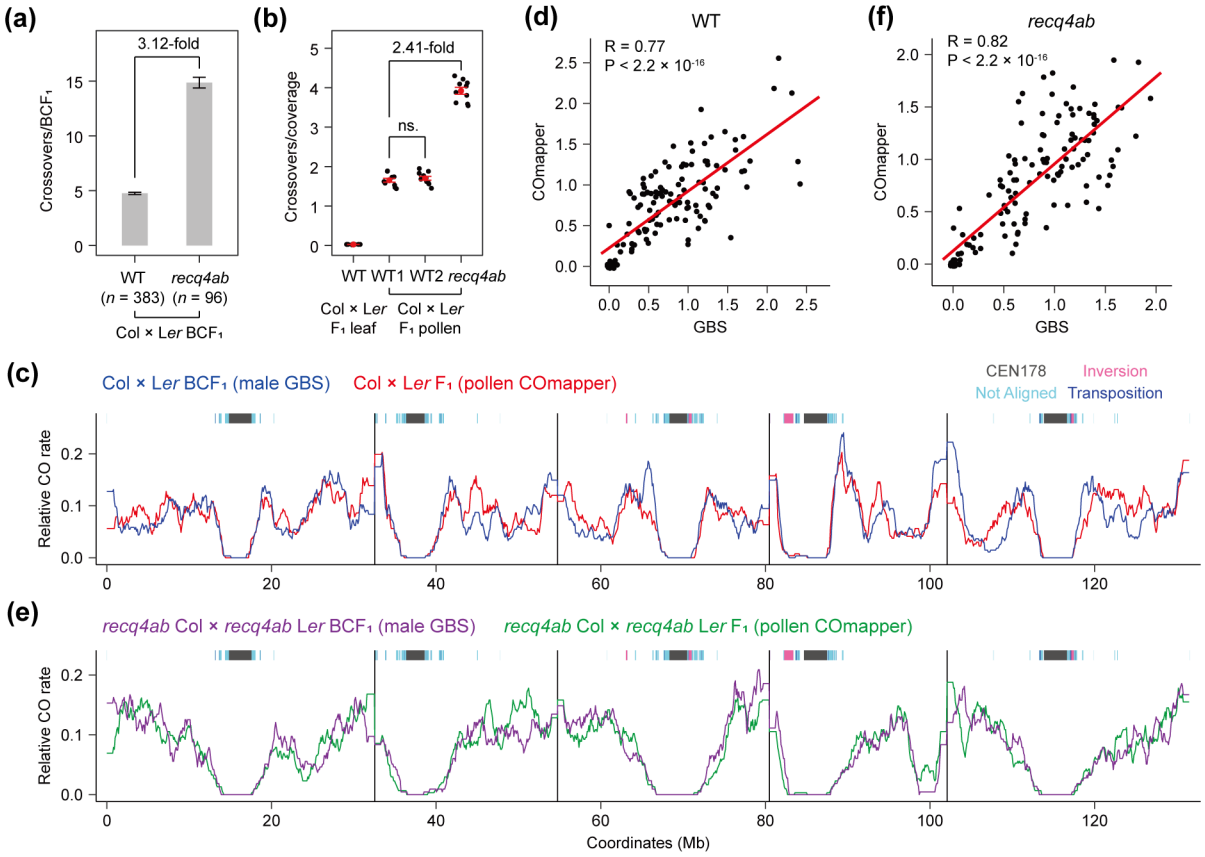

(a)
3.12-fold
Crossovers/BCF₁
WT
recq4ab
(n = 383) (n = 96)
Col × Ler BCF₁
(b)
2.41-fold
ns.
Crossovers/coverage
WT WT1 WT2 recq4ab
Col × Ler F₁ leaf
Col × Ler F₁ pollen
(d)
WT
R = 0.77
P < 2.2 × 10⁻¹⁶
COmapper
GBS
(f)
recq4ab
R = 0.82
P < 2.2 × 10⁻¹⁶
COmapper
GBS
(c)
Col × Ler BCF₁ (male GBS)
Col × Ler F₁ (pollen COmapper)
CEN178
Inversion
Not Aligned
Transposition
Relative CO rate
(e)
recq4ab Col × recq4ab Ler BCF₁ (male GBS)
recq4ab Col × recq4ab Ler F₁ (pollen COmapper)
Relative CO rate
Coordinates (Mb)

