Ecogeographic signals of local adaptation in a wild relative help to identify variants associated with improved wheat performance under drought stress

Moses Nyine[1,2], Dwight Davidson[1,2], Elina Adhikari[1,3], Marshall Clinesmith[4,7], Huan Wang[1,5], Alina Akhunova[1,6], Allan Fritz[7], Eduard Akhunov[1,2]

[1] Department of Plant Pathology, Kansas State University, Manhattan, USA
[2] Wheat Genetics Resource Center, Kansas State University, Manhattan, USA
[3] Bayer, Chesterfield, USA
[4] Syngenta, Junction City, USA
[5] Broad Institute, Cambridge, Boston, USA
[6] Integrated Genomics Facility, Kansas State University, Manhattan, USA
[7]Department of Agronomy, Kansas State University, Manhattan, USA

Author for correspondence

Eduard Akhunov
Tel: +1 (785) 532-1342
Email: eakhunov@ksu.edu

Summary

Prioritizing wild relative diversity for improving crop adaptation to emerging drought-prone environments is challenging. Here, we combined the genome-wide environmental scans (GWES) in wheat diploid ancestor *Aegilops tauschii* with allele testing in the genetic backgrounds of adapted cultivars to identify new diversity for improving wheat adaptation to water-limiting conditions. Evaluation of adaptive allele effects was carried out in *Ae. tauschii*-wheat introgression lines (ILs) phenotyped for multiple agronomic traits under irrigated and water-limiting conditions using both UAS-based imaging and conventional approaches. The GWES showed that climatic gradients alone explain most (57.8%) of genomic variation in *Ae. tauschii*, with many alleles

associated with climatic factors in *Ae. tauschii* being linked with improved performance of ILs under water-limiting conditions. The most significant GWES SNP located on chromosome 4D and associated with temperature annual range was linked with reduced canopy temperature in ILs. Our results suggest that (i) introgression of climate-adaptive alleles from *Ae. tauschii* have potential to improve wheat performance under water-limiting conditions, (ii) variants controlling physiological processes responsible for maintaining leaf temperature are likely among the targets of adaptive selection in a wild relative, and (iii) adaptive variation uncovered by GWES in wild relatives has potential to improve climate resilience of crop varieties.

Introduction

The wild relatives of modern crops are a valuable source of adaptive diversity for developing improved varieties (Gill *et al.* 2006; Sohail *et al.* 2011; Kishii 2019). However, only a small fraction of wild relative diversity from the germplasm collections is utilized in breeding. The size of these collections, which may include thousands of accessions, complicates the selection of the most relevant genotypes for improving traits of interest. The prioritization of genebank germplasm for breeding climate adapted varieties is especially challenging due to the polygenic nature of adaptation to local environments (Araus *et al.* 2007; Exposito-Alonso *et al.* 2019). Therefore, the development of effective strategies, which are aimed at prioritizing wild relative germplasm for specific breeding applications, remains critical (Bohra *et al.* 2022).

The allopolyploid bread wheat, the second most important crop worldwide, originated by the hybridization of three wild grass species from the *Triticum* and *Aegilops* genera (Kihara 1944; Nesbitt and Samuel 1996; Dvorak *et al.* 1998; Tanno and Willcox 2006; Luo *et al.* 2007; Ozkan *et al.* 2011; Avni *et al.* 2017). Since its origin 10,000 years ago, wheat was disseminated by human migration and trade to diverse geographic regions with distinct climatic conditions(Balfourier *et al.* 2019). Archeological records and analyses of ancient DNA samples suggest that wheat reached Britain about 8,000 years ago (YA) (Smith et al., 2015) and China and Africa about 3,000 YA (Shewry 2009). Selection for performance in these diverse environments enriched local wheat populations for alleles contributing to adaptation to new climatic conditions (He *et al.* 2019; Zhao *et al.* 2023). However, the future climate change scenarios predict that climatic conditions in many

wheat growing areas could be outside of the adaptive range of existing genotypes and lead to severe yield reduction (Tack et al., 2015; Ortiz-Bobea et al., 2019). The results of climate modeling suggest that nearly 40% of crop growing areas might require new varieties to sustain crop production (Schlenker and Roberts 2009; Zabel et al., 2021). Leveraging the genetic diversity of multiple wild ancestors of wheat, which are evolved to grow in diverse environments, is one of the promising strategies for broadening the climate adaptive potential of modern wheat.

