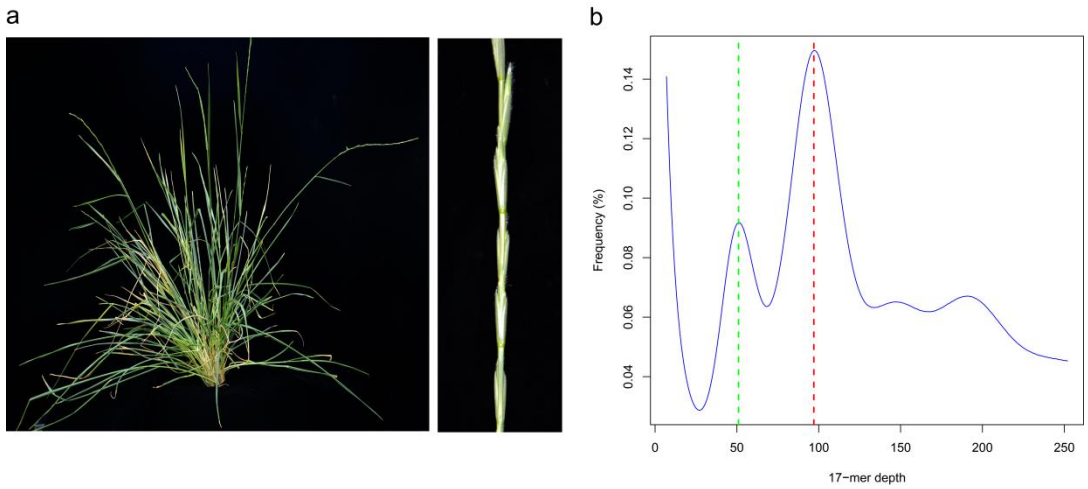
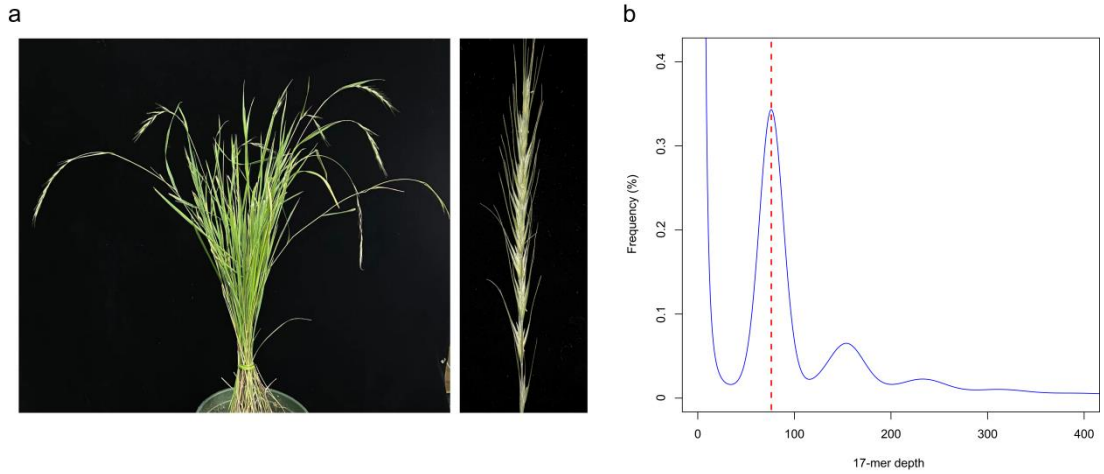
Analysis wheat wild relatives *Thinopyrum intermedium* and *Roegneria kamoji* genomes reveal different polyploid evolution paths

