nature communications

Article

https://doi.org/10.1038/s41467-026-70026-w

Population genomics reveals association of transposable elements variants with climatic adaptation in wild Amur grape

Received: 10 June 2025

Accepted: 14 February 2026

Published online: 26 February 2026

Check for updates

Zhiyao Ma[1,2,3,4,5,9], Xiaodong Xu[1,2,9], Wenjing Peng[1], Tianhao Zhang[1], Zhuyifu Chen[1], Shuo Cao[1], Fan Zhang[1], Yiwen Wang[2,6], Hua Xiao[1], Yingchun Zhang[1], Zhenya Liu[1], Zhongjie Liu[1], Hui Xue[1], Qiming Long[1], Ting Hou[1], Wenrui Wang[1], Yuting Liu[1], Zhongxin Jin[1], Mengyan Zhang[1], Yanling Peng[1], Jun Wen[7], Brandon Gaut[8] & Yongfeng Zhou[1,2]

Amur grape (*Vitis amurensis* Rupr.) is widely recognized for its cold tolerance traits and serves as a valuable genetic resource for breeding climate-resilient grape cultivars. Here, we construct a graph pangenome reference (Vampan_V1.0) and generate a variant map comprising 48,308,434 short variants and 127,094 TE-associated structural variants (TEVs) using deep resequencing data from 330 samples across 31 natural populations covering the species' distribution range. We discover a biased accumulation of SNPs around TEVs and identify 823 candidate adaptive genes associated with environmental variables. Using machine learning-based genetic offset models, we further show that putative adaptive TEVs significantly reduce genetic offsets by 7.3% to 8.2% under future climate scenarios. Our study shows the power of a graph-based pangenome to resolve complex variation and highlights the impact of TEVs on genetic diversity, local adaptation, and resilience to future climate change, providing insights into utilizing crop wild relatives in climate-resilient crop breeding.

Crop wild relatives represent an essential reservoir of genetic diversity for crop improvement and climate-resilient breeding[1–4]. Grapevine (*Vitis vinifera*) is one of the earliest domesticated and most economically significant fruit crops[5–8]. Its wild relatives, species of the *Vitis* genus (*Vitis* L.), are recognized as important germplasm resources[9,10], which play an essential role in grape breeding for resistance against both abiotic and biotic stresses, such as cold[11–14], heat[15], drought[16,17], salt[18,19], powdery mildew[20,21], and Pierce's disease[22,23]. Climate change has a significant impact on viticulture, with estimates suggesting that between 24 and 73% of current vineyard areas may no longer be suitable for grape growing by 2070[24–26]. Consequently, incorporating adaptive traits from wild *Vitis* species through selective breeding will be crucial for the adaptability and sustainability of viticulture in the future. In this study, we investigate the local adaptation and genomic

[1]State Key Laboratory of Tropical Crop Breeding, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, China. [2]State Key Laboratory of Tropical Crop Breeding, Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Key Laboratory of Synthetic Biology, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China. [3]Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, China. [4]State Key Laboratory of Plant Diversity and Specialty Crops, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, China. [5]Yunnan Key Laboratory of Forest Ecosystem Stability and Global Change, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, China. [6]Melbourne Integrative Genomics, School of Mathematics and Statistics, The University of Melbourne, Melbourne, VIC, Australia. [7]Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. [8]Department of Ecology and Evolutionary Biology, University of California, UC, Irvine, CA, USA. [9]These authors contributed equally: Zhiyao Ma, Xiaodong Xu. e-mail: zhouyongfeng@caas.cn

vulnerability to climate change of the wild Amur grape (*V. amurensis* Rupr.) (Fig. 1a). This species is distributed in the north temperate zone of East Asia, and exhibits remarkable cold and drought tolerance[11,13,27,28]. These traits have made *V. amurensis* a valuable genetic resource for grape breeding and an important system for studying stress tolerance in grapevine and other crops. However, despite its recognized agronomic and ecological importance, population-scale assessments of genetic diversity and adaptive potential across its natural range remain limited.

Understanding environmental responses in populations requires a comprehensive characterization of genetic variation across the genome, including complex genomic variants. Transposable elements (TEs), first discovered in maize (*Zea mays* L.) by Barbara McClintock in 1948, are ubiquitous mobile genetic elements, constituting a

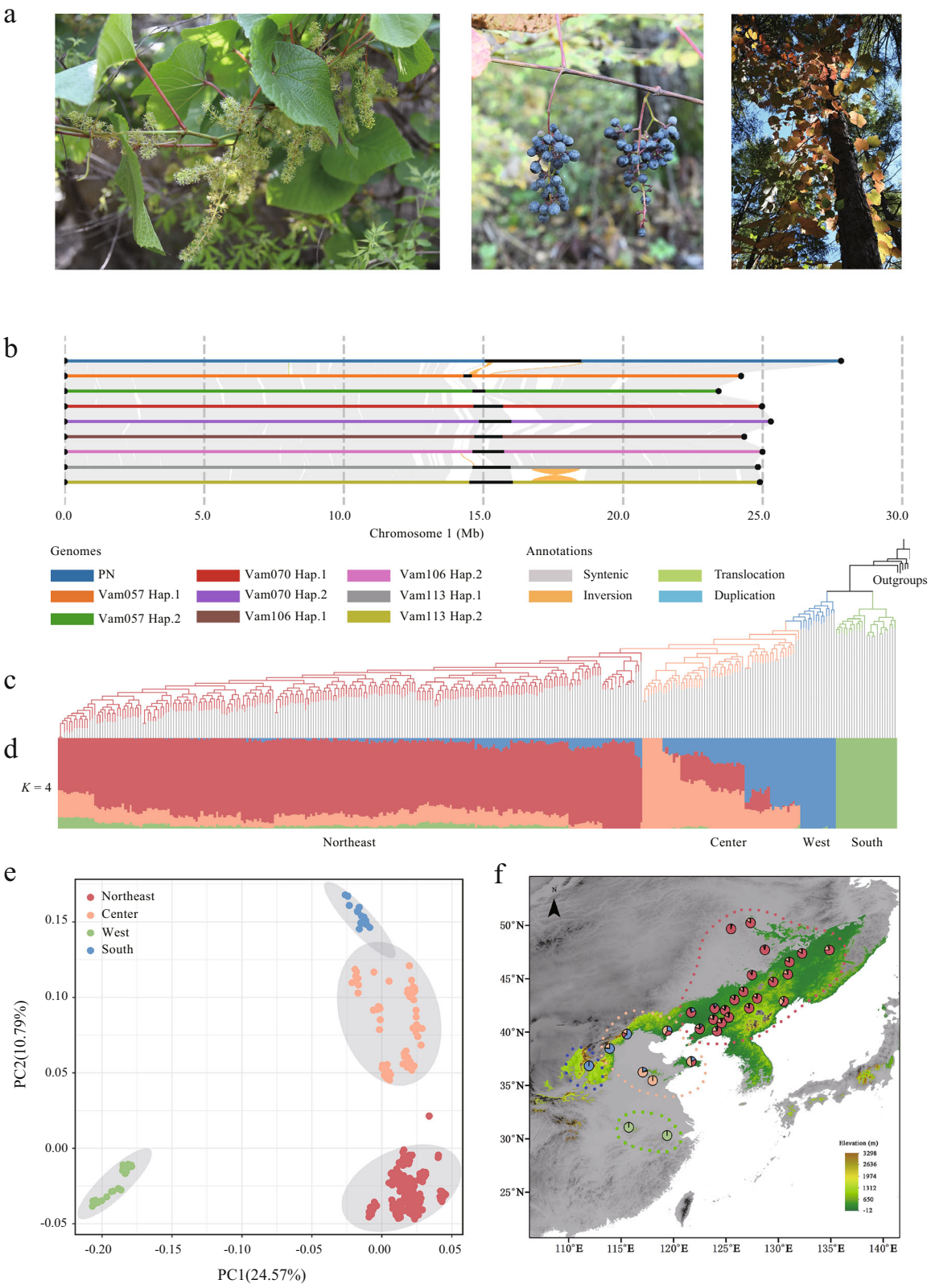

Fig. 1 | Morphology, genome assemblies, and population structure of *V. amurensis*. **a** Flowers, fruits, and natural habitat of *V. amurensis*. **b** Synteny among PN_T2T reference genome and eight haplotype genomes of *V. amurensis* (shown for chromosome 1 as a representative). Black solid dots indicate telomeres, and areas covered by black rectangles indicate centromeric regions. **c** Maximum Likelihood phylogenetic relationship of 330 accessions of *V. amurensis* with whole-genome LD-pruned SNPs. Four major geographic groups were identified: Northeast, Center, West, and South. **d** Model-based population assignment by ADMIXTURE analysis for $K = 4$ with whole-genome LD-pruned SNPs. **e** Principal component analysis (PCA) of all accessions. **f** Geographic distribution ranges of *V. amurensis* predicted by ecological niche models (ENMs). The ancestral components of 31 natural populations (circles) inferred by ADMIXTURE with whole-genome LD-pruned SNPs ($K = 4$). The geographic distribution of the four genome samples sequenced in this study has been marked on the map. The map is based on elevation data from WorldClim v2[159]. Source data are provided as a Source data file.

substantial proportion of eukaryotic genomes and contribute to plant genome evolution[29–35]. In plants, TEs often account for a large fraction of genome content, including approximately 63.9% of the genome in cultivated grape (*V. vinifera* L.)[36–38]. The interaction between TEs and the plant's epigenetic system can lead to the retention of TEs, thereby causing an expansion of plant genome sizes[39]. A number of TEs have been found to affect agricultural traits and cause genetic diseases, which significantly impacts advancements in biology, medicine, and agriculture[40–43]. Additionally, TEs affect plant gene expression and function through cis- and trans-regulatory mechanisms[44–52]. As an important source of genetic diversity and contributor of large-effect mutations, TEs play an important role in promoting genomic structural variation, phenotypic divergence, rapid adaptation, trait evolution, and speciation[53–61]. Population genetic analyses of TE-associated structural variants (TEVs) have revealed that most TEVs might be slightly deleterious or neutral, with rarer alleles compared to non-synonymous single-nucleotide polymorphisms (SNPs) in plant populations[59,62]. It is particularly crucial to explore the role of TEVs in plant genome evolution at the population level, as these evolutionary events are often recent and involve local adaptation[63]. However, the potential role of TEVs in shaping genetic diversity and their associations with climatic variables remain poorly understood.

Global climate change is an increasing threat to biodiversity and food security, potentially leading to extinctions of species, transformations of the global ecosystems, and a reduction in crop production[64,65]. Species may persist under climate change through genetic mutations, increased environmental tolerance via phenotypic plasticity, or migration to suitable locations[66–68]. Due to the relatively weak dispersal ability of plants compared to animals, climatic adaptation appears to be more important for the evolution and species persistence in plants[69]. The rapid development of genomic technologies provides an unprecedented opportunity to comprehensively examine genetic loci, including SNPs, indels, and structural variants (SVs) that have involved in climatic adaptations[70–75]. Detecting potential environment-associated genetic variants, including TEVs, is critical for understanding the genetic mechanisms of local adaptation and for predicting plant adaptability to climate change[72]. In addition, in spatially heterogeneous environments, species exhibit sensitivity to various selective pressures, leading to the evolution of local adaptations[2]. Genotype-environment associations (GEA), integrating genetic offset and niche suitability modeling, provide valuable insights into predicting local adaptive potential and assessing vulnerability to climate change in plants, thereby facilitating effective management and conservation to address the loss of global biodiversity and agricultural security[71,76–78]. Past studies have primarily focused on the impact of SNPs on climatic adaptation and the assessment of future climate-induced vulnerability[79,80], while the possible contribution of TEVs to local adaptation and response to future climate change remains rarely investigated.

In this study, we construct a graph-based pangenome reference from haplotype-resolved genomes of representative *V. amurensis* accessions, providing a comprehensive genomic framework for population-scale analyses. By integrating this pangenome with deep resequencing data of 330 accessions from 31 natural populations across the species' distributional range, we generate a high-resolution variant map including 41,451,941 SNPs, 6,856,493 indels, and 127,094 TEVs. This genomic resource captures genetic diversity at the population scale and provides a robust framework for examining the association of TEVs with climate adaptation. Our study establishes a foundational dataset for investigating climate-related genomic patterns in *V. amurensis* and offers valuable insights into breeding and conservation strategies for this important crop wild relative in the context of rapid global climate change.

