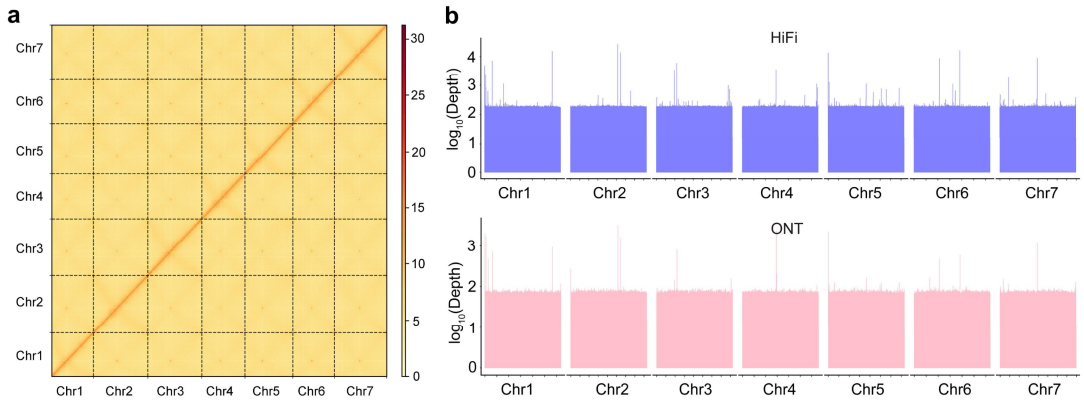
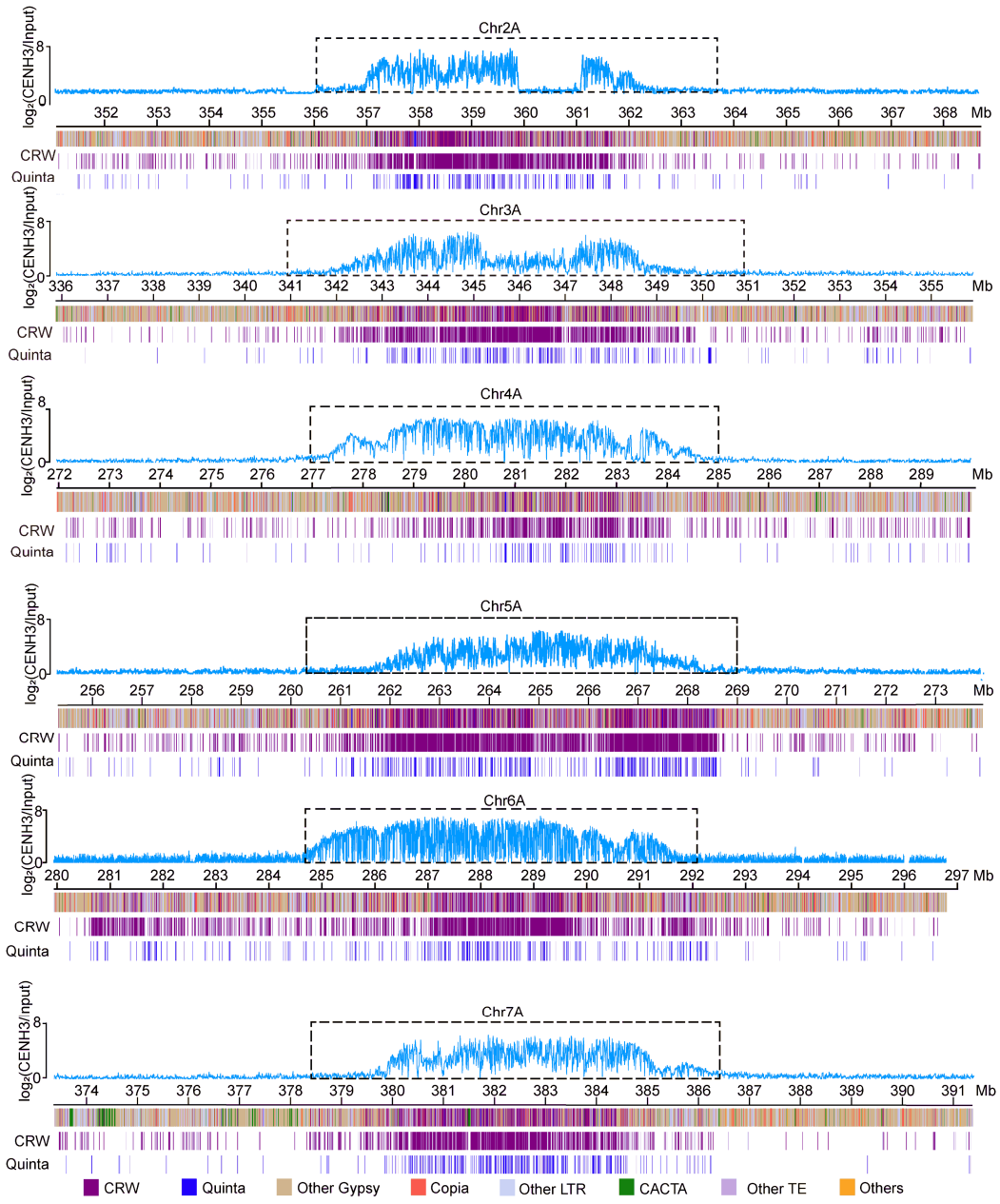
Complete genome and indexed EMS mutant library of einkorn wheat to advance polyploid wheat functional genomics

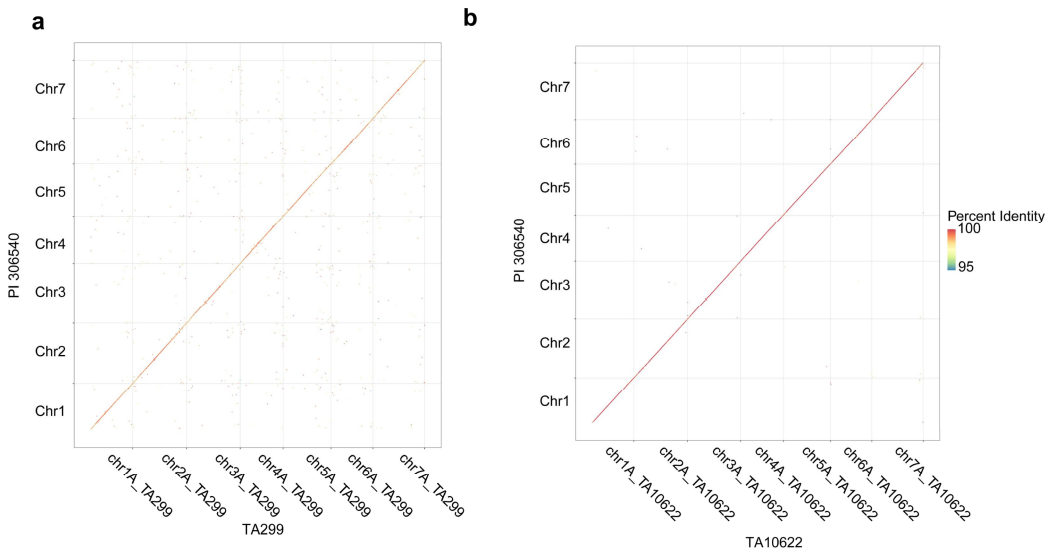
Li *et al.*



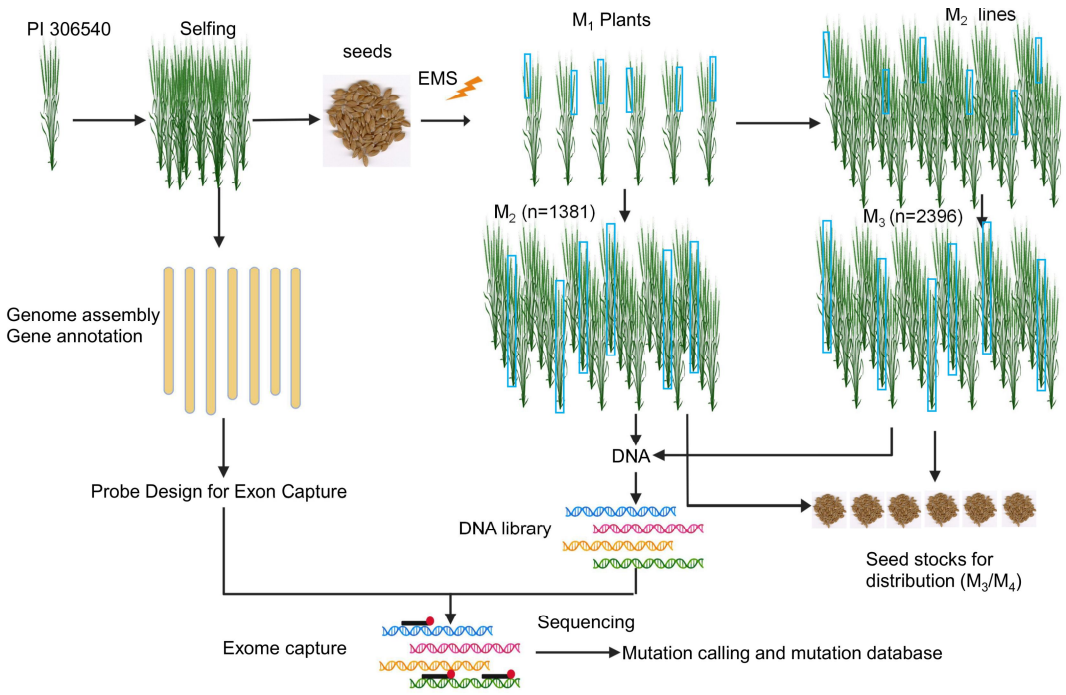
