Plant Communications
Research article

CellPress
Partner Journal

Genome architecture of the allotetraploid wild grass *Aegilops ventricosa* reveals its evolutionary history and contributions to wheat improvement

Zehou Liu[1,8], Fan Yang[2,8], Hongshen Wan[1,8], Cao Deng[3,4,8], Wenjing Hu[5,8], Xing Fan[6], Jirui Wang[6], Manyu Yang[1], Junyan Feng[2], Qin Wang[1], Ning Yang[1], Li Cai[6], Ying Liu[6], Hao Tang[1], Shizhao Li[1], Jiangtao Luo[1], Jianmin Zheng[1], Ling Wu[1], Ennian Yang[1], Zongjun Pu[1], Jizeng Jia[7,*], Jun Li[1,*] and Wuyun Yang[1,*]

[1]Crop Research Institute, Sichuan Academy of Agricultural Sciences, Environment Friendly Crop Germplasm Innovation and Genetic Improvement Key Laboratory of Sichuan Province, Key Laboratory of Wheat Biology and Genetic Improvement on Southwestern China, Key Laboratory of Tianfu Seed Industry Innovation, Chengdu, China

[2]Biotechnology and Nuclear Technology Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu, China

[3]The Key Laboratory of Animal Disease and Human Health of Sichuan Province, College of Veterinary Medicine, Sichuan Agricultural University, Chengdu, China

[4]Departments of Bioinformatics, DNA Stories Bioinformatics Center, Chengdu, China

[5]Lixiahe Institute of Agricultural Sciences, Yangzhou, Jiangsu, China

[6]Triticeae Research Institute, Sichuan Agricultural University, Chengdu, China

[7]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing, China

[8]These authors contributed equally to this article.

***Correspondence: Jizeng Jia (jiajizeng@caas.cn), Jun Li (lijunchd@126.com), Wuyun Yang (yangwuyun@126.com)**

https://doi.org/10.1016/j.xplc.2024.101131

ABSTRACT

The allotetraploid wild grass *Aegilops ventricosa* (2n = 4x = 28, genome $D^vD^vN^vN^v$) has been recognized as an important germplasm resource for wheat improvement owing to its ability to tolerate biotic stresses. In particular, the $2N^vS$ segment from *Ae. ventricosa*, as a stable and effective resistance source, has contributed greatly to wheat improvement. The $2N^vS$/2AS translocation is a prevalent chromosomal translocation between common wheat and wild relatives, ranking just behind the 1B/1R translocation in importance for modern wheat breeding. Here, we assembled a high-quality chromosome-level reference genome of *Ae. ventricosa* RM271 with a total length of 8.67 Gb. Phylogenomic analyses revealed that the progenitor of the D^v subgenome of *Ae. ventricosa* is *Ae. tauschii* ssp. *tauschii* (genome DD); by contrast, the progenitor of the D subgenome of bread wheat (*Triticum aestivum* L.) is *Ae. tauschii* ssp. *strangulata* (genome DD). The oldest polyploidization time of *Ae. ventricosa* occurred ~0.7 mya. The D^v subgenome of *Ae. ventricosa* is less conserved than the D subgenome of bread wheat. Construction of a graph-based pangenome of 2AS/$6N^vL$ (originally known as $2N^vS$) segments from *Ae. ventricosa* and other genomes in the Triticeae enabled us to identify candidate resistance genes sourced from *Ae. ventricosa*. We identified 12 nonredundant introgressed segments from the D^v and N^v subgenomes using a large winter wheat collection representing the full diversity of the European wheat genetic pool, and 29.40% of European wheat varieties inherit at least one of these segments. The high-quality RM271 reference genome will provide a basis for cloning key genes, including the *Yr17-Lr37-Sr38-Cre5* resistance gene cluster in *Ae. ventricosa*, and facilitate the full use of elite wild genetic resources to accelerate wheat improvement.

Key words: *Aegilops ventricosa*, $6N^vL$/2AS translocation, pangenome, introgression, wheat improvement

Liu Z., Yang F., Wan H., Deng C., Hu W., Fan X., Wang J., Yang M., Feng J., Wang Q., Yang N., Cai L., Liu Y., Tang H., Li S., Luo J., Zheng J., Wu L., Yang E., Pu Z., Jia J., Li J., and Yang W. (2025). Genome architecture of the allotetraploid wild grass *Aegilops ventricosa* reveals its evolutionary history and contributions to wheat improvement. Plant Comm. **6**, 101131.

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

INTRODUCTION

Wheat is the most important food crop in the world, and its production is affected by various factors, including biotic and abiotic stresses. *Aegilops ventricosa* Tausch., an allotetraploid (2n = 4x = 28, genome $D^vD^vN^vN^v$) wild Triticeae grass endemic to the Mediterranean region (Yen et al., 2020), is an important germplasm resource for the improvement of bread wheat (*Triticum aestivum* L.) owing to its ability to resist biotic stresses (Doussinault et al., 1983; Delibes et al., 1987; Bonhomme et al., 1995; Badaeva et al., 2008; Gao et al., 2021). Cytogenetic and phylogenetic evidence has suggested that the speciation of *Ae. ventricosa* resulted from hybridization between *Ae. tauschii* (2n = 2x = 14, genome DD) as the female parent and *Ae. uniaristata* (2n = 2x = 14, genome NN) as the male parent (Lilienfeld, 1951; Chen et al., 2020).

In work aiming to transfer resistance genes from *Ae. ventricosa* to bread wheat, the $2N^vS$ segment was introduced into the wheat line VPM1 (Doussinault et al., 1983), which contained the well-known $2N^vS$-2AS translocation and was subsequently widely used to breed numerous excellent wheat cultivars, such as Jagger (Turner et al., 2016) and Renan (Dedryver et al., 2009). This $2N^vS$ segment from *Ae. ventricosa* has made an important contribution to wheat improvement worldwide because of the *Pch1* gene for resistance to eyespot (Mena et al., 1992) and the *Yr17-Lr37-Sr38-Cre5* resistance gene cluster, whose genes confer resistance to yellow, leaf, and stem rusts and cereal cyst nematode, respectively (Tanguy et al., 2005). Subsequently, Tanguy et al. (2005) demonstrated that the same $2N^vS$ segment was present as a preexisting translocation in *Ae. ventricosa* ($2N^vS$-$6N^vL$).

The low genetic diversity of the *T. aestivum* D subgenome has long motivated breeders to recruit diversity from *Ae. tauschii* (Gaurav et al., 2022) by hybridization between bread wheat and a synthetic hexaploid wheat derived from tetraploid wheat and *Ae. tauschii* (Li et al., 2018) or by direct hybridization between bread wheat and *Ae. tauschii* (Gill and Raupp, 1987; Gill et al., 2006). However, it is expected that more D subgenomes in *Aegilops* could be exploited for wheat improvement. Indeed, it has been demonstrated that *Ae. ventricosa* and *T. turgidum* can be hybridized directly, generating an octoploid amphiploid as a bridge for the transfer of valuable alien genes from *Ae. ventricosa* to wheat (Doussinault et al., 1983; Zhang et al., 2020) and extending the source of D subgenomes for wheat improvement. However, thus far, the lack of a reference genome has limited the cloning and further study of resistance genes and the utilization of D-subgenome resources derived from *Ae. ventricosa*.

To pave the way for wheat improvement using *Ae. ventricosa*, we built a comprehensive and high-quality reference genome for *Ae. ventricosa*. The reference-quality genome assemblies provide novel genome resources for deeper evolutionary studies of *Ae. ventricosa* as well as the Triticeae. Our results also shed new light on the evolution of the D lineage, confirming that the progenitor of the D^v subgenome of *Ae. ventricosa* is *Ae. tauschii* ssp. *tauschii* (genome DD); by contrast, the progenitor of the D subgenome of bread wheat (*T. aestivum* L.) is *Ae. tauschii* ssp. *strangulata* (genome DD). We confirmed that the previously well-known $2N^vS$ segment is present as a preexisting translocation in *Ae. ventricosa* ($2N^vS$-$6N^vL$), and by constructing a graph-based pan-genome of $2AS/6N^vL$ segments from *Ae. ventricosa* and other genomes in the Triticeae, we identified six candidate *Ae. ventricosa*-sourced disease-resistance genes. In addition to the classical $6N^vL$ introgression segment, many novel introgression segments derived from the N^v and D^v subgenomes of *Ae. ventricosa* were found. The novel genomic resources provided here can be mined to identify natural variants and genes to meet the ever-increasing global demand for wheat improvement.

RESULTS

Genome assembly and annotation

We assembled a high-quality genome of *Ae. ventricosa* RM271 (RM271 hereafter) using PacBio HiFi circular consensus sequencing (CCS) reads corresponding to ~45-fold coverage (Supplemental Figure 1A and 1B; Supplemental Table 1). The total assembly size of the RM271 genome was 8.67 Gb, and the longest contig and N50 contig lengths were 255.18 and 54.71 Mb, respectively (Table 1). By integrating chromosome conformation capture data (Hi-C, ~150×), ~97.23% (8.43 Gb) of the contigs were assigned to 14 pseudomolecules, with the remainder combined into an unanchored pseudochromosome (~246 Mbp) (Table 1; Supplemental Table 5). We assigned the D^v and N^v subgenomes using SubPhaser (Jia et al., 2022), a robust allopolyploid subgenome phasing method based on subgenome-specific k-mers (Figure 1). The D^v and N^v subgenomes of RM271 were 3.91 and 4.52 Gb, respectively. After we restored three inter-subgenome translocations to their pre-translocation positions (see below), we observed 4.15 and 4.28 Gb of D^v and N^v subgenomes, respectively.

The base-level accuracy consensus quality value and k-mer completeness scores evaluated with Merqury (Rhie et al., 2020) were 63.45% and 97.56%, respectively (Supplemental Figure 2A and 2B). The RM271 genome showed CEGMA (Core Eukaryotic Gene Mapping Approach; Parra et al., 2007) and BUSCO (Benchmarking Universal Single-Copy Orthologs; Simão et al., 2015) completeness levels of 97.18% and 99.18%, respectively, indicating high gene completeness, and its polyploid nature was indicated by 91.97% duplicated core genes (Supplemental Figure 2C). The distribution of subtelomeric and centromeric (Figure 1) homologous sequences from *Ae. tauschii* clones (Zhang et al., 2004) and *T. monococcum* subsp. *aegilopoides* clones (Liu et al., 2008) also revealed the presence of these complex regions in the RM271 genome (Supplemental Figure 3). The long terminal repeat (LTR) Assembly Index (LAI) (Ou et al., 2018) of the RM271 genome assembly ranged from 11.96 to 14.97 for individual chromosomes, and the LAI values were 14.52 and 13.59 for the N^v subgenome and the D^v subgenome, respectively, higher than those for the common wheat cultivar Chinese Spring (International Wheat Genome Sequencing Consortium [IWGSC] et al., 2018) and comparable to those for the synthetic hexaploid wheat-derived cultivar Chuanmai 104 (Liu et al., 2024). We thus assembled a comprehensive and high-quality reference genome for *Ae. ventricosa*.

Using transposable elements (TEs) from Repbase, ClariTeRep, and the TRansposable Elements Platform (TREP), we annotated 83.52% of the RM271 genome as repetitive sequences, including

Genomic feature	Value
Total assembly size	8.67 Gb
Longest contig	255.18 Mb
No. of contigs	5854
N50 contig length	54.71 Mb
L50 contig count	49
Anchored size	8.43 Gb
Anchored size (D^v + N^v)	3.91 Gb + 4.52 Gb
No. of anchored contigs	280
No. of anchored contigs (D^v + N^v)	96 + 186
Repeats (D^v + N^v), %	84.77 + 84.67
Genome completeness, %	99.10
No. of HC genes (total)	69 000
No. of HC genes (D^v + N^v)	30 982 + 36 395

Table 1. Statistics of the RM271 genome assembly and annotation

5.27% tandem repeats, 67.69% retrotransposons, and 14.33% DNA transposons, comparable to numbers in other Triticeae species. In total, 69 000 high-confidence (HC) genes were annotated in the RM271 genome (Table 1), a number close to that reported for wild emmer wheat (65 012 HC genes) and slightly higher than that reported for durum wheat (66 559 HC genes). Among these HC genes, we identified 3787 potential resistance (R) genes, including 332 receptor-like protein genes, 95 coiled-coil (CC) nucleotide-binding region leucine-rich repeat genes, and 516 receptor-like kinase genes (Supplemental Figure 4; Supplemental Table 12). These gene sets (Supplemental Figure 4; Supplemental Tables 11–17) will provide references for improved utilization of *Ae. ventricosa* as a genetic resource, especially for the cloning of resistance genes.

Phylogenomics of the D^v and N^v subgenomes

Phylogenomic analyses of the D^v and N^v subgenomes of *Ae. ventricosa* and the genomes of 16 other Triticeae species (including 29 genomes or subgenomes) revealed that the N^v subgenome of *Ae. ventricosa* was closely related to the *Ae. uniaristata* genome (2n = 2x = 14, genome NN), as revealed by both phylogenomics and sequence similarity (Figure 2A and 2B; Supplemental Table 18). These results were consistent with previous cytogenetic and phylogenetic evidence showing that the speciation of *Ae. ventricosa* resulted from hybridization between *Ae. tauschii* (genome DD) and *Ae. uniaristata* (genome NN) (Lilienfeld, 1951; Chen et al., 2020). Phylogenomic results showed that the genome of *Ae. tauschii* ssp. *strangulata* was clustered with the D subgenome of bread wheat, which confirmed that this genome was the progenitor of the bread wheat D subgenome (Zhou et al., 2021); however, the D^v subgenome of *Ae. ventricosa* was clustered with the genome of *Ae. tauschii* ssp. *tauschii*, forming an alternate clade. Previous population genomics results indicated that *Ae. tauschii* consists of two subspecies (Wang et al., 2013), *Ae. tauschii* ssp. *strangulata* and *Ae. tauschii* ssp. *tauschii*. Sequence similarity analyses showed that the *Ae. tauschii* ssp. *tauschii* genome shared the highest identity (~98.67%; Supplemental Table 19) with the D^v subgenome of *Ae. ventricosa* (Figure 2B and Supplemental Figure 5). These results suggest that *Ae. tauschii* ssp. *tauschii* was the progenitor of the D^v subgenome of *Ae. ventricosa,* which could be an important resource for enriching the genetic diversity of the D subgenome in wheat improvement.

