

# Domestication and improvement reshaped the genomic architecture of abiotic stress tolerance in sunflower

Yavuz Delen<sup>1,2</sup>, Semra Palali-Delen<sup>1,2,3</sup>, Sai Subhash Mahamkali VS<sup>1,3</sup>, Haichuan Wang<sup>4</sup>, Ismail Dweikat<sup>1</sup>, Jinliang Yang<sup>1,3\*</sup> and Gen Xu<sup>1,3\*</sup>

<sup>1</sup> Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

<sup>2</sup> Biotechnology Research Center, Field Crops Central Research Institute, Ankara 06170, Türkiye

<sup>3</sup> Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

<sup>4</sup> USDA-ARS-AMRU, Lincoln, NE 68583, USA

\* Correspondence: [jinliang.yang@unl.edu](mailto:jinliang.yang@unl.edu) (Yang J); [gxu6@unl.edu](mailto:gxu6@unl.edu) (Xu G)

## Abstract

Sunflower (*Helianthus annuus* L.) has undergone extensive domestication and modern improvement, yet the extent to which selection has shaped its stress-adaptive variation remains largely unclear. Using genomic data from diverse accessions, including the wild ancestor of sunflower, we identified widespread balancing selection that maintains historical polymorphisms in abiotic-stress-related genes. Additionally, we identified strong positive selection signals associated with domestication and improvement. Using a cutting-edge DNA language model to characterize deleterious alleles, we found that modern cultivars exhibited an elevated deleterious score, particularly within positively selected regions, consistent with patterns observed in other major domesticates. To cross-validate the functional consequences of the selected regions on stress responses, we assessed salt tolerance in an independent association panel and evaluated ancestral allele effects on salt-responsive traits. Although only a small subset of the selected regions reached genome-wide significance, several ancestral alleles within positively selected sweeps tended to enhance salt tolerance. Together, these results indicate that recent selection may have unintentionally eliminated historically beneficial alleles, reshaping the genomic landscape of sunflower stress responses. This study highlights a breeding strategy to recover ancestral beneficial alleles for enhancing abiotic stress tolerance in modern sunflower cultivars.

**Citation:** Delen Y, Palali-Delen S, Mahamkali VS SS, Wang H, Dweikat I, et al. 2026. Domestication and improvement reshaped the genomic architecture of abiotic stress tolerance in sunflower. *Genomics Communications* 3: e007 <https://doi.org/10.48130/gcomm-0026-0006>

## Introduction

Sunflower (*Helianthus annuus* L.) was domesticated in North America about 4,000 years ago from wild populations of *H. annuus*, according to archeological, morphological, and recent genetic evidence<sup>[1–4]</sup>. Today, it is widely cultivated as a major oilseed crop and used for various purposes, including snack foods, animal feed, and industrial applications<sup>[5]</sup>. As a staple oil source in many countries, seed oil content and quality have been key targets of selection throughout sunflower improvement and remain central goals of modern breeding programs<sup>[6]</sup>.

Sunflowers are highly adaptable and can thrive in environments where many other crops fail, making them particularly valuable in regions affected by soil salinity<sup>[7]</sup>. Its robust root system and inherent tolerance to multiple abiotic stresses enable growth across diverse ecological zones, including areas with high salt levels<sup>[8]</sup>. Soil salinity, affecting roughly 20% of the world's cultivable land, poses major challenges for crop establishment and yield, especially in arid and semi-arid regions where sunflower is commonly grown<sup>[9]</sup>. High salinity disrupts germination, seedling emergence, water uptake, and cellular metabolism, leading to stunted growth, leaf chlorosis, and significant reductions in both seed and oil yield<sup>[10,11]</sup>. Given its natural resilience, sunflower represents a promising salt-tolerant crop for sustaining production in increasingly saline and climate-stressed agricultural landscapes<sup>[12]</sup>.

Like many domesticated crops<sup>[13–16]</sup>, sunflowers experienced a substantial genetic bottleneck during domestication<sup>[17]</sup>, and modern breeding efforts that focused heavily on improving oil content have likely overlooked valuable salt-tolerance alleles. Although sunflower is considered moderately salt-tolerant, substantial variation among

germplasm underscores the need for improved, efficient screening approaches<sup>[18,19]</sup>. With the rapid expansion of genomic resources in sunflowers<sup>[20–22]</sup>, there is now a renewed opportunity to recover naturally occurring salt-tolerant or other stress-resilient alleles from diverse cultivars, native American landraces, and wild relatives.

In this study, we leverage publicly available genomic resources that span wild, landrace, and modern cultivated sunflower accessions to perform a comprehensive population genomic analysis. Through these analyses, we identify historically balanced loci that are associated with abiotic stress responses, including putative salt-tolerance sites. In addition, we detect strong positive selection during domestication and modern improvement, involving genes related to oil biosynthesis, flowering regulation, and stress responses. We further find an enrichment of deleterious variants within selected regions, suggesting that some beneficial ancestral alleles may have been lost during recent selection. Together, these results underscore the need for targeted breeding efforts to restore historically advantageous alleles related to salt tolerance. By providing a clearer picture of salt-related genetic variation and its evolutionary history, this study offers valuable guidance for developing more resilient sunflower cultivars.