Results

Haplotype-resolved genome assemblies and TE landscape of *V. amurensis*

To capture the genetic diversity of *V. amurensis*, we selected four representative accessions based on their geographic distribution (see "Methods"). For de novo assembly of the four newly generated haplotype-resolved *V. amurensis* genomes, we integrated a total of 17.7–28 Gb (~35–56×) of PacBio HiFi CCS reads, and ~58–78 Gb (~116–156×) of Hi-C reads. The final assembly of haplotype-resolved genomes resulted in total lengths ranging from 497.35 to 512.02 Mb, with contig N50 ranging from 15.14 to 24.15 Mb (Supplementary Data 1). The assembly consensus quality values (QVs) for haplotype-resolved genomes ranged from 67.29 to 69.11 based on HiFi data, indicating over 99.99% accuracy (Supplementary Data 1). More than 98.3% of the core embryophyte gene sets from the benchmarking universal single-copy orthologs (BUSCOs) assessment were detected in the haplotype-resolved genomes of each accession, supporting that our assemblies were of high quality and completeness (Supplementary Fig. 1). Furthermore, most of the centromeric and telomeric regions were identified on *V. amurensis* diploid genomes (Fig. 1b, Supplementary Fig. 2a, Supplementary Data 1). Approximately 63.13–64.08% of the haplotype sequences were identified as repetitive elements, including 36.72–37.76% retrotransposons, 16.79–17.25% DNA transposons, and 0.18–0.22% rolling-circle elements (Supplementary Table 1). TE insertion time analysis across the eight haplotype assemblies of *V. amurensis* revealed a broad peak around 5 million years ago (Mya) in the temporal distribution of TE insertions, followed by a sharp decline in more recent insertions, indicating that most TE insertions are relatively ancient (Supplementary Fig. 2b).

Population genetic dynamics of TEVs in *V. amurensis* populations

We collected 330 individuals from 31 populations across the main distribution areas of the *V. amurensis*. Whole-genome resequencing of these accessions generated ca. 7.37 Tb raw data, with an average coverage of ca. 40× per sample (ca. 22.33 Gb per individual; see Supplementary Data 2). We identified 41,451,941 SNPs and 6,856,493 indels based on a single-reference approach by mapping all resequencing reads to the Vam057 haplotype1 genome (see Methods).

We further filtered the SNP set using linkage disequilibrium (LD) pruning (see Methods), resulting in 344,087 independent SNPs for population genetic diversity analyses. The maximum-likelihood (ML) tree of *V. amurensis* supported an early divergence between populations of the southern range and those from the West-Center-Northeast range of the species, while populations distributed in the northeast represented later divergences (Fig. 1c and Supplementary Fig. 3). Population structure and principal component analysis (PCA) revealed four major geographic groups: West, Center, Northeast, and South

(Fig. 1d–f, Supplementary Fig. 4a, Supplementary Data 2). We further confirmed that the four genomes (Vam057, Vam070, Vam106, and Vam113) were assigned to the Northeast, Center, West, and South groups, respectively (Fig. 1 and Supplementary Data 1).

To better detect structural variants, we utilized the PanGenome Graph Builder (PGGB)[81] to build the *V. amurensis* pangenome graph (Vampan_V1.0) using the eight haplotype assemblies, which comprises 54.16 million nodes and 74.84 million edges, with a total sequence length of 832.77 Mb. The Vampan_V1.0 pangenome, generated from phased diploid assemblies, provides a pivotal framework for comprehensive analyses of haplotype diversity and the identification of TEVs on a population scale.

The utilization of a pangenome for *V. amurensis*, combined with the extensive resequencing data generated in this study, facilitated the identification of a broad spectrum of genomic variants. In total, 220,864 structural variants were identified, among which 127,094 were TEVs, which accounted for 57.54% of all structural variants (see "Methods") (Fig. 2a). Population structure analysis based on TEVs yielded results consistent with those obtained from SNPs (Supplementary Fig. 4b). Statistics on the number of TEVs in each accession indicated that 91% of the accessions had at least 36,000 TEVs (Supplementary Fig. 5a). In addition, the main type of TEs in the *V. amurensis* population was retroelements (71.28%), including 16.66% Copia LTRs and 44.77% Gypsy LTRs, compared to DNA transposons (26.04%) (Fig. 2b). Based on high-quality 97,013 extensive TEVs and 4,114,075 SNPs (see "Methods"), the extents of population genetic diversity (π) of SNPs and TEVs over 1 Mb non-overlapping windows were calculated, which demonstrated that there was a significant positive correlation between the two types of variants ($P = 1.19e{-}13$, Fig. 2c). Moreover, the number of SNPs in surrounding regions of TEVs was calculated, and our results suggested that SNP frequency around TEVs is higher than the average SNP frequency across the genome (Fig. 2d). A similar pattern was observed for SVs (including TEVs), with SNPs also exhibiting biased accumulation around structural variants (see Methods, Supplementary Fig. 5b).

The distribution of TEVs in the whole genome indicated that most TEVs were located at intergenic regions of the genome (79.21%). Approximately 26.11% TEVs were identified at the cis-regulatory regions (within 2 kb upstream and downstream of genes), and ca. 23.16% TEVs were distributed in genic regions, including 3′ untranslated region (UTR), 5′ UTR, introns, and coding sequences (CDSs) (Supplementary Table 2, Supplementary Fig. 5c). These reported percentages are not mutually exclusive, as individual TEVs may span multiple genomic features. The frequency of each TEV was calculated as the proportion of haplotypes carrying the variant allele in the population. Most TEVs were observed at allele frequency ≤0.5 across intergenic and genic regions, potentially due to either recent mobilization events or purifying selection which constrained the spread of TEVs throughout the genome (Fig. 2e). TEVs with a population frequency greater than 0.8 in the CDS regions are likely under positive selection in natural populations (Fig. 2e). To explore their potential functional impact, we performed gene ontology (GO) enrichment analysis on genes containing TEVs with an allele frequency >0.8 in CDS regions. A total of 201 genes were annotated with GO biological processes terms and used for enrichment analysis. These genes showed significant enrichment in categories related to defense response ($P = 0.0078$), defense response to other organisms ($P = 1.99e{-}4$), and regulation of innate immune response ($P = 0.0081$) (Fig. 2f). We acknowledge that these gene categories may include members under relaxed selection or colocalized with TEVs without direct functional consequences.

TEVs can regulate gene expression by rearranging gene fragments or disrupting cis-regulatory elements, which are generally considered deleterious. To estimate the effect of TEVs on the genome, we analyzed the unfolded site frequency spectra (SFS), which show how often newly derived variants occur at different frequencies in a population, using TEVs and SNPs across the genome from the four main groups (Supplementary Fig. 5d). This approach has been widely applied in population genomic studies of grapevine[6,59]. In each group, SNP frequencies showed a U-shaped distribution, whereas TEVs demonstrated a lower proportion of fixed variants compared with SNPs (Supplementary Fig. 5d). A total of 39,212 TEVs were shared among all four groups, and the South group contained the highest number of group-specific TEVs (8852), which were not detected in the other groups (Supplementary Fig. 6).

Potential role of TEVs in the local adaptation of *V. amurensis*

To explore the association of TEVs to local adaptation, we identified high-quality 97,013 TEVs, 4,114,075 SNPs, and 677,438 indels (see "Methods") and performed genotype–environment association (GEA) analyses for *V. amurensis* populations. Genomic variants associated with 20 environmental variables, including temperature, precipitation, and elevation-related variables, were identified using Latent factor mixed model (LFMM)[82] (Supplementary Table 3, Supplementary Data 3). Moreover, the importance of each environmental variable for *V. amurensis* was assessed by the Gradient Forests (GF) model (Supplementary Fig. 7a). Through the integration of environmental variable importance rankings and pairwise correlations among twenty variables, six variables with Spearman correlation coefficient $|r| < 0.8$ were identified, including temperature seasonality (bio4), elevation (elev), precipitation seasonality (bio15), isothermality (bio3), precipitation of warmest quarter (bio18), and mean temperature of warmest quarter (bio10) (Supplementary Fig. 7). These six environmental variables can largely explain the adaptive environment-associated genomic variants across the distribution of *V. amurensis* based on the GF model. Temperature seasonality (bio4) was identified as the most important driver of local adaptation of natural populations distributed in the northeastern range (Supplementary Fig. 8). In contrast, precipitation seasonality (bio15) and mean temperature of warmest quarter (bio10) significantly influenced the putative adaptive genomic variants in most of the remaining natural populations of *V. amurensis* (Supplementary Fig. 8).

A total of 22,419 putative adaptive variants were identified using the LFMM analysis (False discovery rate, FDR < 0.01), including 1098 TEVs and 21,321 short variants. Based on the linkage disequilibrium (LD) information of *V. amurensis* (Supplementary Fig. 9a), the upstream and downstream regions of 20 kb of candidate adaptive variants were used to detect environment-associated genes. A total of 823 candidate genes were identified by the LFMM analysis of six environmental variables (Supplementary Data 4–9). We identified a set of precipitation-associated candidate adaptive genes that are homologous to genes in *Arabidopsis*, including *TLP3* (Fig. 3, Supplementary Figs. 9b and 10). The *V. amurensis* homolog of *TLP3* was associated with precipitation-related environmental variables, including precipitation seasonality (bio15) and precipitation of warmest quarter (bio18), and was involved in drought stress signaling (Fig. 3a and Supplementary Fig. 10c). A candidate adaptive TE insertion (TE5_3254586, with a length of 10,631 bp, and classified as a Gypsy LTR) was located in the putative cis-regulatory region of the gene *TLP3* (ca. 1 kb upstream of the 5′UTR) associated with precipitation seasonality (bio15), and there was relatively high linkage disequilibrium within the region surrounding this gene (Fig. 3a). This long TE insertion may affect a transcription factor binding site (chr5: 3255018) predicted using the PlantCARE database, suggesting a potential influence on the transcriptional regulation of *TLP3* (Supplementary Fig. 9b). Our results indicated that there was a positive correlation ($r = 0.67$, $P = 3e{-}5$) between the precipitation seasonality and the allele frequency of the TE among 31 natural populations of *V. amurensis* (Fig. 3b). The geographic distribution of the allele frequencies demonstrated that a higher prevalence of the TE allele in populations of the Center and