To determine the hybridization time of *Ae. ventricosa*, we examined Ks (synonymous substitution rate) distributions of orthologous gene pairs between the D^v and N^v subgenomes and their potential diploid progenitor species, and the results indicated that the oldest polyploidization time of the D^v and N^v subgenomes was from 0.695 to 0.756 mya (Figure 2C). Molecular dating yielded a similar divergence time (0.83 mya) of the *Ae. ventricosa* D^v subgenome and the *Ae. tauschii* ssp. *tauschii* genome (Figure 2A). Therefore, the hybridization time of *Ae. ventricosa* is likely to have been earlier than the time of hybridization between *Ae. tauschii* ssp. *strangulata* and the tetraploid emmer wheat *T. turgidum* (AABB) (~10 000 years ago) (Salamini et al., 2002). However, only three samples from *Ae. tauschii* ssp. *tauschii* were included, and these samples are not likely to represent plants from the actual ancestor populations that gave rise to the D^v subgenome in *Ae. ventricosa*. Therefore, the molecular dating-based estimation between *Ae. tauschii* ssp. *tauschii* and *Ae. ventricosa* is likely to reflect population divergence events before the actual polyploidization time. Further analyses using populations of *Ae. tauschii* ssp. *tauschii* and *Ae. ventricosa* may further improve the accuracy of the polyploidization time estimates.

Evolutionary history of subgenomes

HC genes from 18 species (including 32 genomes or subgenomes) were clustered into gene families using OrthoFinder (Emms and Kelly, 2019). We identified 1684 *Ae. ventricosa*-specific HC genes, which were significantly enriched in Gene Ontology (GO) terms including nucleic acid binding (GO: 0003676, Fisher's exact test, corrected p = 0), zinc ion binding (GO: 0008270, p = 8.92E−11), 3′-5′ exonuclease activity (GO: 0008408, p = 5.69E−9), protein binding (GO: 0005515, p = 1.84E−8), serine-type endopeptidase inhibitor activity (GO: 0004867, p = 1.86E−4), nucleobase-containing compound metabolic process

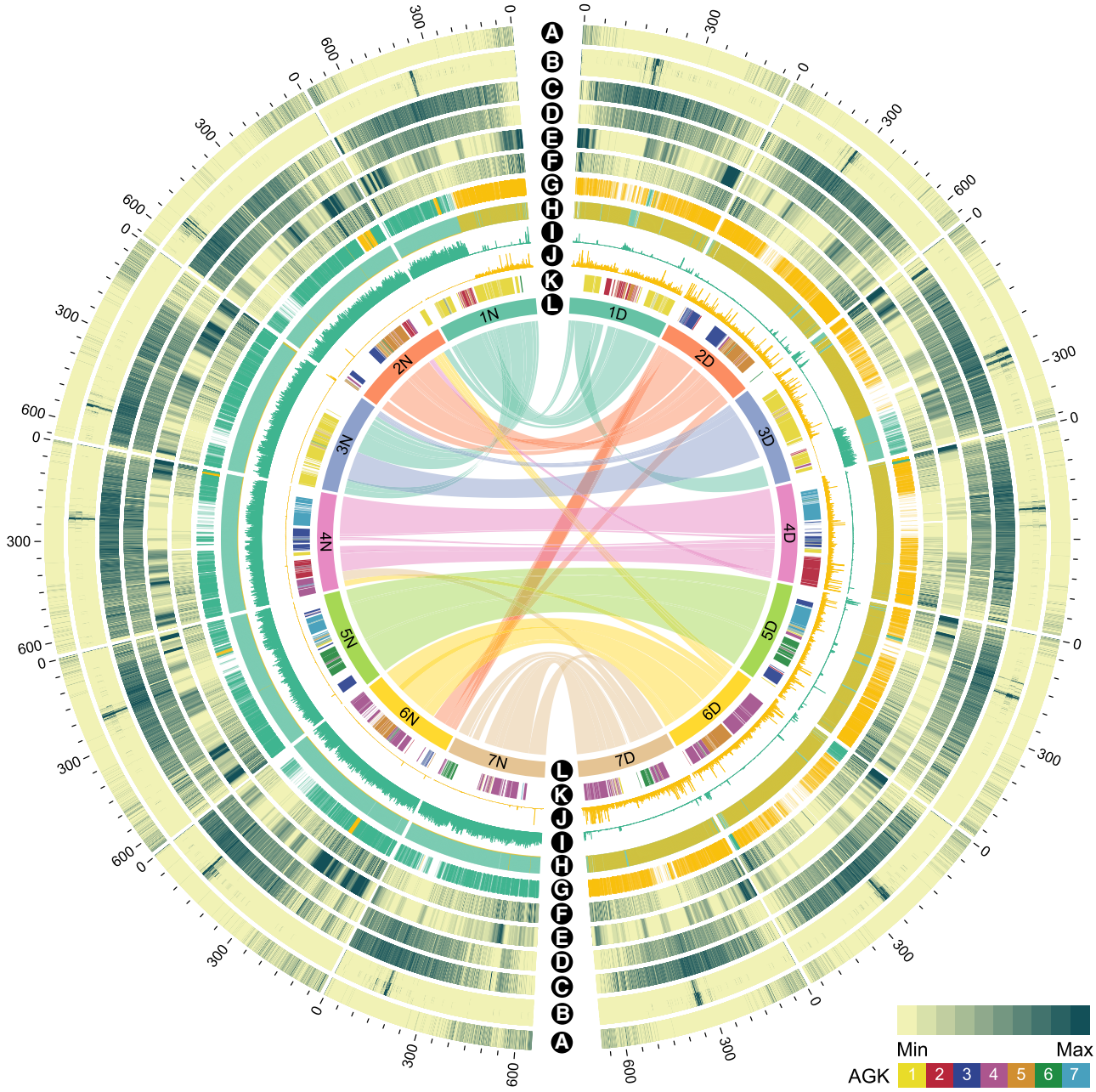

Figure 1. Genomic features of RM271.

(A) Distribution of the *Ae. tauschii* clone A6-10 subtelomeric tandem repeat sequence (AY249980.1).

(B) Distribution of the *Ae. tauschii* clone 6C6-3 (AY249981.1) and 6C6-4 (AY249982.1) and *T. monococcum subsp. aegilopoides* clone BAC TbBAC5 (DQ904440.1) and TbBAC30 (EF624064.1) centromere-specific tandem repeat sequences.

(C–F) Distribution of the tandem repeat density **(C)**, LTR density **(D)**, non-coding RNA gene (microRNA, tRNA, small nucleolar RNA, small nuclear ribonucleoprotein particle, and rRNA) density **(E)**, and HC protein-coding gene density **(F)**.

(G) Assignment of individual HC protein-coding genes to the D^v (gold) and N^v (medium aquamarine) subgenomes.

(H) Significant enrichment of subgenome-specific k-mers identified by SubPhaser (gold for D^v, medium aquamarine for N^v, blank for non-enriched 1-Mbp windows).

(I) Density distribution of the N^v subgenome-specific k-mer set identified by SubPhaser.

(J) Density distribution of the D^v subgenome-specific k-mer set identified by SubPhaser.

(K) RM271 chromosome evolution according to the AGK model.

(L) Chromosome nomenclature; links between chromosomes are collinear blocks identified with MCScanX (Wang et al., 2012), which are colored corresponding to the seven chromosomes of the D^v subgenome.

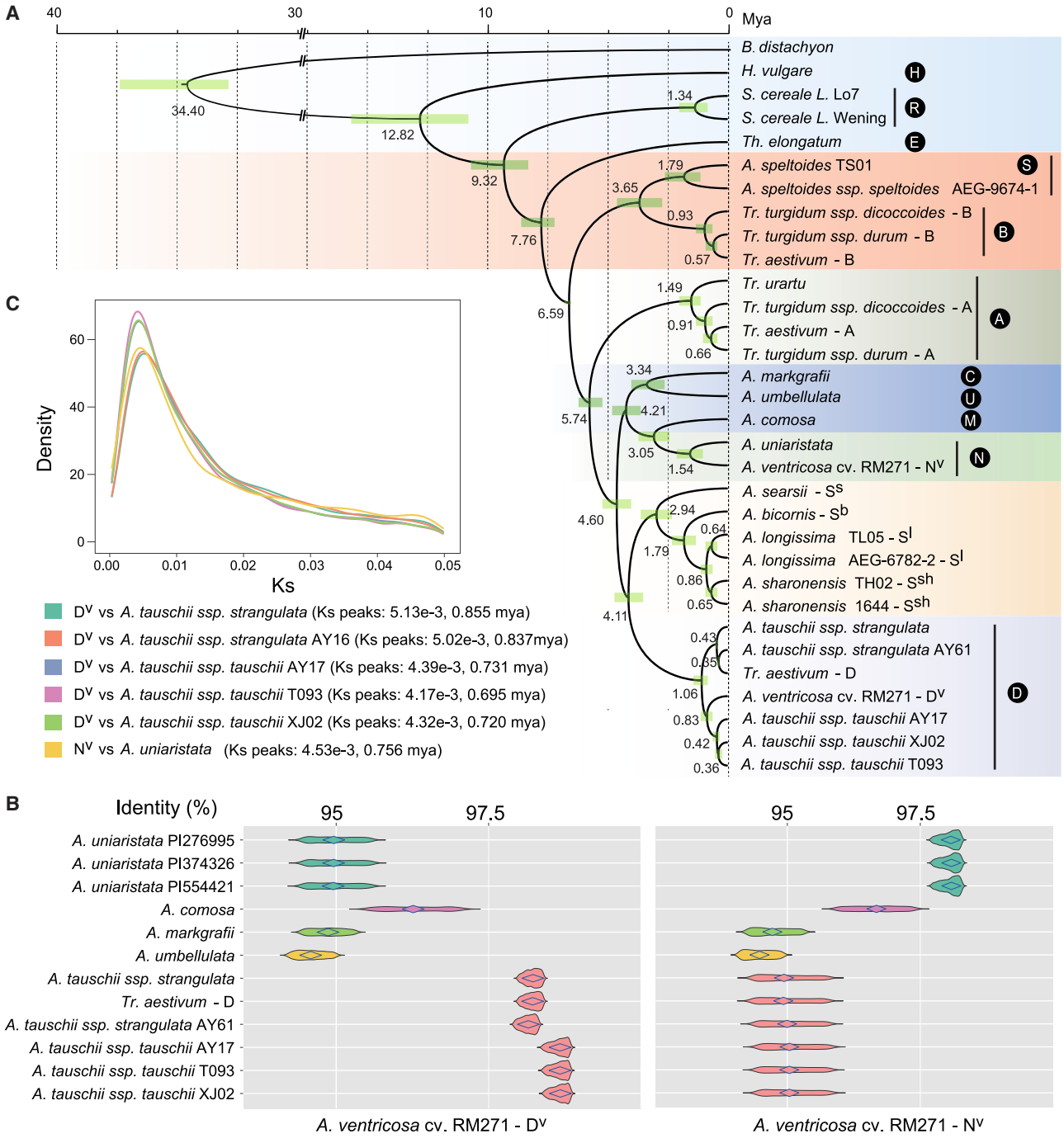

Figure 2. Phylogenomics of the allotetraploid wild grass *Ae. ventricosa*.

(A) Maximum likelihood tree and divergence times of Triticeae species based on single-copy orthologous genes.

(B) Sequence similarities of WGS short reads or genome assemblies from diploid species and polyploid subgenomes that were uniquely mapped to the D^v and N^v subgenomes.

(C) Synonymous substitution rate (Ks) distributions of orthologous gene pairs between the D^v and N^v subgenomes and their potential ancestral diploid species.

(GO: 0006139, p = 1.81E−15), response to wounding (GO: 0009611, p = 1.49E−11), and regulation of DNA-templated transcription (GO: 0006355, p = 0.021981).

The subgenomes of polyploids commonly evolve bias after polyploidization (Zhang et al., 2015; International Wheat Genome Sequencing Consortium [IWGSC] et al., 2018; Ling et al., 2018; Peng et al., 2022). The N^v subgenome (Ka = 0.0022) had a nonsynonymous substitution rate (Ka) similar to that of the D^v subgenome (Ka = 0.0024). However, Ks of the N^v subgenome (Ks = 0.0108) was elevated compared with that of the D^v subgenome (Ks = 0.0055), indicating a higher neutral mutation

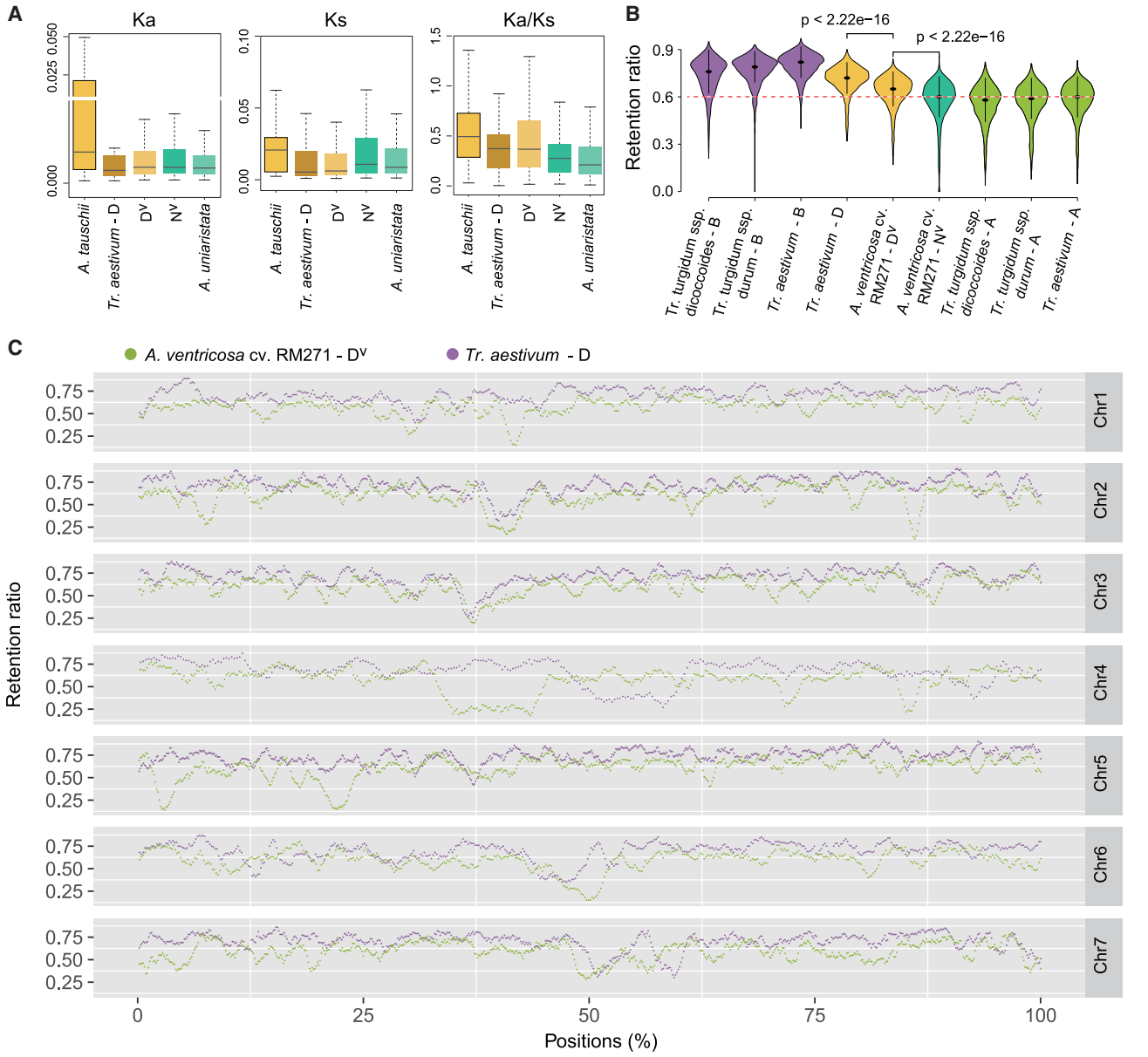

Figure 3. Evolutionary history of the genome of the allotetraploid wild grass *Ae. ventricosa*.