Fig. 4

868 **Fig. 4 Mapping of crossovers from pooled F_1 hybrid pollen DNA by COmapper.**
869 (**a**) Number of male crossovers per BCF_1 individual calculated from GBS data of WT
870 Col × L*er* BCF_1 (*n* = 383) (Kim *et al.,* 2024) and *recq4ab* Col × *recq4ab* L*er* BCF_1 (*n* =
871 96) plants. *n* indicates the number of backcrossed F_1 plants. Data are presented as
872 means ± SD. (**b**) Number of crossovers per genomic coverage of nanopore
873 sequencing data of gDNA from WT Col × L*er* F_1 leaves, WT Col × L*er* F_1 pollen (WT
874 set1, WT set2) and *recq4ab* Col × *recq4ab* L*er* F_1 pollen. Each data point indicates
875 crossovers per genome from a randomly selected million reads. Data are presented
876 as means ± SD from 10 replicates. (**c**) Relative frequency of crossovers mapped by
877 Col × L*er* male GBS (blue, *n* = 1,824 crossovers) and Col × L*er* F_1 pollen COmapper
878 (red, *n* = 1,679 crossovers) along the five *Arabidopsis* chromosomes on a continuous
879 *x*-axis in Col × L*er* male meiosis using Col-CEN complete genome. Fisher′s exact test
880 was applied to each 100-kb bin in every 1 Mb window to assess differences in relative
881 crossover rates. (**d**) Correlation between crossover frequencies in 1-Mb bins of Col ×
882 L*er* BCF_1 individuals (measured by GBS) and Col × L*er* F_1 pollen (measured by
883 COmapper) across the genome, assessed using Pearson′s correlation test. (**e**) As in
884 (**c**), but showing the relative crossover frequency of crossovers mapped by *recq4ab*
885 Col × *recq4ab* L*er* male GBS (blue, *n* = 1,428 crossovers) and *recq4ab* Col × *recq4ab*
886 L*er* F_1 pollen COmapper (red, 1,963 crossovers). (**f**) As in (**d**), but showing the
887 correlation between crossover frequencies of crossovers mapped by *recq4ab* Col ×
888 *recq4ab* L*er* BCF_1 individuals and *recq4ab* Col × *recq4ab* L*er* F_1 pollen.