Supplementary Fig. 1. Genome assembly of PI 306540. a, Genome wide intensity signal heatmap of Hi-C chromosome interaction. **b**, Distribution of coverage depth of HiFi (blue) reads and ONT (pink) reads mapped to the PI 306540 genome.



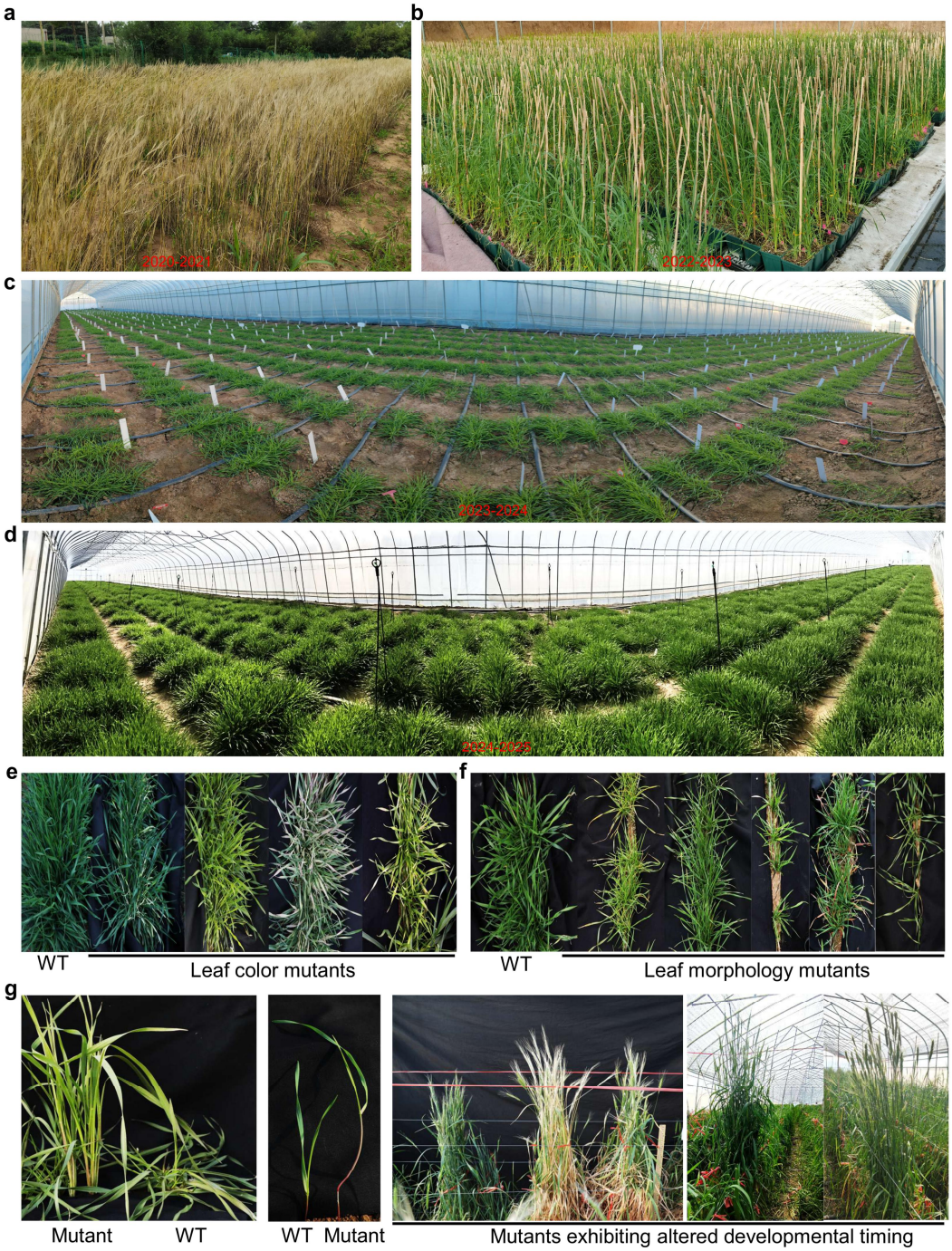
