Genomic insight into the origin, domestication, dispersal, diversi□cation and human selection of Tartary buckwheat

Yuqi He[1†], Kaixuan Zhang[1†], Yaliang Shi[1†], Hao Lin[1†], Xu Huang[1†], Xiang Lu[1†], Zhirong Wang[1†], Wei Li[1], Xibo Feng[2], Taoxiong Shi

Background

The current appearance of crops is the result of the combined action of their natural and cultural environments [1]. During long-term crop domestication, allelic variations with desired qualities in traits such as yield, taste, and cultivation practices were artificial selected [2]. When these domesticated crops spread to broader geographical areas through human migration, only those types adapted to their new environment and of use to humans would be selected, leading to the gradual expansion of the proportion of the allelic variations within the population, and ultimately differentiation into diverse germplasm resources [3, 4]. The diverse germplasm resources also lead to different dietary habits, creating unique cultural environments for human concentrated communities in different regions [1]. Thus, the study of the genetic basis of crop domestication not only helps to promote crop genetic improvement, but also contributes to a comprehensively understanding of the history and development of modern agricultural societies.

Buckwheat belongs to the Polygonaceae family, which is known for its abundant pharmaceutical plants, including [illegible] and [illegible]. These pharmaceutical plants are rich in various bioactive substances with health promoting effects. As the food crop with the closest phylogenetic relationship to these pharmaceutical plants, buckwheat is generally considered to have more abundant bioactive substances than other more widespread main grain crops of the Poaceae [5]. Besides these health promoting effects, these substances are usually present due to their role in plant defense against biotic and abiotic stress [6, 7]. At present, there are two most widely cultivated buckwheat species, including self-pollinated Tartary buckwheat and self-incompatible common buckwheat [8]. The self-pollinated nature of Tartary buckwheat makes it more suitable for genetic diversity research than common buckwheat. Meanwhile, it is generally considered that Tartary buckwheat exhibited greater health protection efficacy and high-altitude adaptability than common buckwheat [9]. According to pharmaceutical classics such as 'Compendium of Materia Medica', 'Qian Jin Yao Fang', and 'Dictionary of Traditional Chinese Medicine', Tartary buckwheat has health beneficial effects such as calming the mind, strengthening the heart, anti-inflammatory bioactivities as well as the ability to promote weight loss. However, compared to wild accessions, domesticated Tartary buckwheat bear as a common set of traits, known as the domestication syndrome, which includes loss of seed shattering, increased seed size and reduced seed dormancy [10]. Along with changes in these visible traits, a lower level of many bioactive compounds has been selected for, likely due to the fact that they are usually bitter in taste [11, 12]. Given this, study of the domestication history of Tartary buckwheat will improve the understanding of the genetic basis of the accumulation of bioactives as well as the utilization of wild buckwheat for molecular breeding.

The unique natural characteristics of Tartary buckwheat and not being a member of the Poaceae distinguish it from the major grain crops, increasing the interest in its domestication history. De Candolle initially speculated that it originated in northern China. However, no one has confirmed the distribution of wild buckwheat in the region, leading to this speculation is not widely accepted [13]. Subsequently, using molecular markers, Ohnishi speculated that Tartary buckwheat originated in the eastern part of Tibet and the neighboring areas of Yunnan and Sichuan [14, 15]. Although the historiography, morphology, reproductive biology and the distribution of wild relatives supports

one representative landrace with an easily-dehulled-phenotype collected from southwest China. By contrast 496 accessions were described in a previous study [8]. We then performed phylogenetic and genetic structure analyses of the Tartary buckwheat population, examining two to six clusters (K) (Fig. 1b). At K = 6, the outgroup forms its own group, and Tartary buckwheat was optimally characterized by the presence of five major clusters. ree clusters are similar to those found previously [8], i.e., accessions collected from the Himalayan region formed Himalayan wild (HW) group, accessions mainly collected from southwestern China formed Southwestern landraces (SL) group, accessions mainly collected from northern China formed Northern landraces (NL)

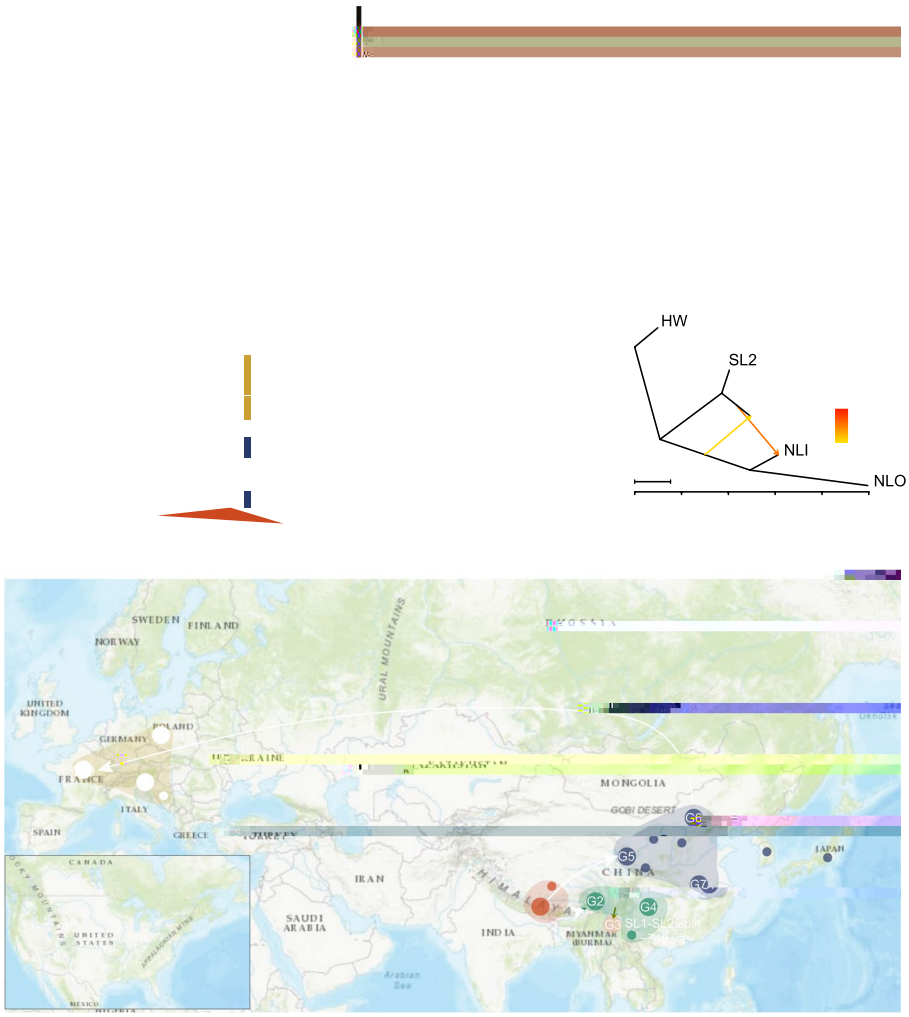
group. In addition, NL landraces splitted into two groups in our analysis (one group of NL within China landraces [NLI] and the newly sequenced NL outside China landraces [NLO]), and the SL group divided into two sub-groups, namely SL1 and SL2. e newly added wild accessions grouped with the HW group. e clustering based on K = 2 illustrated the previously reported strong north-south divide. NLI group divided into two subgroups (K = 5) while merged as one (K = 6). e principal component analysis (PCA) revealed a similar population structure compared to the evolutionary tree analysis (Fig. 1c). e population structure shown here is consistent with that in previous research [8].