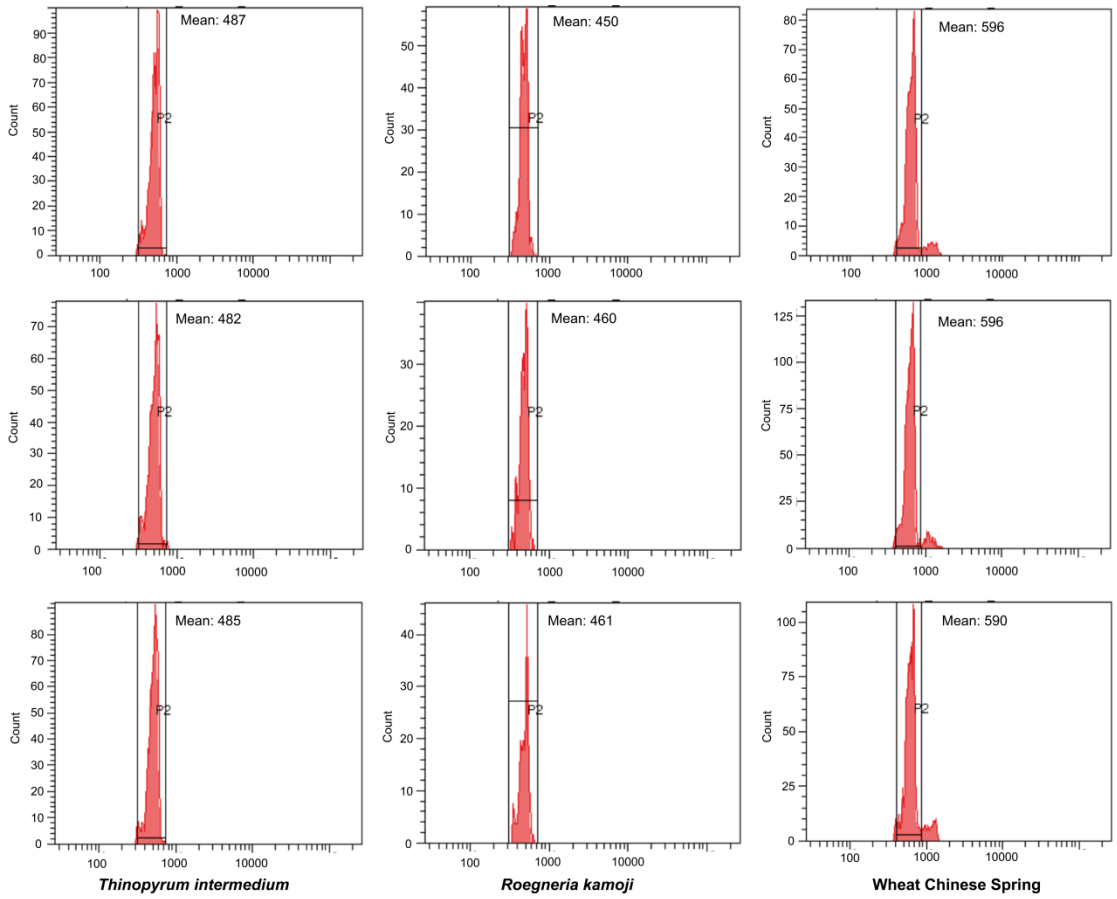
Sun *et al.*



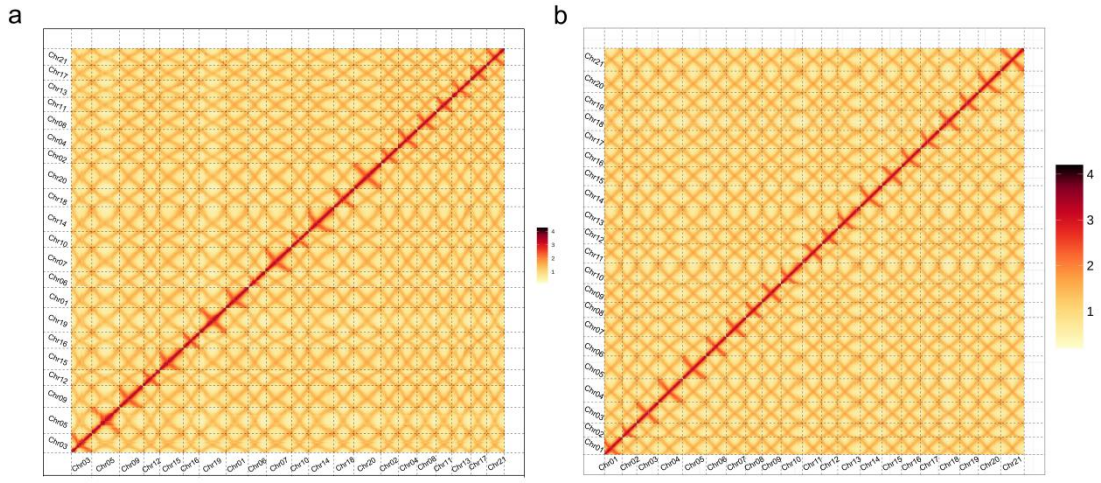
Supplementary Fig. 1. Plant morphology and genome size estimation of *Th. intermedium*. (a) Morphology photographs of *Th. intermedium*. **(b)** 17-mer depth distribution of the *Th. intermedium* genome. The genome size was estimated to be 11.75 Gb, with a heterozygosity rate of 1.35% and a repeat content of 85.8%.



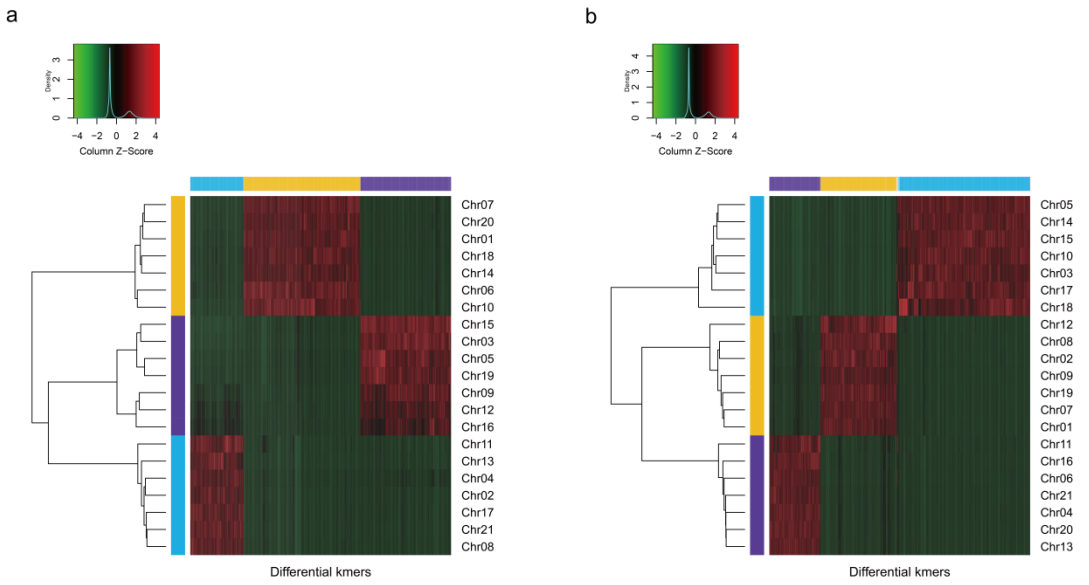
Supplementary Fig. 2. Plant morphology and genome size estimation of *R. kamoji*. (a) Morphology photographs of *R. kamoji*. **(b)** 17-mer depth distribution of the *R. kamoji* genome. The genome size of *R. kamoji* was estimated to be 11.26 Gb, with a heterozygosity rate of 0.17% and a repeat content of 87.42%.



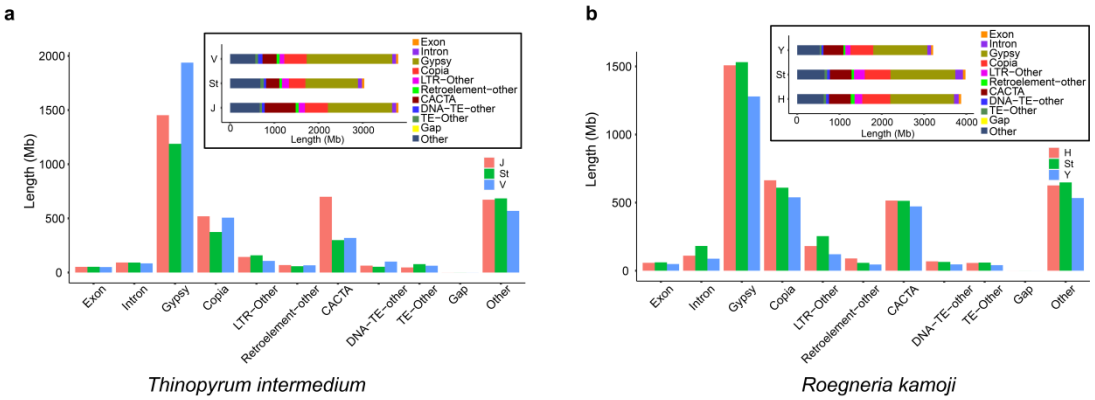
Supplementary Fig. 3. Genome size estimation by flow cytometry for *Th. intermedium*, *R. kamoji*, and the wheat cultivar "Chinese Spring" (CK). Three biological replicates were performed for each sample. Genome size = (mean peak value / peak value of CK) × genome size of CK, where the genome size of "Chinese Spring" is 14.51 Gb. For *Th. intermedium*: (487 + 482 + 485) / (596 + 596 + 590) × 14.51 Gb = 11.84 Gb; For *R. kamoji*: (450 + 460 + 461) / (596 + 596 + 590) × 14.51 Gb = 11.16 Gb.