## Materials and methods

### Genetic materials and variant calling procedure

We downloaded raw sequence data from 332 sunflower accessions, including modern cultivars ( $n = 288$ ), landraces ( $n = 18$ ), and wild relatives ( $n = 26$ ), from a previous study<sup>[23]</sup>. We then processed raw sequencing reads with fastp (v0.20.0) to trim adapter sequences and remove

low-quality bases before downstream analyses<sup>[24]</sup>. Next, the cleaned reads were aligned to the reference genome HanXRQr2.0-SUNRISE<sup>[6]</sup> using Bowtie2 (v2.4.4)<sup>[25]</sup>. Subsequently, we removed the duplicated reads using Picard tools (v2.18) and conducted SNP calling using Genome Analysis Toolkit's (GATK, v4.1) HaplotypeCaller<sup>[26]</sup>, with the following parameters: QD < 2.0, FS > 60.0, MQ < 20.0, MQRankSum < -12.5, and ReadPosRankSum < -8.0, to filter out low-quality SNPs. In addition, we removed SNPs with missing rates > 50%.

## Principal component analysis (PCA) and phylogenetic tree

We performed PCA using the SNP dataset with plink1.9<sup>[27]</sup>. Using the same software, we computed identity-by-state (IBS) values among the 332 sunflower accessions with the parameter "-ibs-matrix". The resulting IBS similarity matrix was converted into a genetic distance matrix (1-IBS), which was then used to construct a neighbor-joining (NJ) phylogenetic tree using the ape package<sup>[28]</sup> in R (v4.3.0). The tree was visualized with a circular layout using the 'ggtree' package<sup>[29]</sup>.

## Salt treatment experiment at the seedling stage

We evaluated salt tolerance at the seedling stage using an independent association panel of 274 accessions<sup>[30]</sup>. All accessions used in this experiment came from the United States Department of Agriculture, Agricultural Research Service (USDA-ARS), North Central Regional Introduction Station (NCRPIS) in Ames, Iowa, USA. For the experiment, we treated seeds with a fungicide and placed them in petri dishes. Under control conditions, we germinated the seeds in pure water, whereas under salt treatment, we germinated them in a 200 mM NaCl solution. We maintained all petri dishes in a 27 °C growth chamber for 4 d. After incubation, we recorded the germination ratio and radicle length for each accession, using three biological replicates per treatment.

We calculated salt-responsive (SR) traits using the following formula:

$$SR = \frac{T_{\text{nosalt}} - T_{\text{salt}}}{T_{\text{salt}}}$$

where,  $T_{\text{nosalt}}$  and  $T_{\text{salt}}$  are the mean values for a given trait measured from non-salted and salted conditions.

## Genome-wide scan for positive and balancing selection

We performed a genome-wide scan for positive selection signals using the latest version of the cross-population composite likelihood ratio (XP-CLR)<sup>[31]</sup>. In the XP-CLR analysis, we used a 50 kb sliding window and a 5 kb step size. To ensure comparability of the composite likelihood score in each window, we fixed the number assayed in each window to 100 SNPs. We then conducted a balancing selection scan with default parameters using BalLeRMix (v2.3)<sup>[32]</sup>. Using the major alleles in wild species as ancestral alleles, we calculated the derived allele frequency (DAF). For both positive and balancing selection, we considered the 1% outliers as the selective sweeps; adjacent sweeps separated by a physical distance of < 50 kb were merged into a single selected sweep. The predicted protein-coding genes were functionally annotated using EggNOG Mapper (v2.1.11)<sup>[33]</sup>. Gene ontology (GO) analysis for the genes under selection was performed using the g:GOST toolset in g:Profiler<sup>[34]</sup>.

## Identification of deleterious mutations using a DNA language model

We used Plant Caduceus<sup>[35]</sup>, a DNA language model that infers the relative fitness effects of nucleotide substitutions through a zero-shot comparison of reference and alternate allele probabilities, to compute

site-specific zero-shot scores (ZSS) and estimate genome-wide deleterious scores across wild, landrace, and improved sunflower groups. Variants were then classified into the top 1%, 5%, and 10% most deleterious categories based on their ZSS distributions, with more negative ZSS values indicating stronger predicted deleterious effects.

To test whether deleterious variants are enriched in genomic regions under selection, we intersected the genomic coordinates of deleterious sites with regions identified as under positive selection or balancing selection. For each category, we calculated the observed number of deleterious sites within selected regions and compared it with a null distribution generated from 1,000 sets of randomly sampled genomic windows matched in size.