(A) Ks, Ka, and Ka/Ks values of single-copy gene families from Figure 2A.

(B and C) Gene retention patterns of the D^v and N^v subgenomes and their close relatives. A sliding window approach with a window size of 100 syntenic genes and a step size of 10 syntenic genes is used to show the percentages of retained genes in polyploid subgenomes using *T. urartu* (for A subgenomes), *Ae. speltoides* ssp. *speltoides* (for B subgenomes), or *Ae. tauschii* (for D and N^v subgenomes). The retention rates of various subgenomes were compared (Wilcoxon rank-sum test) and plotted as violin plots; the red dashed line indicates the median retention rate between N^v and the *Ae. tauschii* genome **(B)**. The distribution of retention rates of the bread wheat D genome and D^v subgenome along the *Ae. tauschii* chromosomes; chromosome sizes are normalized to 100% **(C)**.

rate in the N^v subgenome (Figure 3A). Interestingly, both the nonsynonymous substitution rate (Ka) and Ks values in the D^v and N^v subgenomes of RM271 were higher than those of the D subgenome of Chinese Spring (Ka = 0.0019), implying faster evolutionary rates in allotetraploid *Ae. ventricosa* (Figure 3A). The D^v subgenome retained more ancestral genes than the N^v subgenome (Figure 3B), which is consistent with the D^v subgenome experiencing stronger selection pressure (Figure 3A). However, when the diploid genome of *Ae. tauschii* was used as a reference, the D^v subgenome of RM271 had a lower proportion of retained genes than the D subgenome of bread wheat among various chromosomes (Figure 3C), implying that ploidy level may affect the subgenome balance.

Structural variations among subgenomes

Whole-genome alignments of *Ae. tauschii* ssp. *tauschii* and RM271, as well as synteny analyses using protein-coding genes, revealed extensive chromosomal rearrangements and structural variations (Figure 4A–4C). The most significant rearrangement

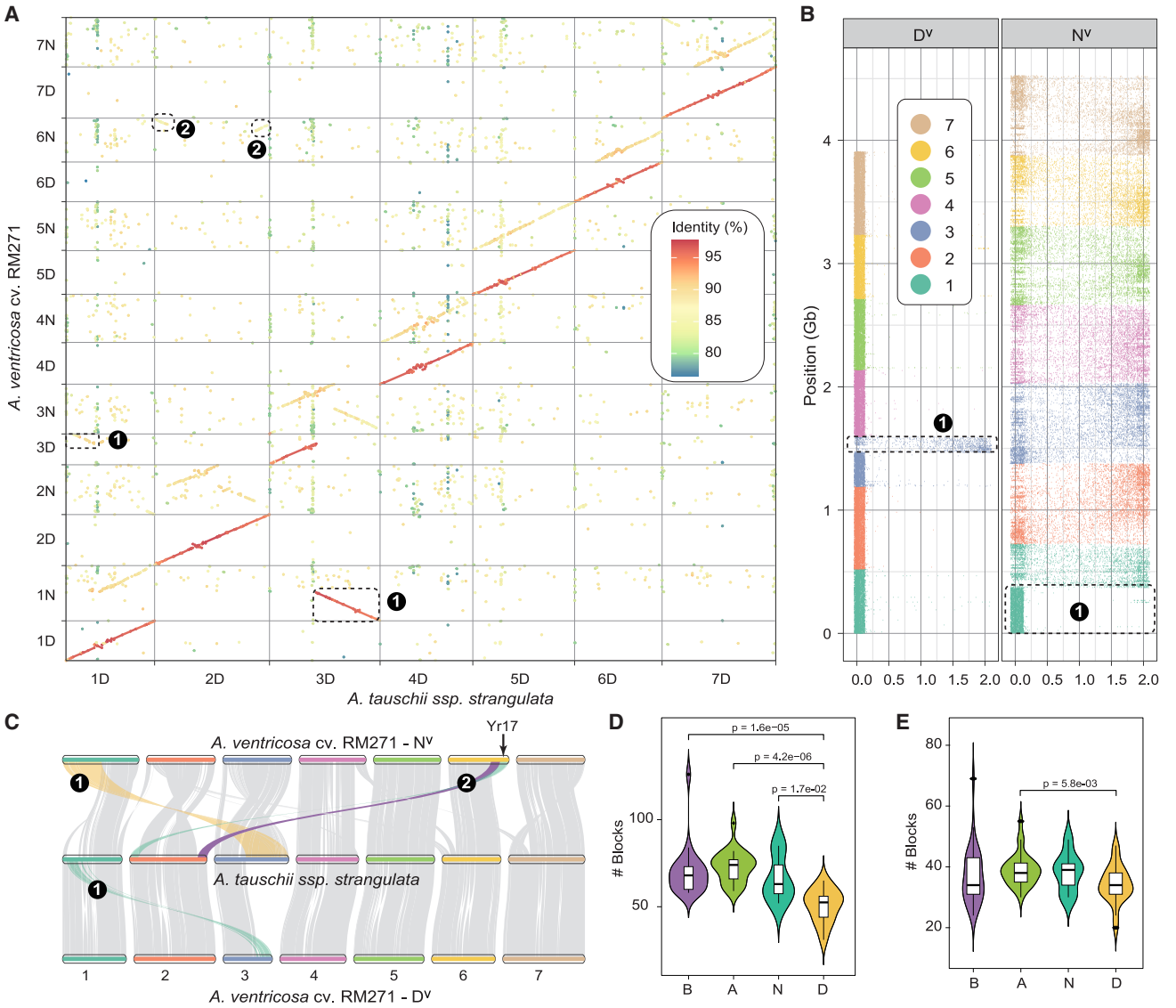

Figure 4. Structural evolution of the RM271 chromosomes.

(A) Whole-genome nucleotide alignments between RM271 and *Ae. tauschii* ssp. *strangulata* produced using mashmap (Jain et al., 2018). Two translocations are illustrated by numbered pairs of dashed-line boxes. The first mark shows the largest translocation event identified, and the second mark includes the *Yr17* gene.

(B) The largest 1N^{v}S-3D^{v}L translocation event is supported by the mapping of NGS reads from the N-genome diploid species *Ae. uniaristata* (PI276995) to the RM271 genome.

(C) Collinear blocks between the RM271 subgenomes and the genome of *Ae. tauschii* ssp. *strangulata* identified using MCScanX (Wang et al., 2012).

(D and E) The numbers of syntenic blocks between diploid species/polyploid subgenomes and ATK **(D)** and AGK **(E)** were compared (Wilcoxon rank-sum test) and plotted as violin plots.

was the chromosome arm exchange between the D^{v} and N^{v} subgenomes (1N^{v}S-3D^{v}L translocation), which generated translocated 1N^{v} and 3D^{v} chromosomes and was supported by the mapping of next-generation sequencing (NGS) reads from the N-genome diploid species *A. uniaristata* (PI276995) to the RM271 genome (Figure 4B). Two moderately sized inter-chromosomal translocations were also observed between the 2N^{v} and 6N^{v} chromosomes, namely 2N^{v}S-6N^{v}L and 2N^{v}L-6N^{v}L translocations (Figure 4A and 4C). A similar 2N^{v}S-6N^{v}L translocation in *Ae. ventricosa* was identified by cytogenetic evidence (Tanguy et al., 2005), and further analyses identified an LTR/Copia element spanning the breakpoint (Supplemental Figure 6). Returning these three rearrangements to their pre-translocation positions, we observed significantly improved synteny (Supplemental Figure 7) compared with the translocated version (Figure 4A). Using the number of conserved blocks of the ancestral Triticeae karyotype (ATK) constructed by DESCHRAMBLER (Kim et al., 2017; Wu et al., 2022; Peng et al., 2023) (see methods) (Figure 4D) and the ancestral grass karyotype (AGK) (Figure 4E) as proxies for chromosomal conservation, we found that the N^{v} subgenome contained more chromosomal rearrangements than the D genome/subgenome in the Triticeae (Figure 4D and 4E and Supplemental Figure 8).

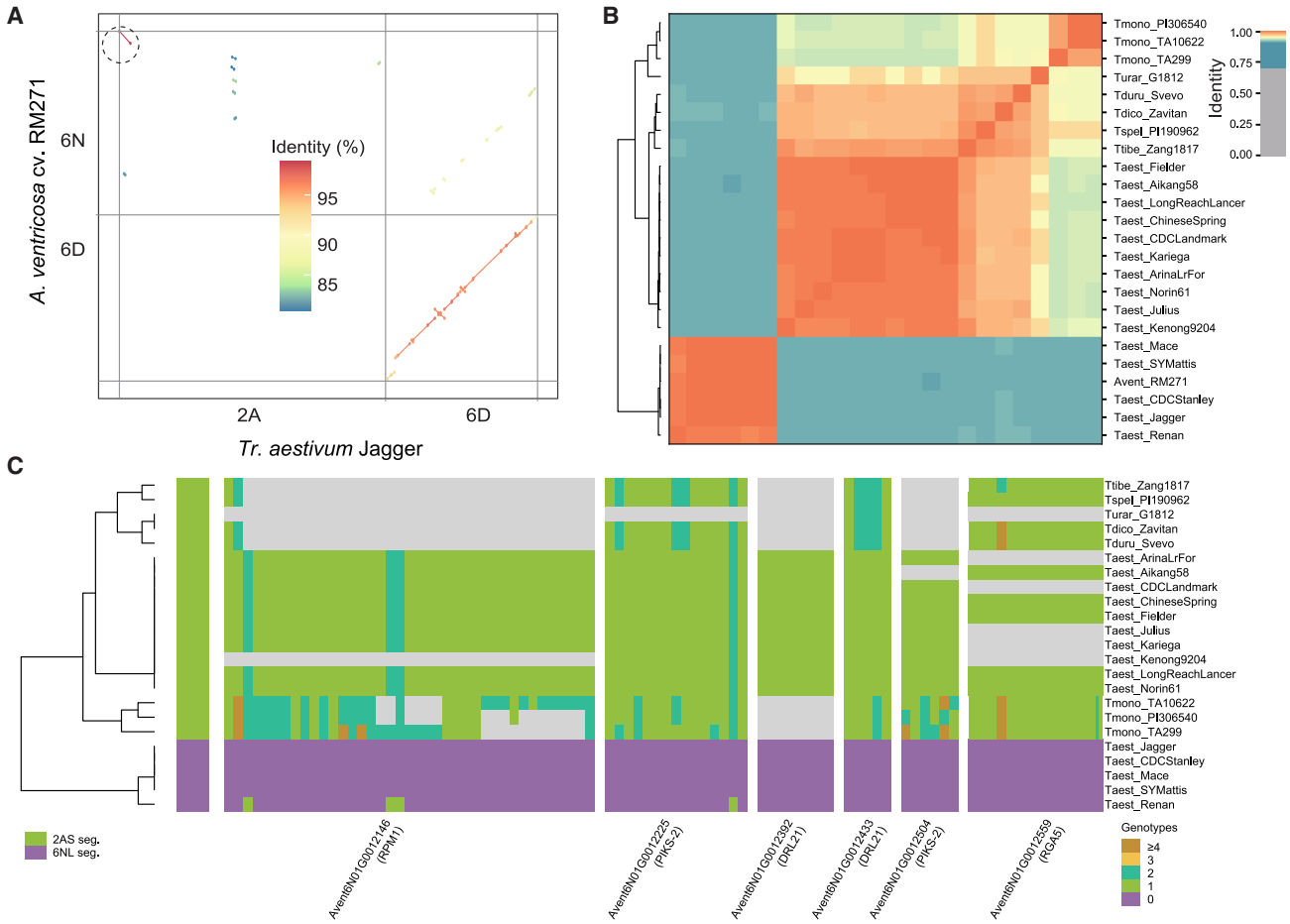

Figure 5. 6N^{v}L segment-assisted identification of disease-resistance genes.

(A) Whole-genome alignments between RM271 and the wheat cultivar Jagger revealed that the 6N^{v}L segment is the same as the previously well-known 2N^{v}S segment in *Ae. ventricosa*.

(B) ANIb analysis of the 6N^{v}L segment in RM271 and its homologous 2AS segments in other published reference genomes. Color scheme: turquoise for identity at 0.9 or below, red for values at 1.0; anything below 0.7 is gray.

(C) Genotypes for the variants in the six R genes located in the 6N^{v}L segment (seg.). The variants were called from the 6N^{v}L/2AS pangenome using RM271 as the reference. The "0" genotype is the reference allele, and the other numbers are the alternative alleles. Gray indicates that no genotyping data were available.

Identification of introgression segments based on genome assemblies

With reference to the high-quality genome of RM271, we identified *Ae. ventricosa*-sourced introgression segments in 17 bread wheat varieties (AK58, ArinaLrFor, Chinese Spring, Fielder, Jagger, Julius, Kariega, KN9204, LongReach Lancer, CDC Landmark, Mace, SY Mattis, Norin 61, Renan, CDC Stanley, PI190962, and Zang1817) (Figure 5A and Supplemental Figure 9; Supplemental Table 20). A total of 10 introgressions were identified, 7 of which have been detected previously (Walkowiak et al., 2020; Aury et al., 2022), confirming the reliability of our identification method. The three novel introgressions included segments derived from the chromosome 7D^{v}L of *Ae. ventricosa* on chromosomes 7DL (1.90 Mb) and 2BS (1.95 Mb) of Jagger and chromosome 7DL (5.51 Mb) of SY Mattis (Supplemental Table 20).

Among these introgressions, the most prevalent is the 2AS segment at the beginning of 2A in the wheat genomes, a segment that is homologous to the 6N^{v}L segment at the terminal of 6N^{v} in RM271. Previous comparisons of chromosomes from reference-quality assemblies confirmed that Jagger, Mace, SY Mattis, CDC Stanley, and Renan carry the 2N^{v}S introgression, which spans ~33 Mb on chromosome 2A (Walkowiak et al., 2020; Aury et al., 2022). We found that the 2N^{v}S segment was present as a preexisting translocation in *Ae. ventricosa* (2N^{v}S-6N^{v}L translocation), as demonstrated by genomic (Figure 4A and 4C) and cytogenetic evidence (Tanguy et al., 2005). Therefore, the 6N^{v}L segment translocated from 2N^{v}S in RM271 was exactly the same as the previously well-known 2N^{v}S segment used for wheat improvement. We next mapped NGS reads from the *Yr17* isogenic line of the bread wheat variety Avocet, which contains the 2N^{v}S segment, to the concatenated genome assemblies from RM271 and Chinese Spring, and the results showed the presence of sequencing coverage at the 6N^{v}L segment at the terminal of 6N^{v} in RM271 (Supplemental Figure 10).