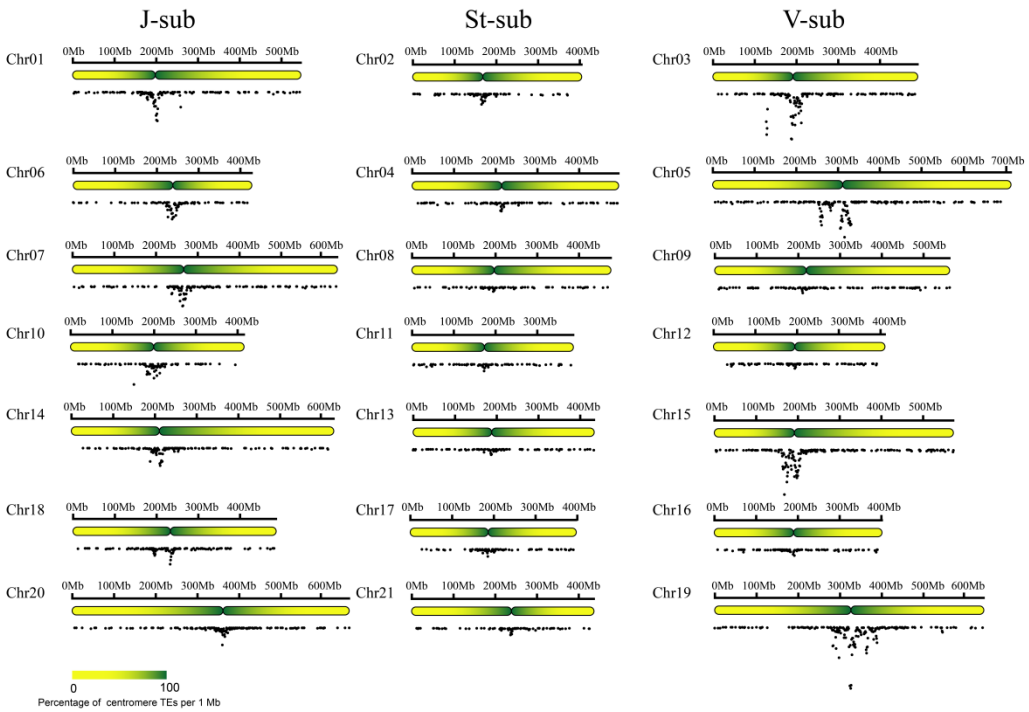
Supplementary Fig. 4. Hi-C interaction heatmap for the pseudo-chromosome assembly of *Th. intermedium and R. kamoji*. **(a)** *Th. inermedium*; **(b)** *R. kamoji*. The resolution used to calculate interaction strength for each bin is 100 kb.



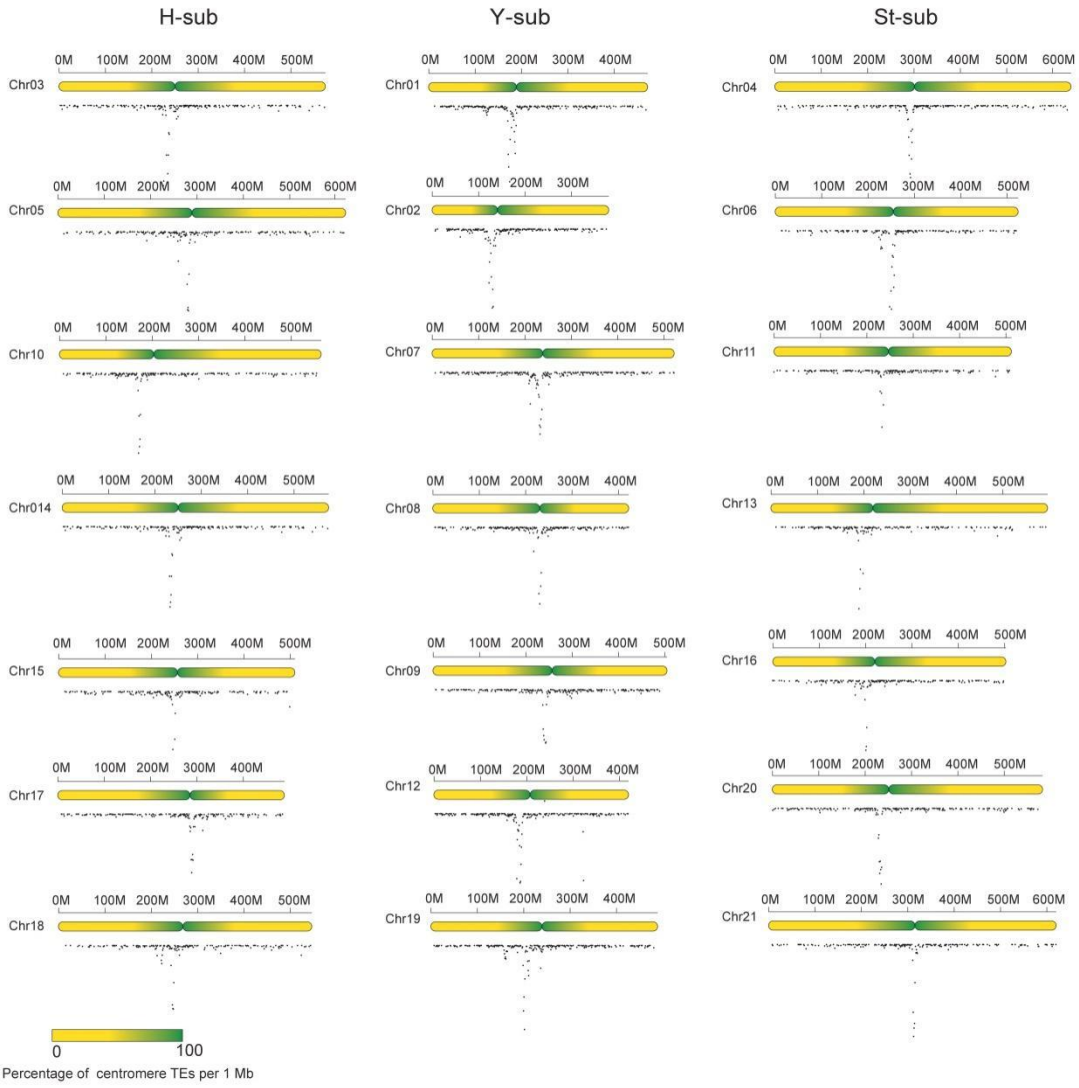
Supplementary Fig. 5. Subgenome phase results of *Th. intermedium and R. kamoji*. (a) *Th. inermedium*; **(b)** *R. kamoji*. Phased results for *Th. intermedium* genome: Subgenome1 (Chr01, Chr06, Chr07, Chr10, Chr14, Chr18, and Chr20), Subgenome2 (Chr03, Chr05, Chr09, Chr12, Chr15, Chr16, and Chr19), and Subgenome3 (Chr02, Chr04, Chr08, Chr11, Chr13, Chr17, and Chr21); Phased results for *R. kamoji* genome: Subgenome1 (Chr17, Chr05, Chr14, Chr18, Chr10, Chr15, and Chr03), Subgenome2 (Chr02, Chr07, Chr19, Chr12, Chr01, Chr08, and Chr09), and Subgenome3 (Chr16, Chr04, Chr20, Chr11, Chr13, Chr06, and Chr21).