## Genome-wide association analysis

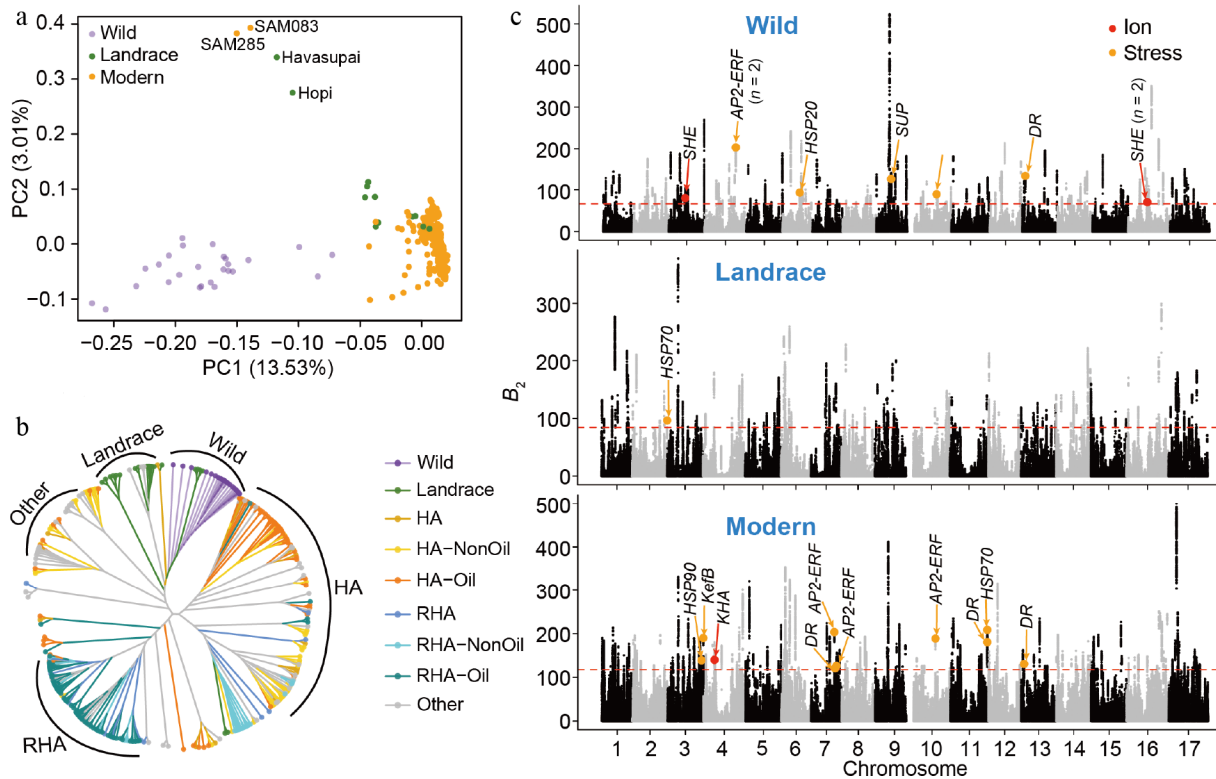
We performed a genome-wide association analysis (GWAS) for salt responsiveness of germination ratio and radicle length using GEMMA<sup>[36]</sup>, based on 226,779 high-quality SNPs generated in our previous study<sup>[30]</sup>. We applied a mixed model Q + K to account for the confounding effects of both population structure (Q) and relatedness (K)<sup>[37]</sup>. The Q matrix consisted of the first three principal components calculated using plink1.9<sup>[27]</sup>, while the K matrix was estimated with TASSEL 5.0<sup>[38]</sup>. To determine significant associations, we used a genome-wide threshold of  $2.9 \times 10^{-5}$ , corresponding to  $1/n$ , where  $n = 33,801$  represents the number of independent SNPs. The number of independent markers was estimated using PLINK's "indep-pairwise" function (window size = 10 kb, step size = 10,  $r^2 \geq 0.1$ ). The significant GWAS loci were then defined by extending 50 kb upstream and downstream of each significant SNP, and the overlapping windows were merged into single genomic intervals. We estimated the sizes of allele effects using the beta coefficients from the GEMMA output. To determine the ancestral allele, we assigned the allele with the highest frequency in the wild group as the ancestral allele. A positive effect of the ancestral allele indicates that the ancestral allele contributes to a higher phenotypic value than the derived allele.

## Results

### Balancing selection maintains historical alleles associated with stress adaptation

To characterize the genetic relationships of the sunflower accessions, we first performed a principal component analysis (PCA) using 11.3 million genome-wide SNPs from 332 accessions<sup>[23]</sup>, including wild ( $n = 26$ ), landrace ( $n = 18$ ), and modern cultivated sunflower lines ( $n = 288$ ) (Supplementary Table S1, Materials and methods). The PCA result clearly separated wild accessions from the cultivated group, with landraces positioned between them, with several outliers (Fig. 1a). This pattern is largely consistent with a gradual transition during domestication and improvement processes observed for other domesticated crops<sup>[14,39–41]</sup>. The outlier accessions clustered closely with wild sunflower, likely reflecting ancestral genomic introgression as previously reported<sup>[23]</sup>.

We further investigated the genetic relationships among modern accessions by building a neighbor-joining phylogenetic tree (Materials and methods). The tree showed a distinct clustering of female (HA) and male (RHA) lines in the cultivated group, corresponding to two heterotic groups used in hybrid breeding programs (Fig. 1b). This separation was less evident in the PCA, including all accessions (Fig. 1a), likely due to the dominant variation among wild, landrace, and modern groups. However, PCA restricted to modern accessions



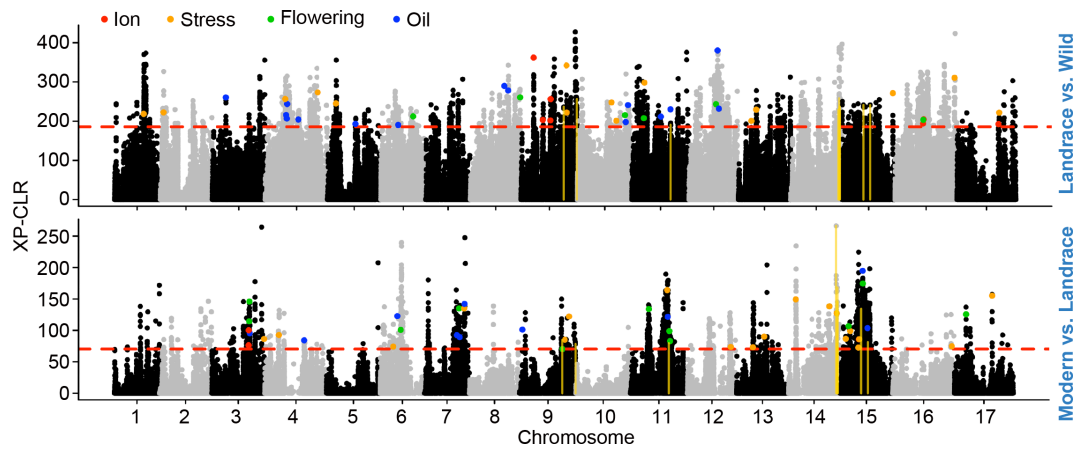
**Fig. 1** Genomic regions under balancing selection and their relationship with salt-stress-responsive loci in sunflower. (a) Principal component analysis (PCA) of the 332 accessions showed clear population differentiation among wild, landrace, and modern accessions. (b) Neighbor-joining phylogenetic tree of the modern accessions constructed from genome-wide SNPs. Colors represent different groups according to the previous study<sup>[23]</sup>. HA and RHA represent female and male lines, respectively. 'Other' represents accessions that do not belong to the defined HA or RHA groups. (c) Genomewide balancing selection signals in each population, with colored dots indicating candidate genes involved in ion transport (red) and stress-responsive pathways (orange). The red horizontal dashed lines represent the significance level of the top 1%. SHE, sodium hydrogen exchanger; AP2-EREBP, ethylene-responsive transcription factor; HSP, heat shock protein; SUP, stress upregulated; DR, disease resistance; KeFB, potassium-efflux system protein; HKA, putative potassium channel.