The 6N^{v}L segment has a length of 37.5 Mb ranging from 537.2 to 574.7 Mb on the 6N^{v}L chromosome of RM271 (hereafter, we refer

to the 6N^{v}L segment as the homolog of the well-known 2N^{v}S segment). This segment covered the physical interval of the 6N^{v}L-2AS translocated segments in CDC Stanley (33.49 Mb) and Jagger (32.53 Mb) (Gao et al., 2021). When compared with its counterparts in the wheat cultivars, the 6N^{v}L segment in RM271 is longer than those in the wheat genome assemblies, and the wheat counterparts also have many rearrangements (Supplemental Figure 11). The 6N^{v}L segment in RM271 is also longer than its homologous 2AS segment in Chinese Spring (29.06 Mb; Supplemental Figure 12). Within the 6N^{v}L segment in RM271, 146 of the 649 HC genes (22.5%) were predicted to be disease-resistance genes or genes associated with disease resistance (cytochrome P450 [CYP450], ABC transporter, or WRKY genes), whereas 130 HC genes were predicted to be related to disease resistance in the homologous 2AS segment of Chinese Spring (Supplemental Figure 12; Supplemental Tables 21 and 22). Among the disease-resistance-related genes in this segment, we found that Avent6N01G0012810.1 (LRR receptor-like serine/threonine-protein kinase RGI1) was specific to *Ae. ventricosa*. In summary, the high-quality reference genome of RM271 can be used to provide a more detailed analysis of the genes introduced into bread wheat by the introgression of the *Ae. ventricosa* 6N^{v}L segment.

To further explore the 6N^{v}L segment, the average nucleotide identity (ANI) metric was used to clarify genomic relationships between 6N^{v}L and 2AS segments (Figure 5B). 6N^{v}L and 2AS segments were further assembled into a graph-based pangenome using pggb (Garrison et al., 2023) (Supplemental Figure 13). According to the pyani (Pritchard et al., 2016) and pggb results, our 24 6N^{v}L/2AS segments were grouped into two groups, exactly corresponding to six 6N^{v}L segments and 18 2AS segments (Figure 5B and Supplemental Figure 13). The graph-based pangenome contained 24 paths corresponding to 24 segments and 5 835 512 nodes with a total length of 143 644 349 bp, which is much longer than the size of the 6N^{v}L segments (37.5 Mb for RM271) and the 2AS segments (29.06 Mb for Chinese Spring), indicating the high divergence of these two groups. Using variants called with pggb and the 6N^{v}L segment in RM271 as the reference, we identified six candidate R genes with different genotypes in the 6N^{v}L segment and the 2AS segment groups (Figure 5C; Supplemental Tables 23 and 24).

All six candidate genes contained the NBS domain (Supplemental Table 24). Among these genes, two (*Avent6N01G0012392* and *Avent6N01G0012433*) were homologous to an *Arabidopsis thaliana* LRR protein (*At3g14460*; *AtLRRAC1*), which encodes an adenylyl cyclase that has a role in the cyclic AMP-dependent pathway in the defense against biotrophic and hemibiotrophic plant pathogens (Bianchet et al., 2019); *Avent6N01G0012146* was homologous to the *RPM1* gene in *A. thaliana*, which encodes an NBS-LRR protein that confers resistance to *Pseudomonas syringae* (Grant et al., 1995; Boyes et al., 1998; Rose et al., 2012; Lee et al., 2015); *Avent6N01G0012559* was homologous to the *RGA5* gene in rice, whose protein recognizes the AVR-Pia effector of the rice blast fungal pathogen *Magnaporthe oryzae* through direct interaction (Ortiz et al., 2017; Varden et al., 2019; Anbu et al., 2023; Cadiou et al., 2023); and two genes (*Avent6N01G0012225* and *Avent6N01G0012504*) were homologous to the *Piks-2* gene in rice, whose protein recognizes the *M. oryzae* effector AVR-Pik (Li et al., 2019; Varden et al., 2019; Anbu et al., 2023; Cadiou et al., 2023; Xiao et al., 2023). When grain yield and resistance to yellow rust were evaluated in a large winter wheat collection, 142 of 533 significantly associated k-mers from whole-genome sequencing (WGS) were absent from the wheat pangenome assemblies (Walkowiak et al., 2020; Schulthess et al., 2022). We mapped these 142 k-mers to our RM271 assemblies, and two k-mers were found close to the *Avent6N01G0012225* loci, suggesting potential *Yr17* candidates.

Identification of introgression segments based on whole-genome resequencing

To further analyze the contribution of *Ae. ventricosa* to existing wheat resources, we extended the above method (Supplemental Figure 10) to a large-scale population (Schulthess et al., 2022) by aligning WGS data between concatenated genome assemblies from RM271 and bread wheat varieties (see the example in Supplemental Figure 14), including 664 worldwide wheat accessions. We obtained accurate physical location and sequencing information for 12 introgressed segments on the 1D^{v}, 2D^{v}, 3D^{v}, 5D^{v}, 7D^{v}, 1N^{v}, and 6N^{v} chromosomes (Supplemental Table 25). Importantly, in addition to the 6N^{v}L introgression segments, many novel introgression segments derived from the N^{v} and D^{v} subgenomes of *Ae. ventricosa* were found (Supplemental Table 25), and a large proportion of wheat accessions contained introgressed segments from *Ae. ventricosa* (Supplemental Tables 26 and 27), including approximately 29.40% (167/568) of the European wheat varieties tested (Figure 6A). Among these European wheat varieties, 26.06% (148/568) contained the well-known 6N^{v}L-2AS translocation (Supplemental Figure 15). Most of the varieties with introgressions were cultivated in the last 20 years (2001–present) (Figure 6B and 6C). These results suggested that *Ae. ventricosa* genetic resources have been involved in the wheat varieties. However, a larger panel of WGS datasets from bread wheat varieties would deepen our understanding of the role of *Ae. ventricosa* in wheat improvement.

DISCUSSION

We assembled a comprehensive and high-quality reference genome of *Ae. ventricosa* RM271, with a total assembly size of 8.67 Gb and an N50 contig length of 54.71 Mb (Figure 1 and Supplemental Figures 1 and 2; Table 1). Contig N50s from previous diploid *Ae. tauschii* (genome DD) assemblies based on short-read sequencing or PacBio continuous long-read sequencing ranged from 50.3 kb to 2.2 Mbp (Zhao et al., 2017; Zhou et al., 2021; Gaurav et al., 2022); thus, the RM271 assembly represents a ~24–1000-fold increase in contiguity. The distribution of subtelomeric and centromeric sequences along the chromosomes also revealed the presence of such complex regions in the RM271 genome (Figure 1 and Supplemental Figure 3). The RM271 D^{v} subgenome (3.91 Gb) was smaller than the diploid genome of *Ae. tauschii* (~4.3 Gb) and the D subgenome (~4.82 Gb) of Chinese Spring wheat (genome AABBDD) owing to the chromosome arm exchange between the D^{v} and N^{v} subgenomes, which generated chimeric 1N^{v} and 3D^{v} chromosomes (Figure 4A–4C).

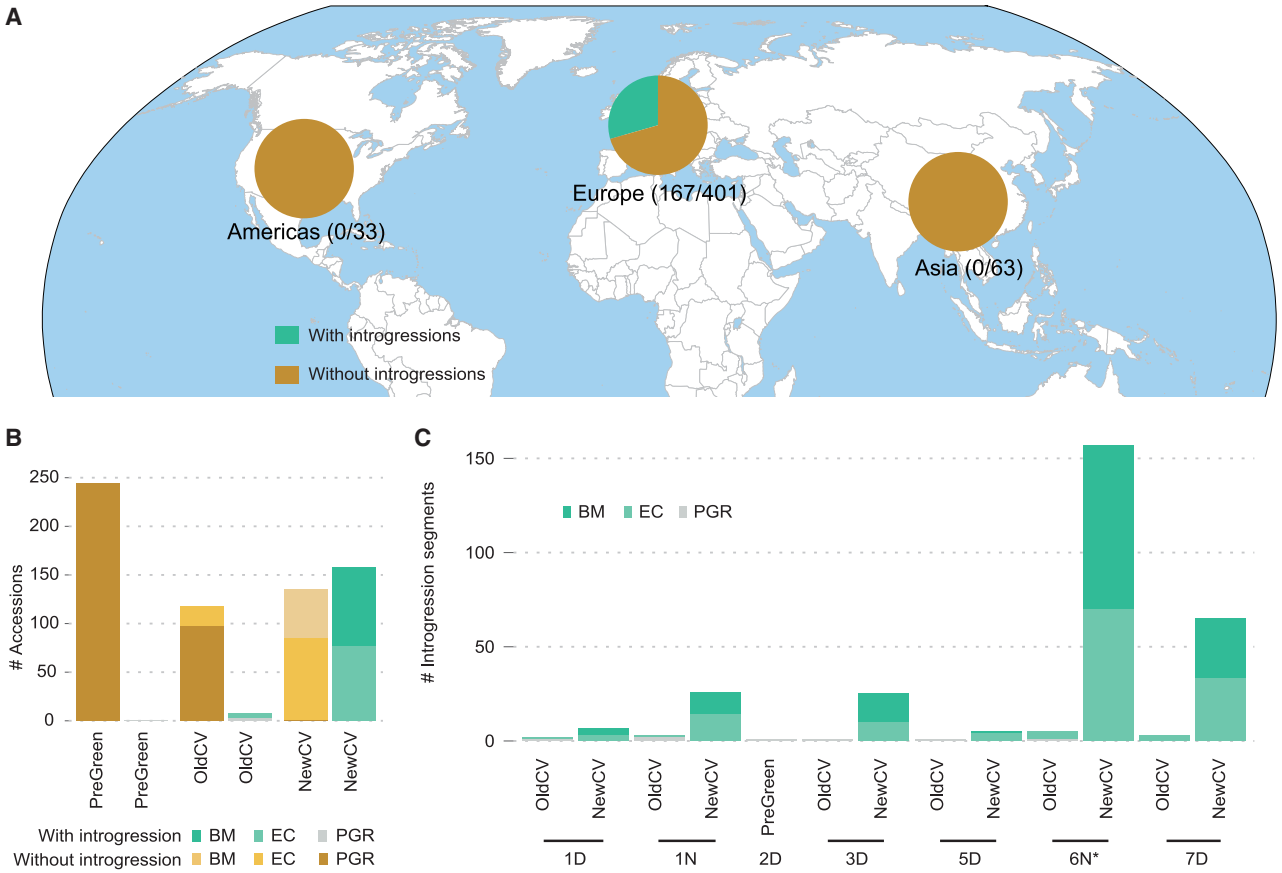

Figure 6. Introgression breeding history identified using WGS datasets.

(A–C) The numbers of accessions with putative introgressions per geographic region **(A)**, introgressions per period per sample source **(B)**, and introgression segments per period per sample source in each chromosome **(C)** were plotted. The three wheat sample sources were elite cultivars (EC), breeding materials (BM), and plant genetic resources (PGR). The three different historic periods were PreGreen (before 1970), OldCV (1971–2000), and NewCV (2001 to present). The asterisk indicates that this 6N^{v} segment is exactly the same as the 2N^{v}S-6N^{v}L translocation segment in Figure 4C and the previously well-known 2N^{v}S segment that contains the *Yr17-Lr37-Sr38-Cre5* gene cluster.

Phylogenomic results suggested that *Ae. tauschii* ssp. *tauschii* was the progenitor of the D^{v} subgenome of *Ae. ventricosa* (Figure 2A and 2B). Previous population genomics results indicated that *Ae. tauschii* consists of two subspecies (Wang et al., 2013), *Ae. tauschii* ssp. *strangulata* and *Ae. tauschii* ssp. *tauschii*, and that *Ae. tauschii* ssp. *strangulata* was the progenitor of the bread wheat D subgenome (Zhou et al., 2021). It has been shown that *Ae. ventricosa*, as a stable and effective resistance source, can be directly hybridized with *T. turgidum* to transfer valuable alien segments from *Ae. ventricosa* to wheat (Doussinault et al., 1983; Zhang et al., 2020). It can also be used to recruit the genetic diversity of the bread wheat D subgenome; therefore, the genome of RM271 could further accelerate the use of *Ae. ventricosa* in wheat improvement breeding.

The oldest polyploidization time of *Ae. ventricosa* occurred from 0.695 to 0.756 mya (Figure 2C); therefore, it was likely earlier than the time of hybridization between *Ae. tauschii* ssp. *strangulata* and the tetraploid emmer wheat *T. turgidum* (AABB) (~10 000 years ago) (Salamini et al., 2002). Molecular dating also yielded a similar divergence time (0.83 mya) of the *Ae. ventricosa* D^{v} subgenome and the *Ae. tauschii* ssp. *tauschii* genome (Figure 2A). However, it is important to note that this study does not include all populations of diploid species with the DD genome. As a result, our estimated times are likely to reflect population divergence events closely related to or before the actual polyploidization event, and additional analyses of more samples from broader populations are needed to improve the accuracy of the polyploidization time estimates.

Analyses comparing the assemblies of *Ae. ventricosa* revealed extensive chromosomal rearrangements and structural variations and showed that the N^{v} subgenome has experienced more chromosomal rearrangements than the D^{v} subgenome (Figure 4A–4C). Among these rearrangements, the 2N^{v}S-6N^{v}L translocation may be the most important for wheat breeders and geneticists (Figure 4A and 4C), as our analyses (Figure 5A and Supplemental Figure 9) revealed that this 6N^{v}L segment is exactly the same as the previously well-known 2N^{v}S segment used for wheat improvement.

The 2N^{v}S-2AS translocation wheat line VPM1 was developed in the 1980s, representing the first application of *Ae. ventricosa* genetic resources to wheat breeding, and VPM1 has since been used in wheat improvement. Jagger, Renan, CDC Stanley, and

a further series of well-known wheat varieties were cultivated successively. *Yr17*, *Lr37*, *Sr38*, *Cre5*, and other resistance genes from the 6N^{v}L segment of *Ae. ventricosa* were then successfully mapped and used in the development of multiple-resistance wheat strains worldwide. Here, we found that the well-known 2N^{v}S segment used for wheat improvement was translocated to the region from 537.2 to 574.7 Mb (a total length of 37.5 Mb [6N^{v}L]) on the 6N^{v}L chromosome of RM271 (Figure 4A and 4B and Supplemental Figure 6; Supplemental Table 20).