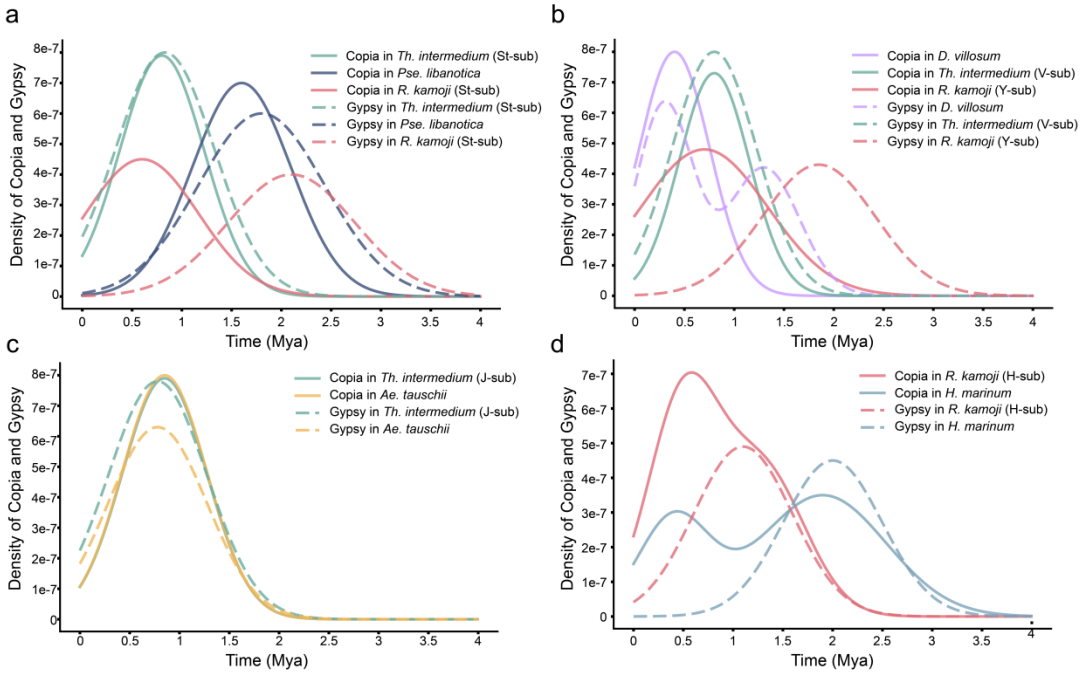
Supplementary Fig. 6. Statistics of genome composition in *Th. intermedium* and *R. kamoji* genome assemblies. (a) *Th. inermedium*; **(b)** *R. kamoji*. Subgenome assignments were based on the phylogenic results.



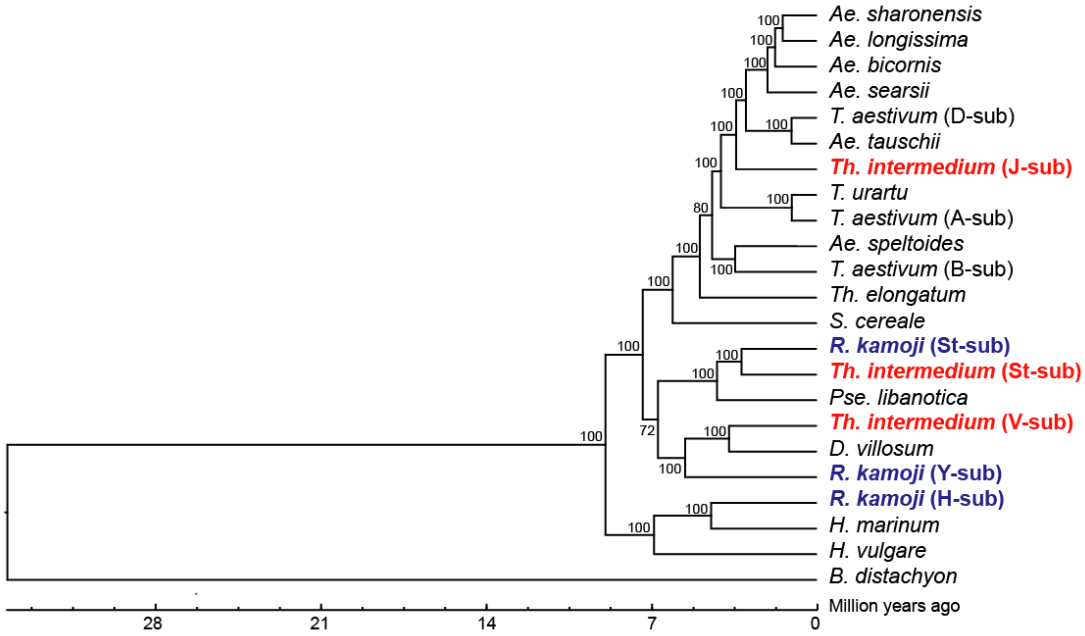
Supplementary Fig. 7. Density distribution of *Cereba* and *Quinta* retrotransposons across each pseudo-chromosome of *Th. intermedium* genome assembly. The resolution used to calculate the density of centromere TEs in each bin was 1 Mb.



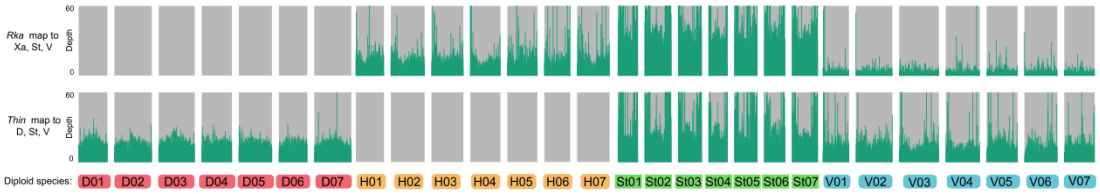
Supplementary Fig. 8. Density distribution of *Cereba* and *Quinta* retrotransposons across each pseudo-chromosome of *R. kamoji* genome assembly. The resolution used to calculate the density of centromere TEs in each bin was 1 Mb.