https://doi.org/10.1038/s41467-026-70026-w

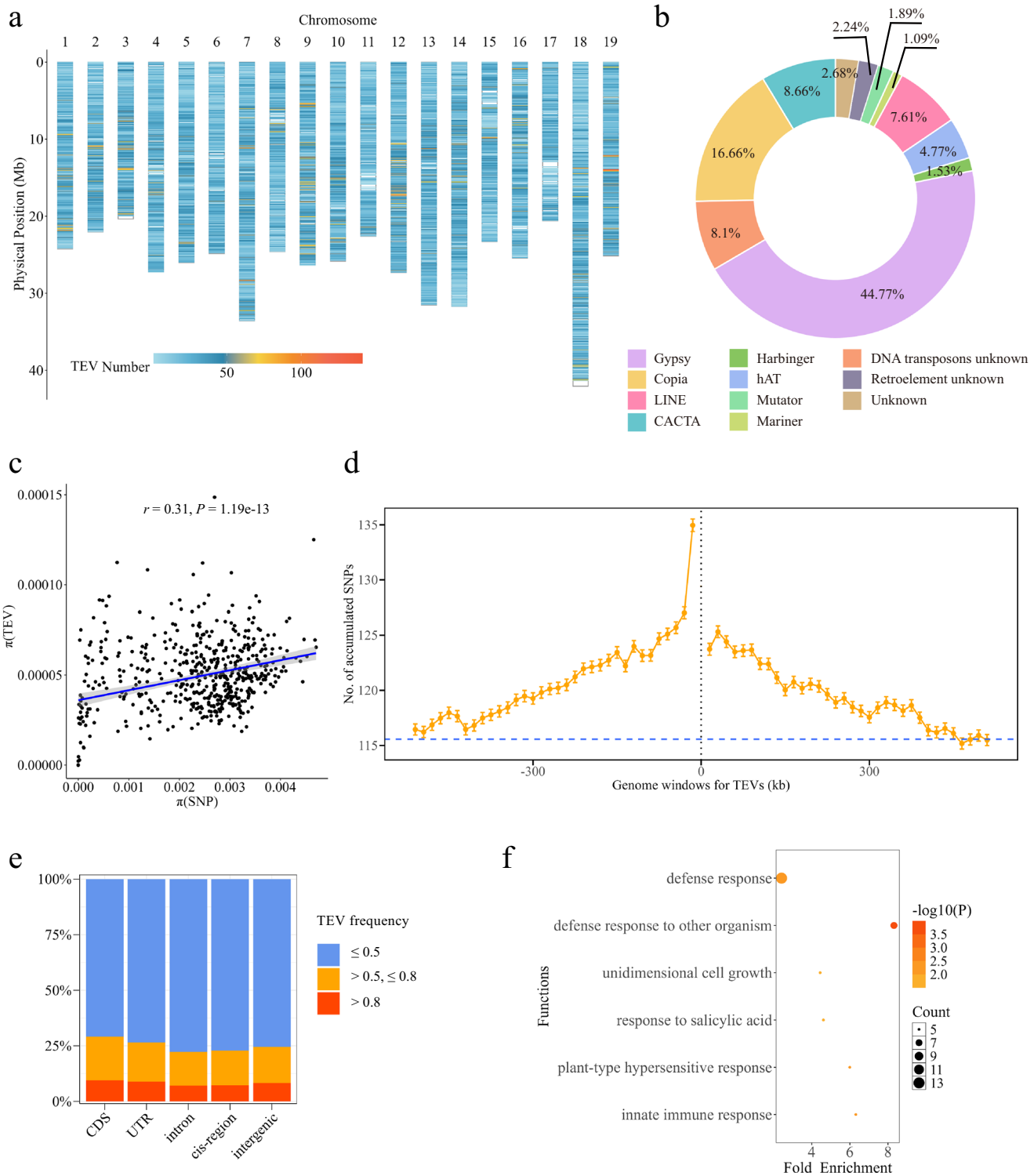

Fig. 2 | Population genetic characteristics of TEVs in *V. amurensis* populations. **a** TEV density across the whole genome, based on the Vam057 haplotype1 reference. **b** Composition of TEVs in the population of *V. amurensis*. **c** Correlation between TEVs and SNPs diversity (π) of *V. amurensis* populations with a window size of 1 Mb. The value of *r* represents the Pearson correlation coefficient between genetic diversity (π) of SNPs and TEVs. The blue line represents the fitted linear regression line based on the least squares method, and the shaded area indicates the 95% confidence interval. **d** A biased accumulation of SNPs around TEVs for 15-kb non-overlapping windows. Each point represents the mean number of SNPs per window flanking 97,013 independent TEVs, and error bars represent the standard error of the mean (SEM). Dashed line: average number of SNPs per 15 kb sliding windows across the genome. **e** Frequency of TEVs in different genomic regions. **f** Gene ontology (GO) analysis of genes containing TEVs (allele frequency > 0.8) in their CDS regions. Statistical significance was determined using a one-sided Fisher's Exact Test. Intergenic region: a region of the genome other than the genic region; Cis-region: within 2 kb upstream and downstream of genes; Genic region: including UTR, CDS, and intron; UTR untranslated region, CDS coding sequences. Source data are provided as a Source data file.

https://doi.org/10.1038/s41467-026-70026-w

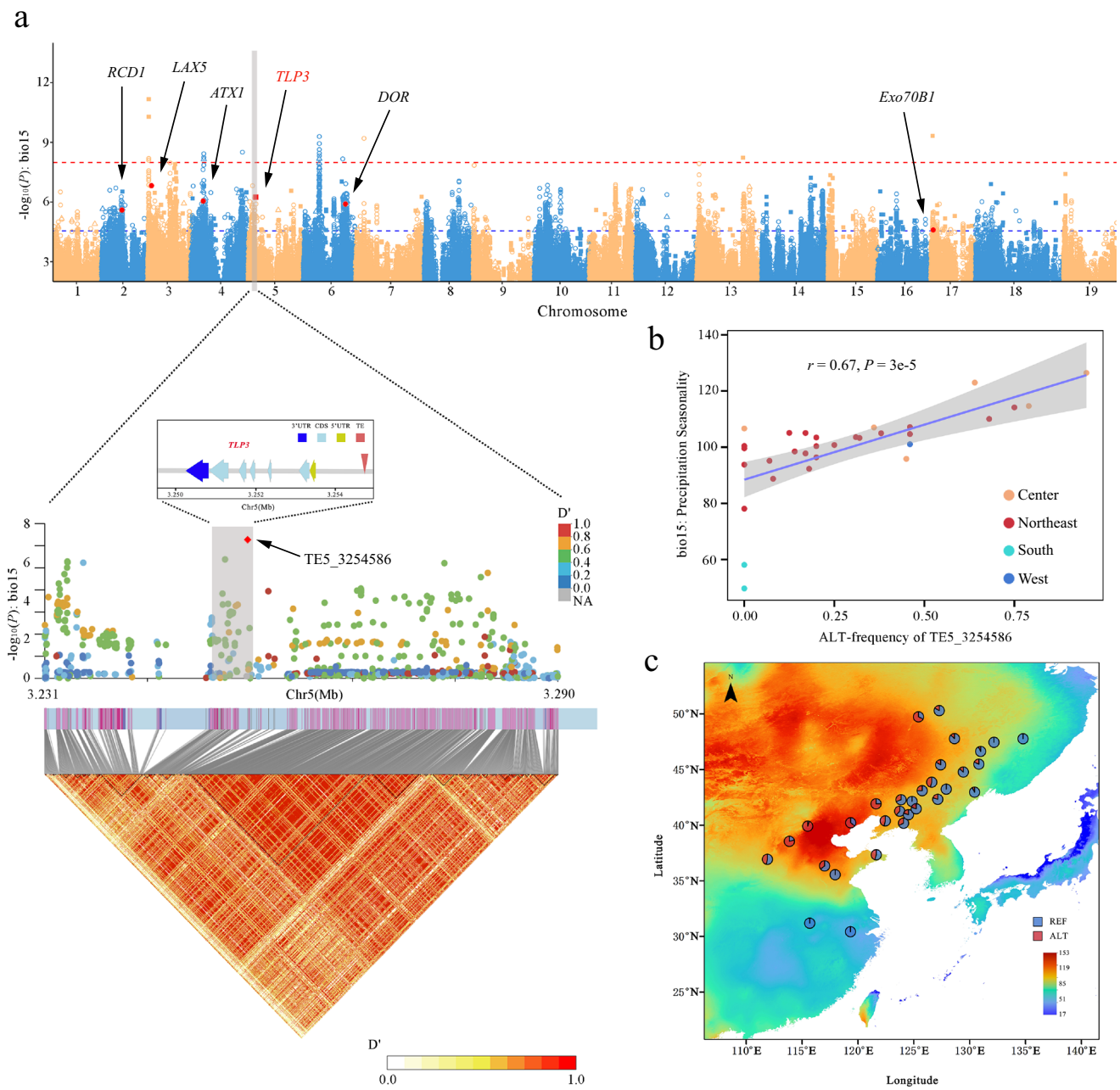

Fig. 3 | Genomic mapping of adaptive variants associated with a precipitation-associated environmental variable. a Manhattan plot for the genotype-environment association of the precipitation seasonality (bio15). Upper panels: the gene structure of *TLP3*. Lower triangle LD heatmaps of the associated regions. **b** Pearson correlation was used to test the association between environment variable and allele frequency. The line represents the fitted linear regression line based on the least squares method, and the shaded area indicates the 95% confidence interval. **c** Allele frequency of candidate adaptive TE variation (TE5_3254586) associated with bio15 across the 31 natural populations. The map is based on bio15 data from WorldClim v2[159]. SNPs (hollow circle), indels (hollow triangle), and TEVs (solid square) are showed in Manhattan plots. Lower and upper dashed horizontal lines represent genome-wide significance thresholds of FDR = 0.01 and the Bonferroni, respectively. Selected candidate genes are labeled in the plot at their respective genomic positions. Source data are provided as a Source data file.

West groups. In contrast, this allele was relatively rare in populations of the South and Northeast groups (Fig. 3c). Numerous candidate genes were identified to be associated with temperature-related environmental variables, such as *ICE1*, *CAMTA2*, *CBF1*, *HHP4*, *COR27*, *BAM3* and *BAG7* (Fig. 4a and Supplementary Fig. 10). It is noteworthy that the gene *TLP3* was also strongly associated with two temperature-related environmental variables, temperature seasonality (bio4) and mean temperature of warmest quarter (bio10) (Fig. 4a and Supplementary Fig. 10b). The tubby-like protein (TLP) gene family has also been reported to play a role in plant responses to extreme temperature stress[83]. Allelic variation in the gene *HHP4* was associated with temperature seasonality (bio4) (Fig. 4a). The heptahelical protein (HHP) gene family has been implicated in multiple abiotic stresses, including extreme temperature and salt[84]. In addition, the gene *ICE1*, a transcriptional activator regulating the cold-induced transcription of CBF/DREB1 genes, was associated with variation in temperature seasonality (Fig. 4a). We selected the candidate adaptive SNP (SNP14_26191757, representing a G-to-A substitution), located in the putative cis-regulatory region (ca. 1 kb upstream of the 5′UTR) of the gene *ICE1*, as the best example to illustrate the relationship between the allele frequency and the environmental variable (Fig. 4a). According to the PlantCARE database, this SNP is located within a predicted

https://doi.org/10.1038/s41467-026-70026-w

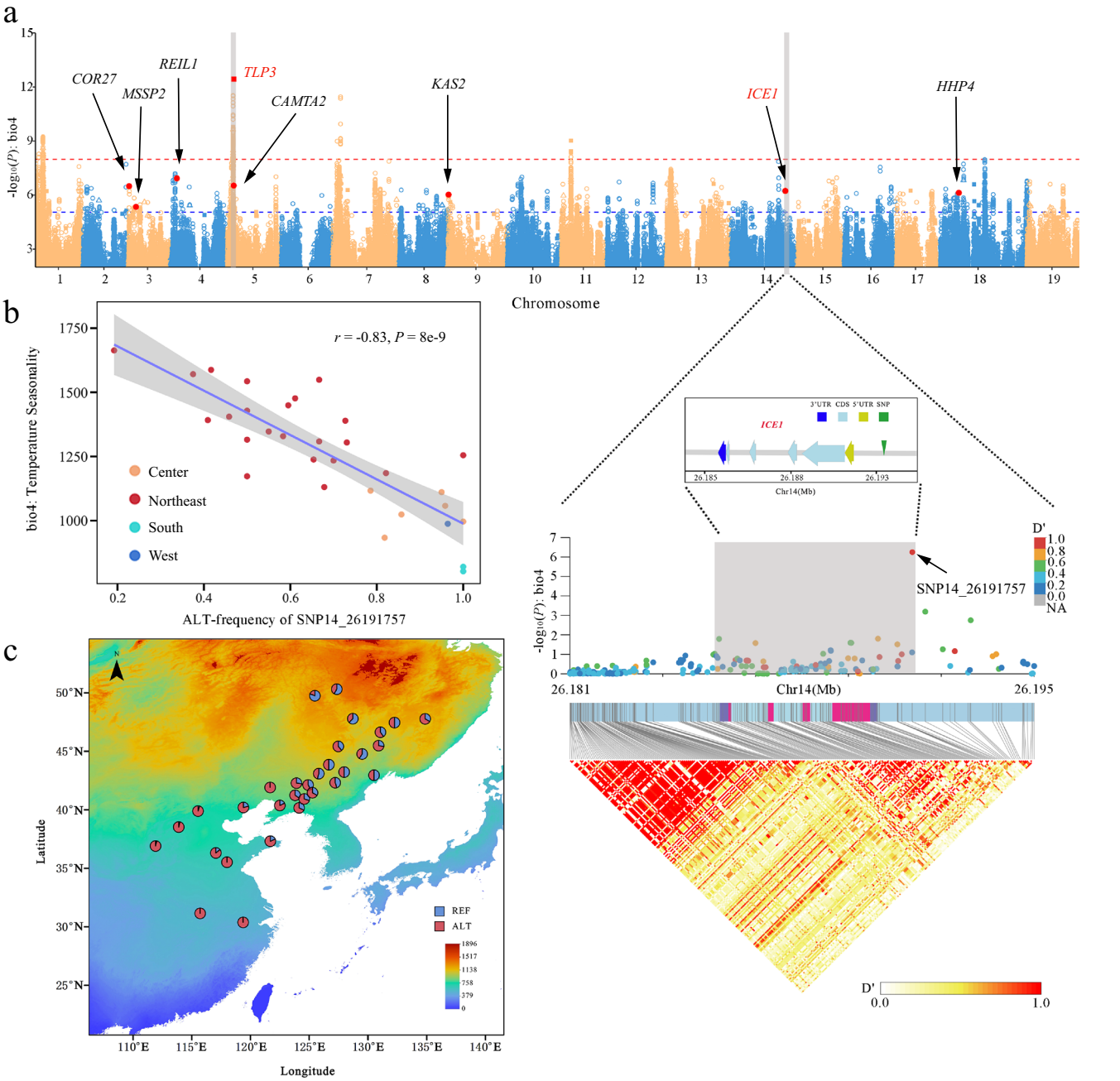

Fig. 4 | Genomic mapping of adaptive variants associated with a temperature-related environmental variable. a Manhattan plot for the genotype-environment association of the temperature seasonality (bio4). Upper panels: the gene structure of *ICE1*. Lower triangle LD heatmaps of the associated regions. **b** Pearson correlation was used to test the association between environment variable and allele frequency. The line represents the fitted linear regression line based on the least squares method, and the shaded area indicates the 95% confidence interval. **c** Allele frequency of candidate adaptive variant (SNP14_26191757) associated with bio4 across the 31 natural populations. The map is based on bio4 data from WorldClim v2[159]. SNPs (hollow circle), indels (hollow triangle), and TEVs (solid square) are showed in Manhattan plots. Lower and upper dashed horizontal lines represent genome-wide significance thresholds of FDR = 0.01 and the Bonferroni, respectively. Selected candidate genes are labeled in the plot at their respective genomic positions. Source data are provided as a Source data file.

transcription factor binding site (chr14: 26191744), which may alter the transcriptional regulation of *ICE1* (Supplementary Fig. 9c). Our results revealed a strong negative correlation ($r = -0.83$, $P = 8e{-}9$) between the temperature seasonality and the allele frequency of this SNP among 31 natural populations of *V. amurensis* (Fig. 4b). The distribution pattern of allele frequencies suggested that the reference allele was predominantly found in populations of Northeast Asia, characterized by relatively high values of temperature seasonality, whereas the alternative allele was fixed in populations located in areas with low values of temperature seasonality (Fig. 4c). In addition, thirteen genes were identified as candidates for elevation-related adaptation, associated with UV-B/photo-oxidative damage protection and freezing tolerance (Supplementary Fig. 10d). The identified candidate genes were subjected to GO enrichment analysis. These genes showed significant enrichment in biological processes such as response to cold ($P = 5.92e{-}11$), response to water deprivation ($P = 9.87e{-}9$), and response to abscisic acid ($P = 1.1e{-}6$) (Supplementary Table 4). These results suggest that some candidate genes may function within coordinated stress-response pathways, collectively contributing to environmental adaptation.