Nucleotide diversity ( ) and population fixation statistics ($_{ST}$) were subsequently estimated in five major groups (Fig. 1d; Additional file 1: Table S3; Additional file 2: Fig. S1, S2). e HW group (Himalayan accessions mainly grouped) exhibited higher genetic diversity compared to SL group (Yunnan and Sichuan province accessions mainly grouped) and NL group (northern China accessions mainly grouped). e $_{ST}$ between NLO and NLI is smaller than that between NLO and other groups, supporting the hypothesis that Tartary buckwheat was spread from northern China to Europe. Linkage disequilibrium (LD) decayed faster in the HW group than other groups (Fig. 1e), which was consistent with the highest in HW, confirming that the Himalayan region is more likely to be the origin center of cultivated Tartary buckwheat compared to northern China and Sichuan or Yunnan province in southern China. e LD in the NLO subgroup decayed slower than that in NLI, which might be expected given that the NLO accessions have been selectively bred and improved, which is consistent with the genetic diversity and population fixation statistics. In summary, these results demonstrate that Tartary buckwheat originated in the Himalayan region, and subsequently domesticated, forming the SL and NL groups, respectively.

Dispersal of Tartary buckwheat followed routes of human migration

Human migration has promoted the spread of many cultivated crops [1]. Population structure analysis suggested a Himalayan origin and divergent selection of Tartary buckwheat (Fig. 1). To further investigate the possible dispersal history of Tartary buckwheat, the population relationship was further analyzed using $_3$ statistics, with other species as the outgroup. e results further confirm the close relationship between SL1 and SL2 and between NLI and NLO and the relatively distant relationship between NL and SL groups (Fig. 2a), in accordance with the population structure (Fig. 1). en, using qpGraph analysis to consider the potential population mixing events (Additional file 2: Fig. S3), similar relationships between subgroups in SL and NL were found, suggesting the reliability of the grouping.

Subsequently, SMC++ was used to estimate the divergence time (Fig. 2b; Additional file 2: Fig. S4) among the five populations. Cultivated accessions diverged from the HW group around 2,028-5,814 years ago, which coincides with the time when the Yi people migrated from Tibet to the Sichuan province [23]. According to the Yi classic 'Southwest Yi Annals', the ancestors of the Yi people migrated from the Himalayan region, seemingly bringing Tartary buckwheat to Sichuan province. Subsequently, the SL and NL groups di erentiated approximately 1,450-4,411 years ago. e SL1/SL2 populations and the NLI/NLO groups diverged at a similar time, ca. 300-1,900 YBP, which was in accordance



with the time of the westward expansion of the Mongol Empire. e result of e ective population size (; Additional file 2: Fig. S5) exhibited similar divergent time. We therefore speculate that Tartary buckwheat spread to Europe with the expansion of the Mongol Empire, which was also illustrated in ' e History of e Mongol Empire'.

To evaluate the accuracy of the candidate dispersal route of Tartary buckwheat, we divided Tartary buckwheat accessions into ten mini-groups based on geographical distribution. e silhouette score based on genotype showed the groups can be well clustered (Additional file 2: Fig. S6). e $_3$ statistics revealed the genetic relationship

between HW and SL is comparable to that between HW and NL, suggesting HW is the common ancestor of SL and NL groups (Fig. 2c). e accessions collected from outside China (G8-G10) have closer genetic relationship with NL group (G5-G7) compared to HW (G1) and SL groups (G2-G4). e phylogenetic tree showed that compared to individuals distributed in northern China (G5-G7) and outside China (G8-G10), individuals in G1 (located in Himalayan region) possess a closer genetic relationship with outgroup (Fig. 2d). And individuals in G5 (located in Qinghai-Gansu) were closer to their ancestors than other individuals in NL group, which was in accordance with the dispersal route of Tartary buckwheat from the Himalayas to northern China. Not only phylogenetic tree (Fig. 2d) but also pairwise fixation index (Fig. 2e) showed that individuals in NLO (G8-G10) have closer relationships with G6 (Inner Mongolia-Hebei) than other mini-groups in NL (G5 and G7, located in Qinghai, Gansu, Hunan, Hubei and Jiangxi province), supporting the hypothesis that Tartary buckwheat spread to Europe through the Mongolian region.

In cases where populations are not geographically isolated admixture and introgression can occur, and in some cases this can be adaptive [36]. TreeMix identified two instances of gene flow among the five subpopulations, namely a substantial migration from SL1 to NLI and a lesser migration from the NLI/NLO ancestor to SL1 (Fig. 2f; Additional file 2: Fig. S7). e analysis additionally reveals that the SL1 population introgressed more genetic components into NLI than NLO (Additional file 1: Table S4; Additional file 2: Fig. S8). D-statistics found that NLI accessions located in Hunan-Hubei-Jiangxi province (G7) were characterized by substantial introgressions from accessions located in Qinghai (G5; $|Z\ score| = 4.09$, $= 4.26\times10^{-5}$) and Inner Mongolia province (G6; $|Z\ score| = 10.2$, $= 2.24\times10^{-24}$), possibly due to the close geographical proximity (Additional file 1: Table S5). Such large-scale gene transfer may enhance the genetic diversity of the accessions.

Subsequently, a pattern diagram displaying the dispersal route of Tartary buckwheat was summarized (Fig. 2g). About 3,300 years ago, possibly with the migration of the Yi people, Tartary buckwheat spread from the Himalayas to southwestern China. Around 3,000 years ago, Tartary buckwheat spread to northern China. Around 1,500 years ago, the SL1 and SL2 populations di erentiated and formed SL1 subgroup with higher domestication degree. Subsequently, possibly with the westward expansion of the Mongol Empire about 1,000 years ago, Tartary buckwheat dispersed from northern China to Europe, ultimately resulting in its current global distribution pattern.

Selection targets during domestication

To identify potential selective signals involved in the domestication of Tartary buckwheat, we performed the cross-population composite likelihood ratio test (XP-CLR) between HW and SL (Fig. 3a) and between HW and NL (Fig. 3b). We identified genomic regions in the top 5% of the distribution of XP-CLR values which revealed 404 sweeps containing 2,909 genes in the HW-SL comparison and 415 sweeps containing 2,793 genes in HW-NL (Additional file 1: Table S6, S7). Among them, 1,282 genes overlapped in both comparisons (Additional file 1: Table S8; Additional file 2: Fig. S9). e remaining 1,627 (56% of the candidate genes) in HW-SL and 1,511 (54%) in HW-NL represent those with divergent histories since the origin of domesticated Tartary buckwheat. Only

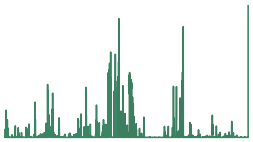
330 genes located in 44 selective sweeps in HW-SL comparison and 317 genes located in 78 selective sweeps in HW-NL comparison were overlapped with previous study. is was because more than half of HW accessions and 10% of the NL accessions were newly added in this study. In addition, de-correlated composite of multiple signals (DCMS) approach was also used to identify selective sweeps. 2,803 genes in 410 selective sweeps were identified in HW-SL comparison, and 3,377 genes in 487 selective sweeps were identified HW-NL comparison (Additional file 1: Table S9, S10). Only 785 genes were overlapped in both comparisons (Additional file 1: Table S11), further confirming the independent domestication process.