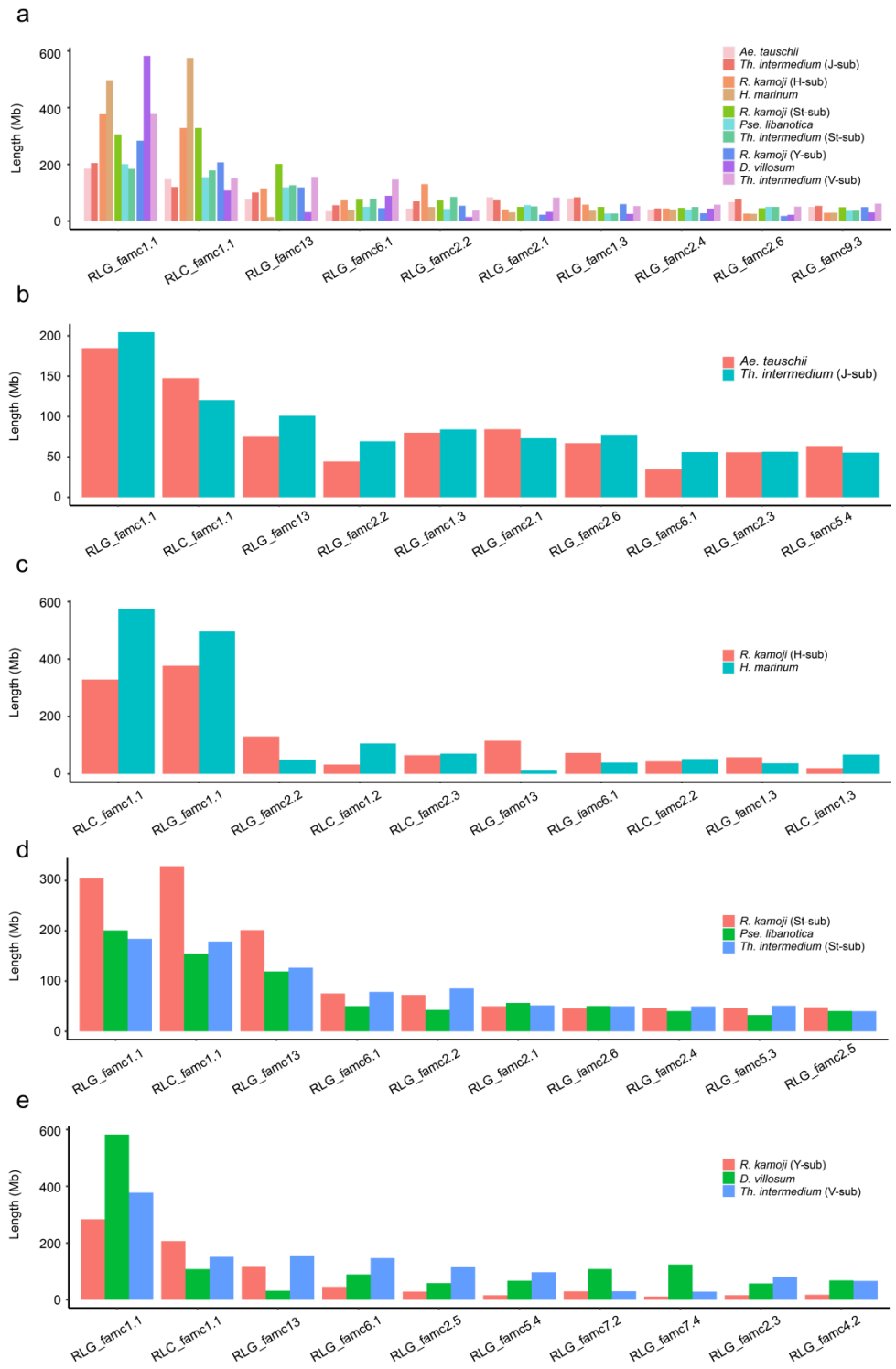
Supplementary Fig. 9. LTR activity in the subgenomes of *Th. intermedium*, *R. kamoji*, and closely related diploid species. (a) Activity of *Copia* and *Gypsy* retrotransposons in *Th. intermedium* (St-sub), *R. kamoji* (St-sub) and *Pse. libanotica*; **(b)** Activity of *Copia* and *Gypsy* retrotransposons in *Th. intermedium* (V-sub), *R. kamoji* (V-sub), and *D. villosum*; **(c)** Activity of *Copia* and *Gypsy* retrotransposons in *Th. intermedium* (J-sub) and *Ae. tauschii*; **(d)** Activity of *Copia* and *Gypsy* retrotransposons in *R. kamoji* (H-sub) and *H. marinum*.



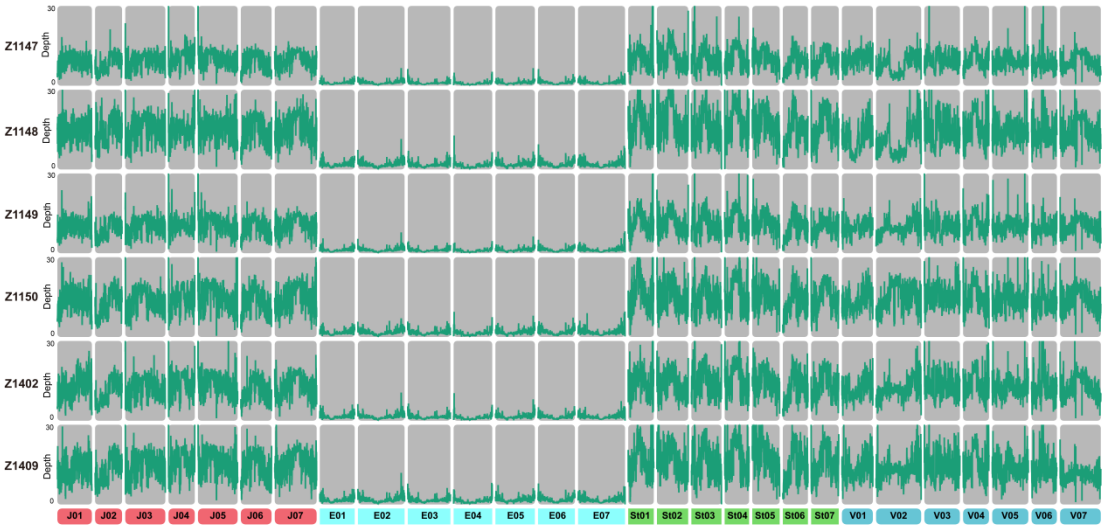
Supplementary Fig. 10. Maximum likelihood phylogenomic tree reconstruction using single-copy orthologous genes extracted from publicly available *Poaceae* genomes and our assemblies. Branch support values were based on 1,000 standard bootstrap replicates.



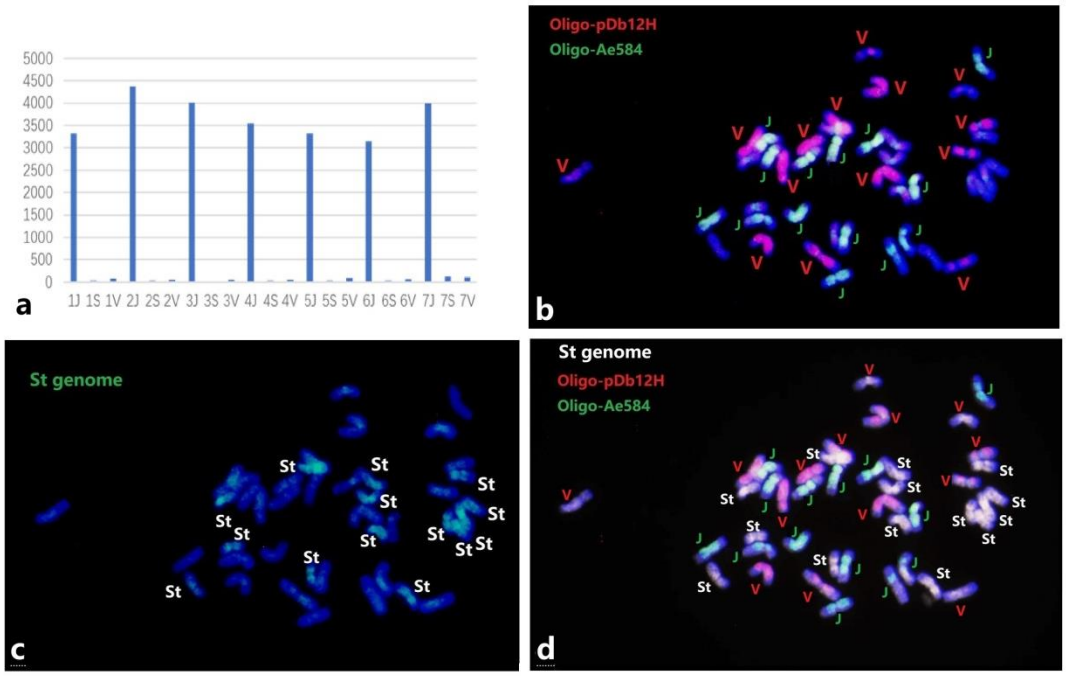
Supplementary Fig. 11. Whole-genome mapping of simulated reads from *Th. intermedium* and *R. kamoji* to the related diploid species. All reads from *Th. intermedium* were mapped to the combined *Ae. tauschii*, *Pse. libanotica*, and *D. villosum* genome assemblies and reads from *R. kamoji* were mapped to the combined *H. marinum*, *Pse. libanotica*, and *D. villosum* assemblies, respectively. Mapped reads depth was calculated within a 1 Mb window. *Rka*: *R. kamoji*; *Thin*: *Th. intermedium*. Xa: *H. marinum*; St: *Pse. libanotica*; V: *D. villosum* and D: *Ae. tauschii*.