clearly resolved the HA–RHA structure (Supplementary Fig. S1), consistent with the phylogenetic tree. Within each group, moderate genetic diversity was observed, reflecting their different breeding origins and partial overlap between the oilseed and confectionery types, consistent with previous results<sup>[42]</sup>.

To explore how long-term evolutionary processes contribute to sunflower adaptive variation, we conducted a balancing selection scan and identified hundreds of genomic regions under balancing selection within each group (Fig. 1c, Supplementary Table S2, Materials and methods). These balancing selection signals were widely distributed throughout the genome, overlapping with several candidate genes annotated as sodium–hydrogen exchangers ( $n = 3$  in the wild population), heat shock proteins ( $n = 1$  in the wild,  $n = 1$  in the landrace, and  $n = 2$  in modern cultivars) and EREBP-type transcription factors ( $n = 3$  in the wild, and  $n = 3$  in modern cultivars), which are associated with ion transport and stress-response pathways (Supplementary Table S3). Such loci may represent allelic variants maintained by balancing selection that contribute to abiotic stress adaptation, such as salt tolerance. To further investigate the functional roles of genes located within regions under balancing selection, we performed Gene Ontology (GO) enrichment analysis (Supplementary Table S4). The results revealed significant enrichment of ion transport-related categories, including sodium proton antiporter activity, potassium proton antiporter activity, sodium ion transport, and inorganic cation transport across plasma membranes.

## Positive selection drives genomic differentiation during historical domestication and recent improvement

Next, we investigated genomic footprints of positive selection during sunflower domestication and improvement processes (Fig. 2). Using the cross-population composite likelihood ratio (XP-CLR) approach (Materials and methods), we identified a total of 652 domestication-related selective sweeps (landraces vs. wild accessions) and 549 improvement-related sweeps (modern accessions vs. landraces), covering 2.3% (67 Mb) and 1.9% (56.9 Mb) of the genome, respectively (Supplementary Table S5). Domestication sweeps ranged from 50 to 910 kb, with an average length of 103 kb and a total of 1,919 genes. During the improvement process, a slightly smaller number of regions was detected, with a comparable mean footprint size (104 kb), which encompassed 2,062 genes. The genes within these regions include key functional categories such as kernel oil biosynthesis (19 genes in domestication; 11 in improvement), flowering-time regulation ( $n = 6$  in domestication;  $n = 13$  in improvement), ion transport ( $n = 9$  in domestication;  $n = 3$  in improvement), and stress-response pathways ( $n = 18$  in domestication;  $n = 20$  in improvement, Supplementary Table S6). The differences in category composition suggest that domestication targeted loci associated with early agronomic traits, particularly oil accumulation and ion homeostasis. In contrast, modern improvement more strongly emphasized flowering regulation and stress-resistance functions. In particular, nine domestication sweeps showed evidence of continued selection during improvement (Fig. 2), suggesting that a subset of key loci remained under strong selection



**Fig. 2** Genomic regions under positive selection during sunflower domestication and improvement. Genome-wide XP-CLR analysis was performed to detect genomic regions under differential selection between landraces and wild accessions (top), and between modern cultivars and landraces (bottom). The red dashed line indicates the top 1% threshold used to define candidate selective sweep regions. Colored points indicate windows overlapping annotated genes, which are further categorized based on their putative biological functions: ion transport (red), stress response (orange), flowering regulation (green), and oil biosynthesis (blue). The gold vertical lines indicate the overlapped domestication and improvement sweeps.

pressure throughout the history of sunflower improvement. In addition, GO enrichment analysis further revealed significant enrichment of signaling-related functional categories among genes under positive selection (Supplementary Table S4).

### Domestication and selection processes shaped deleterious scores in sunflower

To assess the genome-wide burden (or genetic deleterious score) of deleterious mutations in sunflower, using a pretrained DNA language model<sup>[35]</sup>, we calculated zero-shot scores (ZSS) for 11.3 million SNPs (Materials and methods). In the analysis, ZSS is calculated as the logarithmic probability of the alternative allele relative to the reference allele, with more negative values indicating a higher predicted deleteriousness as evidenced by the relationship between the frequency of the minor allele and the ZSS values (Supplementary Fig. S2). The ZSS values exhibited a broad distribution (Fig. 3a). To identify deleterious sites, instead of relying on a single arbitrary threshold, we defined the upper-tail cutoff points of 1%, 5%, and 10% as increasingly relaxed thresholds (Fig. 3a). Using these thresholds, although the total number of deleterious alleles varies, a consistent pattern emerged: modern cultivars carried the highest number of deleterious alleles, followed by landraces, whereas wild accessions harbored the lowest deleterious score (Fig. 3b). This pattern is consistent with the increased accumulation of mildly deleterious mutations observed in other domesticated crops, such as maize<sup>[43]</sup>, wheat<sup>[44]</sup>, and rice<sup>[45]</sup>. In addition, we evaluated the deleterious score within regions under selection. Regions experiencing positive selection during both domestication and improvement showed significantly higher counts of deleterious SNPs than expected from randomly sampled genomic intervals (Fig. 3c). This enrichment was consistent across all three deleteriousness thresholds, supporting the idea that selective sweeps may have carried linked deleterious alleles to higher frequencies. Similarly, regions under balancing selection also exhibited an enrichment of deleterious mutations (Fig. 3d), indicating that long-term maintenance of polymorphism may preserve some slightly deleterious variants.