The direct ancestors of wheat, *Aegilops tauschii* and *Triticum turgidum* ssp. *dicoccoides* (wild emmer), are two most broadly distributed species among the wheat wild relatives (Dvorak *et al*. 1998; Avni *et al*. 2017). These species are also among the most represented in germplasm collections, some of which host thousands of accessions of *Ae. tauschii* and *T. turgidum* (Sharma *et al.* 2021). Because these wild ancestors of wheat share homologous genomes, their chromosomes could easily recombine, facilitating introgression of allelic diversity from *Ae. tauschii* and wild emmer into wheat (Nyine *et al.* 2020). By using the synthetic hexaploid wheat (SHW) lines, which are hybrids of tetraploid wheat and *Ae. tauschii*, the allelic diversity of these ancestors was introduced into multiple international breeding programs from CIMMYT, ICARDA, China, Australia, United Kingdom, and United States (Pestsova et al., 2004; Börner et al., 2015). For example, it was demonstrated that these ancestors of wheat have potential to improve adaptation to water limiting conditions and heat, and increase biomass and harvest index (Singh et al., 2019; Molero et al., 2023). However, considering the broad adaptive potential of these species reflected in their wide geographic distribution, the question remains of how effective these efforts were at capturing adaptive diversity of *Ae. tauschii* and wild emmer.

The prioritization of wild relative accessions for pre-breeding of climate resilient crops remains challenging. Wild relatives could be phenotypically pre-screened for target traits. However, this screening could be performed only for simple traits, and has limited utility for complex adaptive traits if phenotypic evaluation was not performed in the genetic background of adapted cultivars. Another approach to prioritize accessions is the development of “core collections” assembled from a large number of genotypes selected to maximize the genetic diversity of the sample (Frankel 1984). While this strategy could effectively reduce the number of accessions, its major disadvantage in application to large collections is that it targets only common

adaptive alleles, removing rare alleles or allelic complexes. The third approach is based on selection of wild relative accessions based on environmental parameters at the site of accession's origin (Turner et al., 2010; Jones et al., 2012; Lasky et al., 2015) (Bari *et al.* 2012). However, while this strategy could capture alleles contributing to an adaptive phenotype of small number of accessions, its ability to maximize the recovery of adaptive genetic diversity at species-wide level would be limited.

A combination of genome diversity analyses with the geographic patterns of environmental variation for detecting adaptive diversity is another strategy that so far had limited usage in crop breeding. The cost-efficiency of next-generation sequencing (NGS) genotyping approaches made possible generating genome-wide variation for geographically diverse populations. By combining genomic data with eco-geographic variables, it became possible to identify alleles associated with adaptive phenotypes. These approaches, referred to as genome-wide environmental scans (GWES), identify loci involved in local adaptation based on a high correlation between allele frequencies and eco-geographic variables. In an early GWES study, a number of climate-associated alleles (CAA) were mapped in *Arabidopsis* by using 13 climatic variables, among others including extremes and seasonality of temperature and precipitation (Hancock et al., 2011). The CAA identified by GWES allowed for accurate prediction of the relative fitness of *Arabidopsis* accessions in local environments (Turner et al., 2010; Hancock et al., 2011; Frachon et al., 2018). The GWES in sorghum and Mexican white oak detected adaptive variants that also produced reliable phenotypic predictions (Lasky et al., 2015; Martins et al., 2018). These studies suggest that adaptive alleles identified using the GWES have potential to predict agronomic phenotypes in target environments.