We estimated the proportion of environmental variance (PVE) associated with different genetic variants. TEVs were identified as associated with the largest proportion of variance in mean temperature of warmest quarter (bio10, 71.97%), precipitation of warmest quarter (bio18, 87.01%), and elevation (elev, 86.64%) (Supplementary Fig. 11). The unique contribution of TEVs to most environmental variables was larger than that of SNPs or indels when compared individually (Supplementary Fig. 11). TEVs uniquely explained 11.12% of the genetic variation associated with precipitation of warmest quarter, a much larger proportion than that explained by SNPs or indels (Supplementary Fig. 11e). Although the unique variance associated with TEVs was relatively lower than that of SNPs for isothermality (bio3), temperature seasonality (bio4), and precipitation seasonality (bio15), TEVs were still identified as associated with these environmental variables (Supplementary Fig. 11). These results suggest that TEVs may contribute to ecological adaptation in *V. amurensis*, representing a component of adaptive genetic diversity.

The role of TEVs in genomic adaptation to future climate change

To evaluate how populations of *V. amurensis* may respond to future climate change, shared putative adaptive variants identified by LFMM and the redundancy analysis (RDA) that were consistently associated with the six uncorrelated environmental variables (bio4, elev, bio15, bio3, bio18, and bio10) were defined as "core adaptive variants" for local adaptation (Supplementary Fig. 12). These variants were then further filtered (see "Methods"), resulting in 40 LD-pruned adaptive TEVs, along with 206 adaptive SNPs and indels. We predicted the genomic vulnerability of *V. amurensis* to future climate change using the LD-pruned adaptive SNPs and indels under SSP245 and SSP585 scenarios in two periods (Fig. 5 and Supplementary Fig. 13, see "Methods"). This involved estimating the genetic offset, representing the expected degree of allele frequency change required for populations to remain adapted under future climate scenarios. The SSP585 model was regarded as projecting higher emissions and more severe climate change than the SSP245 model. Genetic offset demonstrated an upward trajectory in correlation with escalating emissions when assessed across different emission scenarios (Fig. 5 and Supplementary Fig. 13). Populations with higher genetic offsets are particularly vulnerable to climate change because they require more substantial changes in adaptive allele frequencies to adapt to future climate.

To explore the role of TEVs in shaping the adaptive potential of *V. amurensis* under future climate change, we estimated genetic offsets using LD-pruned adaptive TEVs, SNPs, and indels. We specifically assessed the potential role of TEVs by comparing models that included LD-pruned adaptive TEVs alongside SNPs and indels (adaptive TEVs set) with those using only LD-pruned adaptive SNPs and indels (baseline set). Across all periods and emission scenarios, the baseline set showed significantly higher genetic offsets ($P < 0.003$), whereas inclusion of putative adaptive TEVs reduced offsets by 7.3–8.2% (Supplementary Table 5). Results from different datasets showed similar spatial and temporal genetic offset patterns (Fig. 5 and Supplementary Figs. 13–15). Populations of the Northeast group showed relatively high genetic offsets, indicating greater vulnerability to future climate change (Supplementary Fig. 15). As a control, the average genetic offsets based on random non-adaptive TEVs plus LD-pruned adaptive SNPs and indels (control set) showed no significant difference from the baseline set (Fig. 5c). These findings suggested that putative adaptive TEVs may play a role in future climate adaptation of *V. amurensis* by providing complementary adaptive genetic variation that reduces genetic offset.

To adapt to rapidly changing environments, plants must either change allele frequencies in situ or migrate to other suitable habitats. The role of migration in population-level adaptability to climate change was considered and integrated into this study. We assessed the potential maladaptation based on forward genetic offsets under SSP585 models during the period 2081–2100, considering different maximum dispersal distances (100, 500, and 1000 km), respectively (Supplementary Fig. 16). Restricting the maximum migration distances will increase forward offsets (Supplementary Fig. 16). Compared to analyses that do not consider migration, we inferred that migration benefits *V. amurensis* populations by facilitating TE exchange and enhancing their adaptation to the rapidly changing global climate.

Discussion

This study constructs a graph-based pangenome reference and a comprehensive variants map, revealing a biased accumulation of SNPs around TEVs. In addition, we investigate the potential role of TEVs in local adaptation to present and future climate for *V. amurensis* survival in rapidly changing environments. Firstly, we estimated the possible contribution of different variants to local adaptation and identified environment-associated candidate genes. For example, *TLP3* is involved in drought and cold stress responses, and key genes in the *ICE-CBF-COR* signaling pathway, which are crucial for plant cold stress tolerance. These genes are inferred to promote the ability of *V. amurensis* to withstand freezing temperatures. Secondly, using a machine learning model, we predicted the spatiotemporal adaptability of *V. amurensis* populations to future climate change, with the results suggesting that TEVs may contribute to enhancing adaptive potential to climate change. Overall, we provide new evidence supporting the potential role of TEVs in adaptive evolution and highlight the importance of integrating population genomics approaches in future studies on the breeding and conservation of crops and their wild relatives under climate change.

TEs have long been recognized as important drivers of genetic innovation and genome evolution[85–88]. TEs can create genetic variability through insertions and rearrangements, which can result in the reprogramming of gene expression or the introduction of new functional elements[89,90]. The insertion of TE can enhance genomic instability, modify the surrounding sequence contexts, and affect DNA repair and recombination rates of closely linked genomic regions[29,91,92]. Reduced recombination rates around TEVs may lead to relaxed purifying selection, resulting in the biased accumulation of SNPs in these regions[93]. Alternatively, the observed pattern may result from variant purging in more conserved genomic regions. We infer that these effects collectively influence mutation patterns and may lead to hotspots of genetic variation in regions around TEVs (Fig. 2).

Local adaptation is essential for species survival in rapidly changing ecological environments by influencing species diversity and facilitating the expansion of species' ranges[94]. Exploring the genetic basis of local adaptation is critical for understanding species' adaptive evolution[95]. Previous population genomic studies mainly focused on the impact of SNPs on local adaptation. However, the role of TEVs in adaptive evolution has attracted increasing attention over the past decade[47,96–100]. As a primary cause of structural variation, TE provides significant genetic diversity[40]. Large-effect variants may play a crucial role in driving local adaptation, rather than contributing significantly to global adaptation[101]. It is interesting that the mobilization of TE facilitates the generation of large-effect mutations across the genome and generations, potentially promoting rapid local adaptation[102]. To investigate the diversity of TEVs at the population level and the possible contribution of TEVs to adaptive evolution in the context of rapid environmental change, we identified TEVs based on a graph pangenome and whole-genome resequencing data from 330 accessions of *V. amurensis*. Our findings support the hypothesis suggesting TEVs may contribute to adaptive evolution (Fig. 3 and Supplementary Fig. 10). In this study, TEVs show possible contributions to the environmental adaptation of *V. amurensis*, particularly for precipitation of the warmest quarter (Supplementary Fig. 11). TEVs located near or within genes are recognized to have significant transcriptional effects, and when TEVs insert into introns, they may influence phenotype by

https://doi.org/10.1038/s41467-026-70026-w

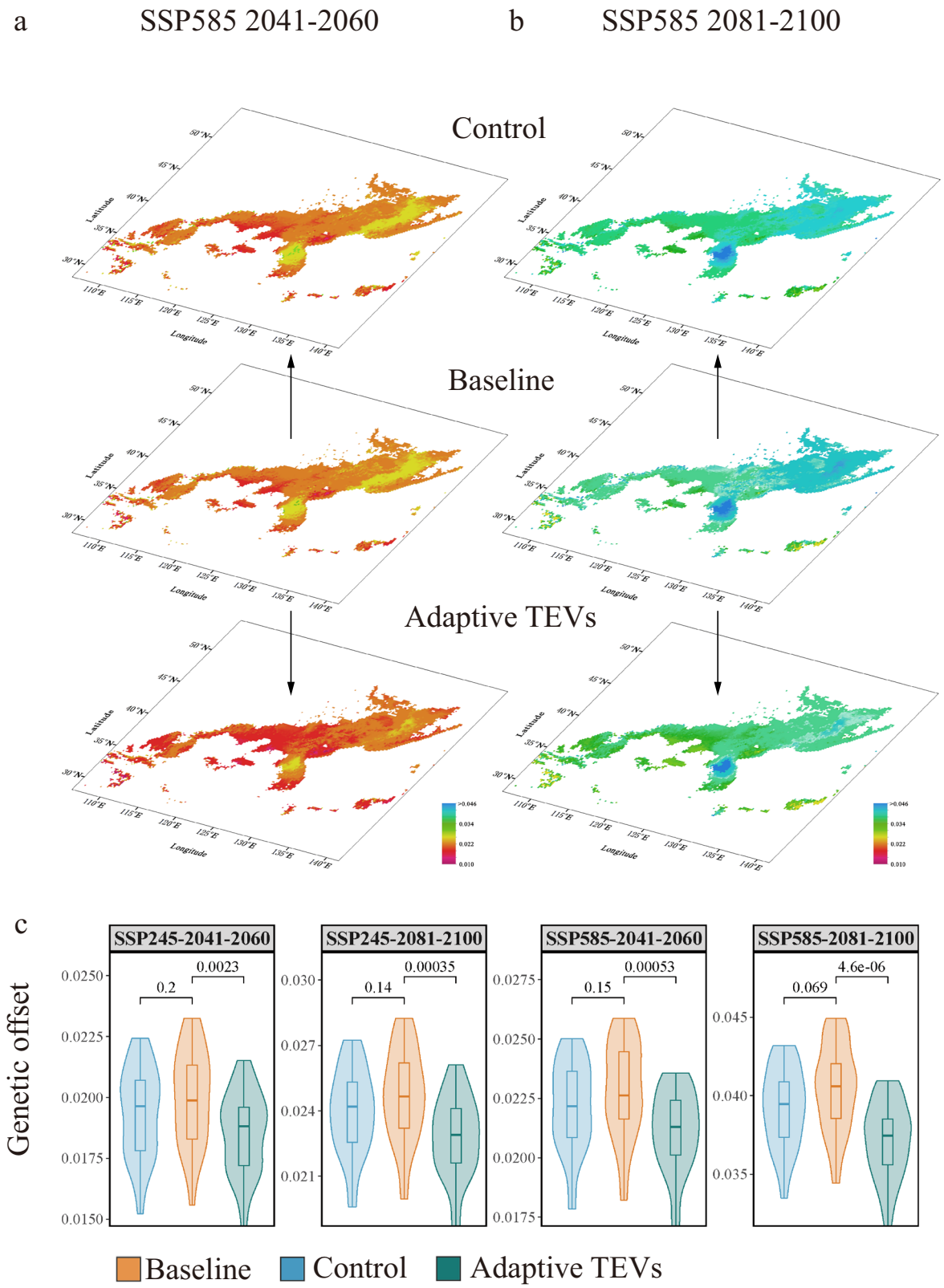

Fig. 5 | Predicted averaged genetic offsets to future climate change. a, **b** Genetic offset maps across the distribution range of *V. amurensis* based on different adaptation-associated variants under SSP585 scenario in 2041–2060 and 2081–2100, respectively. Maps are based on bioclimatic data from WorldClim v2[159]. **c** Genetic offsets of 31 sampled populations based on different adaptation-associated variants. Baseline set: only adaptive SNP and indel variants, without TE variants. Adaptive TEVs set: baseline set combined with the adaptive TE variants. Control set: baseline set combined with random non-adaptive TE variants. The box edges represent the interquartile range (IQR, 25th–75th percentile), the horizontal lines inside boxes represent median values, and whiskers extend to the largest and smallest values within 1.5 × IQR. Points outside this range are shown as outliers. Numbers above the bars indicate *P* values derived from two-sided Student's *t*-tests. Source data are provided as a Source data file.

generating multiple transcript isoforms[103]. Identifying adaptive genes and exploring the effects of TEVs on local adaptation provide valuable insights into establishing breeding strategies in grapes to enhance tolerance to abiotic stresses.