In recent years, *Ae. ventricosa* genome-derived introgressed segments have been detected in the genome assemblies of several important wheat varieties (Walkowiak et al., 2020; Aury et al., 2022). However, because of the lack of an *Ae. ventricosa* reference genome, the homologs of these introgressed segments in *Ae. ventricosa* with ancestral states remain unknown. The high-quality genome assembly of RM271 reported here enabled us to obtain information on introgressed segments from *Ae. ventricosa* present in bread wheat varieties worldwide and to assess the contributions of key genes or chromosome segments from *Ae. ventricosa* genetic resources to wheat improvement in terms of multiple agronomic and resistance traits. Importantly, in addition to the 6N^{v}L introgression segment, many novel introgression segments derived from the N^{v} and D^{v} subgenomes of *Ae. ventricosa* were found (Supplemental Tables 20 and 25), and 29.40% of European wheat varieties inherited at least one of these segments. It is worth noting that the datasets used here were enriched for European wheat varieties and included only a limited number of wheat varieties from outside Europe. Therefore, whether wheat varieties from other continental regions (such as Asia and the Americas) have inherited introgression segments derived from the N^{v} and D^{v} subgenomes of *Ae. ventricosa* will need to be determined using larger panels.

At present, the rapid change in the global climate is having many adverse effects on wheat production. Fortunately, utilization of wild germplasm resources will effectively broaden the genetic basis of modern wheat and provide more favorable and readily available excellent genes to facilitate wheat improvement. As an important *Aegilops* species, *Ae. ventricosa* has been crossed with tetraploid wheat to develop a bridge material (Doussinault et al., 1983; Zhang et al., 2020) with which to breed novel wheat accessions such as VPM1 and other 6N^{v}L segment derivatives. Much evidence has demonstrated that the utilization of key chromosome segments from *Ae. ventricosa* has made important contributions to the improvement of bread wheat worldwide. However, the exploration and utilization of *Ae. ventricosa* are still very limited owing to the lack of a reference genome. To fill these gaps, we developed a high-quality reference genome of *Ae. ventricosa*, found that the D^{v}-subgenome donor of *Ae. ventricosa* was *Ae. tauschii* ssp. *tauschii*, and demonstrated that both the D^{v} and N^{v} subgenomes have already been involved in wheat improvement. The D^{v} subgenome from *Ae. ventricosa* will further enrich the genetic diversity of the D subgenome of bread wheat, accompanied by introgression of more N^{v} segments, in wheat improvement. The high-quality reference genome generated here also lays a foundation for the development of genome-specific screening markers and the cloning of key genes from *Ae. ventricosa*, which will greatly accelerate molecular breeding efforts for wheat improvement. Future work should focus on the graph-based representation of pan-genomes and WGS of large panels for *Ae. ventricosa*. With these multi-layer datasets, we would have powerful tools to increase the rate of wheat improvement.

METHODS

Plant materials

The allotetraploid wild grass *Ae. ventricosa* RM271 (2n = 4x = 28, genome D^{v}D^{v}N^{v}N^{v}) was chosen for WGS. The *Ae. uniaristata* (2n = 2x = 14, genome NN) accessions PI276995, PI374326, and PI554421 and the *Yr17* near-isogenic line of bread wheat Avocet (Avo17), which contains the 2N^{v}S segment, were used to generate NGS reads. Plants were grown in the greenhouse of the Crop Research Institute, Sichuan Academy of Agricultural Sciences.

High-throughput sequencing

Long-read WGS

Genomic DNA (gDNA) was extracted from the seedling leaf tissues of RM271 according to the instructions for the cetyltrimethylammonium bromide method. Three methods were used for DNA quantification and quality testing: (1) NanoDrop 2000 spectrophotometry (Thermo Fisher Scientific), (2) gel electrophoresis, and (3) Qubit fluorometry (Invitrogen). Total DNA was purified with AMPure PB beads (PacBio 100-265-900). High-quality gDNA ($\geq$10 μg, $\geq$100 ng/μL) was prepared for the next step of library construction. PacBio single-molecule real-time (SMRT) library preparation was performed using the SMRTbell Express Template Prep Kit 2.0 (PacBio, PN 101-853-100) in accordance with the manufacturer's instructions (Procedure & Checklist–Preparing HiFi SMRTbell Libraries using SMRTbell Express Template Prep Kit 2.0). In brief: (1) 10 μg high-quality gDNA was evaluated using pulsed-field electrophoresis, and it was confirmed that most DNA fragments were longer than 40 kb. (2) DNA was sheared to a mode size of $\geq$15 kb using a Megaruptor instrument (Diagenode B06010001). (3) AMPure PB Beads (PacBio 100-265-900) were used to concentrate the DNA. (4) The SMRTbell library was prepared using Kit 2.0. Single-strand overhangs were removed, and we performed DNA damage repair, end repair, A-tailing, adapter ligation, and enzymatic digestion. Finally, library size selection was performed with SageELF (Sage Science ELF000) or the BluePippin system (Sage Science BLU0001). (5) After library size selection, the library was prepared for sequencing with a 30-h movie on the Sequel IIe system (PacBio) (Supplemental Figure 1A and 1B; Supplemental Table 1).

Full-length transcriptome sequencing

Total RNA was isolated from the mixed organs of RM271, which included whole-plant organs from seed germination to the two-leaf stage, plant shoots in the seedling stage, and leaves, stems, ears, and seeds from the heading to the late-filling stage. The RNA was isolated according to the instructions for the TRIzol method. The integrity, quality, and concentration of RNA were detected with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) and an Agilent 2100 Bioanalyzer (Agilent Technologies). High-quality RNA (2 μg, 300 ng/μL) was prepared for the next step of library construction. PacBio SMRTbell library preparation was performed using a SMRTbell Express Template Prep Kit 2.0 (PacBio, PN 101-853-100) in accordance with the manufacturer's instructions (Procedure & Checklist – Iso-Seq Express Template Preparation for Sequel and Sequel II Systems). In brief: (1) first-strand cDNA was synthesized by reverse transcription with an oligo(dT) primer to pair with the poly(A) tails. (2) Template switching was performed by adding the other primer paired with template switch oligo. (3) Full-length cDNA was obtained by cDNA amplification with barcoded cDNA PCR primers. (4) Amplified cDNA was purified with ProNex Beads and quantified with a Qubit fluorometer (Invitrogen). (5) The SMRTbell library was prepared using Kit 2.0. (6) Single-strand overhangs were removed, and we performed DNA damage repair, end repair, A-tailing, and adapter ligation. After library purification using 0.45× AMPure PB Beads and 1- to 10-kb size selection, the library was prepared for sequencing with a 30-h movie on the Sequel IIe system (PacBio) at the Berry Genomics Corporation (Beijing, China).

Hi-C sequencing

Hi-C sequencing of RM271 was performed using the protocol of Peng et al. (2022). In brief, plants were grown in a growth chamber for 2 weeks. Samples of 2–4 g tender leaves were harvested, cut into pieces of ~2 cm^2, and transferred to 50-mL tubes containing 15 mL of ice-cold nuclear isolation buffer with 2% formaldehyde, followed by vacuum infiltration (400 mbar) and incubation with a supplemented crosslinking agent for 1 h. Crosslinking was quenched by adding 2 M glycine to a final concentration of 0.125 M with incubation for 5 min under vacuum, followed by fixation on ice. Then, the fixed leaf pieces were washed three times with sterile Milli-Q water, ground in liquid nitrogen, and subjected to nucleus isolation. The isolated nuclei were purified, checked for quality and quantity, and digested with 100 U *DpnII*. The next steps were Hi-C specific, including marking the DNA ends with biotin-14-dATP and performing blunt-end ligation of the crosslinked fragments. After ligation, crosslinking was reversed by overnight incubation with Proteinase K at 65°C. Biotin-14-dATP was removed from non-ligated DNA ends using the exonuclease activity of T4 DNA polymerase. DNA was purified by phenol:chloroform (1:1) extraction, precipitated, and washed as described previously. The purified DNA was physically sheared to a size of 300–600 bp by sonication and was size-fractionated using standard 2% agarose gel electrophoresis to obtain fragments in the range of 300–600 bp. The fragmented ends were blunt-end repaired and A-tailed, followed by purification through biotin-streptavidin-mediated pulldown. PCR amplification was performed using 12–15 cycles to enrich the ligation products. After the quality check, the Hi-C libraries were sequenced on the MGISEQ2000 platform at BGI (Wuhan, China).

Short-read WGS

For the three accessions of *Ae. uniaristata* (PI276995, PI374326, and PI554421) and the Yr17 near-isogenic line of the bread wheat variety Avocet (Avo17), high-quality genomic DNA was isolated from fresh leaf tissue using the Qiagen DNeasy Plant Mini Kit. MGI libraries were constructed using the MGIEasy Universal DNA Library Prep Set (BGI, China). In brief, 1–1.5 μg gDNA was randomly fragmented with a Covaris instrument. Then, fragments with sizes between 200 and 400 bp were selected using an Agencourt AMPure XP-Medium Kit, followed by end repair, 3′ adenylation, and adapter ligation. After PCR enrichment, the PCR products were recovered using the AxyPrep Mag PCR Clean-Up Kit. The double-stranded PCR products were heat denatured and circularized using the splint oligo sequence. Single-stranded circular DNA was formatted as the final library and qualified according to quality control procedures. The qualified libraries were sequenced on the MGISEQ2000 platform at BGI.

Genome assembly and validation

Genome survey

Genome sizes were estimated using three different algorithms, gce (Liu et al., 2013), GenomeScope2 (Ranallo-Benavidez et al., 2020), and findGSE (Sun et al., 2018), with different k-mer sizes (Supplemental Figure 1C–1E; Supplemental Table 2). A total of 11 genome assemblies were produced with hifiasm (Cheng et al., 2021) (version 0.16.1, default parameters) using subsets of the PacBio HiFi data (the number of PacBio SMRT cells ranged from 1 to 11) (Supplemental Figure 1F).

Contig assembly and assembly quality evaluation

All the PacBio HiFi CCS reads were assembled using hifiasm (Cheng et al., 2021) (version 0.16.1, default parameters), which generates three types of assemblies (haplotype 1, haplotype 2, and primary assembly) (Supplemental Table 3). The quality and completeness of the genome assemblies were assessed with Merqury (Rhie et al., 2020), which uses a reference-free, k-mer-based approach (Supplemental Figure 2A and 2B), and BUSCO (Simão et al., 2015) (version 5, Poales), which is based on evolutionarily informed expectations of the near-universal single-copy orthologous gene content (Supplemental Figure 2C and 2D). The comparison to other Triticeae genome assemblies revealed a high NG50 value, implying high connectivity (Supplemental Figure 2E). The primary assembly was subjected to GC depth analysis (Supplemental Figure 2F). LAI (Ou et al., 2018), which evaluates assembly continuity using LTR-RTs, was calculated (Supplemental Figure 2G).

Pseudomolecule construction

The Hi-C reads were incorporated using Juicer tools (Durand et al., 2016) (version 1.6) and EndHiC (Wang et al., 2022). In brief, preprocessing of the Hi-C reads was performed with juicer.sh (Durand et al., 2016) (parameter: -s *DpnII*) (Supplemental Table 4). The output file corresponding to the Hi-C contacts with duplicates removed and mapping quality values larger than 30 was generated as input for EndHiC (Wang et al., 2022). These result files were plotted to visualize the Hi-C map and for manual curation (Supplemental Figure 2G), and they were used to generate the final assembly (14 pseudomolecules and 1 unanchored pseudomolecule) (Supplemental Table 5). To explore the distribution of subtelomeric and centromere-specific repeat sequences, genome assemblies were blasted to the sequences from *Ae. tauschii* clones (Zhang et al., 2004) A6-10 (AY249980.1), *Ae. tauschii* clones 6C6-3 (AY249981.1) and 6C6-4 (AY249982.1), and *T. monococcum* subsp. *aegilopoides* clones (Liu et al., 2008) BAC TbBAC5 (DQ904440.1) and BAC TbBAC30 (EF624064.1) (Supplemental Figure 3).

Organelle assemblies

Using minimap2 (Li, 2018) (version 2.17, default parameters), PacBio HiFi CCS reads were aligned to 11 chloroplast genomes from *Aegilops* (Supplemental Table 6), and GetOrganelle (Jin et al., 2020) (version 1.7.6.1, default parameters) was then used to assemble the chloroplast genome of *Ae. ventricosa*. Using minimap2 (Li, 2018) (version 2.17, default parameters), PacBio HiFi CCS reads were aligned to the mitochondrial genome of *Ae. speltoides*. Then, the mapped reads were extracted and assembled into contigs with hifiasm (Cheng et al., 2021) (version 0.16.0-r369, default parameters). Finally, GetOrganelle (Jin et al., 2020) (version 1.7.6.1, default parameters) was used to assemble the primary contigs from hifiasm (Cheng et al., 2021), generating the mitochondrial genome of *Ae. ventricosa*. Finally, we obtained the chloroplast and mitochondrial genomes of *Ae. ventricosa*, with total lengths of 135 613 and 586 840 bp, respectively (Supplemental Table 7).

Subgenome assignment, validation, and nomenclature

To assign each chromosome to linkage groups and nomenclature, SubPhaser (Jia et al., 2022), a robust allopolyploid subgenome phasing method based on subgenome-specific k-mers, was performed (Figure 1). To validate the correctness of our subgenome assignment, a reference-guided strategy based on subgenome homology was also used to distinguish the subgenomes. We mapped the RM271 genome to the *Ae. tauschii* ssp. *tauschii AY17* genome using mashmap (Jain et al., 2018) (*-f map –perc_identity 80 -s 1000000*). The alignments were plotted and manually checked. This successfully split the 14 chromosomes into 2 homologous groups. The group sharing higher synteny and identity with *Ae. tauschii* ssp. *tauschii AY17* was assigned as the D^v subgenome, and the other group was assigned as the N^v subgenome. The nomenclature system for *Ae. tauschii* ssp. *tauschii AY17* chromosomes was used to name the homologous groups (1–7) of the RM271 genome (Figure 4A).

Repeat annotation

Identification and clustering of tandem repeats

Tandem repeats consist of monomer sequence units arranged directly next to each other. For a simple and pragmatic classification, they are commonly split into three main categories by their unit length: 1- to 9-bp minisatellites, 10- to 99-bp microsatellites, and ≥100-bp satellites. Tandem repeats of all lengths were annotated with TRF version 4.09 (Benson, 1999) under default parameters (Match Mismatch Delta PM PI Minscore MaxPeriod: 2 7 7 80 10 50 500). Overlapping annotations were removed using a priority-based approach, assigning higher-scoring and longer elements first. Elements that overlapped with already-assigned elements were either discarded (>90% overlap) or shortened (≤90% overlap) if their remaining length exceeded 49 bp. The tandem repeat statistics are listed in Supplemental Tables 8 and 9.