Though GWES were shown to be effective at identifying loci contributing to environmental adaptation, it remains unclear whether these loci could be used to prioritize wild relative accessions for introgression into modern crop varieties to improve their adaptive potential in extreme environments. To address this question, we used a diverse collection of *Ae. tauschii* to conduct GWES and identified variants contributing to climatic adaptation. We specifically focused on those variants that correlate with precipitation and temperature gradients during growth season. Then, we selected a geographically diverse set of *Ae. tauschii* accessions to develop introgression populations by crossing them with the adapted wheat varieties. The developed introgression

population was grown for several seasons across diverse environments and agronomic performance of introgression lines (ILs) was assessed by measuring agronomic and physiological traits. The physiological status and growth of ILs were evaluated using the UAS-based phenotyping with the RGB and thermal cameras. The wheat productivity was assessed by measuring yield and yield component traits (thousand grain weight, grain area, grain width and grain length). The relationship between phenotypic data and climate adaptive alleles introgressed from *Ae. tauschii* was investigated to better understand the value of GWES in wild relatives as a tool for selecting wild relative accessions to improve the adaptive potential of wheat varieties.

Materials and methods

Plant materials

A diverse set of 137 geo-referenced *Ae. tauschii* accessions collected over a geographic range of species distribution and representing locations with diverse historic climatic and bioclimatic characteristics was acquired from the USDA NSGC to identify the CAAs (Table S1). A subset of 21 geographically diverse accessions was selected from this population and crossed with hard red winter wheat varieties to generate *Ae. tauschii*-wheat amphiploids. The amphiploids were then crossed with six hard red winter wheat cultivars adapted to grow in the US Great Plains to develop *Ae. tauschii*-wheat ILs (Nyine et al., 2020, Nyine et al., 2021). A total of 351 $BC_1F_{3:5}$ introgression lines that had phenology similar to that of the recurrent parents were used to study the impact of introgressed CAA on the adaptative traits.

Genotyping and imputation

DNA was extracted from two-week old seedling leaf tissues of the diverse *Ae. tauschii* accessions and the derived introgression population using DNeasy 96 Plant DNA extraction kit (Qiagen) following the manufacturer's protocol. The quality and concentration of the DNA was assessed using PicoGreen dsDNA assay kit (Life Technologies). The extracted DNA was normalized to 400 ng (20ul of 20ng/ul) using the Qiagility robot (Qiagen). Genotyping by sequencing (GBS) included a library size selection step performed using the Pippin Prep system

(Sage Scientific) to enrich the library for 270-330 bp fragments, as described in Saintenac et al. (2013). The prepared libraries were sequenced on Illumina NextSeq 500. Variant calling was done using the TASSEL v5.0 GBS v2 pipeline (Glaubitz et al., 2014).

To increase the density of SNP markers in both populations, we re-sequenced the panel of 21 *Ae. tauschii* accessions and the six recurrent hexaploid wheat lines using whole-genome sequencing approach. PCR-free genomic libraries were constructed using Illumina protocol at the Integrated Genomic Facility (IGF) at Kansas State University. Paired-end sequences (2 x 150 bp) were generated using NovaSeq at Kansas University Medical Center and NextSeq 500 at IGF. The data were combined and processed as described by Nyine et al. (2021). Missing and ungenotyped SNPs in the *Ae. tauschii* diversity panel and the introgression population were imputed from the parental genotypes using Beagle v5.0 (Browning and Browning 2013). After imputation and filtering out SNPs with genotype probability below 0.7, we retained 6,365,631 SNPs in *Ae. tauschii* diversity panel and 5,208,054 SNPs in the introgression population.