Rapid climate change is threatening global biodiversity and food security[104]. However, comprehensive predictions of species vulnerability to future climate change remain challenging, due to the incomplete information on the genetic variants at the genome scale. The fluctuation in TE transposition rate may impact the adaptability of plants to future climate change, yet the possible contribution of TEVs to the response to such change has long been ignored[105]. In this study, we identified SNPs and indels based on a single linear reference genome and detected TEVs using a graph-based pangenome in *V. amurensis* populations. Our findings indicate that genetic offsets of populations are significantly decreased ($P < 0.003$) with additional putative adaptive TEVs as covariates at all periods and emission scenarios (Fig. 5 and Supplementary Table 5). We suggest that TEVs may contribute to the adaptation of populations to future climate change.

Cold and drought stresses are important environmental factors that influence plant growth, development, production, and geographic distribution. The mechanism of cold acclimation in temperate crops involves the regulation of a series of transcription factors and proteins. *V. amurensis*, the most cold-resistant wild grape in East Asia, is recognized as a valuable genetic resource and has received considerable attention recently[11,12,14,106]. However, the genetic basis of cold tolerance of *V. amurensis* is still not adequately investigated. In this study, we identify a set of candidate genes potentially associated with environmental adaptation, including the gene *TLP3*, which may be involved in drought, cold, and heat stress responses (Figs. 3 and 4, Supplementary Fig. 10). Tubby-like proteins (TLPs), eukaryotic proteins with a Tubby domain, are important in plant abiotic stress tolerance[83,107–113], such as tolerance to drought[114–118], cold stress[119,120], and high temperature[121,122]. Moreover, the gene *BAM3* has been identified as being associated with adaptation to temperature-related and elevation-related environmental factors (Supplementary Fig. 10). A recent study confirmed that *BAM3* enhances cold tolerance through the regulation of soluble sugar accumulation in *V. amurensis*[123]. The *ICE-CBF-COR* signaling pathway in plants is considered as a significant regulatory mechanism for adapting to cold stress[124,125]. Our results identify several candidate genes, including *COR27*, *ICE1*, *CBF1*, *CAMTAs*, *HHPs*, *CRPK1*, *CRLK1*, and *HOS15*, potentially involved in this pathway, suggesting their important role in freezing adaptation (Fig. 4 and Supplementary Fig. 10). For example, *COR27* is involved in regulating flowering and freezing tolerance of *Arabidopsis thaliana*[126]. *ICE* gene homologs of *V. amurensis*, *VaICE1* and *VaICE2*, have been confirmed to act as key early regulators in the transcriptional cascade controlling freezing tolerance, modulating the expression of multiple cold-related genes in the C-repeat binding factor (CBF) pathway[127–129]. In addition, overexpression of *VaCBF4*, a homolog of *CBF* in *V. amurensis*, can enhance cold tolerance in *Arabidopsis*[106]. Additionally, HHPs interact with the *CBF* upstream regulators, such as *ICEs* and *CAMTA*s, to activate freezing tolerance[130]. We identified numerous putative environment-associated genes and generated testable hypotheses for future functional studies. We acknowledge that the present work does not include experimental validation of these molecular processes, and that such efforts remain essential. Future studies should prioritize functional analyses to clarify the biological roles of these genes in the environmental adaptability of the *V. amurensis*.

Populations from the Northeast area, the distribution center of *V. amurensis*, exhibit higher genetic offsets, including local and forward offsets, compared to those from the central areas under different future climate scenarios (Fig. 5 and Supplementary Figs. 13–16). The results indicated that populations of the Northeast group are highly vulnerable to future climate change, hence it is more challenging for them to survive by migration or adaptive evolution. The Center group of *V. amurensis* is expected to possess a relatively strong adaptive capacity to confront climate change and may be preadapted to future stresses related to temperature and precipitation (Fig. 5, Supplementary Figs. 15 and 16). Furthermore, estimates of forward offsets of populations decrease with increasing maximum dispersal distances, implying that assisted migration to new habitats with suitable environments may help them avoid collapse in the face of climate change (Supplementary Fig. 16). Effective conservation requires strategies based on the genetic and ecological characteristics of the species and its populations. Genomic technologies can help develop conservation decisions and actions[78]. Conservation and management strategies for the economically important *V. amurensis* should maintain a balance between in situ adaptation and assisted migration. For the West and the South groups, assisted migration and restoration may be good conservation strategies, considering their relatively small effective population size. In contrast, we highlight the populations of the Northeast group as a conservation priority, due to their high genetic offset to future climate change and their high value as genetic resources for tolerance to extreme low temperatures.

In this study, we investigate the population demographic history of *V. amurensis* based on a pangenome-referenced variants map. We confirm the possible role of TEVs to local adaptation and highlight the possible impact of TEVs on population adaptation to future climate change. Our findings suggest that comprehensive population genomics research, incorporating various genetic variants, is crucial for breeding and conservation of crops and their wild relatives in the context of climate change.

Methods

Sample collection and genome sequencing

The *V. amurensis* is one of the most cold-tolerant grape species in the world, making it an important genetic resource for grape breeding. We selected four representative *V. amurensis* accessions for genome assemblies, which were collected from Heilongjiang province (voucher number Vam057), Shandong province (voucher number Vam070), Shanxi province (voucher number Vam106), and Zhejiang province (voucher number Vam113), China (Supplementary Data 1). The total genomic DNA was isolated from fresh leaves using the DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. HiFi reads of four *V. amurensis* accessions were sequenced on the PacBio Revio platform in a circular consensus sequencing (CCS) mode, using a 15–20 kb SMRT cell library. For Hi-C sequencing, the library was prepared with chromatin extraction and digestion, and subsequently sequenced on the Illumina NovaSeq 6000 platform.

We collected 330 individuals from 31 natural populations across the natural range of *V. amurensis* (Supplementary Data 2). Each population was sampled with more than four individuals, and most populations (25 out of 31) included eight or more individuals. Individuals within each population were sampled after confirming that they were separated by at least 100 m. The sampling strategy accounted for variants in latitude, altitude, and habitat to ensure a comprehensive representation of genetic diversity. Genomic DNA was extracted from fresh leaves with the DNeasy Plant Mini Kit (Qiagen). Each DNA sample was fragmented by sonication for library preparation, and whole-genome paired-end sequencing was generated using the Illumina NovaSeq 6000 platform with a coverage of ca. 40× per sample. The sequencing data of six accessions from North American wild grapes were downloaded from the NCBI database[9] (Supplementary Table 6). Adapters were removed using Trimmomatic (v0.39) with default parameters[131]. All trimmed reads were filtered by fastp (v0.23.2) with default parameters[132]. After quality control, high-quality clean reads were obtained.

De novo genome assembly of four *V. amurensis* accessions

Following the genome assembly procedure described in ref. 18,38, the initial contigs level of genomes of the four *V. amurensis* accessions

were assembled by Hifiasm (v0.15) with HiFi and HiC reads as input using default parameters (https://github.com/chhylp123/hifiasm)[133]. Furthermore, for chromosome-level scaffolding, Hi-C reads were filtered by fastp (v0.23.2) with default parameters[132], and then used to anchor all contigs with Juicer (v1.6)[134]. The "scaffold" function of the software RagTag (v2.1.0)[135] was used to adjust the order and orientation of the contigs with the genome of the domesticated grapevine (*V. vinifera* ssp. *vinifera*) cultivar PN40024 as reference[38]. Finally, the HiC contact map was visualized by 3d-dna (with the run-assembly-visualizer.sh script and -q 0 parameter)[136], and manually adjusted in Juicebox (v1.11.08)[137] to generate the two haplotype genomes of each accession. To complete the genome assembly, the HiFi reads were mapped on the genomes of each accession using Minimap2 (v2.26)[138]. Gaps were filled and closed using the corresponding HiFi read sequences when more than 3 reads covered the gap region. The results were manually checked in Integrative Genomics Viewer (IGV) (v.2.13)[139] to ensure the filling of the gap was trustworthy. The k-mer-based consensus quality (QV) scores of the four *V. amurensis* accession genomes, which represent the log-scaled probability of error E at each base ($-10\log10^{E}$), were calculated using Merqury (v1.3) with the k-mers counted by Meryl (v1.4.1) based on HiFi data[140].

The genome assembly completeness was assessed by BUSCO (v5.2.2)[141]. To identify telomeric repeat units in the genome, the TIDK (v0.2.0) was used with parameters: -s TTTAGGG (https://github.com/tolkit/telomeric-identifier), according to the telomere repeat unit previously reported in grape[142]. The centromere localizations were explored using TRF[143] with parameters 2 7 7 80 10 80 500, and the results were transformed to gff3 format using TRF2GFF. The centromere region of each chromosome was confirmed based on protein-coding genes and TEs annotation, and the 107 bp unit of tandem repeats reported in the grape reference genome[38]. To analyze genome synteny and structural variations (SVs) among the PN_T2T genome[38] and the eight haplotype assemblies of *V. amurensis*, we performed pairwise genome alignments between adjacent haplotypes using Minimap2 (v2.26)[138]. Genome rearrangement regions were identified using SyRI[144]. The synteny of these three genomes and the SVs were plotted using plotsr (v1.1.5)[145].

Annotation of genes and transposable elements

A genome-wide annotation pipeline was used for the annotation of putative genes of *V. amurensis* haplotype genomes (https://github.com/unavailable-2374/Genome-Wide-Annotation-Pipeline). The genes were initially searched based on transcriptome data from previous studies[38,146] and UniProt sequences (https://www.uniprot.org/) as evidence. Meanwhile, the gene prediction based on the *V. amurensis* genome sequence was conducted using AUGUSTUS53 (v3.4.0)[147]. The predicted putative genes involved in duplicated regions, with total CDS regions length shorter than 90 or not supported by any evidence, were filtered. The missing genes were attempted to be complemented, and the complete genes were subjected to alternative splicing analysis. All the results were checked by hidden Markov models downloaded from the Pfam database to obtain the final gene models. TE families were identified by RepeatModeler2[148] with a pan-*Vitis* database of repeat families[38]. The repeat elements were classified by using deepTE with the Plant model[149]. Finally, we annotated all repetitive sequences using RepeatMasker (v4.1.2)[150]. We estimated TE insertion times for the eight haplotype assemblies. The outputs generated by RepeatMasker were parsed using the Perl script parseRM.pl (https://github.com/4ureliek/Parsing-RepeatMasker-Outputs) with the parameters -l 50,1 -m 0.0025. The parameter -m specifies the substitution rate, following estimates from a previous study[7].

Construction of the graph pangenome

To comprehensively represent the genetic diversity of *V. amurensis*, we utilized the PanGenome Graph Builder (PGGB) to construct the *V. amurensis* graph pangenome with the Vam057 haplotype1 genome serving as the reference, using the settings -s 5000 -l 25,000 -p 80 -n 2 -K 19 -F 0.001[81]. All four genomes, encompassing eight haplotypes, were included in the construction of the graph pangenome. Through this approach, we aimed to capture the structural variants present within the *V. amurensis* populations effectively.