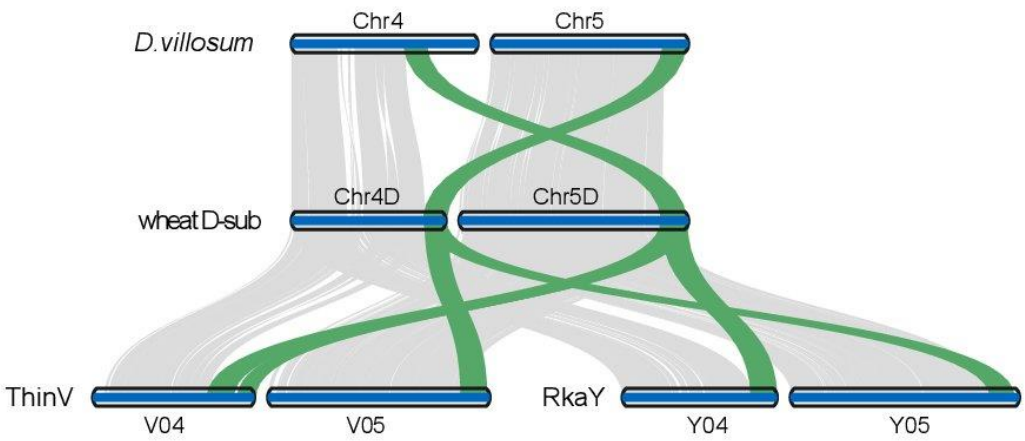
Supplementary Fig. 12. Length distribution of the top 10 LTR-RT subfamilies. (a) Top 10 LTR-RT subfamilies in *Th. intermedium*, *R. kamoji*, and potential diploid progenitors. **(b)** Top 10 LTR-RT subfamilies in *Ae. Tauschii* and *Th. intermedium* (J-sub). **(c)** Top 10 LTR-RT subfamilies in *R. kamoji* (H-sub) and *H. marinum*. **(d)** Top 10 LTR-RT subfamilies in *R. kamoji* (St-sub), *Pse. libanotica*, and *Th. intermedium* (St-sub). **(e)** Top 10 LTR-RT subfamilies in *R. kamoji* (Y-sub), *D. villosum*, and *Th. intermedium* (V-sub).



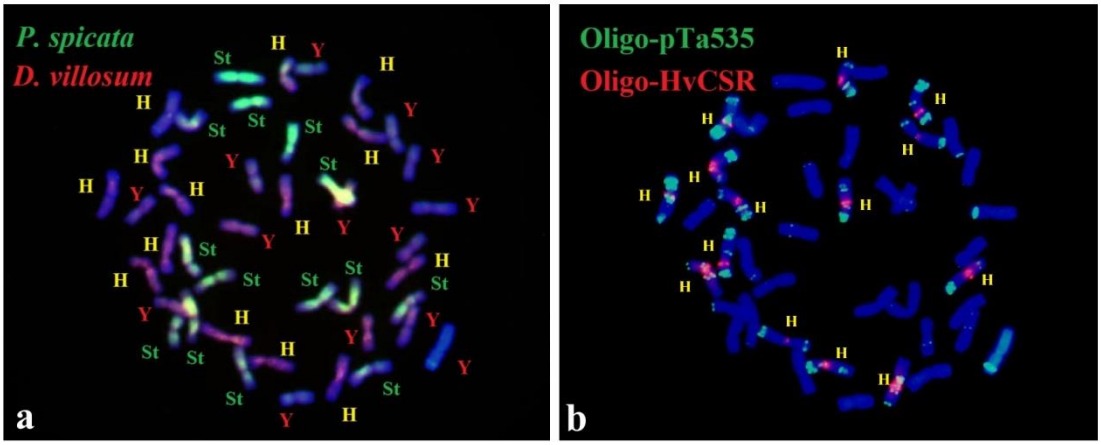
Supplementary Fig. 13. Whole-genome mapping of re-sequencing reads from six *Th. intermedium* accessions. All reads were mapped to the combined *Th. intermedium* and *Th. elongatum* genome assemblies, and reads depth were calculated within a 1 Mb window.



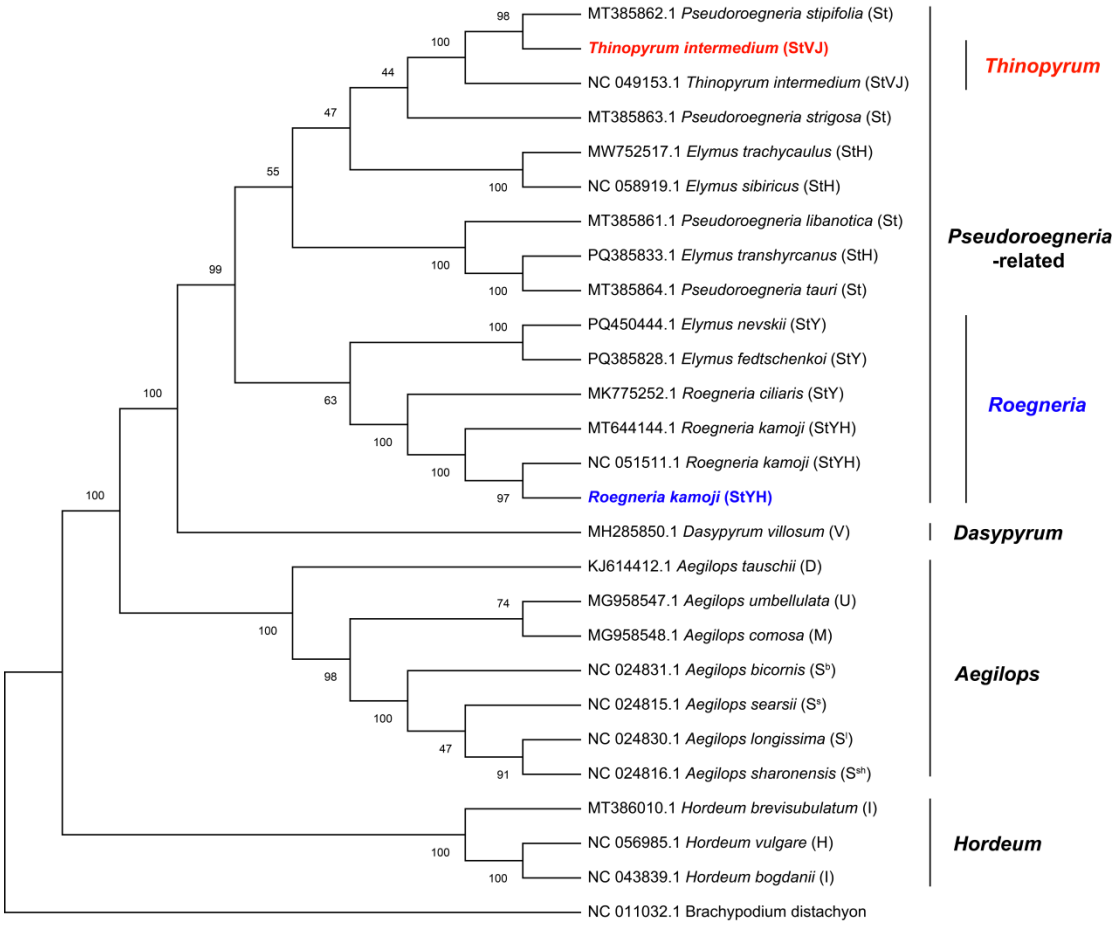
Supplementary Fig. 14. Sequential GISH-FISH patterns for the metaphase of *Th. intermedium*. (a) The predicted copy numbers of probe Oligo-Ae584 in *Th. intermedium* genome. **(b)** FISH results obtained using V-genome specific probe Oligo-pDb12H and *Aegilops* species specific probe Oligo-Ae584. **(c)** GISH results obtained using the total genomic DNA of *Pse. spicata* (St). **(d)** Merged GISH-FISH results obtained using probes Oligo-pDb12H, Oligo-Ae584 and genomic DNA of *Pse. spicata*.