Population structure and variance partitioning of SNP diversity in *Ae. tauschii*

To understand the level of genetic diversity within the *Ae. tauschii* population and how both geography and climate shaped the SNP variation in the population, we pruned the 6.3 million SNPs based on linkage disequilibrium (LD) using PLINK v1.9 and retained 109,627 SNPs that had $r^2 < 0.5$ in 50 kb sliding window with step size of 5 kb. The proportion of ancestry shared between accessions was estimated from the LD pruned SNPs and the geographical coordinates for the accessions' collection sites using the tess3r R package (Caye et al. 2016). The maximum number of ancestral populations tested was eight (K = 1:8). Each K was run 10 times for 200 iterations (rep = 10, max.iteration = 200) and the spatial projection of ancestral coefficients was based on least squares method (method = "projected.ls"). The optimal number of ancestral populations selected based on cross-validation scores was K = 4 because it split the *Ae. tauschii* population into two lineages and four sub-lineages that coincided with previous findings by Wang et al., (2013). A plot showing the population admixture and the spatial distribution of accessions from different subspecies at the sites of sample collection was generated.

The proportion of SNP variance in the *Ae. tauschii* accessions was partitioned into those explained by geographic distance and climate using the 'varpart' function in R package 'vegan'. The geographic distances were calculated using the 'distVicentyEllipsoid' function in R package 'geosphere' using the GPS coordinates from the accessions' collection sites and the results were presented in a Venn diagram.

Redundancy analysis

The diverse set of *Ae. tauschii* accessions used in this study came from a wide range of geographical locations with distinct climatic and bioclimatic conditions suggesting that certain genetic factors are involved in local adaptation. Based on this hypothesis we modeled the relationship between response variables (SNPs) and explanatory variables (climatic and bioclimatic factors, and geographic distance) using the redundancy analysis (RDA) (Van den Wollenberg 1977; Lasky et al., 2015) and mixed linear models to identify variants contributing to local adaptation. Variables were ranked to identify those that contribute most to SNP diversity in the *Ae. tauschii* population. To achieve this, the ordiR2step function was applied on the RDA results with adjusted R^2 using the forward selection method from 10,000 permutations. Type 1 error was minimized during the selection of the most important factors contributing to SNP diversity by following the rules proposed by Blanchet et al., (2008). The full RDA model based on all climatic and bioclimatic variables with the calculated adjusted R^2 values was evaluated to determine variables that (1) significantly improved the explained variation of SNP diversity distribution in *Ae. tauschii* population at alpha 0.05, and (2) whose total adjusted R^2 did not exceed the adjusted R^2 value of the full model. Based on the aforementioned conditions, the most important variables identified were projected on the first two principal components as a biplot. To illustrate the variation of temperature annual range (BIO7) from the *Ae. tauschii* accessions' collection sites, a heatmap was plotted using the 'heat_point' function provided in R package 'autoimage' and an overview of the variation of the most important variables at the collection sites for the sub-lineages was compared using boxplots. All variables were scaled to range between 0 and 1 by dividing with the highest value within the dataset and then squaring them to eliminate the negatives before generating the boxplots with ggplot2 in R.

206

207

208 **Identification of climate associated alleles (CAAs) in *Ae. tauschii***

209 To determine the genetic basis of local adaptation in *Ae. tauschii*, we used both RDA and
210 GWAS to identify SNPs that were significantly associated with geographic, climatic and
211 bioclimatic variables. We extracted the first three RDA loadings for each SNP from the RDA
212 model described above and transformed them to Z-scores. The mean Z-score was calculated from
213 all SNPs and any SNP with three standard deviations from the mean was considered a candidate
214 CAA SNP. Pearson's correlation coefficients were used to determine the variable with the
215 strongest association to each SNP, thus each CAA SNP was assigned to only one variable. To
216 capture most CAA, we also performed a GWAS using a compressed mixed linear model in GAPIT
217 (Lipka et al., 2012). The geographic, climatic and bioclimatic variables were used as phenotypes.
218 The population structure was accounted for by including PCAs calculated from the marker data as
219 covariates in the model. Multiple test correction was performed using the Benjamini-Hochberg's
220 method (FDR ≤ 0.05).