Variant maps of *V. amurensis*

To ensure the reliability of the short variant dataset and maintain analytical consistency across samples, we employed a widely used and well-validated linear reference-based approach. High-quality clean NGS reads were mapped to the Vam057 haplotype1 genome using the BWA-MEM algorithm (v0.7.17)[151]. BAM files were filtered using samtools view -bS -F 0 × 100 -q 10 to remove secondary alignments and low-quality reads. The filtered BAM files were then sorted with SAMtools (v1.13)[152], and PCR duplicates were identified and marked using Picard (v2.23.0) MarkDuplicates (http://broadinstitute.github.io/picard/). These steps ensured that only reliably mapped reads were retained for variant calling. The SNP and indel calling was performed using GATK (v4.2.6.0)[153], functions of HaplotypeCaller, GenomicsDBImport, GenotypeGVCFs, SelectVariants, and VariantFiltration. Specifically, variants were called per sample using HaplotypeCaller in GVCF mode to generate individual GVCF files. All GVCF files were combined with GenomicsDBImport to build a genomic database, and joint genotyping across all samples was performed using GenotypeGVCFs to produce cohort-level VCF files for each chromosome. The SNPs and indels were extracted from the joint VCF files using SelectVariants with the options -select-type SNP and -select-type INDEL, respectively. These variant sets were then filtered using VariantFiltration. SNPs were filtered based on "QD < 2.0 || FS > 60.0 || MQRankSum <- 12.5 || ReadPosRankSum <- 8.0 || SOR > 3.0 || MQ < 40.0", while Indels were filtered using "QD < 2.0 || FS > 200.0 || MQRankSum <- 12.5 || ReadPosRankSum <- 8.0 || SOR > 10.0". The resulting variants were filtered by VCFtools (v0.1.15) using parameters: -max-missing 0.8[154].

The PGGB pangenome was used to generate the structural variant set among *V. amurensis* populations. The re-genotyping tool PanGenie (v2.1.0) was used to determine unique *k*-mers in the graph with PanGenie index and generate structural variant VCF files of each accession, using the setting -r Vam057_haplotype1_genome along with default parameters[155]. These VCF files were further combined into a merged structural variant map using BCFtools (v1.13)[156]. We searched homologous nucleotide sequences of structural variants in the database of TE families mentioned above to identify TEVs using NCBI-BLASTN with default parameters (evalue < 1e−5, percentage identity >70%)[38,157]. These TEVs were further filtered using VCFtools (v0.1.15) with parameters: -max-missing 0.8[154].

Ecological niche modeling

To infer the potential geographic distribution of *V. amurensis*, we performed the ecological niche modeling using Maxent (v3.4.4) with 10% random test and 10 replicates[158]. The occurrence records of *V. amurensis* were collected from the Global Biodiversity Information Facility (GBIF, www.gbif.org) and fieldwork. A total of 20 current environmental factors, including 19 bioclimatic and 1 elevation variables, were used with a resolution of 2.5 arcmin from the WorldClim (http://worldclim.org/)[159].

Phylogenetic, principal component, and population structure analyses

Firstly, a linkage disequilibrium pruned SNP set was generated using PLINK (v1.90) software with parameters: --indep-pairwise 50 10 0.2 --vcf-min-gq 20 --maf 0.05 --geno 0.1[160], which included 344,087 independent SNPs to be used in the following analyses. The phylogenetic relationship among populations of *V. amurensis* was reconstructed by IQ-TREE (v1.6.12)[161] using the GTR model and 1000

bootstrap using six wild grapes from North America as outgroups. Principal component analysis was performed using the PLINK (v1.90) software with parameter "--pca"[160]. Population structures were examined to estimate individual admixture proportions from ancestors using ADMIXTURE (v1.3.0)[162] with LD-pruned SNPs set and TEs set mentioned above, respectively. The number of clusters (*K*) was set from 2 to 6.

TEVs dynamics in *V. amurensis* population

To assess the potential functional impact of TEVs, we conducted gene ontology (GO) analysis using the web tool DAVID on genes containing TEVs with an allele frequency >0.8 in their CDS regions. GO terms were assigned based on UniProt functional annotations (https://www.uniprot.org/).

To calculate the derived site frequency spectrum (SFS) of each population, we used the superSFS script (https://github.com/xhchauvet/superSFS) with six North American wild grapes as outgroups (Supplementary Table 6). Moreover, PopLDdecay (v3.42) was used to calculate the decay of linkage disequilibrium between SNPs and TEVs[163].

Structural and SNP variants were further filtered by VCFtools (v0.1.15) with parameters "--maf 0.03" and "--maf 0.05", respectively[154]. After filtering, we obtained high-quality 177,789 structural variants (including 97,013 TEVs) and 4,114,075 SNPs for the following analyses. Population genetic diversity (π) of SNPs and TEVs was generated by VCFTools (v0.1.15) in sliding windows of 1 Mb, respectively,[154]. Pearson correlation analysis was performed to evaluate the relationship between the π values of SNPs and TEVs. In order to explore the relationship between structural variants/TEVs and SNPs, the number of SNPs in the surrounding regions of each filtered structural variant/TEV was counted with sliding windows of 15 kb. And the average number of SNPs per 15 kb sliding window across the genome was calculated.

Detection of genotype-environment associations

To identify environment-associated genotype variants, including SNPs, indels, and TEVs, two different approaches were employed. TE variants were filtered with the parameter "--maf 0.03", while both indel and SNP variants were filtered with "--maf 0.05" using VCFtools (v0.1.15)[154]. A total of 97,013 TEVs, 4,114,075 SNPs, and 677,438 indels were used in these analyses. Firstly, the univariate LFMM implemented in the R package LEA (v3.11.4)[82] was performed with the 19 bioclimatic and 1 elevation variables (resolution 2.5 arcmin). According to the population structures inferred by ADMIXTURE (v1.3.0)[162], four latent factors were used in LFMM to account for the number of ancestry clusters in the genotype data. We performed five independent MCMC runs for each environmental variable, with a burn-in of 5000 iterations, followed by 10,000 iterations for sampling. The mean *P* values obtained from all five runs were subsequently utilized for further analysis, and the significant cutoff was set at a False discovery rate (FDR) of 0.01. The LDBlockShow was used to show the regional linkage disequilibrium of candidate genes[164]. Gene ontology enrichment analysis of candidate genes was performed using the web tool DAVID. Secondly, six uncorrelated environmental variables (bio4, elev, bio15, bio3, bio18, and bio10) with Spearman correlation coefficient $|r| < 0.8$ were selected to perform the redundancy analysis using the R package vegan (v2.6–4) to identify environment-associated genetic variants with a standard deviation cutoff of 3.5 along one or multiple RDA axes[165].

To evaluate the potential regulatory impact of environment-associated variants, we analyzed the genomic regions surrounding these variants using the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The predicted regulatory elements, including promoter motifs and transcription factor binding sites, were identified and compared between the reference and variant alleles. This analysis was used to assess whether the variants could disrupt or modify regulatory motifs, thereby potentially influencing the transcriptional regulation and expression of adjacent genes.

The contribution of environmental variables

The principal component analysis was performed with PLINK (v1.90) software using adaptive candidates of SNPs, indels, and TEVs based on the LFMM analysis (FDR < 0.01). The first two principal components were used for the variance partitioning analysis. The varpart function, using adjusted in RDA from the R package vegan (v2.6-4), was performed to estimate the proportion of environmental variance explained by different genetic variants[165].

Genetic offset analysis

Future bioclimatic variables for the periods 2041–2060 and 2081–2100, under SSP245 and SSP585 scenarios, were obtained from three different climate models (ACCESS-CM2, CanESM5, and GISS-E2-1-G; resolution 2.5 arcmin), using the WorldClim version 2.1 database[159]. The overlapping candidate adaptive variants identified by both LFMM and RDA approaches were defined as "core adaptive variants" for local adaptation. To reduce biases arising from linkage disequilibrium, we employed PLINK (v1.90) software with the parameters (--indep-pairwise 50 10 0.5)[160] based on core adaptive variants to obtain LD-pruned 40 adaptive TEVs, and 206 adaptive SNPs and indels for the following analyses. We used the gradient forest model, which is a machine learning regression tree approach, to predict the spatiotemporal adaptability of *V. amurensis* to future environmental change with calculating population genetic offset using the R package gradientForest[166]. Genetic offset represented the mismatch between contemporary and future spatial environment-associated genotype composition to quantify the maladaptation of populations. Populations with higher genetic offsets imply that they are greater vulnerability to future climate change due to the requirement for more pronounced changes in adaptive allele frequencies to ensure successful adaptation[71,167]. We incorporated predictions of genetic offsets from three different future climate models by considering distinctions of these models using the gradient forest approach. The mean genetic offset values from the three models were used to estimate local maladaptation under future climate conditions. Moreover, two different emission scenarios (SSP245 and SSP585) and two defined periods (2041–2060 and 2081–2100) were used for the predicted patterns of genetic offsets throughout the distribution range of *V. amurensis*.

To estimate the possible contribution of adaptive TEVs to the future climate adaptability of *V. amurensis* populations, we followed a systematic approach. First, we calculated the genetic offset using LD-pruned adaptive SNPs and indels as the baseline. Next, we included all LD-pruned adaptive variants in our genetic offset calculation. We then randomly selected the same number of non-adaptive TEVs along with the LD-pruned adaptive SNPs and indels to estimate genetic offsets, repeating this process 100 times to obtain average genetic offsets for a robust genetic vulnerability analysis. Finally, we compared these genetic offsets of 31 sampled populations to determine statistical significance.

Additionally, migration plays an important role in allowing species to maintain populations under rapidly changing climate conditions and to avoid extirpation[167]. To consider the contribution of migration, we evaluated the sensitivity of the forward offset of 2080–2100 under the SSP585 scenario by calculating the average values across all climate models with limiting the maximum allowable migration to different distance classes, such as 100, 500, and 1000 km.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The eight haplotype-resolved genome assemblies of *V. amurensis* have been deposited in the National Center for Biotechnology Information (NCBI) under the BioProject accession PRJNA1272278, PRJNA1272279, PRJNA1272280, PRJNA1272281, PRJNA1272282, PRJNA1272283, PRJNA1272284, and PRJNA1272285, which are also available in Zenodo [https://doi.org/10.5281/zenodo.16534920]. The pangenome and genome annotations are available in Zenodo [https://doi.org/10.5281/zenodo.16534920]. The associated HiFi and Hi-C sequencing data have been deposited in NCBI under BioProject accession PRJNA1308837. In addition, whole-genome resequencing data for the 330 individuals have been deposited in NCBI under BioProject accession PRJNA1089859 and the National Genomics Data Center under accession PRJCA046381. Source data are provided with this paper.