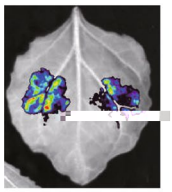
Many genes selected during domestication in both SL and NL are potentially involved in domestication related traits (Additional file 1: Table S8). For instance, a receptor-like protein kinase [37] was a key gene regulating plant height. ioredoxin [38], pathogenesis-related protein [39] and remorin [40] were well-known plant disease resistance associated genes while some homologous of GRAS transcription factors [41] has previously been defined as being involved in grain weight regulation. e identification of these domestication trait related genes provides a genetic basis for the mechanism underlying Tartary buckwheat domestication.

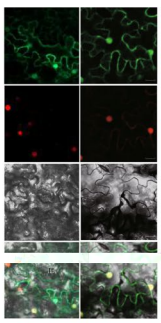
AG4-HGI 3 is a devastating soil-borne pathogen that seriously threatens Tartary buckwheat cultivation [7]. Previous research demonstrated the content of metabolites associated with disease resistance decreased during Tartary buckwheat domestication [12]. We therefore investigated whether genes responsible for resistance to underwent selection during Tartary buckwheat domestication. Notably, one significant locus identified by GWAS of disease resistance [7] was found to have undergone selection during domestication of the NL and SL groups (Fig. 3c; Additional file 1: Table S12; Additional file 2: Fig. S10). Haplotype analysis identified two variants located at 833 bp and 530 bp in the promoter of a gene encoding L-gulonolactone oxidase (,), which is involved in ascorbate biosynthesis (Fig. 3d) [42, 43]. Phylogenetic analysis demonstrates this gene is an orthologue of L-gulonolactone oxidase in other species (Additional file 2: Fig. S11). Accessions

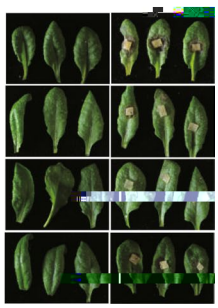
(See figure on next page.)

F , .3 Variation of *FtGULO* controls disease resistance during Tartary buckwheat domestication. **A**-**B** Selective sweeps identified through comparisons between HW and SL (**A**) and HW and NL (**B**) using XP-CLR (cross-population composite likelihood-ratio test). The dashed line represents the top 5% of values therefore scores in these regions were regarded as selective sweeps. **C** Local Manhattan plot of GWAS signals on Chr 8 for resistance to *R. solani* AG4-HGI 3. The dashed line represents the threshold ($-\log_{10}P = 5$). **D** Schematic diagram of *FtGULO* gene structure. Two SNPs in the promoter of *FtGULO* were marked as red letters and result in haplotypes (Hap) A and T. **E** Box plots show disease index in plants carrying the two haplotypes (Hap). $n_{Hap\text{-}A} = 8$, $n_{Hap\text{-}T} = 234$. *P* values were calculated using a two-tailed *t*-tests. **F** Expression of *FtGULO* in accessions harboring the two haplotypes. Error bars indicate the ± s.d., $n = 6$. Significance was tested using one-way ANOVA. **G** Transcription activity of *FtGULO* promoters with two haplotypes. **H** Disease index of accessions among HW, NL and SL groups. $n_{HW} = 10$, $n_{NL} = 96$, $n_{SL} = 140$. Significant was tested using two-tailed *t*-tests. *, $P < 0.05$. **I** Frequencies of the two haplotypes in the HW, NL and SL groups. **J** Subcellular localization of *FtGULO*-GFP fusion protein transient expression in *N. benthamiana* leave cells. Scale bars, 10 μm. (K-L) Relative expression levels of *FtGULO* during *R. solani* infection (**K**) and MeJA treatment (**L**). Histone H3 was used as the internal reference. **M** Disease index of *Arabidopsis* lines heterologously expressing *FtGULO*. Significant di erences were identified using one-way ANOVA. $n = 6$. **N** Phenotypes of *Arabidopsis* WT lines and lines heterologously expressing *FtGULO* with and without infection with *R. solani* AG4-HGI 3. Scale bars, 1 cm









harboring the A-haplotype exhibited higher disease resistance and higher expression compared to those harboring the T-haplotype (Fig. 3e, f), suggesting is an important locus underlying resistance to AG4-HGI 3 in Tartary buckwheat. Transient activation assays demonstrate that higher LUC expression in leaves transient expressing promoters of the A-haplotype compared to those of the T-haplotype, confirming the natural variations in the promoter of were involved in Tartary buckwheat disease resistance (Fig. 3g). e disease index was significantly greater in the SL and NL groups compared to HW (Fig. 3h), confirming disease resistance decreased during Tartary buckwheat domestication. Moreover, the resistant haplotype

was almost completely absent from the SL and NL groups (Fig. 3i; Additional file 2: Fig. S12). Subcellular localization experiments demonstrated that was located in both the nucleus and cytoplasm (Fig. 3j), while the expression of was induced by infection and methyl jasmonate (MeJA) treatment (Fig. 3k, l), suggesting might be involved in jasmonate-mediated disease responses. Heterologous expression of in (Additional file 2: Fig. S13) demonstrated that the three overexpression lines exhibited enhanced disease resistance compared to the wild type (Fig. 3m, n). In summary, these results illustrate that the natural variation in the promoter of was involved in disease resistance reduction during Tartary buckwheat domestication through regulating expression.

Selection targets during Ecogeographic adaptation in China

Environmental di erence caused by varied geographical distribution are important reasons for crop divergence [33]. e northern and southern regions of China have highly di erent climates and soil, leading to the formation of locally adapted germplasm. To characterize the genetic basis of Tartary buckwheat di erentiation caused by environmental adaptation, we used the XP-CLR and DCMS test to compare the SL and NL groups (Additional file 1: Table S13, S14). A total of 430 selective sweeps containing 2,968 genes were found that showed evidence of selection (Fig. 4a). Among them, methyl-cpg-binding domain protein was responsible for Arabidopsis flowering time regulation [44], while histone deacetylase [45] and ABA 8' hydroxylase [46] were involved in plant response to drought stress, heat shock proteins [47] were involved in Arabidopsis heat tolerance. Gene ontology (GO) and KEGG analysis revealed enrichment of categories involved in hormone, chemical, and auxin response, suggesting that response to divergent environments played a significant role in the divergence and evolution of these two groups (Additional file 2: Fig. S14).