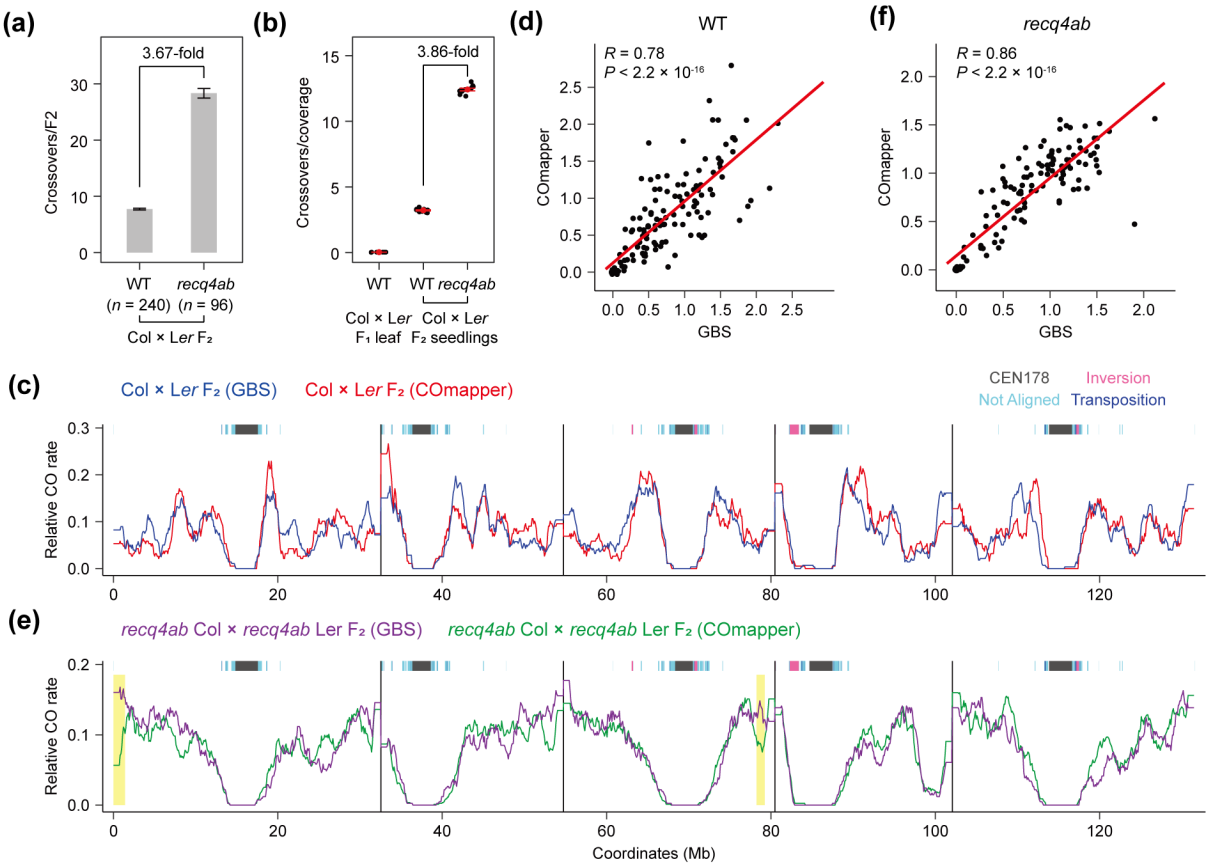

(a)
3.67-fold
Crossovers/F2
WT
recq4ab
(n = 240) (n = 96)
Col × Ler F2
(b)
3.86-fold
Crossovers/coverage
WT
WT recq4ab
Col × Ler F1 leaf
Col × Ler F2 seedlings
(d)
WT
R = 0.78
P < 2.2 × 10-16
COmapper
GBS
(f)
recq4ab
R = 0.86
P < 2.2 × 10-16
COmapper
GBS
(c)
Col × Ler F2 (GBS)
Col × Ler F2 (COmapper)
CEN178
Inversion
Not Aligned
Transposition
Relative CO rate
(e)
recq4ab Col × recq4ab Ler F2 (GBS)
recq4ab Col × recq4ab Ler F2 (COmapper)
Relative CO rate
Coordinates (Mb)

Fig. 5

891 **Fig. 5 Mapping of crossovers from pooled F_2 recombinant seedlings by**
892 **COmapper.** (**a**) Number of crossovers per F_2 individual calculated from GBS data of
893 WT Col × L*er* F_2 (*n* = 240) (Kim *et al.,* 2024) and *recq4ab* Col × *recq4ab* L*er* F_2 (*n* = 96)
894 plants. *n* indicates the number of F_2 individual plants. Data are presented as means ±
895 SD. (**b**) Number of crossovers per genome of nanopore sequencing data from gDNA
896 of Col × L*er* F_1 leaves, Col × L*er* F_2 and *recq4ab* Col × *recq4ab* L*er* F_2 seedlings (*n* =
897 1,000). Each data point indicates the number of crossovers per genome from a
898 randomly selected 1 million reads. Data are presented as means ± SD from 10
899 replicates. (**c**) Relative frequency of crossovers mapped by GBS (blue, *n* = 1,852
900 crossovers) and COmapper (red, *n* = 1,390 crossovers) along the five *Arabidopsis*
901 chromosomes on a continuous *x*-axis in WT Col × L*er* F_2 plants (Col-CEN). Fisher′s
902 exact test was applied to each 100-kb bin to assess differences in relative crossover
903 rates. Bins with false discovery rate (FDR) below 0.05 are highlighted in yellow. (**d**)
904 Correlation between crossover frequencies in 1-Mb bins of Col × L*er* F_2 individuals
905 (measured by GBS) and Col × L*er* F_2 seedlings (measured by COmapper) across the
906 genome, assessed using Pearson′s correlation test. (**e**) As in (**c**), but showing the
907 relative frequency of crossovers mapped by GBS in *recq4ab* Col × *recq4ab* L*er* F_2
908 plants (purple, *n* = 2,731 crossovers) or by COmapper in *recq4ab* Col × *recq4ab* L*er*
909 F_2 seedlings (green, *n* = 5,760 crossovers). (**f**) As in (**d**), but showing the correlation
910 between crossover frequencies of crossovers mapped by *recq4ab* Col × *recq4ab* L*er*
911 F_2 individuals and *recq4ab* Col × *recq4ab* L*er* F_2 seedlings.