221

222 **Phenotyping of *Ae. tauschii* introgression population**

223 The population of ILs was phenotyped under field conditions for three seasons between
224 2018 and 2020 to evaluate the adaptive potential of CAAs introgressed in the winter wheat. In
225 2018 and 2019, phenotyping was done at Colby (Kansas, USA) under irrigated and non-irrigated
226 conditions. In 2020, phenotyping was done at Ashland (Kansas, USA) under non-irrigated
227 conditions. The experimental layout at all locations followed an augmented design with six
228 recurrent hexaploid wheat parents and three additional winter wheat lines adapted to Kansas
229 weather as controls. Experimental plots were 2.5 m x 0.5 m consisting of three rows separated by
230 18 cm. During planting, granular 18-46-0 diammonium phosphate (DAP) fertilizer was applied at a
231 rate of 168.1 kg/ha and liquid 28-0-0 urea ammonium nitrate (UAN) was applied at a rate of 67.3
232 kg/ha in the spring to supply additional nitrogen to the plants. The lateral irrigation system was
233 used to maintain the soil moisture in the irrigated block.

The ILs were phenotyped for yield and the component traits such as spikelet number per spike (SNS), thousand grain weight (TGW), grain area (GA), grain width (GW) and grain length (GL). During the growing season, remote sensing data including RGB, NDVI and canopy temperature (CT) were collected at multiple time points during growth seasons using unmanned aerial system (UAS) mounted with specific sensors for each data type to evaluate the physiological status and growth trend of the introgression lines. The RGB and NDVI imagery data were processed in Agisoft software (version) to generate orthomosaics and digital elevation models (DEM) whereas the thermal data were processed in Pix4D to generate the CT orthomosaics. The raster files generated by Agisoft and Pix4D were imported into QGIS v3.4 software for plot level data extraction. Shape files consisting of rectangular polygons that overlaid each plot in the experimental block were created and the mean pixel values for each color band within the polygon were calculated using raster zonal statistics tools and saved as a comma separated values (csv) file. Other indices such as visible atmospherically resistant index (VARI) and triangular greenness index (TGI) were derived from the RGB data whereas NDVI was derived from near infrared and red color bands using the following equations:

$$VARI = \frac{G - R}{G + R - B} \quad \text{eqn. 1}$$

$$TGI = G - 0.39 * R - 0.61 * B \quad \text{eqn. 2}$$

$$NDVI = \frac{1.236 * NIR - 0.188 * R}{NIR + 0.044 * R} \quad \text{eqn. 3}$$

where R, G, B and NIR are the mean pixel values for the red, green, blue and near infrared color bands.

Heading data were collected in 2020 at Ashland and validated in 2022 at RockyFord, Manhattan, Kansas USA. Heading date was recorded when 50 % of spikes fully emerged from the flag leaf. The number of days to heading (DTH) were calculated by subtracting the planting date from the heading date. To understand how much the heading date for the introgression lines varies from the controls, we calculated the mean DTH for the controls and subtracted the DTH for each introgression to generate the deviation in days to heading (DDTH).

Best linear unbiased predictions (BLUPs)

Best linear unbiased predictions for yield and yield component traits were obtained from a mixed linear model implemented in R package. Given that the experimental layout followed an augmented design all controls were given a code 1, and the test introgression lines were assigned a 0. Each introgession line was assigned a unique numeric code which was used as a group identifier in all experiments whereas the controls were assigned a 999 regardless of the accession as the group identifier. A mixed linear model was run for different traits. For example, BLUPs for the number of days to heading, used in GWAS analysis were estimated from the following model:

DTH~Loc+Check, random= ~ Acc + Acc:range/row + Acc:LocName,

where DTH is number of days to heading, Loc is the field trial location, Check defined lines whether they are controls or test lines and Acc is the accessions.