Supplementary Fig. 2. Centromere identification on chromosomes 2A-7A via CENH3 ChIP-seq. The track shows the CENH3 $\log_2$(ChIP/Input) signals (blue peaks) in 50-bp windows. The distributions of centromeric retrotransposons CRW and Quinta are displayed below. The x-axis represents genomic coordinates, and the y-axis shows the enrichment level of CENH3.



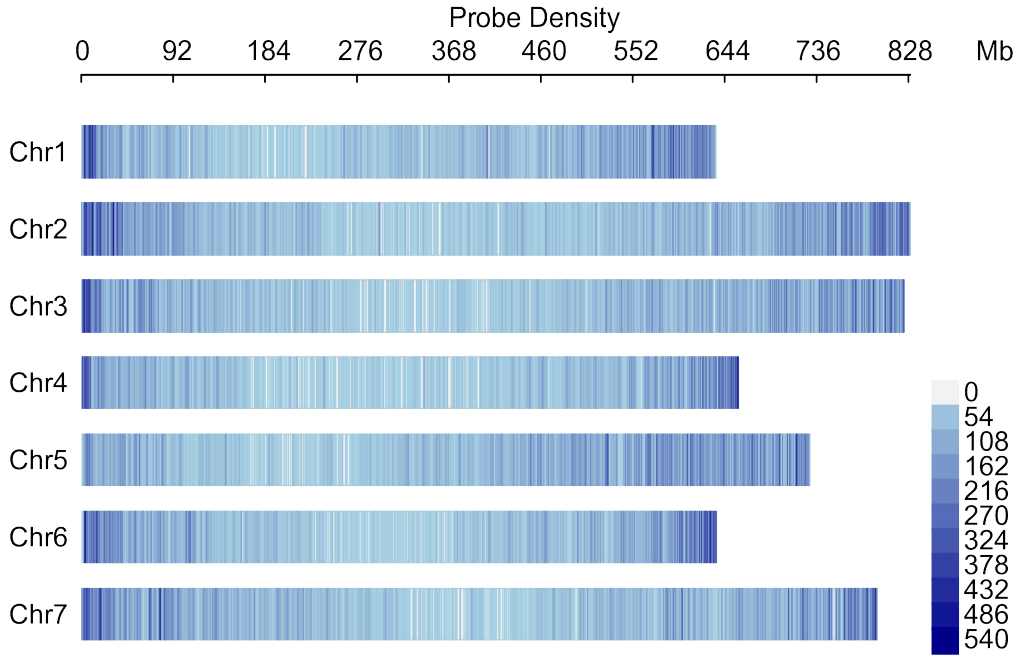
Supplementary Fig. 3. Pairwise whole-genome dot-plots between PI 306540, TA299, and TA10622. a, Dot-plot comparison between PI 306540 and TA299. **b**, Dot-plot comparison between PI 306540 and TA10622.