### Salt-response variation is shaped by selection and reflected in population-differentiated allele effects

Given that numerous genes related to stress, particularly salt-response, are located within regions under balancing or positive selection, we

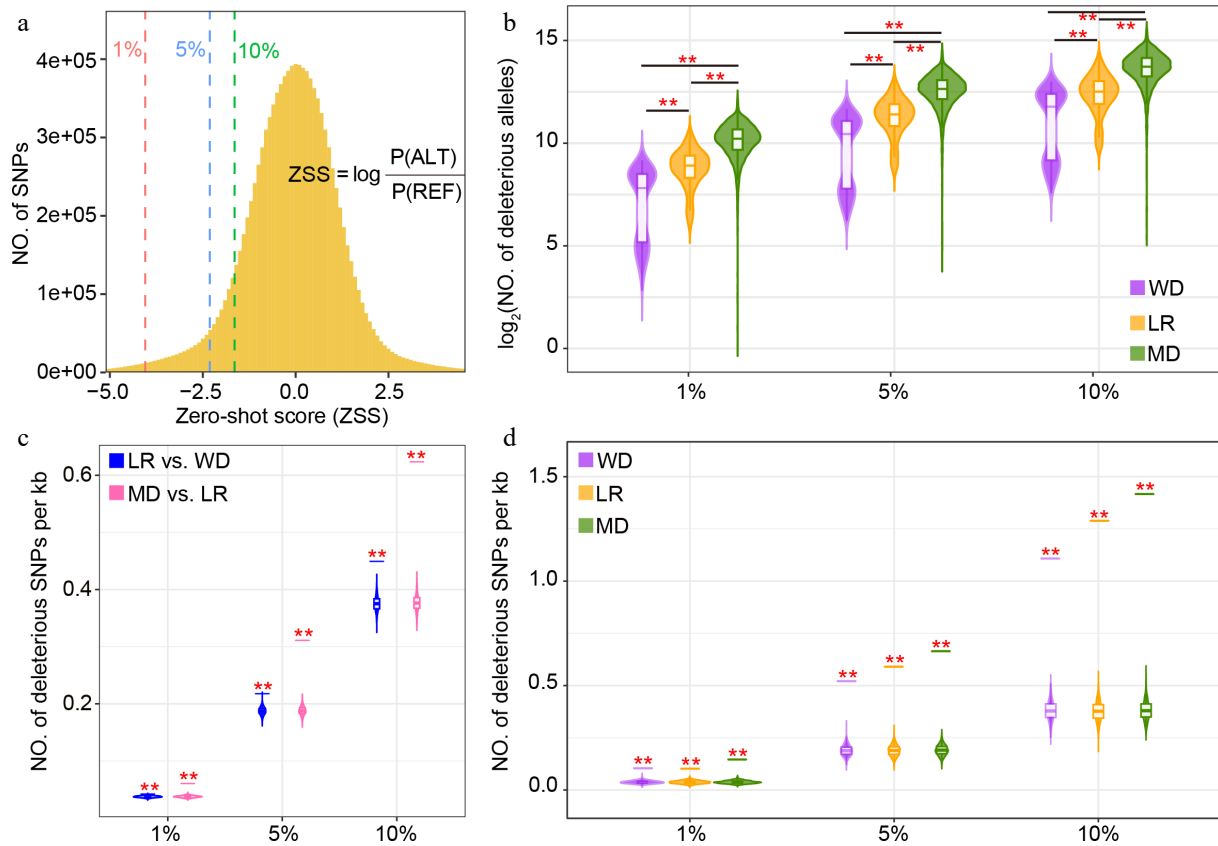
hypothesize that adaptive responses to abiotic stress have played an important role throughout the history of sunflower improvement. To directly connect these evolutionary signatures with functional phenotypes, we performed a salt-stress experiment using an independent association panel and quantified genotype-specific responses (Materials and methods). As expected, both germination ratio (GR) and the radicle length (RL) decreased significantly with salt treatment (Supplementary Fig. S3a, S3b). Beyond raw traits, we also computed salt-responsive (SR) values following a previously described method<sup>[46,47]</sup>. In general, RL exhibited a stronger salt response than GR (Supplementary Fig. S3c) (Wilcoxon test,  $p = 5.2 \times 10^{-66}$ ).

We next conducted GWAS for the transformed SR traits using a linear mixed model (see Materials and methods). For GR, we identified 42 significant trait-associated loci (TALs), six (15%) of which were located within regions under positive selection. For RL, 115 TALs were detected, including three within balancing-selection regions and 14 within positive-selection regions, predominantly overlapping improvement sweeps (Fig. 4a, Supplementary Table S7).