References

1. Aguirre-Liguori, J. A., Morales-Cruz, A. & Gaut, B. S. Evaluating the persistence and utility of five wild *Vitis* species in the context of climate change. *Mol. Ecol.* **31**, 6457–6472 (2022).
2. Zhang, X. et al. The genomic and epigenetic footprint of local adaptation to variable climates in kiwifruit. *Hortic. Res.* **10**, uhad031 (2023).
3. Zhao, X. B. et al. Population genomics unravels the Holocene history of bread wheat and its relatives. *Nat. Plants* **9**, 403–419 (2023).
4. Wu, S. et al. A *Citrullus* genus super-pangenome reveals extensive variations in wild and cultivated watermelons and sheds light on watermelon evolution and domestication. *Plant Biotechnol. J.* **21**, 1926–1928 (2023).
5. Myles, S. et al. Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. USA* **108**, 3530–3535 (2011).
6. Xiao, H. et al. Adaptive and maladaptive introgression in grapevine domestication. *Proc. Natl. Acad. Sci. USA* **120**, e2222041120 (2023).
7. Zhou, Y., Massonnet, M., Sanjak, J. S., Cantu, D. & Gaut, B. S. Evolutionary genomics of grape (*Vitis vinifera* ssp. *vinifera*) domestication. *Proc. Natl. Acad. Sci. USA* **114**, 11715–11720 (2017).
8. Dong, Y. et al. Dual domestications and origin of traits in grapevine evolution. *Science* **379**, 892–901 (2023).
9. Ma, Z. Y. et al. Phylogenomics, biogeography, and adaptive radiation of grapes. *Mol. Phylogenet Evol.* **129**, 258–267 (2018).
10. Ma, Z. Y. et al. Phylogenetic relationships, hybridization events, and drivers of diversification of East Asian wild grapes as revealed by phylogenomic analyses. *J. Syst. Evol.* **61**, 273–283 (2023).
11. Sun, X. M. et al. The ethylene response factor VaERF092 from Amur grape regulates the transcription factor VaWRKY33, improving cold tolerance. *Plant J.* **99**, 988–1002 (2019).
12. Patel, S. et al. Draft genome of the Native American cold hardy grapevine *Vitis riparia* Michx. 'Manitoba 37'. *Hortic. Res.* **7**, 92 (2020).
13. Courbier, S. Bundling up: VaPAT1 forms a complex with VaIDD3 to activate cold tolerance in Amur grape calli. *Plant Physiol.* **186**, 1373–1374 (2021).
14. Wang, Z. M. et al. GRAS-domain transcription factor PAT1 regulates jasmonic acid biosynthesis in grape cold stress response. *Plant Physiol.* **186**, 1660–1678 (2021).
15. Hu, Y. et al. Identification and expression analysis of heat shock transcription factors in the wild Chinese grapevine (*Vitis pseudoreticulata*). *Plant Physiol. Biochem.* **99**, 1–10 (2016).
16. Azri, W. et al. Physiological and proteomic responses to drought stress in leaves of two wild grapevines (*Vitis sylvestris*): a comparative study. *Plant Growth Regul.* **91**, 37–52 (2020).
17. Cui, X. Y., Zhang, P. Y., Chen, C. C. & Zhang & J. X. VyUSPA3, a universal stress protein from the Chinese wild grape *Vitis yeshanensis*, confers drought tolerance to transgenic *V. vinifera*. *Plant Cell Rep.* **42**, 181–196 (2023).
18. Zhang, T. H. et al. Population genomics highlights structural variations in local adaptation to saline coastal environments in woolly grape. *J. Integr. Plant Biol.* **7**, 1408–1426 (2024).
19. Cochetel, N. et al. A super-pangenome of the North American wild grape species. *Genome Biol.* **24**, 290 (2023).
20. Yu, Y. H. et al. Grape (*Vitis davidii*) *VdGATA2* functions as a transcription activator and enhances powdery mildew resistance via the active oxygen species pathway. *Sci. Hortic.* **267**, 109327 (2020).
21. Liu, W. et al. *VqMAPK3*/*VqMAPK6*, *VqWRKY33*, and *VqNSTS3* consist of a regulatory node in enhancing resistance to powdery mildew in grapevine. *Hortic. Res.* **10**, uhad116 (2023).
22. Morales-Cruz, A. et al. Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* sps.) and its implications within a changing climate. *Commun. Biol.* **6**, 580 (2023).
23. Morales-Cruz, A. et al. Introgression among North American wild grapes (*Vitis*) fuels biotic and abiotic adaptation. *Genome Biol.* **22**, 1–27 (2021).
24. Hannah, L. et al. Climate change, wine, and conservation. *Proc. Natl. Acad. Sci. USA* **110**, 6907–6912 (2013).
25. Morales-Castilla, I. et al. Diversity buffers winegrowing regions from climate change losses. *Proc. Natl. Acad. Sci. USA* **117**, 2864–2869 (2020).
26. Zhang, M. et al. The dynamics of wild Vitis species in response to climate change facilitate the breeding of grapevineand its rootstocks with climate resilience. *Hortic. Res.* **12**, uhaf104 (2025).
27. Xu, W. R. et al. Role of an Amur grape *CBL*-interacting protein kinase *VaCIPK02* in drought tolerance by modulating ABA signaling and ROS production. *Environ. Exp. Bot.* **172**, 103999 (2020).
28. Zhao, Y. et al. Analysis of the cold tolerance and physiological response differences of amur grape (*Vitis amurensis*) germplasms during overwintering. *Sci. Hortic.* **259**, 108760 (2020).
29. Bourque, G. et al. Ten things you should know about transposable elements. *Genome Biol.* **19**, 1–12 (2018).
30. Deng, P. J. et al. Structural RNA components supervise the sequential DNA cleavage in R2 retrotransposon. *Cell* **186**, 2865–2879 (2023).
31. Goerner-Potvin, P. & Bourque, G. Computational tools to unmask transposable elements. *Nat. Rev. Genet.* **19**, 688–704 (2018).
32. Guio, L. & Gonzaléz, J. New insights on the evolution of genome content: population dynamics of transposable elements in flies and humans. *Methods Mol. Biol.* **1910**, 505–530 (2019).
33. Wells, J. N. & Feschotte, C. A feld guide to eukaryotic transposable elements. *Annu Rev. Genet.* **54**, 539–561 (2020).
34. Nurk, S. et al. The complete sequence of a human genome. *Science* **376**, 44–53 (2022).
35. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* **36**, 344–355 (1950).
36. Tenaillon, M. I., Hufford, M. B., Gaut, B. S. & Ross-Ibarra, J. Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxurians*. *Genome Biol. Evol.* **3**, 219–229 (2011).
37. Anderson, S. N. et al. Transposable elements contribute to dynamic genome content in maize. *Plant J.* **100**, 1052–1065 (2019).
38. Shi, X. et al. The complete reference genome for grapevine (*Vitis vinifera* L.) genetics and breeding. *Hortic. Res.* **10**, uhad061 (2023).
39. Roessler, K., Bousios, A., Meca, E. & Gaut, B. S. Modeling interactions between transposable elements and the plant epigenetic response: a surprising reliance on element retention. *Genome Biol. Evol.* **10**, 803–815 (2018).
40. Lisch, D. How important are transposons for plant evolution? *Nat. Rev. Genet.* **14**, 49–61 (2013).

41. Fueyo, R., Judd, J., Feschotte, C. & Wysocka, J. Roles of transposable elements in the regulation of mammalian transcription. *Nat. Rev. Mol. Cell Biol.* **23**, 481–497 (2022).
42. Greenberg, M. V. C. & Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* **20**, 590–607 (2019).
43. Payer, L. M. & Burns, K. H. Transposable elements in human genetic disease. *Nat. Rev. Genet.* **20**, 760–772 (2019).
44. Chuong, E. B., Elde, N. C. & Feschotte, C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017).
45. Lisch, D. Epigenetic regulation of transposable elements in plants. *Annu Rev. Plant Biol.* **60**, 43–66 (2009).
46. Cui, X. K. & Cao, X. F. Epigenetic regulation and functional exaptation of transposable elements in higher plants. *Curr. Opin. Plant Biol.* **21**, 83–88 (2014).
47. Stuart, T. et al. Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *Elife* **5**, e20777 (2016).
48. Martin, G. T., Seymour, D. K. & Gaut, B. S. CHH methylation islands: a nonconserved feature of grass genomes that is positively associated with transposable elements but negatively associated with gene-body methylation. *Genome Biol. Evol.* **13**, evab144 (2021).
49. Jiang, N., Bao, Z. R., Zhang, X. Y., Eddy, S. R. & Wessler, S. R. Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **431**, 569–573 (2004).
50. Juretic, N., Hoen, D. R., Huynh, M. L., Harrison, P. M. & Bureau, T. E. The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res.* **15**, 1292–1297 (2005).
51. Du, C., Fefelova, N., Caronna, J., He, L. M. & Dooner, H. K. The polychromatic landscape of the maize genome. *Proc. Natl. Acad. Sci. USA* **106**, 19916–19921 (2009).
52. Muyle, A. et al. Gene capture by transposable elements leads to epigenetic conflict in maize. *Mol. Plant* **14**, 237–252 (2021).
53. Carvalho, C. M. B. & Lupski, J. R. Mechanisms underlying structural variant formation in genomic disorders. *Nat. Rev. Genet.* **17**, 224–238 (2016).
54. Van't Hof, A. E. et al. The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **534**, 102–105 (2016).
55. Niu, X. M. et al. Transposable elements drive rapid phenotypic variation in *Capsella rubella*. *Proc. Natl. Acad. Sci. USA* **116**, 6908–6913 (2019).
56. Xia, B. et al. On the genetic basis of tail-loss evolution in humans and apes. *Nature* **626**, 1042–1048 (2024).
57. Wlodzimierz, P. et al. Cycles of satellite and transposon evolution in *Arabidopsis* centromeres. *Nature* **618**, 557–565 (2023).
58. Gaut, B. S., Seymour, D. K., Liu, Q. & Zhou, Y. Demography and its effects on genomic variation in crop domestication. *Nat. Plants* **4**, 512–520 (2018).
59. Zhou, Y. F. et al. The population genetics of structural variants in grapevine domestication. *Nat. Plants* **5**, 965–979 (2019).
60. Zhou, Y. et al. Graph pangenome captures missing heritability and empowers tomato breeding. *Nature* **606**, 527–534 (2022).
61. Liu, Z. et al. Grapevine pangenome facilitates trait genetics and genomic breeding. *Nat. Genet.* **56**, 2804–2814 (2024).
62. Kou, Y. X. et al. Evolutionary genomics of structural variation in Asian rice (*Oryza sativa*) domestication. *Mol. Biol. Evol.* **37**, 3507–3524 (2020).
63. Li, Z. W. et al. Transposable elements contribute to the adaptation of *Arabidopsis thaliana*. *Genome Biol. Evol.* **10**, 2140–2150 (2018).
64. Urban, M. C. Accelerating extinction risk from climate change. *Science* **348**, 571–573 (2015).
65. Lesk, C., Rowhani, P. & Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **529**, 84–87 (2016).
66. Chen, I. C., Hill, J. K., Ohlemüller, R., Roy, D. B. & Thomas, C. D. Rapid range shifts of species associated with high levels of climate warming. *Science* **333**, 1024–1026 (2011).
67. Dyderski, M. K., Paz, S., Frelich, L. E. & Jagodzinski, A. M. How much does climate change threaten European forest tree species distributions? *Glob. Change Biol.* **24**, 1150–1163 (2018).
68. Waldvogel, A. M. et al. Evolutionary genomics can improve prediction of species' responses to climate change. *Evol. Lett.* **4**, 4–18 (2020).
69. Savolainen, O., Lascoux, M. & Merilä, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14**, 807–820 (2013).
70. Bernatchez, L., Ferchaud, A. L., Berger, C. S., Venney, C. J. & Xuereb, A. Genomics for monitoring and understanding species responses to global climate change. *Nat. Rev. Genet.* **25**, 165–183 (2024).
71. Aguirre-Liguori, J. A., Ramírez-Barahona, S. & Gaut, B. S. The evolutionary genomics of species' responses to climate change. *Nat. Ecol. Evol.* **5**, 1350–1360 (2021).
72. Sang, Y. P. et al. Genomic insights into local adaptation and future climate-induced vulnerability of a keystone forest tree in East Asia. *Nat. Commun.* **13**, 6541 (2022).
73. Zhang, F. et al. Evolutionary genomics of climatic adaptation and resilience to climate change in alfalfa. *Mol. Plant* **17**, 867–883 (2024).
74. Hämälä, T. et al. Impact of whole-genome duplications on structural variant evolution in *Cochlearia*. *Nat. Commun.* **15**, 5377 (2024).
75. Huang, X., Huang, S., Han, B. & Li, J. The integrated genomics of crop domestication and breeding. *Cell* **185**, 2828–2839 (2022).
76. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).
77. Chen, Y. L. et al. The combination of genomic offset and niche modelling provides insights into climate change-driven vulnerability. *Nat. Commun.* **13**, 4821 (2022).
78. Hung, T. H. et al. Range-wide differential adaptation and genomic offset in critically endangered Asian rosewoods. *Proc. Natl. Acad. Sci. USA* **120**, e2301603120 (2023).
79. Yuan, S. et al. Genomic vulnerability to climate change in *Quercus acutissima*, a dominant tree species in East Asian deciduous forests. *Mol. Ecol.* **32**, 1639–1655 (2023).
80. Razgour, O. et al. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proc. Natl. Acad. Sci. USA* **116**, 10418–10423 (2019).
81. Garrison, E. et al. Building pangenome graphs. *Nat. Methods* **21**, 2008–2012 (2024).
82. Frichot, E., Schoville, S. D., Bouchard, G. & Francois, O. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* **30**, 1687–1699 (2013).
83. Wang, M., Xu, Z. C. & Kong, Y. Z. The tubby-like proteins kingdom in animals and plants. *Gene* **642**, 16–25 (2018).
84. Hsieh, M. H. & Goodman, H. M. A novel gene family in *Arabidopsis* encoding putative heptahelical transmembrane proteins homologous to human adiponectin receptors and progestin receptors. *J. Exp. Bot.* **56**, 3137–3147 (2005).
85. Colonna Romano, N. & Fanti, L. Transposable elements: major players in shaping genomic and evolutionary patterns. *Cells* **11**, 1048 (2022).
86. Schrader, L. & Schmitz, J. The impact of transposable elements in adaptive evolution. *Mol. Ecol.* **28**, 1537–1549 (2019).