In the arid and semi-arid regions of northern China, due to the low precipitation and the high evaporation, salt dissolved in the water is prone to accumulate on the soil surface, resulting in higher salt content in the soil [48]. To study the molecular basis of Tartary buckwheat adaption to this soil salinity di erence, salt tolerance of 151 Tartary buckwheat accessions was investigated (Additional file 1: Table S15). A genome wide association study (GWAS) with salt tolerance as the phenotype (Fig. 4b; Additional file 1: Table S16) identified a significant association on chromosome 2, which overlapped with a selective sweep identified in the SL-NL XP-CLR test. Haplotype analysis identified two variants in the promoter of a gene encoding a protein kinase (; ; Fig. 4c). Phylogenetic analysis demonstrates this gene is an orthologue of protein kinase in other species (Additional file 2: Fig. S15). Accessions with Hap-l exhibited greater salt tolerance and expression compared to that with Hap-2 (Fig. 4d, e), and the frequency of Hap-1 in high soil Electrical Conductivity (ECE) was higher than that in low soil ECE conditions (Additional file 2: Fig. S16), suggesting might play a positive role in Tartary buckwheat salt tolerance. Transient activation assays demonstrated that promoters with Hap-1 possess higher transcription activity compared to those with Hap-2, further confirming the natural variations in the promoter of were involved in Tartary buckwheat salt tolerance (Fig. 4f). e salt tolerance was significantly lower in the SL groups than the NL group (Fig. 4g), further confirming the di erentiation of

salt tolerance in populations located in northern and southern China. e frequency of Hap-1 was greater in NL than SL (Fig. 4h; Additional file 2: Fig. S17). Subcellular localization experiments demonstrated that

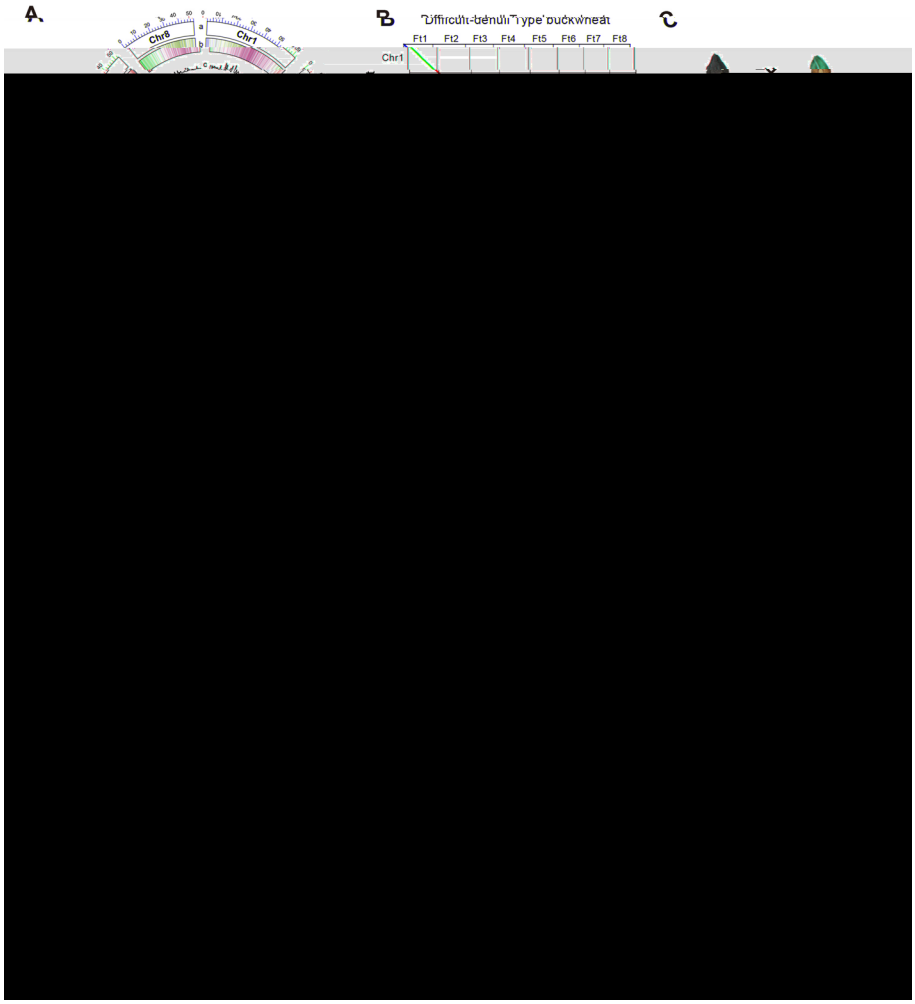


Figure 5 Structural variation of *FtXIP* controls the domestication of easily-dehulled type Tartary buckwheat. **A** Genome features of EDT. The outermost circle represents each chromosome of the genome. The second to fifth circles indicate gene density, SNPs density, deletion density, and insertion density, respectively, using a window size of 500-kb. **B** Gene dot map between easily-dehulled type buckwheat (EDT) and difficult-dehulled type (DDT) Tartary buckwheat. **C** Diagram representing the generation of the EDT x DDT recombinant inbred lines (RILs). **D** Genome wide Δ(SNP index) plot of the population derived from a cross between EDT and EDT. The black lines indicates tricube-smoothed Δ(SNP index), and the gray lines indicate corresponding two-sided 99% confidence intervals. **E** Insertions and deletions larger than 50 bp and within 5 kb of genes in the chr 2 QTL intervals. **F** Expression of genes with insertions and deletions in the QTL intervals in the seed coats of EDT and DDT at the 20-day after pollination (DAP) stage. Each small square represents the differentially expression level of a gene between EDT and DDT. Square with gene ID exhibited the differentially expressed genes. The red gene ID represents *FtXIP*. **G** Schematic diagram showing the deletion of 1,140 bp in the promoter region of *MqXIP* gene. **H** Transient expression assay was conducted to compare the transcription activity of *MqXIP* and an empty vector. **I** The expression level of *XIP* in DDT and EDT Tartary buckwheat. The error bars indicate the ± s. d, $n = 6$. The *P* value was calculated using two-tailed *t*-tests. $P < 0.05$

population, derived from a cross between EDT and DDT buckwheat, was constructed and, along with the parental lines, subjected to Illumina HiSeq2500-based re-sequencing [51]. Among the 221 F7 lines, 79 lines were predominantly EDT, and the remaining 142 lines were predominantly DDT (Fig. 5c). Quantitative trait locus (QTL) analysis identified one major QTL controlling the easily dehulled phenotype on Chr2 (Fig. 5d), which

was consistent with the region identified previously [51]. Analyzing the insertions and deletions > 50 bp within the QTL interval, 54 genes that exhibited structural variants within the 5-kb range upstream and downstream were identified (Fig. 5e).

Subsequently, expression of these genes in EDT and DDT seeds was quantified [52]. Eleven genes displayed > 2-fold expression di erences between EDT and DDT at the 20-day after pollination (DAP) stage of seed development (Fig. 5f; Additional file 1: Table S21). By combining the gene function annotations, a gene encoding a xylanase inhibitor ( ) that suppresses xylan degradation in the plant cell wall [53] was identified which could plausibly contribute to the easily dehulled trait. Compared to DDT, EDT exhibited a 1,140 bp deletion in the region 3-kb upstream of the start codon of (Fig. 5g). A transient activation assay demonstrated that the 1,140 bp sequence in the promoter resulted in significantly higher activity compared to the empty vector (Fig. 5h), and this region exhibited many cis-acting elements (Additional file 1: Table S22). And the expression of is higher in DDT compared to EDT, suggesting this region could significantly up-regulate gene expression in developing DDT seeds. Hence, we speculated that the SV in the promoter region may have resulted in reduced expression of , ultimately leading to the easily dehulled trait in EDT.