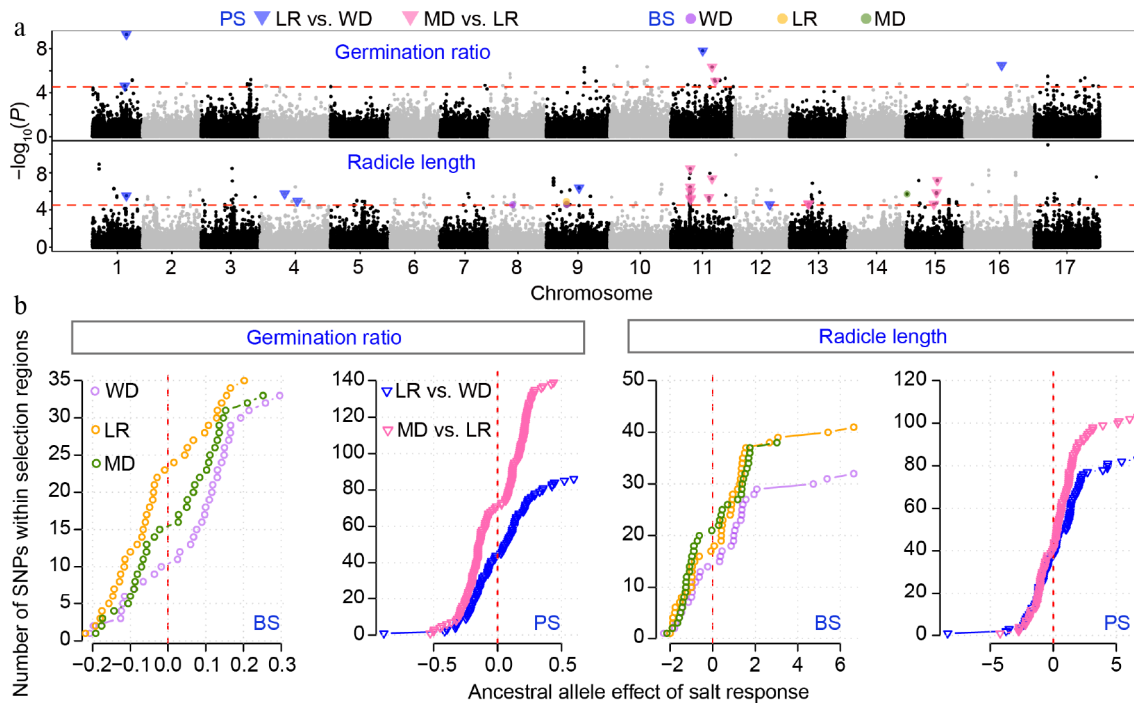
We then examined the ancestral allele effect of the SNPs located within positive selection and balancing selection regions (Fig. 4b). In general, ancestral alleles in positive selection regions showed a wider range of effects, with several SNPs exhibiting relatively large salt-response effects. In contrast, the ancestral alleles in the balancing selection regions showed a weaker salt-responsive effect. In the balanced regions, ancestral alleles tend to yield higher SR values in wild accessions than in landrace and modern groups. This suggests that, although individual SNP effects vary considerably, balanced sites contain alleles whose salt-response effects shift across populations.

## Discussion

Our findings show that sunflower evolution has been shaped by two major selective forces operating at different timescales. Ancient balancing selection preserved allelic diversity at genes related to ion transport, stress signaling, and AP2-EREBP transcriptional regulation—gene functions that likely allowed wild sunflower to cope with heterogeneous environments, consistent with previous reports of high stress-adaptive diversity in wild *Helianthus*<sup>[17,48]</sup>. In contrast,



**Fig. 3** Identification of deleterious mutations in sunflower. (a) Zero-shot score (ZSS) distribution of 11.3 million SNPs in the sunflower population. The 1%, 5%, and 10% cutoffs indicate increasingly relaxed deleteriousness thresholds, with 1% representing the most deleterious variants. (b) Distribution of deleterious score (sum of ZSS) across wild (WD), landrace (LR), and modern (MD) accessions under the three deleteriousness thresholds (1%, 5%, and 10%). Deleterious mutations located within regions under (c) positive, and (d) balancing selection. Violin plots show the distribution generated from 1,000 randomly selected genomic regions, and short horizontal lines indicate the observed values. Asterisks denote significance from permutation tests (\*\*  $p \leq 0.01$ ).



**Fig. 4** GWAS results for salt-responsive traits of germination ratio and radicle length. (a) Stacked Manhattan plots for salt-responsive (SR) traits. The red horizontal dashed line denotes the genome-wide significance threshold. Colored points and triangles indicate significant SNPs that overlap with regions under balancing selection (BS) or positive selection (PS). (b) Effects of the ancestral allele for SNPs located within BS or PS regions. Trait responsive values are shown across wild (WD), landrace (LR), and modern (MD) accessions.

domestication and improvement introduced strong positive selection sweeps in flowering time, oil biosynthesis, and disease-response pathways, aligning with earlier genomic studies of sunflower breeding<sup>[6,23]</sup>. The coexistence of long-maintained polymorphisms and human-driven selective sweeps create a mosaic genomic architecture that reflects both ecological adaptation and modern agronomic selection.

A notable consequence of these directional sweeps is the increased accumulation of deleterious mutations in cultivated sunflowers. Similar to patterns reported in maize, rice, and wheat<sup>[43–45]</sup>, regions under strong positive selection exhibited significantly elevated deleterious scores, likely due to hitchhiking during rapid allele fixation and reduced efficacy of purifying selection<sup>[49–51]</sup>. Comparable enrichment of deleterious mutations in selective sweep regions has also been observed in other species, including *Drosophila*<sup>[52]</sup> and domesticated animals such as dogs<sup>[53]</sup> and horses<sup>[54]</sup>, suggesting that this represents a general consequence of linked selection rather than a crop-specific phenomenon. Regions under balancing selection also retained slightly deleterious mutations. This pattern is consistent with theoretical expectations that balancing selection can maintain linked mildly deleterious mutations through mechanisms such as heterozygote advantage or spatially heterogeneous selection<sup>[55–57]</sup>. Empirical evidence from humans and livestock further supports this interpretation, showing that mutations maintained under long-term balancing selection are often enriched for functional and potentially deleterious polymorphisms<sup>[58,59]</sup>. Together, these observations suggest that long-term maintenance of polymorphism may tolerate small-effect deleterious mutations while preserving adaptive variation over evolutionary timescales.