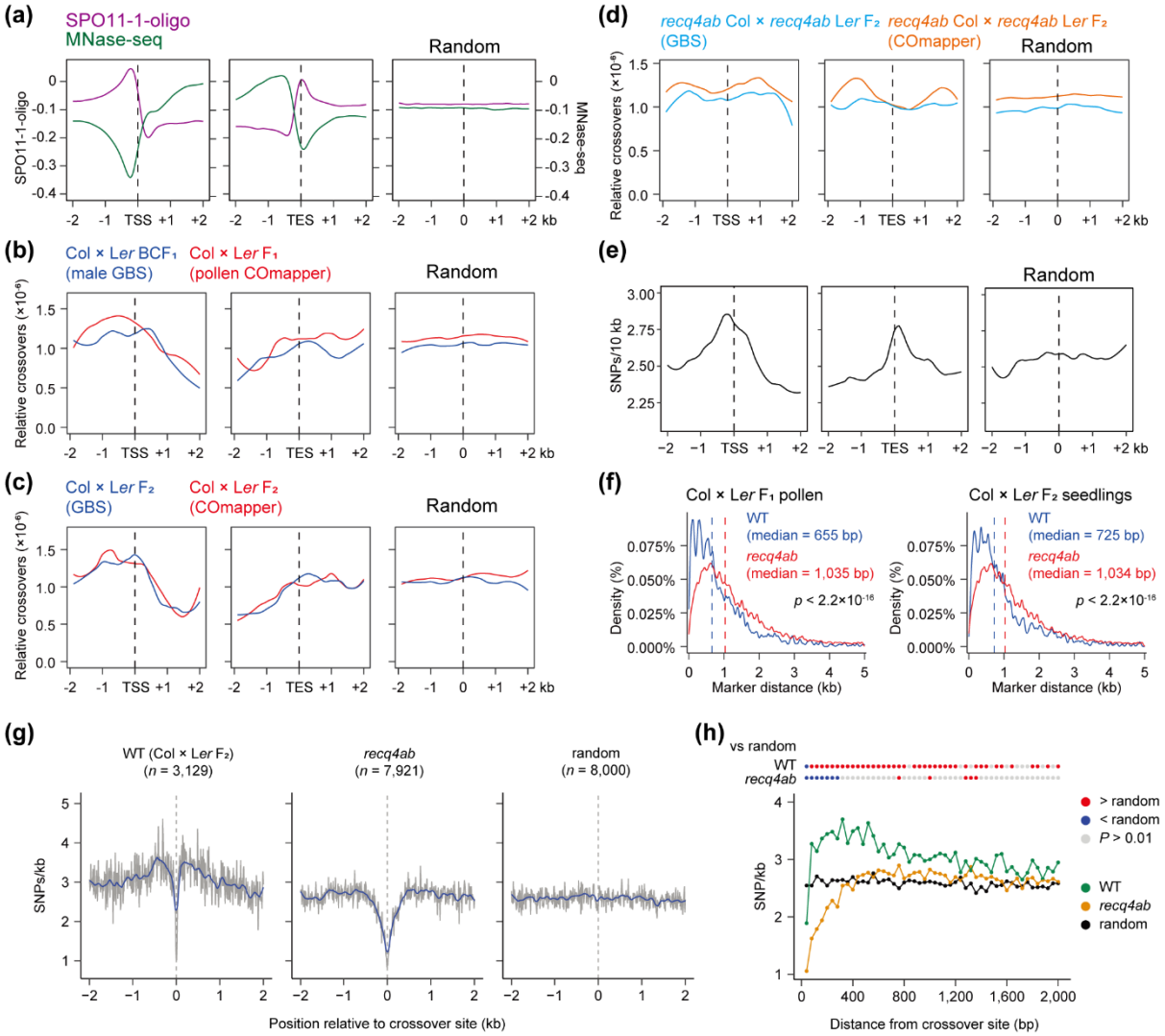

(a)
SPO11-1-oligo
MNase-seq
Random
SPO11-1-oligo
MNase-seq
-2 -1 TSS +1 +2
-2 -1 TES +1 +2
-2 -1 0 +1 +2 kb
(b)
Col × Ler BCF₁ (male GBS)
Col × Ler F₁ (pollen COmapper)
Random
Relative crossovers (×10⁻⁶)
(c)
Col × Ler F₂ (GBS)
Col × Ler F₂ (COmapper)
Random
Relative crossovers (×10⁻⁶)
(d)
recq4ab Col × recq4ab Ler F₂ (GBS)
recq4ab Col × recq4ab Ler F₂ (COmapper)
Random
Relative crossovers (×10⁻⁶)
(e)
Random
SNPs/10 kb
(f)
Col × Ler F₁ pollen
Col × Ler F₂ seedlings
WT (median = 655 bp)
recq4ab (median = 1,035 bp)
p < 2.2×10⁻¹⁶
WT (median = 725 bp)
recq4ab (median = 1,034 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)

Fig. 6

915 **Fig. 6 Analysis of crossovers and SNPs around genes.** (**a**) Metaplots showing the
916 density of SPO11-1–oligonucleotides (SPO11-1–oligo) and nucleosomes from MNase
917 sequencing (MNase-seq) around gene transcription start sites (TSS) and transcription
918 end sites (TES) in a 2-kb window (Choi *et al.,* 2018). (**b**) As in (**a**), but showing the
919 relative rate of crossovers mapped by GBS (blue, *n* = 1,824 crossovers) described
920 (Rowan *et al.,* 2019; Kim *et al.,* 2024) and COmapper (red, *n* = 1,679 crossovers) in
921 *Arabidopsis* Col × L*er* male meiosis. (**c**) As in (**b**), but showing the relative rate of
922 crossovers mapped by GBS (blue, 1,852 crossovers) and COmapper (red, *n* = 1,391