Canopy temperature data were collected over multiple time points (aka flights) in the two years. Spatial correction was performed using SpATS implemented in MrBean, a shiny based R package. Variance due to genotype and environment were estimated as well as narrow sense heritability. After excluding outliers, BLUEs were predicted for the test lines based on the variance in the controls. All flights and blocks with heritability less than 0.2 were excluded from the linear mixed model. The remaining data were used to estimate genotype BLUPs across flights, treatment blocks and years. In the model, experimental treatment and flights were considered as fixed effects whereas genotypes were considered as random effects.

Association between CAAs and phenotypic traits in introgression population

The frequency spectra of CAAs derived from *Ae. tauschii* was estimated for groups of ILs that have trait values falling into the tails of phenotype distributions. Our expectation was that if the CAAs are associated with variation in phenotypic traits in the introgression population, the phenotypic value tails should show CAA frequency spectra distinct from the CAA frequency spectra for the whole population. For this purpose, we ranked the ILs based on trait values and compared the CAA frequency spectrum in the whole population (WP) and the lower and upper 5th percentile of the phenotype distributions. Lines were considered to belong to the lower or upper

tail groups if they were ranked as outliers in at least two trials. The frequency spectrum was built using 5,675 CAAs in the introgression population by counting alleles in 9 allele frequency bins.

Allele frequency of all CAA sites was calculated for the whole population and for the introgression lines that ranked in the lower and upper 5% tails of phenotype distribution using vcftools. Differentiation in allele frequency was determined by calculating the fold change (FC) in allele frequency between the introgression lines in the tails and the whole population. SNPs with FC ≥ 2 were considered to strongly differentiated. For some traits, where the allele frequency FC was less than two across all sites, the threshold was adjusted accordingly.

To determine the relationship between recombination rate gradient and allelic differentiation, we split the chromosome arms into three equal parts. The number of differentiated CAAs within each chromosome segment was counted using the 'bedmap' function of BEDOPs tools. The redundant CAA sites showing the high levels of LD were removed using PLINK. We retained only those SNPs that had $r^2 < 0.5$ within the 50 kb window with a step size of 5 kb. The total number of differentiated alleles was aggregated for all traits and chromosomes.

To confirm the contribution of CAAs to adaptation traits a linear regression of yield on CT was performed to determine the proportion of variance in yield explained by the variation in CT. Introgression lines that ranked in the 5th and 95th percentiles of CT distribution were compared for yield performance relative to the recurrent parents. GWAS was performed on the traits phenotyped in the $BC_1F_{3:5}$ *A. tauschii*-wheat introgression population including CT, heading date, yield and component traits to determine loci with significant associations. Multiple GWAS models were tested on each trait with varying number of principal components to correct for population structure. GWAS analysis was implemented in GAPIT v3.0. CAAs significantly associated with traits in the introgression population at the FDR value 0.05 were considered adaptive in the winter wheat background.

Results

Environmental scans in *Aegilops tauschii*

The genetic basis of *Ae. tauschii* adaptation to diverse climatic conditions across a broad geographic range extending from Eastern Europe to China remains poorly understood. To identify genetic loci contributing to adaptation, we conducted genotype-environment association analyses using 109,627 SNPs identified in a geographically diverse panel of 137 accessions. This set of SNPs was selected by LD-based pruning from a larger set including 6,365,631 genotyped and imputed SNPs. Using this data, we explored the population structure of our samples and its correspondence to the previously identified four main lineages (L1W, L1E, L2E, L2W) of *Ae. tauschii* (Wang et al. 2013) (Fig. 1A). The inferred population structure of *Ae. tauschii* accessions was consistent with the results of previous studies (Wang et al. 2013), showing the split between L1 and L2 lineages, where L1 was composed of *Ae. tauschii* ssp. *tauschii* and L2 included accessions of *Ae. tauschii* ssp. *strangulata*, the closest ancestor of the wheat D genome (Wang et al. 2013), (Fig. S1A). The first two principal components, separating L1 and L2 lineages, accounted for 78.4% of the variation in our samples (Fig. S1B). The split between the L1W and L1E and between the L2E and L2WE was also obvious in our panel.