Transposon annotation based on homology with a TE library

LTR_FINDER (Xu and Wang, 2007) (version 1.05) and LTR_harvest (Ellinghaus et al., 2008) (version 1.5.10) were used for LTR identification, and the results were processed with LTR_retriever (Ou and Jiang, 2018) (version 2.8) to generate a species-specific LTR library. The libraries obtained from 2825 complete plant TE sequences (TREP) (Wicker et al., 2002), wheat TE sequences from ClariTeRep (Daron et al., 2014; jdaron, 2021), species-specific LTR libraries, and plant TE sequences from Repbase (Jurka et al., 2005) were merged to generate the TE library. Transposons were detected and classified by a homology search against the combined TE library. The program vmatch (www.vmatch.de), a fast and efficient matching tool suitable for large and highly repetitive genomes, was used for this computationally intensive task with the following parameters: identity $\geq$70%, minimal hit length 75 bp, and seed length 12 bp (parameters: *-d -p -l 75 -identity 70 -seedlength 12 -exdrop 5*). To obtain an overlap-free annotation, the output was filtered for redundant hits by a priority-based approach from a score-sorted match list. Higher-scoring matches were assigned first and shortened if they overlapped with an already-assigned element and if their remaining length was at least 50 bp and $\geq$10% of their hit length. All other matches were discarded. A summary of the transposons in the RM271 genomes is provided in Supplemental Tables 8 and 10.

Protein-coding gene annotation

Gene model prediction

Gene model prediction was performed according to a previously published method (Mascher et al., 2021) with minor modifications, combining transcriptomics data, protein homology, and *ab initio* prediction. (1) Iso-Seq data were mapped using minimap2 (Li, 2018) (version 2.17-r941; parameters: *-ax splice -uf –secondary = no -C5*), and the redundant isoforms were further collapsed into transcript loci using cDNA_Cupcake (Tseng, 2022) (http://github.com/Magdoll/cDNA_Cupcake; parameter: –dun-merge-5-shorter). TransDecoder (Hass, 2023) (version 5.5.0; https://github.com/TransDecoder/TransDecoder) was used to find potential open reading frames and to predict protein sequences within the candidate transcript set. (2) For protein homology evidence, we projected the gene annotations of Triticeae species, including *Ae. tauschii*, *T. turgidum* ssp. *dicoccoides*, *T. turgidum* L. ssp. *durum*, *T. aestivum* Chinese Spring, *T. urartu*, *Ae. speltoides*, and *Hordeum vulgare* Morex, to the RM271 assemblies using liftoff (Shumate and Salzberg, 2021) with default parameters. All of the translated proteins from the above species and the Triticeae protein sequences downloaded from the SwissProt database (Boeckmann et al., 2003) were also mapped against the assembly using GenomeThreader (Gremme et al., 2005) (version 1.7.1; parameters: *-startcodon -finalstopcodon -species rice -gcmincoverage 70 -prseedlength 7 -prhdist 4 -gff3out*). (3) We produced *ab initio* gene predictions using AUGUSTUS (Stanke et al., 2006) (version 3.4.0), GeneMark-ET (Besemer and Borodovsky, 2005) (version 4.38) and GeneID (Blanco et al., 2007) (version 1.4). In brief, AUGUSTUS (Stanke et al., 2006) gene prediction was performed using a model specifically trained from software and a hints file generated using the previously mentioned Iso-Seq predictions. GeneMark-ET (Besemer and Borodovsky, 2005) was used with the option -ET, and the intron coordinates were calculated using the above Iso-Seq alignments. GeneID (Blanco et al., 2007) was run with a model specifically trained from software (*-GP taestivum.param*). We used EVidenceModeler (Haas et al., 2008) (version 1.1.1) to join all of the gene evidence from transcriptomics, protein alignments, and *ab initio* predictions.

Confidence assignment

Gene models were classified as HC or low confidence according to criteria used by the International Wheat Genome Sequencing Consortium, with minor modifications (International Wheat Genome Sequencing Consortium [IWGSC] et al., 2018). In brief, protein-coding gene models were considered complete when start and stop codons were present. A comparison with PTREP (a database of hypothetical proteins deduced from the nonredundant database of TEs within the TREP database: https://trep-db.uzh.ch/), UniPoa (Poaceae database of annotated proteins from the UniProt database), and UniMag (Boeckmann et al., 2003) (validated Magnoliophyta proteins from SwissProt) was performed using DIAMOND (Buchfink et al., 2021) (version 2.0.9; parameters: -e 1e−10 –query-cover 80 –subject-cover 80). Gene candidates were further classified using the following criteria: an HC gene model was complete with a hit in the UniMag (Boeckmann et al., 2003) database and/or in UniPoa and not PTREP; the remaining gene models were classified as low-confidence genes.

Functional annotation of protein-coding genes

Functional assignments for the predicted protein-coding genes were obtained with BLAST (Camacho et al., 2009) by aligning the coding regions to sequences in public protein databases, including the trEMBL (Boeckmann et al., 2003), RefSeq (Pruitt et al., 2014), and SwissProt (Boeckmann et al., 2003) databases. Putative domains and GO (Ashburner et al., 2000) terms of the predicted proteins were identified using InterProScan (Quevillon et al., 2005) with the default settings. Putative orthologs in the Kyoto Encyclopedia of Genes and Genomes database (Kanehisa et al., 2012) were identified using KoFamScan (Aramaki et al., 2020).

Functional enrichment of genes

GO terms for genes were extracted from IprScan (Quevillon et al., 2005). Significant terms were identified using topGO (Alexa and Rahnenfuhrer, 2023) and were reported up to an adjusted p value of $\leq$0.05.

R genes

R genes were identified in the RM271 genome and other genomes using DRAGO2 (Osuna-Cruz et al., 2018) (version DRAGO2-API: https://github.com/sequentiabiotech/DRAGO-API) (Supplemental Tables 11–13). This program detects LRR, kinase, NBS, and Toll-interleukin receptor domains from 60 hidden Markov model (HMM) modules created for this purpose using the HMMER version 3 package, and it is also able to detect CC and transmembrane (TM) domains using the COILS 2.2 (Lupas et al., 1991) and TMHMM 2.0c programs. Nucleotide-binding and leucine-rich repeat (NLR) signatures were annotated using NLR-Annotator (Steuernagel et al., 2020) (https://github.com/steuernb/NLR-Annotator) with the default parameters (Supplemental Tables 14–16). Another three types of potential resistance-related gene families, including the CYP superfamily, the ABC transporter family, and the WRKY transcription factors (PF00067, PF00005, and PF03106, respectively), were extracted from the InterProScan (Quevillon et al., 2005) results if Pfam (Bateman et al., 2004) hits existed (Supplemental Table 17). The distributions of disease-resistance-related genes were plotted across the whole RM271 genome (Supplemental Figure 4).

Phylogenomics and allopolyploidization history of RM271

Transcript assembly and coding sequence prediction

To understand the evolutionary positions of Dv- and N^v-subgenome species within the Triticeae, we first assembled transcripts of diploid Triticeae species that lacked available genome assemblies but had available RNA sequencing (RNA-seq) datasets, including *Ae. comosa* (DRR202423–DRR202425), *Ae. markgrafii* (DRR202414–DRR202422), *Ae. umbellulata* (DRR119232–DRR119243), and *Ae. uniaristata* (DRR202426–DRR202428). Raw sequencing reads were filtered using the following steps: read pairs with adapter contamination, read pairs with N contents >3%, and read pairs with >20% low-quality bases (Phred Q scores <20) were first removed. Reads with potential low-quality regions were then trimmed with Trimmomatic (Bolger et al., 2014). Reads with a Q score <15 at both ends were also trimmed, and reads containing 3′ or 5′ ends with an average quality score that dropped below Q20 in a 4-bp sliding window were trimmed. Finally, all reads <32 bp were excluded to obtain clean data for further analyses. The clean reads were assembled *de novo* using Trinity (Haas et al., 2013) (version 2.0.3) with the default parameters. Coding sequences (CDSs) and proteins were predicted using TransDecoder (Hass, 2023) (version 5.5.0). To remove redundant sequences, we used BLASTN (Camacho et al., 2009) for similarity searches of the proteins against *Ae. tauschii* ssp. *strangulata* AL8/78. Top hits were filtered if the identity was <80% or the subject coverage was <80%; then, for each locus in AL8/78, only the hit with the highest score was retained.

Phylogenetic tree construction and divergence time estimation

Proteomes of Triticeae species with genome assemblies, the above species with assembled transcripts, and *Brachypodium distachyon* were used to identify conserved single-copy genes. Proteins of these species with a sequence length of <50 amino acids were first removed, and for genes with alternative splice variants, only the longest transcript was selected. Single-copy orthologous gene clusters from these species were then identified using OrthoFinder (Emms and Kelly, 2019) (version 2.2.7). In this step, the subgenomes of the polyploid species were split and used as independent entries. In total, 598 single-copy orthologous gene clusters shared by all 32 subgenomes from 18 species were identified. For each of these single-copy gene clusters, protein sequences of the homologous genes were aligned using MUSCLE (Edgar, 2004) (version 3.8.31) with the default parameters, and the alignments were back-translated into CDSs using an in-house Perl script. The conserved CDS alignments were extracted using Gblocks (Castresana, 2000) (version 0.9b), and the retained CDS alignments of each family were used for further phylogenetic analyses.

For phylogenetic tree construction, the CDS alignments of each single-copy gene family were concatenated to generate a supermatrix of 601 593 unambiguously aligned nucleotide positions and analyzed with RAxML (Stamatakis, 2014) (version 8.2.7) to generate a maximum likelihood tree with the GTR+I+Γ model.

Divergence times were estimated under a relaxed clock model using the MCMCTree program with the "Independent rates model (clock = 2)" and "JC69" model in the PAML (Yang, 2007) (version 4.7) package. Considering that evolutionary rates at different codon positions vary, the three codon positions of the concatenated supergene were treated as three different partitions. The Markov chain Monte Carlo (MCMC) process was run for 6 000 000 iterations after a burn-in of 2 000 000 iterations. We ran the program twice for each data type to confirm that the results were similar between runs. The chronogram was produced using FigTree (Rambaut, 2023) (version 1.4.0) (http://tree.bio.ed.ac.uk/) with the first run (Figure 2A).

Nucleotide similarity among subgenomes

Each genome/subgenome assembly was divided into 150-bp fragments; for each accession of *Ae. uniaristata*, *Ae. markgrafii*, *Ae. umbellulata*, and *Ae. comosa*, ~2× clean short reads were randomly extracted from the WGS data. These fragments/reads were mapped to the D and N subgenomes using BWA (Li, 2013) with the default parameters. Uniquely mapped reads were extracted using SAMtools (Li et al., 2009) and were blasted against the D and N subgenomes using blastn (Camacho et al., 2009). The best hit for each chunk/read was retained when the BLASTN score was 15 points greater than that of the suboptimal hit and the query coverage was over 145 bp. The average identity over a sliding window of 5 Mb with a step size of 1 Mb was calculated and plotted along the chromosomes of the RM271 assemblies (Figure 2B and Supplemental Figure 5; Supplemental Tables 18 and 19).

The timing of allo-tetraploidization

To date the time of tetraploid origin, we obtained orthologous gene pairs between the *Ae. tauschii* ssp. *tauschii* genome and the D^v subgenome of *Ae. ventricosa* and calculated the Ks values of orthologous gene pairs using the yn00 module of the PAML version 4.7 package (Yang, 2007). The average substitution rate (r) of 6.5×10^{-9} substitutions per synonymous site per year for grasses was used to calibrate the ages of the orthologous gene pairs (Gaut et al., 1996; SanMiguel et al., 1998). The time (T) was estimated using the formula T = Ks/r (Figure 2C).

Analyses of synonymous and nonsynonymous substitution rates

We obtained one-to-one orthologous gene sets from the OrthoFinder (Emms and Kelly, 2019) results and used them for the Ka and Ks calculations (Figure 3A). For this analysis, the orthologous gene pair list was used as the input, and the protein sequences from each gene pair were aligned using MUSCLE (Edgar, 2004). PAL2NAL (Suyama et al., 2006) was used to convert the peptide alignment to a nucleotide alignment, and the Ka and Ks values were computed between the gene pairs using codeml from PAML (version 4.7) (Yang, 2007) in free-ratios mode. All estimates with Ks < 0.01 were excluded from the analysis. The average Ka/Ks values of all genes were used to represent the Ka/Ks values for the corresponding subgenomes. The significance of differences in Ka/Ks values between genomes (subgenomes) was estimated using the Wilcoxon rank-sum test for nonnormal distributions in R.

Gene retention

Orthologs between *Ae. tauschii* and the D^v subgenome of *Ae. ventricosa* were identified using reciprocal best hits-based methods. A sliding window approach with a window size of 100 genes, a step size of 10 genes, and *Ae. tauschii* as the reference was used to reveal the percentage of retained genes in the D^v subgenome of *Ae. ventricosa*. The gene retention rates of the bread wheat D subgenome were calculated and plotted using the same methods (Figure 3B and 3C).

Karyotype evolution

The AGK genome, which includes 7 protochromosomes and 7010 ordered protogenes, was downloaded (Murat et al., 2017), and the protein sequences of RM271 and other targeted species were aligned with the AGK protogenes. Syntenic blocks, defined on the basis of the presence of at least five syntenic gene pairs, were identified using the MCScanX package (Wang et al., 2012) with the default settings. The syntenic blocks were then used to deduce the homologous relationships between the AGK marker genes and the protein sequences of RM271 and other Triticeae species (Supplemental Figure 8D).

Syntenic blocks of chromosomes were identified by MCScanX (Wang et al., 2012), and the ATK was then constructed with DESCHRAMBLER (Kim et al., 2017; Wu et al., 2022; Peng et al., 2023) (Supplemental Figure 8A and 8B). The genome of *Ae. tauschii* ssp. *strangulata* was selected as the reference genome, and Triticeae species (except Triticinae species) served as the outgroup taxa. The homologous relationships between the ATK marker genes and the protein sequences of RM271 and other Triticeae species were then analyzed as described above (Supplemental Figure 8C).

Identification of introgression segments based on genome assemblies

To identify introgressed segments from *Ae. ventricosa* present in bread wheat varieties worldwide, we collected reference-quality genomes from 17 wheat varieties (AK58, ArinaLrFor, Chinese Spring, Fielder, Jagger, Julius, Kariega, KN9204, LongReach Lancer, CDC Landmark, Mace, SY Mattis, Norin 61, Renan, CDC Stanley, PI190962, and Zang1817). We mapped the genome assemblies from each wheat variety to the RM271 genome using mashmap (Jain et al., 2018) with a mapping segment length of 1 000 000 bp. Alignments with an identity less than 98% were removed, and adjacent segments were manually merged into longer blocks (Supplemental Table 20). For visualization purposes, we realigned the RM271 genome to the wheat cultivar Jagger using mashmap (Jain et al., 2018) with a mapping segment length of 1 000 000 bp and plotted synteny as dot plots.