923 crossovers) in *Arabidopsis* Col × L*er* F_2 plants. (**d**) As in (**c**), but showing the relative
924 rate of crossovers mapped by GBS (skyblue, *n* = 2,731 crossovers) and COmapper
925 (orange, *n* = 5,760 crossovers) in *Arabidopsis recq4ab* Col × *recq4ab* L*er* F_2 plants.
926 (**e**) As for (**a**), but showing the density of SNPs between Col and L*er* accessions. (**f**)
927 Density and physical distance between two consecutive SNPs at crossover sites in
928 WT Col × L*er* and *recq4ab* Col × *recq4ab* L*er* for F_1 pollen and F_2 progeny. The vertical
929 dashed line indicates median distance (bp) between two consecutive SNPs at the
930 crossover site. Significance between genotypes was examined using a Wilcoxon test.
931 (**g**) Metaplots showing SNP density around crossovers mapped by GBS and
932 COmapper in WT Col × L*er* F_2 and *recq4ab* Col × *recq4ab* L*er* F_2 progeny compared
933 to random positions. (**h**) As in (**g**), but showing the significance of changes in SNP
934 density according to distance from the crossover site. Each dot represents the SNP
935 density per 40-bp bin. Significance between genotype and a random position per
936 interval was examined using a Wilcoxon test. Red and blue dots indicate higher and
937 lower SNP density than random, respectively; gray dots indicate non-significant
938 differences.