D

As human societies around the world transitioned to agriculture, crop plants began the long-term process of domestication [54]. e only food crop in the Polygonaceae family, buckwheat is thought to have had its origin in south-eastern China [14–17]. However, due to the limited sampling and methods, more molecular evidence is needed to confirm this hypothesis. Previously, we attempted to validate the center of origin of Tartary buckwheat [8], however, the wild resources of Tartary buckwheat are mainly distributed in high-altitude areas of the Himalayas, posing serious challenges for the acquisition of this wild material. Here, we obtained 19,321,018 SNP from the genome re-sequencing data of 567 Tartary buckwheat accessions collected from throughout the world. Both the sampling representativeness and the variations are greater than previous studies [14, 15]. We found the HW group (Himalayan accessions enriched) exhibited higher nucleotide diversity ( ) and faster LD decay compared to SL group (Yunnan and Sichuan accessions enriched) and NL group (northern China accessions enriched), confirming that Tartary buckwheat indeed originate from the Himalayan region, which is di erent from the center of origin of other grain crops of the Poaceae. As one of the youngest and loftiest mountain chains in the world, the Himalayas has unique climatic environments caused by large altitude variations, resulting in abundant plant diversity [55]. us, the confirmation of the Himalayan origin of Tartary buckwheat not only helps to protect the genetic diversity in its center of origin, thus promoting the use of wild germplasm resources for molecular breeding, but also has unique significance for the development of agricultural civilization, the protection of the global plant diversity.

Human migration has changed the face of the world, including the appearance and distribution of crops [56]. Due to the excellent environment for Tartary buckwheat cultivation, the Yi people, an ethnic minority of southwestern China, were the earliest people planting Tartary buckwheat where it is traditionally regarded as a staple food [23]. According to the Yi language classic 'Southwest Yi Annals', the ancestors of

the Yi people came from 'outside the yak field', suggesting that the Yi people migrated from the Himalayan region. According to pollen abundance of Tartary buckwheat, the ancestors of the Yi people began planting Tartary buckwheat about 4,000 years ago [23]. By analyzing the genetic relationships and the timing of divergence between modern groups, we found that Tartary buckwheat in the southwest region spread from the Himalayas around 3,000~4,000 years ago, in exact accordance with the migration of the Yi people. ere is a custom that brides bring their own Tartary buckwheat seeds as a dowry to their new homes, when the Yi people get married, which may promote the spread of Tartary buckwheat. Linguistic evidence suggested that European Tartary buckwheat is closely related to the Mongols. According to ' e History of e Mongol Empire', Tartary buckwheat spread to Europe with the expansion of the Mongol Empire. European historical data shows that Tartary buckwheat was introduced into Europe in the Middle Ages [21, 22]. A close phylogenetic relationship was found between accessions from northern China and outside China, indicating that Tartary buckwheat was introduced to Europe potentially only once from northern China [8]. However, due to only a few accessions used in our analysis which came from outside China, this conclusion needs further verification. e predicted divergence time suggested Tartary buckwheat was introduced to Europe around 1,000 years ago, which closely mirrors the time of the Mongols westward expansion. ese results are of great significance not only for genetic improvement of Tartary buckwheat, but also for the understanding of the development of human cultures. In addition, as phylogeny showed individuals distributed in Qinghai-Gansu province (G5) were closer to their ancestors than other individuals distributed in Inner Mongolia-Hebei province (G6) and Hunan-Hubei-Jiangxi province (G7), and D-statistics exhibited a week gene flow ($Z < 3$) from individuals distributed in Qinghai-Gansu province to that distributed in Inner Mongolia-Hebei province, implying gene transfer between individuals in Qinghai-Gansu and Inner Mongolia-Hebei province.

Compared to wild germplasm resources, domesticated crops usually exhibit increased yield, better taste, and a plant architecture more suitable for cultivation. However, resistance to biotic or abiotic stress is often decreased during domestication, resulting in vulnerability to diseases and extreme weather and as such bringing severe yield losses [57]. Previous research demonstrated disease resistance associated metabolites are reduced in content in domesticated Tartary buckwheat relative to the wild accessions [12]. Here, by identifying selective sweeps between domesticated groups and the wild group, candidate genes responsible for domestication and diversification were identified. By combining genome-wide association studies with disease index of Tartary buckwheat collected worldwide, transcriptomics of Tartary buckwheat response to infection and MeJA treatment, , a gene involved in ascorbate biosynthesis [42] was found to be responsible for decreased disease resistance in domesticated Tartary buckwheat. Only 25% resistant haplotype were identified in HW group, which might be due to that it is a newly generated haplotype in HW group and has not yet introgression into the domesticated group. But this speculation needs to be proved by further study. e exploration of such domestication genes will help transform wild plants into cultivated crops in a relatively short time by precisely changing key genes of important domestication traits [58].

Di erent genetic adaptations drive the formation of di erent ecotypes, and there are significant di erences in the precipitation and temperature between northern and southern China, resulting in higher soil salinity in northern China compared to southern China [48]. We provide multiple lines of evidence that the increased frequency of a haplotype of ⁄ with high expression is responsible for the greater salt tolerance of Tartary buckwheat from northern China than those from southern China. is suggest that ⁄ plays an essential role in salt tolerance, which is according to the function of its houmologous [59, 60]. Besides the natural environment, the cultural environment will also generate unique germplasm resources that adapt to the dietary habits of local people [1]. e easily-dehulled type Tartary buckwheat is a unique landrace used for steaming as a staple food, wine- and tea- making in areas settled by the Wa people. Its easily dehulled nature of EDT allows local Wa people to use ancient artificial wooden mortars and pestles to dehull Tartary buckwheat and steam together with rice as staple food to prevent lysine deficiency. Comparative genomics and QTL analyses identified a xylanase inhibitor, a gene inhibiting the degradation of xylan, the main component of hemi-cellulose [53], was involved in the easily-dehulled phenotype. Not only do the results of this study demonstrate the center of origin and domestication history of Tartary buckwheat but the identification of genes responsible for important traits to productivity and cultivation that di erentiate the groups, therefore providing important tools for the genetic improvement of this important dual use food and medicinal crop.

C ,

In conclusion, our genomic studies provide valuable insights into the domestication, dispersal, and diversification of Tartary buckwheat. rough the analysis of wild and domesticated germplasm, we have unraveled the complex evolutionary history of this crop. e identification of selective sweeps, population relationship, and genetic markers associated with traits like salt tolerance has shed light on how adaptive processes and cultivation practices have shaped Tartary buckwheat. Additionally, the discovery of candidate genes, such as ⁄ , has highlighted the molecular mechanisms underlying important agronomic traits. Further research and genetic investigations are necessary to fully comprehend the complexities and dynamics of its evolutionary journey.