Integrating evolutionary signatures with salt-response phenotypes revealed that adaptive stress-tolerance alleles have decreased in frequency during improvement. Ancestral alleles, particularly those maintained by long-term balancing selection tended to confer stronger salt responsiveness in wild accessions, but became rare in modern cultivars, paralleling observations in maize, sorghum, and tomato, where domestication reduced abiotic-stress tolerance<sup>[14,16,40]</sup>. These findings suggest a practical breeding strategy: reintroducing or prioritizing ancestral beneficial alleles through genomic selection with advanced phenomics technology<sup>[60]</sup>, targeted introgression, or genome editing<sup>[61]</sup>. Our results, therefore, provide a framework for leveraging evolutionary information to guide future improvements in sunflowers under increasingly challenging environments.

## Author contributions

The authors confirm their contributions to the paper as follows: designed this work: Xu G, Delen Y, Dweikat I, Yang J; conducted the cross-validation experiment: Delen Y, Palali-Delen S; analyzed the data: Xu G, Mahamkali VS SS, Yang J, Delen Y; provided conceptual advice: Dweikat I, Wang H; wrote the manuscript: Xu G, Delen Y, Yang J. All authors reviewed the results and approved the final version of the manuscript.

## Data availability

Supplementary data and code are available on the GitHub repository: [https://github.com/jyanglab/Sunflower\\_domestication](https://github.com/jyanglab/Sunflower_domestication).

## Acknowledgements

This research was funded by the Agriculture and Food Research Initiative (Grant Nos 2022-67013-36560) from the USDA National

Institute of Food and Agriculture. Financial support for Yavuz Delen was provided by the Ministry of Education of the Republic of Turkey.

## Conflicts of Interests

The authors declare that they have no conflict of interest.

**Supplementary information** accompanies this paper online at: <https://doi.org/10.48130/gcomm-0026-0006>.

## Dates

Received 4 December 2025; Revised 2 April 2026; Accepted 7 April 2026; Published online 24 April 2026

## References

- [1] Smith BD. 1989. Origins of agriculture in eastern North America. *Science* 246:1566–1571
- [2] Harter AV, Gardner KA, Falush D, Lentz DL, Bye RA, et al. 2004. Origin of extant domesticated sunflowers in eastern North America. *Nature* 430:201–205
- [3] Heiser CB, Jr. 1951. The sunflower among the North American Indians. *Proceedings of the American Philosophical Society* 95:432–448
- [4] Smith BD. 2014. The domestication of *Helianthus annuus* L. (sunflower). *Vegetation History and Archaeobotany* 23:57–74
- [5] Adeleke BS, Babalola OO. 2020. Oilseed crop sunflower (*Helianthus annuus*) as a source of food: nutritional and health benefits. *Food Science & Nutrition* 8:4666–4684
- [6] Badouin H, Gouzy J, Grassa CJ, Murat F, Staton SE, et al. 2017. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* 546:148–152
- [7] Shi D, Sheng Y. 2005. Effect of various salt-alkaline mixed stress conditions on sunflower seedlings and analysis of their stress factors. *Environmental and Experimental Botany* 54:8–21
- [8] Rauf S, Shahzad M, Teixeira da Silva JA, Noorka IR. 2012. Biomass partitioning and genetic analyses of salinity tolerance in sunflower (*Helianthus annuus* L.). *Journal of Crop Science and Biotechnology* 15:205–217
- [9] Gupta B, Huang B. 2014. Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization. *International Journal of Genomics* 2014:701596
- [10] Xiong L, Zhu JK. 2002. Molecular and genetic aspects of plant responses to osmotic stress. *Plant, Cell & Environment* 25:131–139
- [11] Pessaraki M. 2019. *Handbook of plant and crop stress*. 4<sup>th</sup> Edition. Boca Raton: CRC press. 974 pp. doi: 10.1201/9781351104609
- [12] Wahid A, Masood I, Javed IUH, Rasul E. 1999. Phenotypic flexibility as marker of sodium chloride tolerance in sunflower genotypes. *Environmental and Experimental Botany* 42:85–94
- [13] Huang X, Kurata N, Wei X, Wang ZX, Wang A, et al. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490:497–501
- [14] Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia JM, et al. 2012. Comparative population genomics of maize domestication and improvement. *Nature Genetics* 44:808–811
- [15] Qi J, Liu X, Shen D, Miao H, Xie B, et al. 2013. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nature Genetics* 45:1510–1515
- [16] Lin T, Zhu G, Zhang J, Xu X, Yu Q, et al. 2014. Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics* 46:1220–1226
- [17] Tang S, Knapp SJ. 2003. Microsatellites uncover extraordinary diversity in native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics* 106:990–1003
- [18] Li W, Zhang H, Zeng Y, Xiang L, Lei Z, et al. 2020. A salt tolerance evaluation method for sunflower (*Helianthus annuus* L.) at the seed germination stage. *Scientific Reports* 10:10626
- [19] Chen F, Xiao L, Huang Q, Xiang L, Li Q, et al. 2024. Physiological evaluation of salt tolerance in sunflower seedlings across different genotypes. *Agronomy* 14:2995