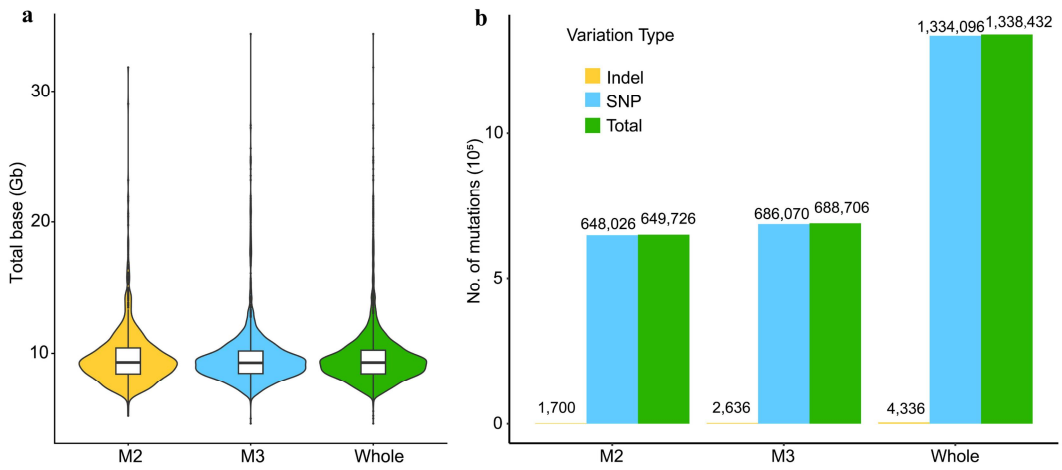
Supplementary Fig. 4. Overview of the development of the sequenced mutant populations. High-quality seeds were subjected to EMS mutagenesis. A total of 1,381 M_2 and 2,396 M_3 plants (one plant per line) were used for exome capture sequencing, followed by mutation calling to establish the mutation database. Seed stocks from M_3/M_4 lines were maintained for distribution. The figure was created with BioRender (https://www.biorender.com/).



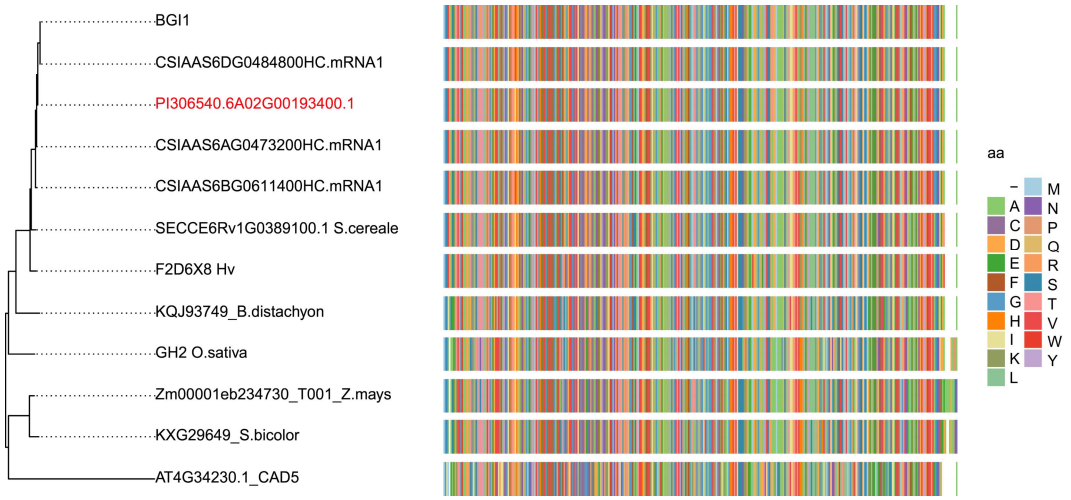
Supplementary Fig. 5. Creation of an EMS-induced mutant population in diploid wheat. **a-d**, Field and greenhouse propagation of mutant lines during the 2020-2025 growing seasons. M_1 plants were grown in the field, and a single M_2 plant derived from each M_1 plant was grown in the greenhouse. For phenotypic screening, 15-20 seeds from each line were sown in single-row plots. **e**, Leaf color mutants. **f**, Leaf morphology mutants. **g**, Mutants exhibiting altered plant architecture, internode length, and developmental timing.