M , ,

Genome re-sequencing, SNP calling and population structure analysis

A total of 567 Tartary buckwheat accessions, including 501 cultivated accessions and 66 wild accessions, were used in this study. Among them, 474 accessions were collected from China, and 93 accessions were collected from the other 16 countries (Additional file 1: Table S1). 489 accessions were re-sequenced in previous research [8, 61], and 78 accessions were newly re-sequenced in this study. Genomic DNA was extracted using cetyltrimethylammonium bromide (CTAB) as previously described [8]. Genomes were re-sequenced using Illumina NovaSeq 6000 platform. Raw reads in fastq file were trimmed to remove poor quality bases and adapters using Trimmomatic v0.33 [62] based on the manufacturer's adapter sequences. A total of 7.7 Tb of clean data (i.e., after removing adapters, reads containing poly-N, and low-quality reads) was obtained. Clean reads were then mapped to the reference genome of Tartary buckwheat variety Pinku1

[63] using BWA-MEM [64]. After sorting by samtools, duplicated reads were removed using MarkDuplicates in Picard v1.13 (http://broadinstitute.github.io/picard/). Average depth was ~27.5× and mapping rate > 90% for each Tartary buckwheat accession. SNPs and small indels (1–50 bp) were called using the GATK pipeline [65]. Variants were called using GATK HaplotypeCaller, and then a joint-genotyping analysis of the gVCFs was performed on all merged samples. SNPs were filtered based on parameters previously used [8]. Population genetic structure was analyzed using the program ADMIXTURE v1.23 [66] with the putative number of populations (K values) from two to six. A maximum likelihood-based phylogenetic tree analysis was performed using IQ-TREE v1.6.6 [67]. Principal component analysis (PCA) was performed as previously described [8]. e nucleotide diversity ( ) was calculated using VCFtools in 20-kb sliding windows with a 10-kb step. e fixation statistics ($_{ST}$) between di erent populations were calculated using a set of Python scripts (https://github.com/simonhmartin/genomics_general/popgeneWindows.py) with the parameters set as -w 100000, -s 10000, -f haplo.

Identi□cation of selective sweeps

To detect putative selective sweeps among di erent groups, the cross-population composite likelihood ratio test was performed using XP-CLR v1.1 [68]. Genome regions with top 5% XP-CLR values were considered as selected regions. Four statistics including XP-CLR, , eta and Tajima D were combined into a single DCMS framework [69]. Genome regions with $\nu < 0.05$ were considered as selective regions.

GWAS analysis

Only SNPs with MAF $\geq$ 0.01 [70–72] and missing rate $\leq$ 0.1 in a population were used for GWAS. E cient Mixed-Model Association eXpedited program (EMMAx) was used for GWAS analysis [73]. e significance threshold was set at $\nu = 1\times10^{-5}$.

Admixture graph modeling and introgression analysis

e SNP dataset was filtered using '-mac 1 -max-alleles 2' in VCFtools [74] and '-indep-pairwise 50 5 0.3' in plink [75], and the convert program from AdmixTools was used to produce eigenstrat format data files. In order to measure allele sharing of three or four sets of subpopulations and to report the |Z|-score between predicted and observed values, the $_{3}$ and $_{ST}$ statistics were computed using ADMIXTOOLS 2.0 (https://uqrmaie1.github.io/admixtools) [76]. A heuristic algorithm to iteratively fit increasingly complex models, qpbrute (https://github.com/ekirving/qpbrute) filtered 1,183 possible admixture graph models and recorded ten graphs that left no $_{4}$ outliers (|Z| < 3) [77]. qpBayes [77] was then used to test the best-fit graph and compute the marginal likelihood of models and their Bayes factors. Analysis using qpGraph to detect the demographic graphs, and the best fitting model (no $_{4}$ outliers, |z|>=3) was carried out to assess putative population relationship under potential admixture events.

To remove the confounding e ect from unclear subpopulation classification, we tested refined populations with additional silhouette filtering (Silhouette score >0) according to the methods described previously [78]. After filtering out monomorphic SNPs and those with missing data (missing rate $\leq$ 0.01), gene flow between the five population were estimated using Treemix v1.13 [79]. To refine the introgressed genomic regions, f_{dM}

statistics were calculated along the whole genome using python scripts (https://github.com/simonhmartin/genomics_general) with 50-kb sliding windows and a 50k step. Geographic subsets of accessions were clustered using latitude and longitude coordinates by the K-means cluster method [80] with range extension less than 5 radius. After the filtering of multidimensional scaling analysis and silhouette examine of pairwise identity-by-state (IBS) distance matrix, ten representative groups consisting of 239 accessions were selected based on distinct population classification and sample size. en the f-statistics and D-statistics were implemented using software referred as above. For D-statistics, only |Z score| >3 were considered as significant [31, 33, 81, 82].

Estimation of divergence time and demographic history

e split function in SMC++ [83] was used to estimate the divergence times and the e ective population size among di erent subpopulations. For normalizing population size, we randomly selected ten di erent samples of each subpopulation per time and ran 20 repeats that covered all samples. e mutation rate was set as 7×10^{-9} per synonymous site for each generation, and split time was calculated using one generation per year.

Genome assembly and comparative genome analysis

e easily-dehulled type (EDT) genomes was assembled using PacBio HiFi reads and the hifiasm [84] assembly method. e Hi-C data was mapped to the corresponding contigs using the Juicer v1.6.2 pipeline [85]. Primary sca olds were constructed using 3D-DNA v180922 [86] with default parameters. e assembly was visually inspected and manually curated using Juicebox Assembly Tools v1.9.8 [87]. Another round of sca olding was performed using 3D-DNA v180922 to generate the final pseudo-chromosomes. To assess the completeness of the assembled genome, Benchmarking Universal Single-Copy Orthologous gene analysis (BUSCO) [88] was conducted using the conserved genes of the Embryophyta_odb10 as a reference. e SyRI v1.1 [89] comparison tool was used to identify SNP and SV between EDT and DDT using minimap2 v2.17 [90]. Structural variants were divided into four types: insertion, deletion, inversion and translocation.

The genetic basis of the easily-dehulled phenotype and candidate genes prediction

To identify candidate mutations associated with the easily dehulled trait, an F7 population was generated from a cross between EDT and DDT accessions. e RILs (Recombinant Inbred Lines) in the population were classified into two groups based on their hull phenotype: easily-dehulled type or di cult-dehulled type. To identify variants between the parental genomes, SNPs (Single Nucleotide Polymorphisms) were calculated using the R package QTLseqr [91], resulting in a ΔSNP index. Each RIL individual was subjected to re-sequencing, and subsequently, individuals of the same dehulled type were merged. e resulting vcf file used for QTLseq analysis included four SNP datasets: EDT, DDT, EDT-RIL, and DDT-RIL. e genomic regions with a ΔSNP index exceeding the 99% confidence interval were considered candidate regions. Genes within these regions are putatively associated with the easily dehulled trait.

Dual-luciferase assay

In the dual-luciferase assay, the promoter constructions were inserted into the pGreenII 0800-LUC vector for analysis. e GV3101 strains carrying the respective promoter constructs were cultured overnight at 28 °C. e cultures were then diluted to an OD600 of 0.6 using resuspension bu er containing 10 mM $MgCl_2$, 10 mM MES, and 100 mM acetosyringone. Separate leaves were injected with carrying the construct. e injected leaves were incubated in the dark for 1 day and then exposed to 2 days of light/dark cycles (23 /22 , 16 h day/8 h night), after which the injected leaves were detached and sprayed with a solution of 1 mM D-Luciferin sodium salt and 0.01% Triton X-100 in ddH_2O. e luminescence of the luciferase activity in the infiltrated area was captured using LB983 Nightowl II.