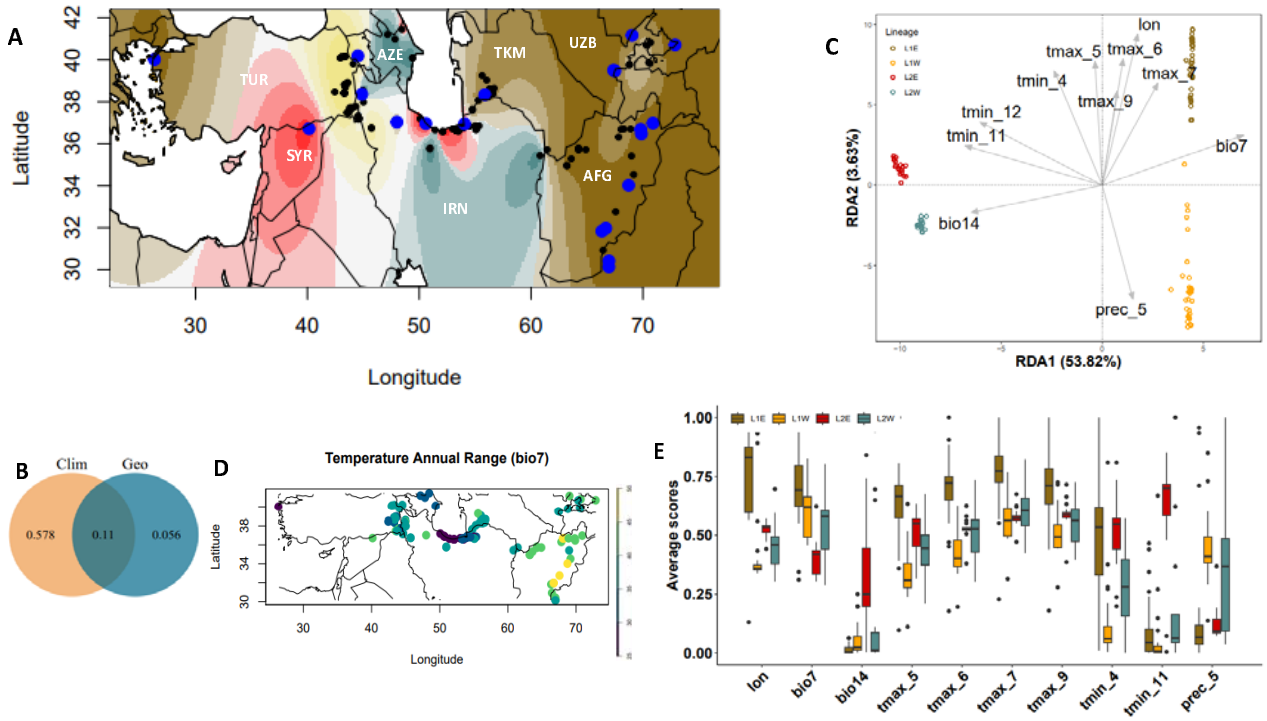


Fig. 1 Ecogeographic distribution of 137 diverse *Ae. tauschii* accessions**. A)** Map shows the geographic locations and the ancestry coefficients of 137 accessions. Twenty-one accessions used to develop the *Ae. tauschii*-wheat introgression population are shown in blue. **B)** Proportion of SNP variance explained by climate (Clim) and geographical distance (Geo) between accessions. **C)** Redundancy analysis (RDA) plot showing the 11 best explanatory variables for SNP variance in *Ae. tauschii* accessions including temperature annual range (bio7), precipitation of driest month (bio14), precipitation in May (prec_5), minimum temperature in April, November and December (tmin_4, tmin_11 and tmin_12, respectively), maximum temperature in May, June, July and September (tmax_5, tmax_6, tmax_7 and tmax_9, respectively) and longitude (lon). **D)** Heatpoint map showing the variation in temperature annual range (bio7) at the sampling locations of the *Ae. tauschii* accessions. **E)** Boxplots showing the variation in lon, bio7, bio14, tmax_5, tmax_6, tmax_7, tmax_9, tmin_4, tmin_11 and prec_5 in the ecogeographic locations of different *Ae. tauschii* lineages. Lineage 1 East (L1E) and Lineage 1 West (L1W) belong to *Ae. tauschii* ssp. *tauschii* whereas Lineage 2 East (L2E) and Lineage 2 West (L2W) belong to *Ae. tauschii* ssp. *strangulata*.

To identify variants contributing to local adaptation, we modeled the relationship between response variables (SNPs) and explanatory variables (climatic and bioclimatic factors, and geographic distance) using redundancy analysis (RDA) (Van den Wollenberg 1977; Mcardle and Anderson 2001; Lasky et al. 2015). For this purpose, we used historical data for the bioclimatic and climatic factors estimated for the geographic locations at the accession collection sites. The total SNP variance explained by both geographic distance and climatic and bioclimatic factors was 85.2% (Fig.1B), with the adjusted R^2 value being 69.6%. Climate alone accounted for 57.8% of the SNP variation in *Ae. tauschii*, whereas geographic distance between accessions and the interaction between geographic distance and climate accounted for 5.6% and 11% of SNP variation, respectively. These results indicate that the distribution of SNP variation among *Ae. tauschii* accessions is primarily driven by gradient in climatic and bioclimatic factors rather than by *Ae. tauschii* geographic dispersal.