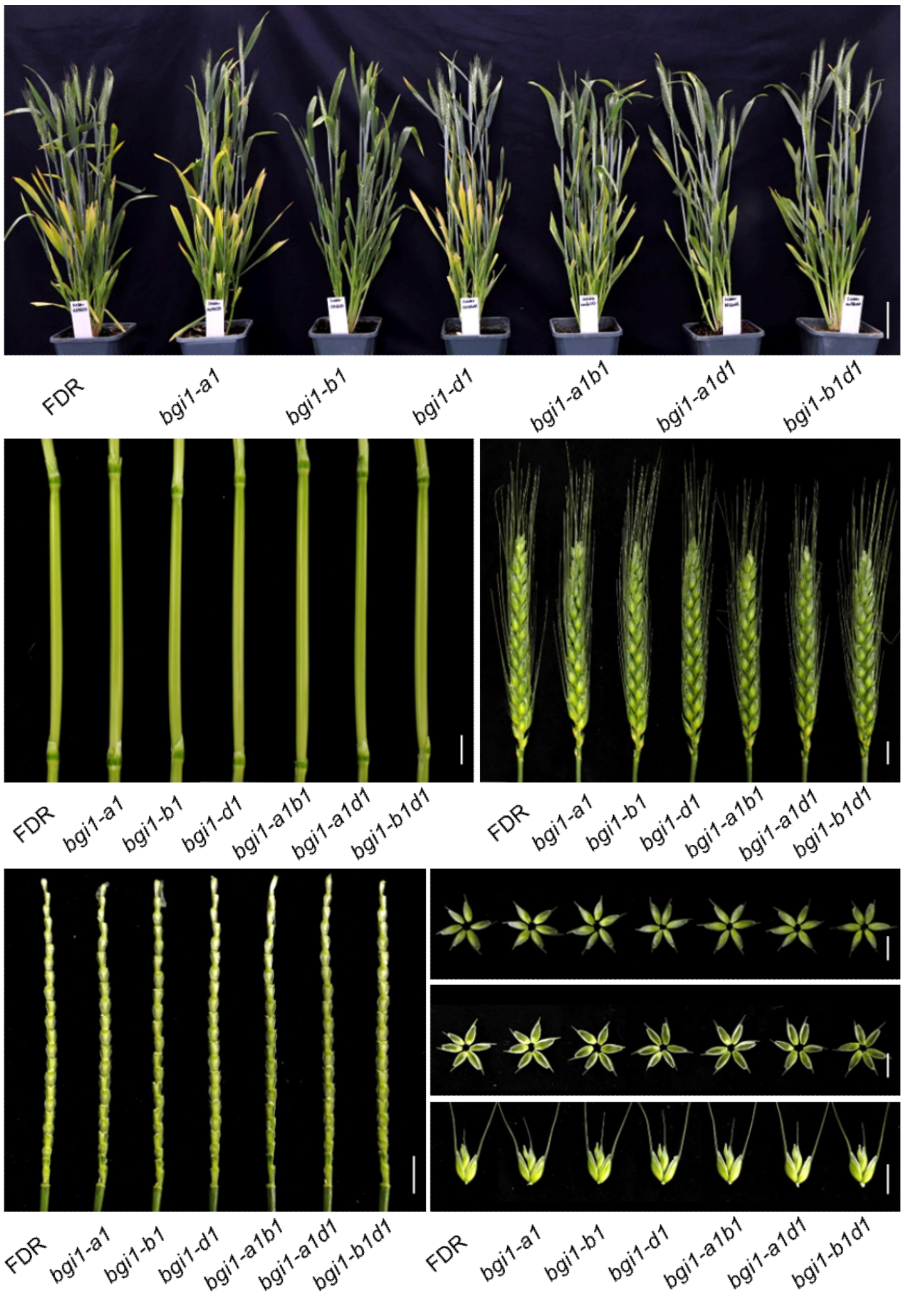
Supplementary Fig. 6. Genome-wide distribution of probe density along chromosomes. Probe density was calculated using 1-Mb non-overlapping windows across the seven chromosomes of PI 306540 and visualized using CMplot.