Real-time quantitative PCR (qRT-PCR)

Total RNA was isolated from plant material using a plant RNA extraction kit (Aidlab, Beijing, China). e extracted RNA was reverse transcribed into cDNA by TRUEscript RT MasterMix PCR (Aidlab, Beijing, China). Primer sequences are listed in Additional file 1: Table S23. BnActin/AtActin was used as the reference and SYBR Green (Takara, Kyoto, Japan) was used as the fluorochrome. e amplification reactions were performed using a Line Gene K thermal cycler (BioRad, USA) under standard conditions.

Transgenic plant construction and phenotype assay in *Arabidopsis thaliana*

Total RNA was extracted by using an RNApre Pure Plant Plus kit (Tiangen, Beijing, China). First-strand cDNA was synthesized with a HiScript III RT SuperMix for qPCR (Vazyme, Nanjing, China). e coding sequence was cloned into pCAMBIA-1302. e overexpression lines were conducted and generated by GV3101 mediated transformation [92]. ree biological replicates were used, and the experiments were performed three times. Primer sequences are given in Additional file 1: Table S23. All genotypes were grown at 22 °C (day/night) under long-day conditions (16-h light/8-h dark). Disease index evaluation was conducted as previously described [7, 93]. Root length and physiological and biochemical assays of were used to evaluate the salt tolerance of transgenic plants. e e ect of NaCl on root length of was studied. Five-day-old Col-0 and trans-genic Arabidopsis seedlings were transferred to 1/2MS Agar medium containing 50 mM NaCl, and root length was measured and photographed after vertical culture for 7 days. e determination of malondialdehyde (MDA) content and peroxidase (POD) activity were performed according to methods described previously [94]. ree biological replicates were conducted and the experiments were performed three times.

e phylogenetic tree of GULOs and PKs were conducted using MEGA X based on the neighbor-joining method [95, 96].

Salt tolerance assay in Tartary buckwheat germplasm resources

To a petri dish covered with two layers of filter paper was added 5mL water and 20 seeds were evenly placed on the filter paper and cultured at 25 (±1) with 12 hours

daylength. Experiments were repeated three times. e germination rate, germination index and membership function value were calculated according to methods illustrated in the previous research [97]. GWAS was performed using membership function value. e Electrical Conductivity (ECE) was searched in Harmonized World Soil Database v 1.2 (HWSD v1.2) based on the longitude and latitude information of the location where accession obtained. Accessions with ECE < 0.2 were regards as samples from low-salinity land, and those with ECE > 1.9 were regards as samples from high-salinity land.

Subcellular Localization

Full-length cDNAs of and were amplified (primer sequences in Additional file 1: Table S23) and inserted into the pCAMBIA1300-GFP vector. p2300-35s-H2B-mCherry was used as a nuclear marker. e plasmid was transferred into leaves using GV3101-mediated transient infiltration [92]. Subcellular localization was observed using a laser scanning confocal microscope (Zeiss LSM900) with the wavelengths of 488 (excitation)/500 to 530 nm (emission) for GFP and 561 (excitation)/590 to 640 nm (emission) for mCherry.

The online version contains supplementary material available at https://doi.org/10.1186/s13059-024-03203-z.

Additional le 1: Table S1. Summary of the 567 Tartary buckwheat accessions and 7 outgroup accessions used in this study with the mapping rate, depth and coverage. **Table S2.** Distribution of SNPs into various genomic regions in Tartary buckwheat. **Table S3.** Statistical di erences between pairwise F_{ST} values between the di erent populations. **Table S4.** Genomic fragments with evidence for introgression based on F_d and F_{dM}.**Table S5.** Interpopulation introgression based on D- statistics. **Table S6.** Putative selective sweeps between HW and SL based on XP-CLR analysis. **Table S7.** Putative selective sweeps between HW and NL based on XP-CLR analysis.**Table S8.** Genes in selective sweeps in both between HW and SL and between HW and NL based on XP-CLR analysis. **Table S9.** Putative selective sweeps between HW and SL based on DCMS analysis. **Table S10.** Putative selective sweeps between HW and NL based on DCMS analysis. **Table S11.** Genes in selective sweeps in both between HW and SL and between HW and NL based on DCMS analysis. **Table S12.** Candidate genes associated with disease index identified by GWAS with MAF >= 0.01. **Table S13.** Putative selective sweeps between SL and NL based on XP-CLR. **Table S14.** Putative selective sweeps between SL and NL based on DCMS. **Table S15.** Salt tolerance of Tartary buckwheat accessions. **Table S16.** Candidate genes associated with salt tolerance identified by GWAS. **Table S17.** Sequencing reads used for assembly of EDT. **Table S18.** Assembly statistics of the EDT genome. **Table S19.** BUSCO analysis of the EDT genome. **Table S20.** Genetic di erences between EDT and DDT. **Table S21.** The expression of genes with insertions and deletions within the QTL inverval. **Table S22.** Summary of *FtXIP* insertion promoter *cis*-acting elements prediction in 1,400 bp region. **Table S23.** Primers used in this study.

Additional le 2: Figure S1. Nucleotide diversity of HW, SL1, SL2, NLI and NLO groups. **Figure S2.** F_{ST} between di erent groups. **Figure S3.** Assessment of graph model of Tartary buckwheat accessions. **Figure S4.** The range of estimated divergence times between the populations. **Figure S5.** Divergence time between HW and SL, HW and NL, SL1 and SL2, NLI and NLO groups predicted with SMC++. **Figure S6.** Individuals of ten mini-groups based on geographical distribution was carried out using silhouette scoring. **Figure S7.** Silhouette scores of individuals used for Treemix analysis was carried out using silhouette scoring. **Figure S8.** Evaluation of introgression components between di erent population. **Figure S9.** Gene ontology analysis and KEGG enrichment analysis of genes in selective sweeps identified in both HW vs. SL and HW vs. NL comparisons. **Figure S10.** Local XP-CLR plot of the locus *FtGULO* located. **Figure S11.** FtGULO phylogenetic based on the neighbor-joining method tree using full-length amino acid sequences of orthologues genes in Tartary buckwheat and other plants. **Figure S12.** Geographic distribution of the Hap-A and Hap-T Tartary buckwheat accessions. **Figure S13.** PCR analysis of *Arabidopsis* lines heterologously expressing *FtGULO*. **Figure S14.** Gene ontology analysis and KEGG enrichment analysis of gene in regions of selective sweeps between SL and NL. **Figure S15.** FtPK phylogenetic tree based on the neighbor-joining method tree using full-length amino acid sequences of orthologues genes in Tartary buckwheat and other plants. **Figure S16.** Frequencies of the two haplotypes in the low ECE and high ECE groups. **Figure S17.** Geographic distribution of the Hap-1 and Hap-2 Tartary buckwheat accessions. **Figure S18.** PCR analysis of *Arabidopsis* lines heterologously expressing *FtPK*. **Figure S19.** MDA content and POD activity in *Arabidopsis* heterologously expressing *FtPK* compared to WT with and without a salt treatment. **Figure S20.** Hi-C contact matrix of the high-quality chromosome-scale genome assembly of EDT. **Figure S21.** Collinearity among the assembly of EDT, the genetic map of the RIL population, the HERA version assembly (DDT genome used in this study) and the V2 version assembly of Tartary buckwheat

variety Pinku1 reference genome. **Figure S22.** The distribution of deletions and insertions between EDT and DDT on eight chromosomes of Tartary buckwheat.