- [20] Bercovich N, Genze N, Todesco M, Owens GL, Légaré JS, et al. 2022. HelianHOME, a public and centralized database of phenotypic sunflower data. *Scientific Data* 9:735
- [21] Huang K, Jahani M, Gouzy J, Legendre A, Carrere S, et al. 2023. The genomics of linkage drag in inbred lines of sunflower. *Proceedings of the National Academy of Sciences of the United States of America* 120:e2205783119
- [22] Yi L, Bao H, Wu Y, Mu Y, Du C, et al. 2025. Chromosome-level genome assemblies of sunflower oilseed and confectionery cultivars. *Scientific Data* 12:24
- [23] Hübner S, Bercovich N, Todesco M, Mandel JR, Odenheimer J, et al. 2019. Sunflower pan-genome analysis shows that hybridization altered gene content and disease resistance. *Nature Plants* 5:54–62
- [24] Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890
- [25] Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. *Nature Methods* 9:357–359
- [26] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20:1297–1303
- [27] Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, et al. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:s13742-015-0047-8
- [28] Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35:526–528
- [29] Yu G, Smith DK, Zhu H, Guan Y, Lam TT. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 8:28–36
- [30] Delen Y, Palali-Delen S, Xu G, Neji M, Yang J, et al. 2024. Dissecting the genetic architecture of morphological traits in sunflower (*Helianthus annuus* L.). *Genes* 15:950
- [31] Chen H, Patterson N, Reich D. 2010. Population differentiation as a test for selective sweeps. *Genome Research* 20:393–402
- [32] Cheng X, DeGiorgio M. 2020. Flexible mixture model approaches that accommodate footprint size variability for robust detection of balancing selection. *Molecular Biology and Evolution* 37:3267–3291
- [33] Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution* 38:5825–5829
- [34] Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, et al. 2023. g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research* 51:W207–W212
- [35] Zhai J, Gokaslan A, Schiff Y, Berthel A, Liu ZY, et al. 2025. Cross-species modeling of plant genomes at single-nucleotide resolution using a pretrained DNA language model. *Proceedings of the National Academy of Sciences of the United States of America* 122:e2421738122
- [36] Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* 44:821–824
- [37] Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38:203–208
- [38] Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, et al. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635
- [39] Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, et al. 2015. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature Biotechnology* 33:408–414
- [40] Wu X, Liu Y, Luo H, Shang L, Leng C, et al. 2022. Genomic footprints of sorghum domestication and breeding selection for multiple end uses. *Molecular Plant* 15:537–551
- [41] Qiu J, Zhou Y, Mao L, Ye C, Wang W, et al. 2017. Genomic variation associated with local adaptation of weedy rice during de-domestication. *Nature Communications* 8:15323
- [42] Mandel JR, Dechaine JM, Marek LF, Burke JM. 2011. Genetic diversity and population structure in cultivated sunflower and a comparison to its wild progenitor, *Helianthus annuus* L. *Theoretical and Applied Genetics* 123:693–704
- [43] Lozano R, Gazave E, dos Santos JPR, Stetter MG, Valluru R, et al. 2021. Comparative evolutionary genetics of deleterious load in sorghum and maize. *Nature Plants* 7:17–24
- [44] Fu YB, Horbach C. 2025. Elevated mutation burdens in Canadian oat and wheat cultivars released over the past century. *Cells* 14:844
- [45] Lu J, Tang T, Tang H, Huang J, Shi S, et al. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. *Trends in Genetics* 22:126–131
- [46] Liu Y, Wang H, Jiang Z, Wang W, Xu R, et al. 2021. Genomic basis of geographical adaptation to soil nitrogen in rice. *Nature* 590:600–605
- [47] Palali Delen S, Xu G, Velazquez-Perfecto J, Yang J. 2023. Estimating the genetic parameters of yield-related traits under different nitrogen conditions in maize. *Genetics* 223:iyad012
- [48] Seiler GJ, Qi LL, Marek LF. 2017. Utilization of sunflower crop wild relatives for cultivated sunflower improvement. *Crop Science* 57:1083–1101
- [49] Yang J, Mezouk S, Baumgarten A, Buckler ES, Guill KE, et al. 2017. Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. *PLoS Genetics* 13:e1007019
- [50] Smith JM, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genetics Research* 23:23–35
- [51] Chun S, Fay JC. 2011. Evidence for hitchhiking of deleterious mutations within the human genome. *PLoS Genetics* 7:e1002240
- [52] Wang Y, Allen SL, Reddix AJ, Chenoweth SF. 2024. The impacts of positive selection on genomic variation in *Drosophila serrata*: insights from a deep learning approach. *Molecular Ecology* 33:e17499
- [53] Marsden CD, Ortega-Del Vecchyo D, O'Brien DP, Taylor JF, Ramirez O, et al. 2016. Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proceedings of the National Academy of Sciences of the United States of America* 113:152–157
- [54] Schubert M, Jónsson H, Chang D, Der Sarkissian C, Ermini L, et al. 2014. Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proceedings of the National Academy of Sciences of the United States of America* 111:E5661–E5669
- [55] Charlesworth D. 2006. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics* 2:e64
- [56] Hedrick PW. 2006. Genetic polymorphism in heterogeneous environments: the age of genomics. *Annual Review of Ecology, Evolution, and Systematics* 37:67–93
- [57] Delph LF, Kelly JK. 2014. On the importance of balancing selection in plants. *New Phytologist* 201:45–56
- [58] Aqil A, Speidel L, Pavlidis P, Gokcumen O. 2023. Balancing selection on genomic deletion polymorphisms in humans. *eLife* 12:e79111
- [59] Derks MFL, Steensma M. 2021. Review: balancing selection for deleterious alleles in livestock. *Frontiers in Genetics* 12:761728
- [60] Jin H, Park A, Sreedasyam A, Li G, Ge Y, et al. 2025. Nitrogen response and growth trajectory of sorghum CRISPR-Cas9 mutants using high-throughput phenotyping. *Genomics Communications* 2:e010
- [61] Qin J, Wang P, Ma W, Zhang CJ. 2025. Genome editing in polyploid crops: progress, challenges, and prospects. *Genomics Communications* 2:e022



Copyright: © 2026 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.