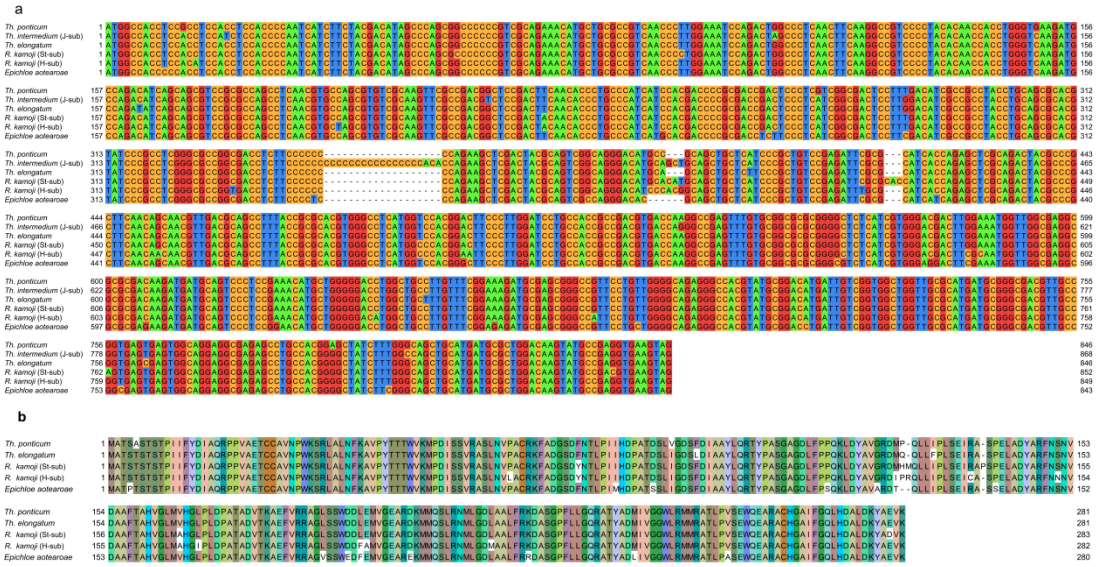
Supplementary Fig. 15. The 4V-5V translocation events among *D. villosum*, *Th. intermedium*, and *R. kamoji*. Green lines represent the translocation blocks, and gray lines represent the synteny blocks. ThinV: V-sub of *Th. intermedium*; RkaY: Y-sub of *R. kamoji*.



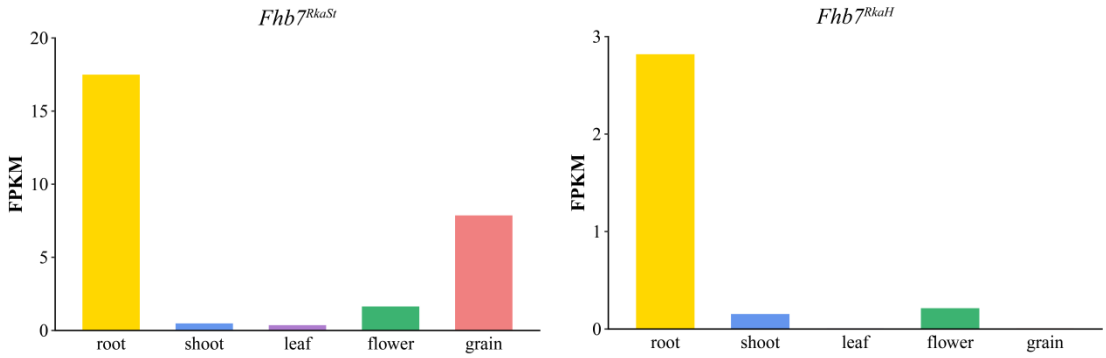
Supplementary Fig. 16. Sequential GISH-FISH patterns for the metaphase of *R. kamoji*. (a) GISH results obtained using the total genomic DNA of *Pse. spicata* (St) and *D. villosum* (V). **(b)** FISH results obtained using both the *Hordeum*-specific centromeric repeat probe Oligo-HvCSR and the high-copy tandem repeat probe Oligo-pTa535.



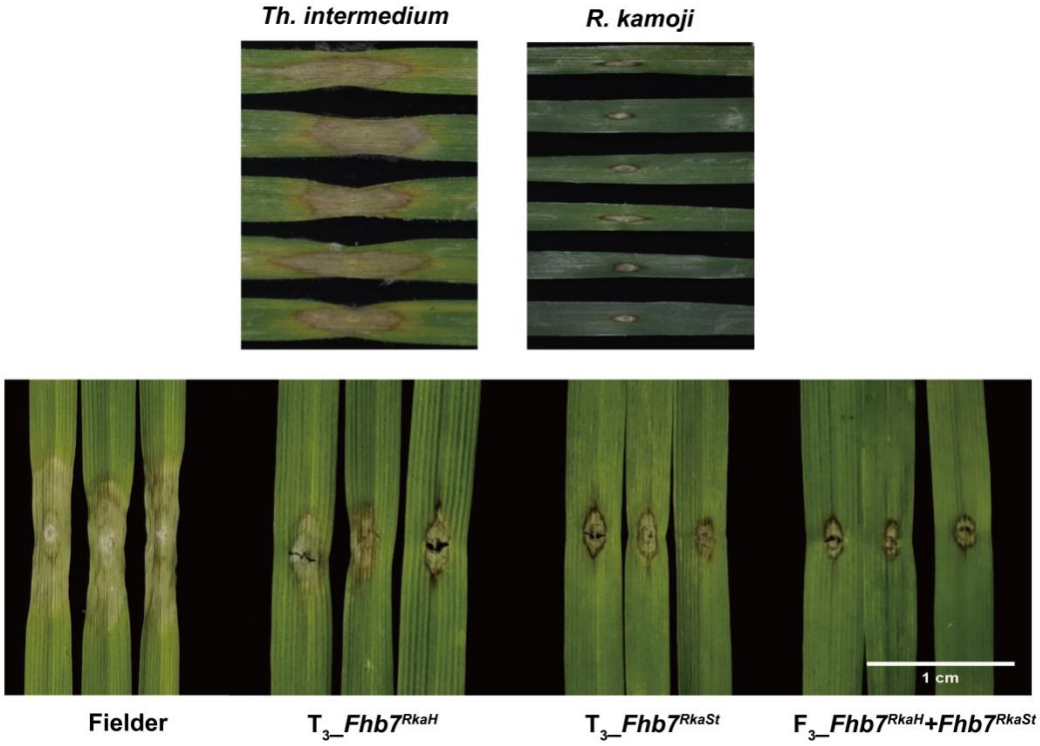
Supplementary Fig. 17. Maximum likelihood phylogenetic tree construction based on the publicly available chloroplast sequences and our assembled *Th. intermedium* and *R. kamoji* chloroplast sequences. All sequences were merged and aligned using MAFFT software, and the phylogenetic tree was constructed using IQ-TREE2 software. Branch support values were based on 1000 standard bootstrap replicates.