Additional □le 3. Review history.

Peer review information

eightZ.L3, I.Kyomoso3, Der3., Vvidoertstrebutioečsbe9q. YYy1(er30(butimoerdatarD)alysi inser) -1.2 Td [(eightfi(e S-4(wtiosi99 .n.)50(Y)8)-3H3,)50(Y)8Wy1(er30(bu

2. Chen YH, Gols R, Benrey B. Crop domestication and its impact on naturally selected trophic interactions. Annu Rev Entomol. 2015;60:35–58.
3. Huang X, Huang S, Han B, Li J. The integrated genomics of crop domestication and breeding. Cell. 2022;185:2828–39.
4. Wang Z, Miao L, Chen Y, Peng H, Ni Z, Sun Q, Guo W. Deciphering the evolution and complexity of wheat germplasm from a genomic perspective. J Genet Genomics. 2023;50:846–60.
5. Huda MN, Lu S, Jahan T, et al. Treasure from garden: Bioactive compounds of buckwheat. Food Chem. 2021;335:127653.
6. Schenke D, Utami HP, Zhou Z, et al. Suppression of UV-B stress induced flavonoids by biotic stress: Is there reciprocal crosstalk? Plant Physiol. Biochem. 2019;134:53–63.
7. He Y, Zhang K, Li S, et al. Multi-omics analysis reveals the molecular mechanisms underlying virulence in *Rhizoctonia* and jasmonic acid-mediated resistance in Tartary buckwheat (*Fagopyrum tataricum*). Plant Cell. 2023;35:2773–98.
8. Zhang K, He M, Fan Y, et al. Resequencing of global Tartary buckwheat accessions reveals multiple domestication events and key loci associated with agronomic traits. Genome Biol. 2021;22:1–23.
9. Zhu F. Chemical composition and health e□ects of Tartary buckwheat. Food Chem. 2016;203:231–45.
10. Hunt HV, Shang X, Jones MK. Buckwheat: a crop from outside the major Chinese domestication centres? A review of the archaeobotanical, palynological and genetic evidence. Veg Hist Archaeobot. 2018;27:493–506.
11. Alseekh S, Scossa F, Wen W, et al. Domestication of crop metabolomes: Desired and unintended consequences. Trends Plant Sci. 2021;26:650–61.
12. Zhao H, He Y, Zhang K, et al. Rewiring of the seed metabolome during Tartary buckwheat domestication. Plant Biotechnol J. 2023;21:150–64.
13. De Candolle, A. (1883). Origine des Plantes Cultivées; G. Baillière et cie: Paris, France, Volume 43.
14. Tsuji K, Ohnishi O. Phylogenetic relationships among wild and cultivated Tartary buckwheat (*Fagopyrum tataricum* Gaert.) populations revealed by AFLP analyses. Genes Genet Syst. 2001;76:47–52.
15. Ohnishi O. Search for the wild ancestor of buckwheat III. The wild ancestor of cultivated common buckwheat, and of Tatary buckwheat. Econ Bot. 1998;52:123–33.
16. Ohnishi O, Konishi T. Cultivated and wild buckwheat species in eastern Tibet. Fagopyrum. 2001;18:3–8.
17. Fan Y, Ding M, Zhang K, et al. Overview and utilization of wild germplasm resources of the genus *Fagopyrum* Mill. In Chinese. J Plant Genet Resour. 2020;21:1395–406.
18. Bradley D. Proto-Tibeto-Burman grain crops. Rice. 2011;4:134–41.
19. Weisskopf A, Fuller DQ. Buckwheat: origins and development. In: Smith Claire, editor. Encyclopedia of Global Archaeology. New York: Springer; 2014. p. 1025–8.
20. Tang, Y., Ding, M., Tang, Y., et al. Germplasm resources of buckwheat in China. In Molecular Breeding and Nutritional Aspects of Buckwheat, Meiliang Zhou et al., ed. (Academic Press), 2016. pp. 13–20.
21. Boivin N, Fuller DQ, Crowther A. Old world globalization and the columbian exchange: comparison and contrast. World Archaeol. 2012;44:452–69.
22. Hughes, D.H., and Henson, R.E. Crop production principles and practices. (The Macmillan Company). 1934.
23. Yao YF, Song XY, Xie G, et al. New insights into the origin of buckwheat cultivation in southwestern China from pollen data. New Phytol. 2023;237:2467–77.
24. Smith BD. Documenting plant domestication: The consilience of biological and archaeological approaches. Proc Natl Acad Sci USA. 2001;98:1324–6.
25. Zeder MA, Emshwiller E, Smith BD, Bradley DG. Documenting domestication: the intersection of genetics and archaeology. Trends Genet. 2006;22:139–55.
26. Huang X, Kurata N, Wei X, et al. A map of rice genome variation reveals the origin of cultivated rice. Nature. 2012;490:497–501.
27. van Andel TR, Meyer RS, Aflitos SA, et al. Tracing ancestor rice of Suriname Maroons back to its African origin. Nat Plants. 2016;2:16149.
28. Wang W, Mauleon R, Hu Z, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature. 2018;557:43–9.
29. Hu□ord MB, Xu X, van Heerwaarden J, et al. Comparative population genomics of maize domestication and improvement. Nat Genet. 2012;44:808–11.
30. Chen L, Luo J, Jin M, et al. Genome sequencing reveals evidence of adaptive variation in the genus *Zea*. Nat Genet. 2022;54:1736–45.
31. Kang L, Qian L, Zheng M, et al. Genomic insights into the origin, domestication and diversification of *Brassica juncea*. Nat Genet. 2021;53:1392–402.
32. Wei T, van Treuren R, Liu X, et al. Whole-genome resequencing of 445 *Lactuca* accessions reveals the domestication history of cultivated lettuce. Nat Genet. 2021;53:752–60.
33. Dong Y, Duan S, Xia Q, et al. Dual domestications and origin of traits in grapevine evolution. Science. 2023;379:892–901.
34. Bellucci E, Benazzo A, Xu C, et al. Selection and adaptive introgression guided the complex evolutionary history of the European common bean. Nat Commun. 2023;14:1901–8.
35. Varshney RK, Roorkiwal M, Sun S, et al. A chickpea genetic variation map based on the sequencing of 3,366 genomes. Nature. 2021;599:622–7.
36. Suarez-Gonzalez A, Lexer C, Cronk Q. Adaptive introgression: a plant perspective. Biol Lett. 2018;14:20170688.
37. Cai W, Hong J, Liu Z, et al. A receptor-like kinase controls the amplitude of secondary cell wall synthesis in rice. Curr Biol. 2023;33:498–506.
38. Wang P. Battle for survival: the role of plant thioredoxin in the war against *Barley stripe mosaic virus*. Plant Physiol. 2022;189:1199–201.
39. Breen S, Williams SJ, Outram M, et al. Emerging insights into the functions of pathogenesis-related protein 1. Trends Plant Sci. 2017;22:871–9.