Graph-based pangenome analyses of $6N^vL/2AS$ segments in the Triticeae

To construct a graph-based pangenome of $6N^vL/2AS$ segments in the Triticeae, we first collected 23 reference-quality genomes from 17 wheat varieties, including AK58 (GWH: GWHANRF00000000) (Jia et al., 2023), ArinaLrFor (Ensembl Plants release 54) (Walkowiak et al., 2020), Fielder (Ensembl Plants release 54) (Walkowiak et al., 2020), Jagger (Ensembl Plants release 54) (Walkowiak et al., 2020), Julius (Ensembl Plants

release 54) (Walkowiak et al., 2020), LongReach Lancer (Ensembl Plants release 54) (Walkowiak et al., 2020), CDC Landmark (Ensembl Plants release 54) (Walkowiak et al., 2020), Mace (Ensembl Plants release 54) (Walkowiak et al., 2020), SY Mattis (Ensembl Plants release 54) (Walkowiak et al., 2020), Norin 61 (Ensembl Plants release 54) (Walkowiak et al., 2020), CDC Stanley (Ensembl Plants release 54) (Walkowiak et al., 2020), PI190962 (Ensembl Plants release 54) (Walkowiak et al., 2020), Chinese Spring (refseqv2.1) (International Wheat Genome Sequencing Consortium [IWGSC] et al., 2018), Kariega (Kariega_v1) (Athiyannan et al., 2022), KN9204 (GWH: GWHBJWI00000000) (Shi et al., 2022), Renan (Renan_v2.1) (Aury et al., 2022), and Zang1817 (NCBI: GCA_014338645.1) (Guo et al., 2020); three *T. monococcum* ssp. *monococcum* accessions, TA299 (v1.1) (Ahmed et al., 2023), TA10622 (v1.1) (Ahmed et al., 2023), and PI306540 (GWH: GWHCBHO00000000) (Wang et al., 2024); *T. urartu* G1812 (GenBank: MKGO00000000) (Ling et al., 2018); *T. turgidum* ssp. *dicoccoides* Zavitan (GenBank: LSYQ00000000) (Avni et al., 2017); and *T. turgidum* ssp. *durum* Svevo (NCBI: GCA_900231445.1) (Maccaferri et al., 2019). We then mapped each genome assembly to the $6N^vL$ segment from 537.2 to 574.7 Mb on the $6N^vL$ chromosome of RM271 using mashmap (Jain et al., 2018) with a mapping segment length of 1 000 000 bp, and the homologous segments in the 23 reference-quality genomes were manually checked. We used pggb (version 0.5.4) (Garrison et al., 2023) to construct a graph-based pangenome of $6N^vL$/2AS segments in the Triticeae and the vg toolkit (Hickey et al., 2020) to call variants using the $6N^vL$ segment from RM271 as the reference. ANIb (ANI blast) analysis of the $6N^vL$ segment in RM271 and its homologous segments in the 23 reference-quality genomes was performed using pyani (Pritchard et al., 2016). According to the pyani and pggb results, our 24 $6N^vL$/2AS segments were grouped into two groups corresponding to six $6N^vL$ segments and 18 2AS segments. To identify $6N^vL$-sourced disease-resistance genes, we extracted the variants that distinguished the $6N^vL$ segments and the 2AS segments in the pangenome using pggb (to avoid the sequencing errors in the genome assemblies, we allowed at most one violation in the $6N^vL$ segments). Variants were then annotated using snpEff (Cingolani et al., 2012). Finally, we extracted the R genes with variants that altered amino acids in the $6N^vL$ segments (Supplemental Tables 23 and 24) by combining R genes from DRAGO2 (Supplemental Table 11) and NLR-Annotator (Steuernagel et al., 2020) (Supplemental Table 14).

Identification of introgression segments based on whole-genome resequencing

To further analyze the contribution of *Ae. ventricosa* to existing wheat resources, we analyzed WGS datasets from 664 wheat accessions that are available in the ENA database under accession numbers PRJEB48738 and PRJEB48988 (Schulthess et al., 2022). NGS reads were mapped to the concatenated genome assemblies from RM271 and Chinese Spring using BWA (Li and Durbin, 2009). The resulting BAM files were sorted with SAMtools (Li et al., 2009), and PCR duplicates were removed using Picard (https://github.com/broadinstitute/picard). For each sample, mean depths in each window (size: 100 000 bp) were calculated with mosdepth (Pedersen and Quinlan, 2018) and then normalized by the average depths of all windows from the Chinese Spring A subgenome. Normalized depths were visualized with R. To identify introgressions in the wheat genome, we extracted windows with normalized depths ranging from 0.75 to 1.25 in the RM271 genome. Regions with at least five adjacent candidate windows were retained as introgression segment candidates.

To remove potential false positives, we extracted the reads located in the introgression segment candidates and remapped them to the concatenated 11 species (12 genomes): *Ae. ventricosa* RM271 (genome DDNN), *Ae. umbellulata* TA1851 (genome UU), *Ae. searsii* TE01 (genome S^sS^s), *Ae. bicornis* TB01 (genome S^bS^b), *Ae. longissima* AEG-6782-2 (genome S^lS^l), *Ae. sharonensis* TH02 (genome $S^{sh}S^{sh}$), *Ae. tauschii* ssp. *strangulata* AL8/78 (genome DD), *Ae. tauschii* ssp. *strangulata* AY61 (genome DD), *Ae. tauschii* ssp. *tauschii* AY17 (genome DD), *Ae. tauschii* ssp. *tauschii* T093 (genome DD), and *Ae. tauschii* ssp. *tauschii* XJ02 (genome DD). After the extraction of unique mapped alignments, we used the same pipeline described above to generate normalized depth plots. We then manually checked the plots and number of alignments in each introgression segment candidate to remove false candidates. Finally, we obtained 12 introgression segments (Supplemental Tables 25–27).

DATA AND CODE AVAILABILITY

- All the raw sequencing data generated in this study are available in the National Genomics Data Center (https://ngdc.cncb.ac.cn/?lang=en) under BioProject numbers PRJCA018801, PRJCA018802 and PRJCA018803 for *Ae. ventricosa* RM271, three *Ae. uniaristata* accessions (PI276995, PI374326, and PI554421), and the Yr17 near-isogenic line of bread wheat Avocet (Avo17), respectively.
- All software and pipelines were executed according to the manuals and protocols of published tools. No custom code was generated for these analyses.

FUNDING

This publication is based on work supported by the Sichuan Provincial Finance Department Project of China (1 + 3 ZYGG001), the Sichuan Province Science and Technology Department Project of China (2021YFYZ0020, 2022NSFSC0161, 2023NSFSC1925), the Program of Strategic Scientist Studio of Sichuan Academy of Agricultural Sciences, and the Program of Chinese Agriculture Research System (CARS-03).

ACKNOWLEDGEMENTS

We are grateful to Prof. Lihui Li from the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, for providing *Ae. ventricosa* RM271 material. We also thank Prof. Zehong Yan from Sichuan Agricultural University for providing the materials of *Ae. uniaristata*. We are grateful for the sequencing support of the Berry Genomics Corporation (Beijing, China) and BGI (Wuhan, China). We thank Springer Nature for providing the English-language editing services. No conflict of interest is declared.

AUTHOR CONTRIBUTIONS

W.Y., J.L., and J.J. designed the study. W.Y. and J. Li provided the funding. Z.L., F.Y., and H.W. collected and prepared the plant tissue samples for sequencing. Z.L., F.Y., H.W., C.D., and W.H. led the bioinformatics analyses. Z.L. and C.D. drafted the manuscript. W.Y., J. Li, J.J., F.Y., H.W., W.H., X.F., J.W., L.W., E.Y., and Z.P. revised the manuscript. M.Y., J.F., Q.W., N.Y., L.C., Y.L., and H.T. contributed to the genome analyses. S.L., J. Luo, and J.Z. contributed to the analyses of the whole-genome resequencing datasets. All authors discussed the results and approved the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

Received: April 13, 2024
Revised: August 14, 2024
Accepted: September 6, 2024
Published: September 10, 2024

REFERENCES

Ahmed, H.I., Heuberger, M., Schoen, A., Koo, D.-H., Quiroz-Chavez, J., Adhikari, L., Raupp, J., Cauet, S., Rodde, N., Cravero, C., et al. (2023). Einkorn genomics sheds light on history of the oldest domesticated wheat. Nature **620**:830–838.

Alexa, A., and Rahnenfuhrer, J. (2023). topGO: Enrichment Analysis for Gene Ontology Advance Access published 2023. https://doi.org/10.18129/B9.bioc.topGO.

Anbu, S., Swart, V., and van den Berg, N. (2023). Unmasking the invaders: NLR-mal function in plant defense. Front. Plant Sci. **14**:1307294.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., and Ogata, H. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinforma. Oxf. Engl. **36**:2251–2252.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. Nat. Genet. **25**:25–29.

Athiyannan, N., Abrouk, M., Boshoff, W.H.P., Cauet, S., Rodde, N., Kudrna, D., Mohammed, N., Bettgenhaeuser, J., Botha, K.S., Derman, S.S., et al. (2022). Long-read genome sequencing of bread wheat facilitates disease resistance gene cloning. Nat. Genet. **54**:227–231.

Aury, J.-M., Engelen, S., Istace, B., Monat, C., Lasserre-Zuber, P., Belser, C., Cruaud, C., Rimbert, H., Leroy, P., Arribat, S., et al. (2022). Long-read and chromosome-scale assembly of the hexaploid wheat genome achieves high resolution for research and breeding. GigaScience **11**:giac034.

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S.O., Gundlach, H., Hale, I., Mascher, M., Spannagl, M., Wiebe, K., et al. (2017). Wild emmer genome architecture and diversity clarify wheat evolution and domestication. Science **357**:93–97.

Badaeva, E.D., Dedkova, O.S., Koenig, J., Bernard, S., and Bernard, M. (2008). Analysis of introgression of Aegilops ventricosa Tausch. genetic material in a common wheat background using C-banding. Theor. Appl. Genet. **117**:803–811.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.L., et al. (2004). The Pfam protein families database. Nucleic Acids Res. **32**:D138–D141.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27**:573–580.

Besemer, J., and Borodovsky, M. (2005). GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res. **33**:W451–W454.

Bianchet, C., Wong, A., Quaglia, M., Alqurashi, M., Gehring, C., Ntoukakis, V., and Pasqualini, S. (2019). An Arabidopsis thaliana leucine-rich repeat protein harbors an adenylyl cyclase catalytic center and affects responses to pathogens. J. Plant Physiol. **232**:12–22.

Blanco, E., Parra, G., and Guigó, R. (2007). Using geneid to identify genes. Curr. Protoc. Bioinforma. Chapter 4:Unit 4.3.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. **31**:365–370.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114–2120.

Bonhomme, A., Gale, M.D., Koebner, R.M., Nicolas, P., Jahier, J., and Bernard, M. (1995). RFLP analysis of an Aegilops ventricosa chromosome that carries a gene conferring resistance to leaf rust (Puccinia recondita) when transferred to hexaploid wheat. Theor. Appl. Genet. **90**:1042–1048.

Boyes, D.C., Nam, J., and Dangl, J.L. (1998). The Arabidopsis thaliana RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. Proc. Natl. Acad. Sci. **95**:15849–15854.

Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods **18**:366–368.

Cadiou, L., Brunisholz, F., Cesari, S., and Kroj, T. (2023). Molecular engineering of plant immune receptors for tailored crop disease resistance. Curr. Opin. Plant Biol. **74**:102381.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinf. **10**:421.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17**:540–552.

Chen, N., Chen, W.-J., Yan, H., Wang, Y., Kang, H.-Y., Zhang, H.-Q., Zhou, Y.-H., Sun, G.-L., Sha, L.-N., and Fan, X. (2020). Evolutionary patterns of plastome uncover diploid-polyploid maternal relationships in Triticeae. Mol. Phylogenet. Evol. **149**:106838.

Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat. Methods **18**:170–175.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) **6**:80–92.

Daron, J., Glover, N., Pingault, L., Theil, S., Jamilloux, V., Paux, E., Barbe, V., Mangenot, S., Alberti, A., Wincker, P., et al. (2014). Organization and evolution of transposable elements along the bread wheat chromosome 3B. Genome Biol. **15**:546.

Dedryver, F., Paillard, S., Mallard, S., Robert, O., Trottet, M., Nègre, S., Verplancke, G., and Jahier, J. (2009). Characterization of genetic components involved in durable resistance to stripe rust in the bread wheat "Renan.". Phytopathology **99**:968–973.

Delibes, A., Lopez-Braña, I., Mena, M., and García-Olmedo, F. (1987). Genetic transfer of resistance to powdery mildew and of an associated biochemical marker from Aegilops ventricosa to hexaploid wheat. Theor. Appl. Genet. **73**:605–608.

Doussinault, G., Delibes, A., Sanchez-Monge, R., and Garcia-Olmedo, F. (1983). Transfer of a dominant gene for resistance to eyespot disease from a wild grass to hexaploid wheat. Nature **303**:698–700.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. **3**:95–98.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797.

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinf. **9**:18.

Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. **20**:238.

Gao, L., Koo, D.-H., Juliana, P., Rife, T., Singh, D., Lemes da Silva, C., Lux, T., Dorn, K.M., Clinesmith, M., Silva, P., et al. (2021). The Aegilops ventricosa 2NvS segment in bread wheat: cytology, genomics and breeding. Theor. Appl. Genet. **134**:529–542.

Garrison, E., Guarracino, A., Heumos, S., Tattini, L., Hagmann, J., Marco-Sola, S., Kubica, C., Thorell, K., Rusholme-Pilcher, R.L., Rudbeck, E., et al. (2023). Building pangenome graphs. Preprint at bioRxiv. https://doi.org/10.1101/2023.04.05.535718.

Gaurav, K., Arora, S., Silva, P., Sánchez-Martín, J., Horsnell, R., Gao, L., Brar, G.S., Widrig, V., John Raupp, W., Singh, N., et al. (2022).

Population genomic analysis of Aegilops tauschii identifies targets for bread wheat improvement. Nat. Biotechnol. **40**:422–431.

Gaut, B.S., Morton, B.R., McCaig, B.C., and Clegg, M.T. (1996). Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc. Natl. Acad. Sci. USA **93**:10274–10279.

Gill, B.S., and Raupp, W.J. (1987). Direct Genetic Transfers from Aegilops squarrosa L. to Hexaploid Wheat1. Crop Sci. **27**:445–450.

Gill, B.S., Friebe, B., Raupp, W.J., Wilson, D.L., Cox, T.S., Sears, R.G., Brown-Guedira, G.L., and Fritz, A.K. (2006). Wheat Genetics Resource Center: The First 25 Years. In Advances in Agronomy (Academic Press), pp. 73–136.

Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W., and Dangl, J.L. (1995). Structure of the Arabidopsis RPM1 Gene Enabling Dual Specificity Disease Resistance. Science **269**:843–846.

Gremme, G., Brendel, V., Sparks, M.E., and Kurtz, S. (2005). Engineering a software tool for gene structure prediction in higher organisms. Inf. Softw. Technol.. Advance Access published 2005.

Guo, W., Xin, M., Wang, Z., Yao, Y., Hu, Z., Song, W., Yu, K., Chen, Y., Wang, X., Guan, P., et al. (2020). Origin and adaptation to high altitude of Tibetan semi-wild wheat. Nat. Commun. **11**:5085.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. **9**:R7.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. **8**:1494–1512.

Hass, B. (2023). TransDecoder Advance Access.

Hickey, G., Heller, D., Monlong, J., Sibbesen, J.A., Sirén, J., Eizenga, J., Dawson, E.T., Garrison, E., Novak, A.M., and Paten, B. (2020). Genotyping structural variants in pangenome graphs using the vg toolkit. Genome Biol. **21**:35.

Jain, C., Koren, S., Dilthey, A., Phillippy, A.M., and Aluru, S. (2018). A fast adaptive algorithm for computing whole-genome homology maps. Bioinformatics **34**:i748–i756. jdaron (2021). CLARI-TE Advance Access published February 6, 2021.

Jia, K.-H., Wang, Z.-X., Wang, L., Li, G.-Y., Zhang, W., Wang, X.-L., Xu, F.-J., Jiao, S.-Q., Zhou, S.-S., Liu, H., et al. (2022). SubPhaser: a robust allopolyploid subgenome phasing method based on subgenome-specific k-mers. New Phytol. **235**:801–809.

Jia, J., Zhao, G., Li, D., Wang, K., Kong, C., Deng, P., Yan, X., Zhang, X., Lu, Z., Xu, S., et al. (2023). Genome resources for the elite bread wheat cultivar Aikang 58 and mining of elite homeologous haplotypes for accelerating wheat improvement. Mol. Plant **16**:1893–1910.

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C.W., Yi, T.-S., and Li, D.-Z. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. **21**:241.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. **110**:462–467.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. **40**:D109–D114.

Kim, J., Farré, M., Auvil, L., Capitanu, B., Larkin, D.M., Ma, J., and Lewin, H.A. (2017). Reconstruction and evolutionary history of eutherian chromosomes. Proc. Natl. Acad. Sci. USA **114**:E5379–E5388.

Lee, D., Bourdais, G., Yu, G., Robatzek, S., and Coaker, G. (2015). Phosphorylation of the Plant Immune Regulator RPM1-INTERACTING PROTEIN4 Enhances Plant Plasma Membrane H^+-ATPase Activity and Inhibits Flagellin-Triggered Immune Responses in Arabidopsis. Plant Cell **27**:2042–2056.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM Advance Access published 2013.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics **34**:3094–3100.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics **25**:1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics **25**:2078–2079.

Li, A., Liu, D., Yang, W., Kishii, M., and Mao, L. (2018). Synthetic Hexaploid Wheat: Yesterday, Today, and Tomorrow. Engineering **4**:552–558.

Li, J., Wang, Q., Li, C., Bi, Y., Fu, X., and Wang, R. (2019). Novel haplotypes and networks of AVR-Pik alleles in Magnaporthe oryzae. BMC Plant Biol. **19**:204.

Lilienfeld, F.A. (1951). Genome-analysis in Triticum and Aegilops. Cytologia **16**:101–103.

Ling, H.-Q., Ma, B., Shi, X., Liu, H., Dong, L., Sun, H., Cao, Y., Gao, Q., Zheng, S., Li, Y., et al. (2018). Genome sequence of the progenitor of wheat A subgenome Triticum urartu. Nature **557**:424–428.

Liu, Z., Yue, W., Li, D., Wang, R.R.-C., Kong, X., Lu, K., Wang, G., Dong, Y., Jin, W., and Zhang, X. (2008). Structure and dynamics of retrotransposons at wheat centromeres and pericentromeres. Chromosoma **117**:445–456.

Liu, B., Shi, Y., Yuan, J., Hu, X., Zhang, H., Li, N., Li, Z., Chen, Y., Mu, D., and Fan, W. (2013). Estimation of genomic characteristics by analyzing k-mer frequency in de novo genome projects Advance Access. Preprint at arXiv. https://doi.org/10.48550/arXiv.1308.2012.

Liu, Z., Yang, F., Deng, C., Wan, H., Tang, H., Feng, J., Wang, Q., Yang, N., Li, J., and Yang, W. (2024). Chromosome-level assembly of the synthetic hexaploid wheat-derived cultivar Chuanmai 104. Sci. Data **11**:670.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science **252**:1162–1164.

Maccaferri, M., Harris, N.S., Twardziok, S.O., Pasam, R.K., Gundlach, H., Spannagl, M., Ormanbekova, D., Lux, T., Prade, V.M., Milner, S.G., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. Nat. Genet. **51**:885–895.

Mascher, M., Wicker, T., Jenkins, J., Plott, C., Lux, T., Koh, C.S., Ens, J., Gundlach, H., Boston, L.B., Tulpová, Z., et al. (2021). Long-read sequence assembly: a technical evaluation in barley. Plant Cell **33**:1888–1906.

Mena, M., Doussinault, G., Lopez-Braña, I., Aguaded, S., García-Olmedo, F., and Delibes, A. (1992). Eyespot resistance gene Pch-1 in H-93 wheat lines. Evidence of linkage to markers of chromosome group 7 and resolution from the endopeptidase locus Ep-D1b. Theor. Appl. Genet. **83**:1044–1047.

Murat, F., Armero, A., Pont, C., Klopp, C., and Salse, J. (2017). Reconstructing the genome of the most recent common ancestor of flowering plants. Nat. Genet. **49**:490–496.

Ortiz, D., de Guillen, K., Cesari, S., Chalvon, V., Gracy, J., Padilla, A., and Kroj, T. (2017). Recognition of the Magnaporthe oryzae Effector

AVR-Pia by the Decoy Domain of the Rice NLR Immune Receptor RGA5. Plant Cell **29**:156–168.

Osuna-Cruz, C.M., Paytuvi-Gallart, A., Di Donato, A., Sundesha, V., Andolfo, G., Aiese Cigliano, R., Sanseverino, W., and Ercolano, M.R. (2018). PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. Nucleic Acids Res. **46**:D1197–D1201.

Ou, S., and Jiang, N. (2018). LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. Plant Physiol. **176**:1410–1422.

Ou, S., Chen, J., and Jiang, N. (2018). Assessing genome assembly quality using the LTR Assembly Index (LAI). Nucleic Acids Res. **46**:e126.

Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics **23**:1061–1067.

Pedersen, B.S., and Quinlan, A.R. (2018). Mosdepth: quick coverage calculation for genomes and exomes. Bioinforma. Oxf. Engl. **34**:867–868.

Peng, Y., Yan, H., Guo, L., Deng, C., Wang, C., Wang, Y., Kang, L., Zhou, P., Yu, K., Dong, X., et al. (2022). Reference genome assemblies reveal the origin and evolution of allohexaploid oat. Nat. Genet. **54**:1248–1258.

Peng, C., Wu, D.-D., Ren, J.-L., Peng, Z.-L., Ma, Z., Wu, W., Lv, Y., Wang, Z., Deng, C., Jiang, K., et al. (2023). Large-scale snake genome analyses provide insights into vertebrate development. *Cell* Advance Access **186**:2959–2976.e22, published June 2023. https://doi.org/10.1016/j.cell.2023.05.030.

Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., and Toth, I.K. (2016). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal. Methods **8**:12–24.

Pruitt, K.D., Brown, G.R., Hiatt, S.M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O., Farrell, C.M., Hart, J., Landrum, M.J., McGarvey, K.M., et al. (2014). RefSeq: an update on mammalian reference sequences. Nucleic Acids Res. **42**:D756–D763.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: protein domains identifier. Nucleic Acids Res. **33**:W116–W120.

Rambaut, A. (2023). Figtree Advance Access.

Ranallo-Benavidez, T.R., Jaron, K.S., and Schatz, M.C. (2020). GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat. Commun. **11**:1432.

Rhie, A., Walenz, B.P., Koren, S., and Phillippy, A.M. (2020). Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. **21**:245.

Rose, L., Atwell, S., Grant, M., and Holub, E.B. (2012). Parallel Loss-of-Function at the RPM1 Bacterial Resistance Locus in Arabidopsis thaliana. Front. Plant Sci. **3**:287.

Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R., and Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. Nat. Rev. Genet. **3**:429–441.

SanMiguel, P., Gaut, B.S., Tikhonov, A., Nakajima, Y., and Bennetzen, J.L. (1998). The paleontology of intergene retrotransposons of maize. Nat. Genet. **20**:43–45.

Schulthess, A.W., Kale, S.M., Liu, F., Zhao, Y., Philipp, N., Rembe, M., Jiang, Y., Beukert, U., Serfling, A., Himmelbach, A., et al. (2022). Genomics-informed prebreeding unlocks the diversity in genebanks for wheat improvement. Nat. Genet. **54**:1544–1552.

Shi, X., Cui, F., Han, X., He, Y., Zhao, L., Zhang, N., Zhang, H., Zhu, H., Liu, Z., Ma, B., et al. (2022). Comparative genomic and transcriptomic analyses uncover the molecular basis of high nitrogen-use efficiency in the wheat cultivar Kenong 9204. Mol. Plant **15**:1440–1456.

Shumate, A., and Salzberg, S.L. (2021). Liftoff: accurate mapping of gene annotations. Bioinformatics **37**:1639–1643.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31**:3210–3212.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30**:1312–1313.

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res. **34**:W435–W439.

Steuernagel, B., Witek, K., Krattinger, S.G., Ramirez-Gonzalez, R.H., Schoonbeek, H.-J., Yu, G., Baggs, E., Witek, A.I., Yadav, I., Krasileva, K.V., et al. (2020). The NLR-Annotator Tool Enables Annotation of the Intracellular Immune Receptor Repertoire. Plant Physiol. **183**:468–482.

Sun, H., Ding, J., Piednoël, M., and Schneeberger, K. (2018). findGSE: estimating genome size variation within human and Arabidopsis using k-mer frequencies. Bioinformatics **34**:550–557.

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. **34**:W609–W612.

Tanguy, A.-M., Coriton, O., Abélard, P., Dedryver, F., and Jahier, J. (2005). Structure of Aegilops ventricosa chromosome 6Nv, the donor of wheat genes Yr17, Lr37, Sr38, and Cre5. Genome **48**:541–546.

International Wheat Genome Sequencing Consortium IWGSC, Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C.J., Choulet, F., Distelfeld, A., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science **361**:eaar7191.

Tseng, E. (2022). cDNA_Cupcake Advance Access published 2022.

Turner, M.K., Jin, Y., Rouse, M.N., and Anderson, J.A. (2016). Stem Rust Resistance in 'Jagger' Winter Wheat. Crop Sci. **56**:1719–1725.

Varden, F.A., Saitoh, H., Yoshino, K., Franceschetti, M., Kamoun, S., Terauchi, R., and Banfield, M.J. (2019). Cross-reactivity of a rice NLR immune receptor to distinct effectors from the rice blast pathogen Magnaporthe oryzae provides partial disease resistance. J. Biol. Chem. **294**:13006–13016.

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M.T., Brinton, J., Ramirez-Gonzalez, R.H., Kolodziej, M.C., Delorean, E., Thambugala, D., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. Nature **588**:277–283.

Wang, Y., Tang, H., DeBarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.h., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. **40**:e49.

Wang, J., Luo, M.-C., Chen, Z., You, F.M., Wei, Y., Zheng, Y., and Dvorak, J. (2013). Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. New Phytol. **198**:925–937.

Wang, S., Wang, H., Jiang, F., Wang, A., Liu, H., Zhao, H., Yang, B., Xu, D., Zhang, Y., and Fan, W. (2022). EndHiC: assemble large contigs into chromosome-level scaffolds using the Hi-C links from contig ends. BMC Bioinf. **23**:528.

Wang, X., Li, H., Shen, T., Wang, X., Yi, S., Meng, T., Sun, J., Wang, X., Qu, X., Chen, S., et al. (2024). A near-complete genome sequence of einkorn wheat provides insight into the evolution of wheat A subgenomes. Plant Commun **5**:100768.

Wicker, T., Matthews, D.E., and Keller, B. (2002). TREP: a database for Triticeae repetitive elements. Trends Plant Sci. **7**:561–562.

Wu, W., Gao, Y.-D., Jiang, D.-C., Lei, J., Ren, J.-L., Liao, W.-B., Deng, C., Wang, Z., Hillis, D.M., Zhang, Y.-P., et al. (2022). Genomic adaptations for arboreal locomotion in Asian flying treefrogs. Proc. Natl. Acad. Sci. **119**:e2116342119.

Xiao, G., Wang, W., Liu, M., Li, Y., Liu, J., Franceschetti, M., Yi, Z., Zhu, X., Zhang, Z., Lu, G., et al. (2023). The Piks allele of the NLR immune receptor Pik breaks the recognition of AvrPik effectors of rice blast fungus. J. Integr. Plant Biol. **65**:810–824.

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. **35**:W265–W268.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. **24**:1586–1591.

Yen, C., Yang, J., Yuan, Z., Ning, S., and Liu, D. (2020). Taxonomy of the Genus Triticum. In Biosystematics of Triticeae: Volume I. Triticum-Aegilops Complex, C. Yen and J. Yang, eds. (Springer), pp. 93–168.

Zhang, P., Li, W., Fellers, J., Friebe, B., and Gill, B.S. (2004). BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. Chromosoma **112**:288–299.

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., Zhang, J., Saski, C.A., Scheffler, B.E., Stelly, D.M., et al. (2015). Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nat. Biotechnol. **33**:531–537.

Zhang, J., Yang, F., Jiang, Y., Guo, Y., Wang, Y., Zhu, X., Li, J., Wan, H., Wang, Q., Deng, Z., et al. (2020). Preferential Subgenome Elimination and Chromosomal Structural Changes Occurring in Newly Formed Tetraploid Wheat—Aegilops ventricosa Amphiploid (AABBDvDvNvNv). Front. Genet. **11**:330.

Zhao, G., Zou, C., Li, K., Wang, K., Li, T., Gao, L., Zhang, X., Wang, H., Yang, Z., Liu, X., et al. (2017). The Aegilops tauschii genome reveals multiple impacts of transposons. Nat. Plants **3**:946–955.

Zhou, Y., Bai, S., Li, H., Sun, G., Zhang, D., Ma, F., Zhao, X., Nie, F., Li, J., Chen, L., et al. (2021). Introgressing the Aegilops tauschii genome into wheat as a basis for cereal improvement. Nat. Plants **7**:774–786.