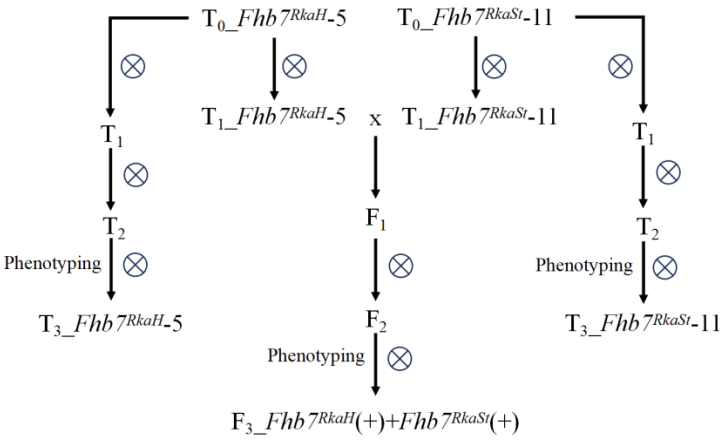
Supplementary Fig. 18. Sequence alignment among *Fhb7* homologs. (a) Alignment of the DNA sequences of *Fhb7* homologs in *Th. ponticum*, *Th. elongatum*, *Th. intermedium* (J-sub), *R. kamoji* (St-sub), *R. kamoji* (H-sub), and *Epichloe aotearoae*. **(b)** Alignment of protein sequences of Fhb7 in *Th. ponticum*, *Th. elongatum*, *R. kamoji* (St-sub), *R. kamoji* (H-sub), and *E. aotearoae*.



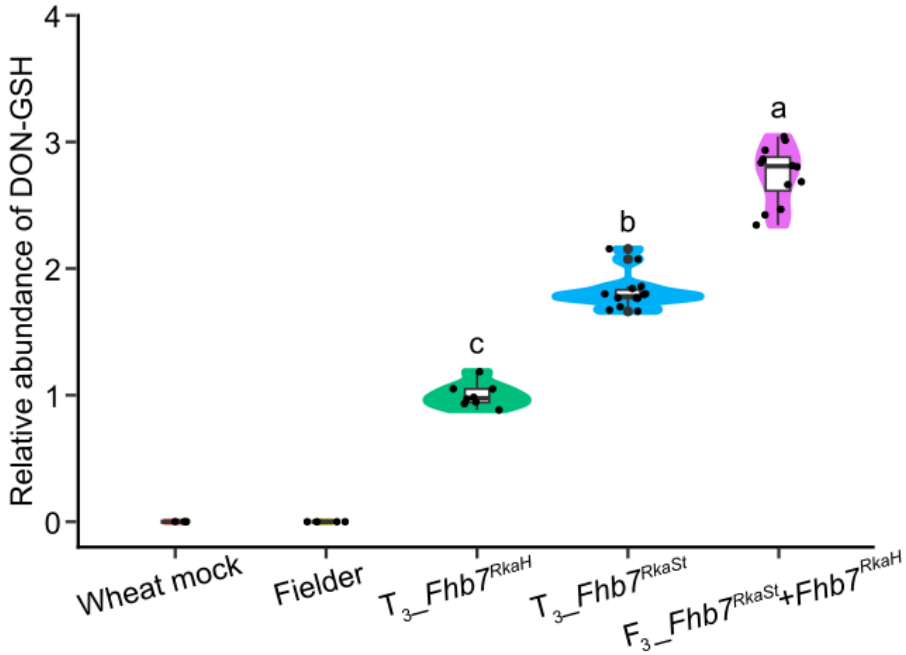
Supplementary Fig. 19. Tissue expression patterns of *Fhb7*^{RkaSt} and *Fhb7*^{RkaH} in *R. kamoji*. The y-axis represents the gene expression values calculated from transcriptome data. FPKM: Fragments Per Kilobase of exon model per Million mapped reads.



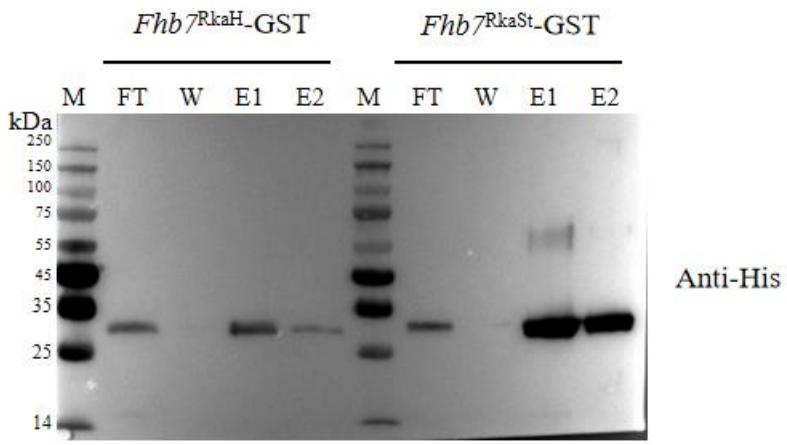
Supplementary Fig. 20. *Fusarium* head blight resistance evaluation using detached leaves.



Supplementary Fig. 21. Construction procedure of transgenic lines carrying *Fhb7*^{RkaH} and *Fhb7*^{RkaSt}.



Supplementary Fig. 22. Quantification of DON-GSH in wheat lines. The mean value of T_3_*Fhb7*^{RkaH} was used as reference. The central box denotes the interquartile range (25th to 75th percentiles), with the median marked by an internal line. Whiskers indicate the standard deviation. Group differences were assessed using Two-sided Wilcoxon test, with significance levels: $P < 0.05$. The letter at the top of boxes indicates significant differences among groups. Sample sizes: Wheat mock (n = 8), Fielder (n = 8), T_3_*Fhb7*^{RkaH} (n = 8), T_3_*Fhb7*^{RkaSt} (n = 14) and F_3_*Fhb7*^{RkaSt}+*Fhb7*^{RkaH} (n = 12). Source data are provided as a Source Data file.



Supplementary Fig. 23. Protein expression and purification of Fhb7^{RkaH}-GST and Fhb7^{RkaSt}-GST in *E. coli.* Protein analyses were performed using western blot. M: marker; FT: flow-through fraction; W: wash fraction; E1, E2: elution fractions. Source data are provided as a Source Data file.