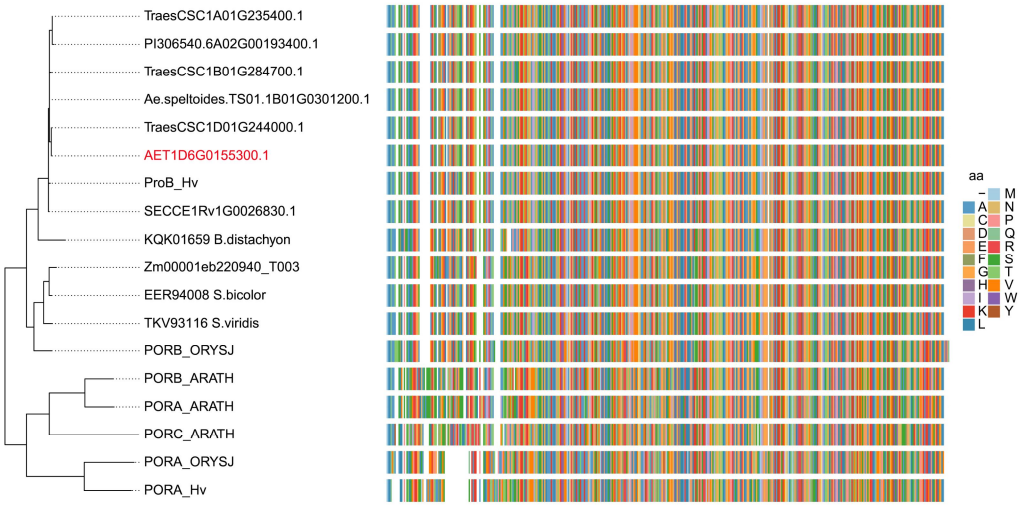
Supplementary Fig. 7. Characterization of the exome sequencing data and mutation profiles in the EMS mutant library. a, Distribution of sequencing data volume (Gb) for the M_2 (n = 1,381; yellow), M_3 (n = 2,396; blue), and combined (whole; n = 3,777; green) populations. The width of violin indicates the density of the data at different values. In each box plot, the horizontal line indicates the median; box edges represent the 25th and 75th percentiles; whiskers extend to the minimum or maximum values within 1.5× IQR. **b**, Number of mutations identified in the M_2, M_3, and the combined (whole) population. Indel (insertions/deletions) is shown in yellow, SNP (single nucleotide polymorphism) in blue, and total (InDel + SNP) in green.



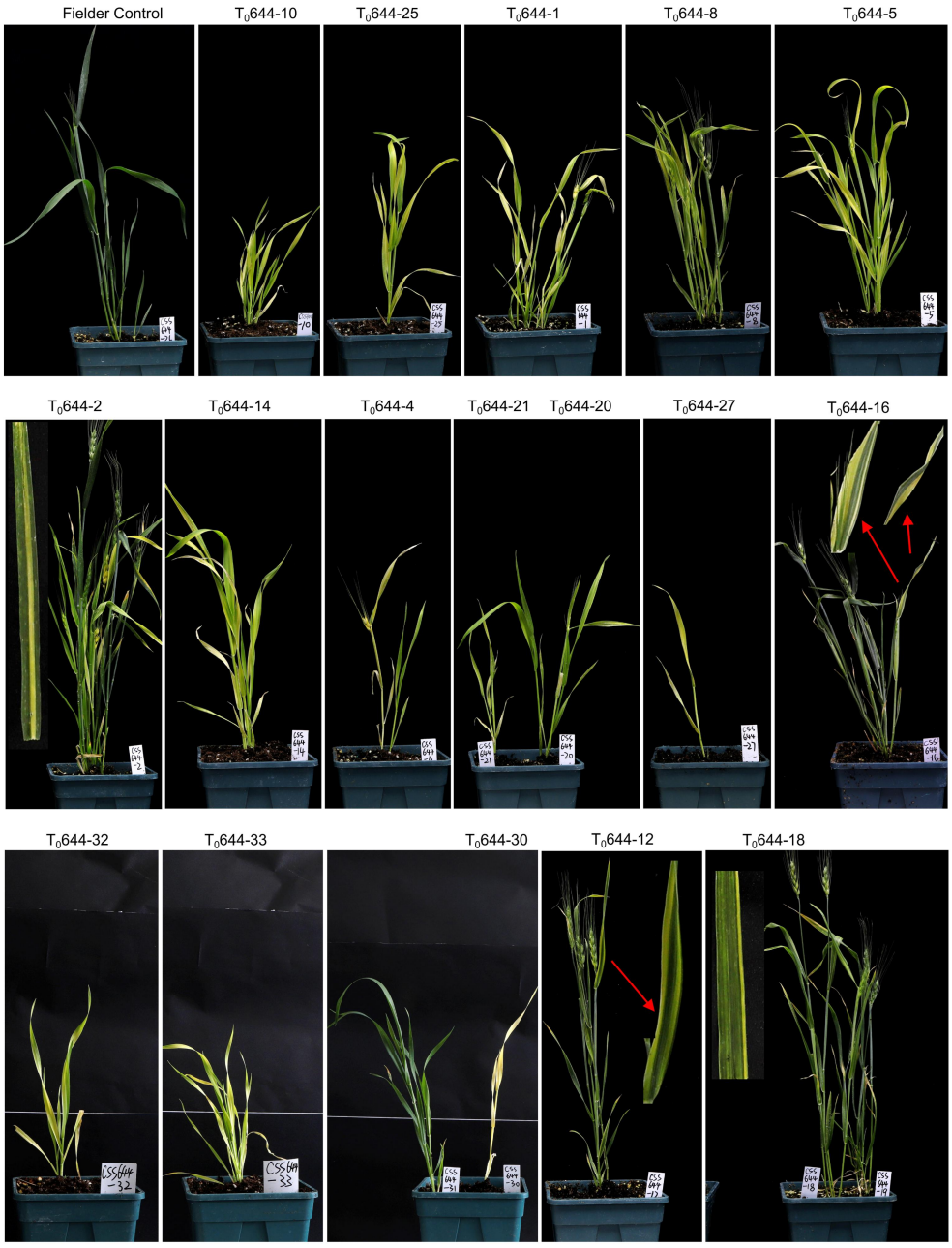
Supplementary Fig. 8. Phylogenetic tree analysis and multiple sequence alignment of PI306540.6A02G00193400. Protein sequences from *Triticum aestivum*, *Hordeum vulgare*, *Secale cereale*, *Aegilops tauschii*, *Zea mays*, *Arabidopsis thaliana*, *Oryza sativa*, and *Sorghum bicolor* were used to construct this phylogenetic tree using the neighbor-joining method in MEGA7. Multiple sequence alignment was performed using ggmsa. PI306540.6A02G00193400.1 is highlighted in red.



Supplementary Fig. 9. Phenotypic comparison of Fielder (FDR) and CRISPR/Cas9-generated single and double mutant plants. No obvious phenotypic alterations were observed in any of the single or double mutant plants. Photographs were taken at the flowering stage.



Supplementary Fig. 10. Phylogenetic tree analysis and multiple sequence alignment of AET1D6G0155300. Protein sequences from *Aegilops tauschii*, *Triticum aestivum*, *Hordeum vulgare*, *Secale cereale*, *Zea mays*, *Arabidopsis thaliana*, *Oryza sativa*, *Sorghum bicolor*, and *Setaria viridis* were used to construct this phylogenetic tree using the neighbor-joining method in MEGA7. Multiple sequence alignment was performed using ggmsa. AET1D6G0155300.1 is highlighted in red.



Supplementary Fig. 11. Phenotypic comparison between Fielder and CRISPR/Cas9-edited transgenic T_0 plants. A total of 17 T_0 transgenic plants (51.52%) carried mutations at the target site and exhibited yellow or yellow-green leaf phenotypes. Enlarged views of representative leaves are indicated by red arrows.