# **Supplemental Figure Legends**

**Figure S1.** **Workflow diagram for genomic DNA library preparation using Nextera LITE.** Gel verification step is optional, but recommended. Workflow as presented requires about 2 - 3 hours of time. Multiple plates can be processed at the same time. Before sequencing, determine which plates have complementary sets of indices (see File S1) and can therefore be pooled into a single lane. Mix each individual 96-sample pool at an equimolar ratio in the final tube containing all libraries to be sequenced in a single lane. Image credits: 96-well plate: <https://doi.org/10.6084/m9.figshare.928645.v1>; Tube: <https://dlpng.com/png/417627>

**Figure S2. Read coverage distributions among samples.** A) Coverage plots for each 96-sample pool for all of the genomic DNA libraries for the 1,920 Col X Ler F2 individuals produced for this study. B) Coverage plots for publicly available read data for additional F2 individuals derived from the same parental backgrounds categorized by their accession numbers at ArrayExpress. Data points are displayed in 0.2x coverage bins. Cross bar indicates the mean.

**Figure S3. CO resolution.** The interval window for CO breakpoint estimation (calculated by the distance between flanking markers determined by TIGER). Shown are the intervals for 16,709 of 17,077 total COs, with intervals > 20 kb omitted for ease of visualization.

**Figure S4. CO per pair relative to chromosome length.** The equation of the line is 5.14\*10-8*x* + 0.355 (R2 = 0.96). The gray shaded area represents the 95% confidence interval.

**Figure S5. Positions of *cis* double crossovers along chromosomes.** A) The positions of the first and second COs for pairs of double COs are plotted along the chromosomes. The chromosome number is indicated in the gray boxes to the right of the plots. B) Gamma distribution fits (red lines) to the observed inter-crossover distances (top row) compared with a random sampling of double crossovers (bottom row) for each chromosome.

**Figure S6. CO frequencies adjacent to a 170-kb inversion on chromosome 3.** CO frequencies are plotted in 50-kb bins around the inversion (light gray shaded region).

**Figure S7. Distance to nearest CO position from the borders of SVs relative to the size of the variant region.** Linear regressions revealed no significant correlation between variant size and the distance to the nearest CO position for any of the variants shown in the plots.

**Figure S8. CO rates in the 50-, 100-, and 200-kb windows up- and downstream of structural variants.** Boxplots show the distributions of CO rates in windows of the indicated sizes that are surrounding inversions, insertions, deletions, transpositions (intrachrom), translocations (interchrom), copy number variations (CNV) and all windows genome-wide (All). Boxplots display the median and the first and third quantiles.

**Figure S9. Mean CO rates in the 200-kb flanking SVs.** Linear regressions revealed a significant correlation between SV size and the mean flanking CO rates only for copy number SVs (p = 0.05 with an R2 value of 0.05).

**Figure S10. Relationships among structural variants (SVs), “CO deserts”, and DNA methylation at CG positions.** A) Overlaps between CO-depleted regions “CO deserts” and SVs. Vertical black line indicates mean number of overlaps expected in 5000 random permutations. Vertical red line indicates the number of overlaps where p = 0.05. Vertical green line indicates the observed number of overlaps. Double-headed arrow highlights the difference between the mean of 5000 permutations and the observed number. B) Box plots showing the percentage of DNA methylation at CG positions in 10-kb windows that overlap with “CO deserts” (Desert) and those that do not (Outside) either with centromeric regions included or excluded. C) Box plots showing the percentage of DNA methylation at CG positions in 10-kb windows that overlap with structural variants (SV) and those that do not (Outside) either with centromeric regions included or excluded. Boxplots display the median and the first and third quantiles.

**Figure S11. Permutation tests of the overlaps between COs and NLR genes.** Permutation tests for all NLR loci (A) and subsets where the locus structure is the same between Col and Ler (B) and different between Col and Ler (C). For all plots, the vertical black line indicates mean number of overlaps expected in 5,000 random permutations. Vertical red line indicates the number of overlaps where p = 0.05. Vertical green line indicates the observed number of overlaps. Double arrow highlights the difference between the mean from 5,000 permutations and the observed number.

**Figure S12. COs in NLR genes relative to the number of copies in the locus.** Boxplots with individual data points (light blue dots) display the CO rates within each locus type for loci where Col and Ler have the same number of copies in the locus (A) and loci where Col and Ler have different copy numbers (B). Boxplots display the median and the first and third quantiles.

**Figure S13. Locations of COs in six focal NLR loci.** These loci were also studied in [Choi *et al*. (2016)](https://paperpile.com/c/3UiuMY/zfIMz).

**Figure S14. COs relative to ATAC-seq sites.** A) Permutation test of overlaps between COs and ATAC-seq sites. The vertical black line indicates mean number of overlaps expected in 1,000 random permutations. Vertical red line indicates the number of overlaps where p = 0.05. Vertical green line indicates the observed number of overlaps. Double-headed arrow highlights the difference between the mean of 1,000 permutations and the observed number. B) Distribution of distances from CO breakpoints to the nearest ATAC-seq site border. Vertical dashed green line indicates the median. The top 5% of distances are omitted from this plot for ease of visualization.

**Figure S15. Crossovers (COs) associated with sequence features.** A) CO rates (black) in comparison with Col/Ler SNP counts (gray shading) in 100-kb windows along the chromosomes and disease resistance genes (red ticks). CO rates are tallied as the number of COs per window divided by the total number of individuals. Horizontal dashed grey line indicates the genome-wide mean CO rate. B) Percentage of AT (red) and GC (blue) content across the genome. The green shading shows the CO rates presented in (A) and the same axis scale applies. Black ticks show disease resistance genes. The horizontal dashed red line indicates the genome-wide mean %AT content and the horizontal dashed blue line indicates the genome-wide mean %GC content. The position information shown in (B) also applies to (A). Solid vertical lines indicate chromosome boundaries and dashed vertical lines represent the mid-points of the centromeres in both A and B.

**File S1:** **Oligos used for Nextera Indexing**. All P7 oligos and P5 oligos S501-S508 contain standard Illumina index sequences. P5 oligos S509-S524 are from [(Baym *et al*. 2015)](https://paperpile.com/c/3UiuMY/VA9NK).

**File S2: Flanking marker positions and CO positions**. In the header, “chr” refers to the chromosome number, “block1.end” indicates the last marker of a genotype block, “block2.start” indicates the first marker of a new genotype block, “co.position” is the mid -point between the two markers flanking the CO. Of the 373 individuals that had been pre-selected by screening for lack of COs in the *420* reporter interval, 20% of individuals still had a CO in the region. Since this fraction was not substantially reduced compared with the fraction of non-selected individuals that had a CO in *420* (26%), we did not exclude them from our analysis. They are indicated in the column “select.420”.

**File S3: List of genes in the CO “hotspot” regions spanning the top 5% of CO rates within 50-kb sliding windows.**

**File S4: List of genes in the CO “desert” regions spanning the top 5% longest genomic intervals without any COs.**

**File S5: Comparison of the structures of NLR and other defense gene loci between Col and Ler.**