87. Jiang, J. et al. Forces driving transposable element load variation during Arabidopsis range expansion. *Plant Cell* **36**, 840–862 (2024).
88. Huang, Y. et al. Pangenome analysis provides insight into the evolution of the orange subfamily and a key gene for citric acid accumulation in citrus fruits. *Nat. Genet.* **55**, 1964–1975 (2023).
89. Hanada, K. et al. The functional role of pack-MULEs in rice inferred from purifying selection and expression profile. *Plant Cell* **21**, 25–38 (2009).
90. Butelli, E. et al. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* **24**, 1242–1255 (2012).
91. Catlin, N. S. & Josephs, E. B. The important contribution of transposable elements to phenotypic variation and evolution. *Curr. Opin. Plant Biol.* **65**, 102140 (2022).
92. Krasileva, K. V. The role of transposable elements and DNA damage repair mechanisms in gene duplications and gene fusions in plant genomes. *Curr. Opin. Plant Biol.* **48**, 18–25 (2019).
93. Cheng, L. et al. Leveraging a phased pangenome for haplotype design of hybrid potato. *Nature* **640**, 408–417 (2025).
94. Tiffin, P. & Ross-Ibarra, J. Advances and limits of using population genetics to understand local adaptation. *Trends Ecol. Evol.* **29**, 673–680 (2014).
95. Fournier-Level, A. et al. A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86–89 (2011).
96. Quadrana, L. et al. The *Arabidopsis thaliana* mobilome and its impact at the species level. *Elife* **5**, 15716 (2016).
97. Uzunović, J., Josephs, E. B., Stinchcombe, J. R. & Wright, S. I. Transposable elements are important contributors to standing variation in gene expression in *Capsella grandiflora*. *Mol. Biol. Evol.* **36**, 1734–1745 (2019).
98. Baduel, P., Quadrana, L., Hunter, B., Bomblies, K. & Colot, V. Relaxed purifying selection in autopolyploids drives transposable element over-accumulation which provides variants for local adaptation. *Nat. Commun.* **10**, 5818 (2019).
99. Carpentier, M. C. et al. Retrotranspositional landscape of Asian rice revealed by 3000 genomes. *Nat. Commun.* **10**, 24 (2019).
100. Rogivue, A. et al. Genome-wide variation in nucleotides and retrotransposons in alpine populations of *Arabis alpina* (Brassicaceae). *Mol. Ecol. Resour.* **19**, 773–787 (2019).
101. Yeaman, S. Evolution of polygenic traits under global *vs* local adaptation. *Genetics* **220**, iyab134 (2022).
102. Quadrana, L. et al. Transposition favors the generation of large effect mutations that may facilitate rapid adaption. *Nat. Commun.* **10**, 3421 (2019).
103. Domínguez, M. et al. The impact of transposable elements on tomato diversity. *Nat. Commun.* **11**, 4058 (2020).
104. Dawson, T. P., Jackson, S. T., House, J. I., Prentice, I. C. & Mace, G. M. Beyond predictions: biodiversity conservation in a changing climate. *Science* **332**, 53–58 (2011).
105. Baduel, P. et al. Genetic and environmental modulation of transposition shapes the evolutionary potential of *Arabidopsis thaliana*. *Genome Biol.* **22**, 138 (2021).
106. Li, J. T., Wang, N. A., Xin, H. P. & Li, S. H. Overexpression of *VaCBF4*, a transcription factor from *Vitis amurensis*, improves cold tolerance accompanying increased resistance to drought and salinity in *Arabidopsis*. *Plant Mol. Biol. Rep.* **31**, 1518–1528 (2013).
107. Lai, C. P. et al. Molecular analyses of the *Arabidopsis* TUBBY-like protein gene family. *Plant Physiol.* **134**, 1586–1597 (2004).
108. Lai, C. P. et al. Functional diversification of the Tubby-like protein gene families(TULPs) during eukaryotic evolution. *Biocatal. Agr. Biotech.* **1**, 2–8 (2012).
109. Bano, N., Aalam, S. & Bag, S. K. Tubby-like proteins (TLPs) transcription factor in different regulatory mechanism in plants: a review. *Plant Mol. Biol.* **110**, 455–468 (2022).
110. Dong, M. Y., Fan, X. W., Pang, X. Y. & Li, Y. Z. Decrypting tubby-like protein gene family of multiple functions in starch root crop cassava. *AoB Plants* **11**, plz075 (2019).
111. Li, Z. S. et al. Improvement of plant tolerance to drought stress by cotton tubby-like protein 30 through stomatal movement regulation. *J. Adv. Res.* **42**, 55–67 (2022).
112. Du, F., Xu, J. N., Zhan, C. Y., Yu, Z. B. & Wang, X. Y. An obesity-like gene *MdTLP7* from apple (*Malus* x *domestica*) enhances abiotic stress tolerance. *Biochem. Biophys. Res. Commun.* **445**, 394–397 (2014).
113. Xu, J. N. et al. The expression of a tubby-like protein from *Malus domestica* (*Md*TLP7) enhances abiotic stress tolerance in *Arabidopsis*. *BMC Plant Biol.* **19**, 1–8 (2019).
114. Wardhan, V. et al. Overexpression of CaTLP1, a putative transcription factor in chickpea (*Cicer arietinum* L.), promotes stress tolerance. *Plant Mol. Biol.* **79**, 479–493 (2012).
115. Bano, N., Fakhrah, S., Mohanty, C. S. & Bag, S. K. Genome-wide identification and evolutionary analysis of *Gossypium* tubby-like protein (TLP) gene family and expression analyses during salt and drought stress. *Front. Plant Sci.* **12**, 667929 (2021).
116. Xu, H. R. et al. Comprehensive profiling of Tubby-Like proteins in soybean and roles of the *GmTLP8* gene in abiotic stress responses. *Front. Plant Sci.* **13**, 844545 (2022).
117. Jain, N., Khurana, P. & Khurana, J. P. AtTLP2, a Tubby-like protein, plays intricate roles in abiotic stress signalling. *Plant Cell Rep.* **42**, 235–252 (2023).
118. Li, S. et al. A tubby-like protein CsTLP8 acts in the ABA signaling pathway and negatively regulates osmotic stresses tolerance during seed germination. *BMC Plant Biol.* **21**, 340 (2021).
119. Hong, M. J., Kim, D. Y. & Seo, Y. W. Interactions between wheat Tubby-like and SKP1-like proteins. *Genes Genet. Syst.* **90**, 293–304 (2015).
120. Li, S. T. et al. Genome-wide identification, characterization, and expression analysis of Tubby-like protein (TLP) gene family members in woodland strawberry (*Fragaria vesca*). *Int. J. Mol. Sci.* **23**, 11961 (2022).
121. Chen, Y. L., Dai, W., Sun, B. M., Zhao, Y. & Ma, Q. Genome-wide identification and comparative analysis of the TUBBY-like protein gene family in maize. *Genes Genom.* **38**, 25–36 (2016).
122. Zhang, J. J. et al. Expression analysis and functional characterization of tomato Tubby-like protein family. *Plant Sci.* **324**, 111454 (2022).
123. Meng, L. et al. *VaWRKY65* contributes to cold tolerance through dual regulation of soluble sugar accumulation and reactive oxygen species scavenging in *Vitis amurensis*. *Hortic. Res.* **12**, uhae367 (2025).
124. Liu, Y., Dang, P., Liu, L. & He, C. Cold acclimation by the CBF-*COR* pathway in a changing climate: Lessons from *Arabidopsis thaliana*. *Plant Cell Rep.* **38**, 511–519 (2019).
125. Hwarari, D. et al. *ICE-CBF-COR* signaling cascade and its regulation in plants responding to cold stress. *Int. J. Mol. Sci.* **23**, 1549 (2022).
126. Li, X. et al. Blue light- and low temperature-regulated *COR27* and *COR28* play roles in the *Arabidopsis* circadian clock. *Plant Cell* **28**, 2755–2769 (2016).
127. Xu, W. R. et al. Chinese wild-growing *Vitis amurensis ICE1* and *ICE2* encode *MYC*-Type *bHLH* tanscription activators that regulate cold tolerance in *Arabidopsis*. *PLoS ONE* **9**, e102303 (2014).
128. Ma, X. L. L., Zhao, F. Y., Su, K., Lin, H. & Guo, Y. S. Discovery of cold-resistance genes in *Vitis amurensis* using bud-based quantitative trait locus mapping and RNA-seq. *BMC Genom.* **23**, 551 (2022).
129. Ren, C. et al. Characterization of chromatin accessibility and gene expression upon cold stress reveals that the RAV1 transcription factor functions in cold response in *Vitis amurensis*. *Plant Cell Physiol.* **62**, 1615–1629 (2021).

130. Lee, H. G. & Seo, P. J. The MYB96-HHP module integrates cold and abscisic acid signaling to activate the CBF-COR pathway in *Arabidopsis*. *Plant J.* **82**, 962–977 (2015).
131. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
132. Chen, S. F., Zhou, Y. Q., Chen, Y. R. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, 884–890 (2018).
133. Cheng, H. Y., Concepcion, G. T., Feng, X. W., Zhang, H. W. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
134. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
135. Alonge, M. et al. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 1–19 (2022).
136. Dudchenko, O. et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
137. Robinson, J. T. et al. Juicebox.js provides a cloud-based visualization system for Hi-C data. *Cell Syst.* **6**, 256–258 (2018).
138. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
139. Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A. & Mesirov, J. P. Variant review with the integrative genomics viewer. *Cancer Res.* **77**, E31–E34 (2017).
140. Rhie, A., Walenz, B. P., Koren, S. & Phillippy, A. M. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245 (2020).
141. Manni, M., Berkeley, M. R., Seppey, M., Simao, F. A. & Zdobnov, E. M. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
142. Castro, C., Carvalho, A., Gaivão, I. & Lima-Brito, J. Evaluation of copper-induced DNA damage in *Vitis vinifera* L. using Comet-FISH. *Environ. Sci. Pollut. R.* **28**, 6600–6610 (2021).
143. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
144. Goel, M., Sun, H., Jiao, W. B. & Schneeberger, K. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**, 277 (2019).
145. Goel, M. & Schneeberger, K. plotsr: visualizing structural similarities and rearrangements between multiple genomes. *Bioinformatics* **38**, 2922–2926 (2022).
146. Wang, Y. et al. The genome of Shanputao (*Vitis amurensis*) provides a new insight into cold tolerance of grapevine. *Plant J.* **105**, 1495–1506 (2021).
147. Stanke, M. et al. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–W439 (2006).
148. Flynn, J. M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. USA* **117**, 9451–9457 (2020).
149. Yan, H. D., Bombarely, A. & Li, S. DeepTE: a computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics* **36**, 4269–4275 (2020).
150. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **25**, 4.10.11–14.10.14 (2009).
151. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
152. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
153. McKenna, A. et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
154. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
155. Ebler, J. et al. Pangenome-based genome inference allows efficient and accurate genotyping across a wide spectrum of variant classes. *Nat. Genet.* **54**, 518–525 (2022).
156. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
157. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
158. Phillips, S. J. & Dudik, M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161–175 (2008).
159. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
160. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
161. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
162. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
163. Zhang, C., Dong, S. S., Xu, J. Y., He, W. M. & Yang, T. L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**, 1786–1788 (2019).
164. Dong, S. S. et al. LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Brief. Bioinform.* **22**, bbaa227 (2021).
165. Oksanen, J. *vegan: community ecology package. R package version 2.6–4*, https://cran.r-project.org/web/packages/vegan/ (2022).
166. Ellis, N., Smith, S. J. & Pitcher, C. R. Gradient forests: calculating importance gradients on physical predictors. *Ecology* **93**, 156–168 (2012).
167. Gougherty, A. V., Keller, S. R. & Fitzpatrick, M. C. Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nat. Clim. Chang.* **11**, 166–171 (2021).

Acknowledgements

We thank Prof. Xiuqun Liu (Huazhong Agricultural University) for providing materials. We thank Prof. Jing Wang (Sichuan University) and Mr. Yupeng Sang (Sichuan University) for suggestions on genetic offset inference. We thank members of the Zhou Lab at the Agricultural Genomics Institute at Shenzhen (AGIS) for helpful discussions and comments on the project. This work was supported by the Science Fund Program for Excellent Young Scholars of the National Natural Science Foundation of China (Overseas) to Yongfeng Z., the National Natural Science Foundation of China (Grants No. 32300191 to Z.M. and No. 32372662 to Yongfeng Z), the National Key Research and Development Program of China (No. 2023YFD2200700 to Yongfeng Z.), the Project of National Key Laboratory of Tropical Crop Breeding (No. NKLTCB-RC202501 to Yongfeng Z.), the Project of State Key Laboratory of Tropical Crop Breeding (No. SKLTCBYWF202507 to Yongfeng Z. and SKLTCBBSH202501 to X.X.).

Author contributions

Y.Z. and Z.M. designed the project. Z.M. and X.X. carried out the bioinformatics analyses and other major work described in the article. Z.M.,

X.X., T.Z., and Z.C. carried out the material collection. Z.M., X.X., T.Z., W.P., and S.C. performed genome assembly and annotation. Z.M., X.X., F.Z., Y.W., H.X., Y.Z., Z.L., and Y.L. made contributions to data analysis and visualization. Y.Z., Z.M., and X.X. wrote the paper. Z.L., H.X., Q.L., T.H., W.W., Z.J., M.Z., Y.P., J.W., and B.G. revised the manuscript. All authors approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-70026-w.

Correspondence and requests for materials should be addressed to Yongfeng Zhou.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.