Depending on their impact on adaptive traits, individual climatic factors could have distinct effects of SNP variation among accessions (Hancock et al., 2011; Lasky et al., 2015; Li et al.,

2021; Chang et al., 2022). The first two RDAs accounted for 22.58% of the total SNP variation in the population. A triplot with two RDAs shows separation of the population into four distinct groups (Fig. S1C) coinciding with the previously detected split between the L1E, L1W, L2E and L2W subpopulations. Among the geographic variables, longitude and altitude showed the strongest effect on SNP distribution between the two subspecies of *Ae. tauschii* followed by latitude (Fig. S1D). The environmental variables contributing most to SNP variation were determined using the 'ordiR2step' function in R package 'vegan' using the forward selection method and 10,000 permutations (Blanchet et al., 2008). A total of 11 variables were detected, including temperature annual range (bio7), precipitation of driest month (bio14), precipitation in May (prec_5), minimum temperature in April, November and December (tmin_4, tmin_11 and tmin_12, respectively), and maximum temperature in May, June, July and September (tmax_5, tmax_6, tmax_7 and tmax_9, respectively) and longitude (lon) (Table 1, Fig. 1C). The tmin_11, tmin_12, bio7, and bio_14 contributed most to RDA1 that explains most of the genetic differentiation between the two subspecies of *Ae. tauschii*. The prec_5, tmin_4, tmax_5, tmax_6, tmax_7 and tmax_9 factors contributed to RDA2 that explains most of the genetic differentiation between the Eastern (L1E, L2E) and Western (L1W, L2W) populations of the two *Ae. tauschii* lineages. These results indicate that temperature and precipitation gradients during the growth periods coinciding with flowering, grain filling and maturation were the main factors that shaped SNP diversity in *Ae. tauschii* and likely contributed to genetic differentiation among the four lineages.

Two subspecies of *Ae. tauschii*, ssp. *strangulata* and ssp. *tauschii* appear to show different levels of adaptation to distinct climatic conditions. We compared the distribution of the main climatic and geographic factors (lon, bio7, bio14, tmax_5, tmax_6, tmax_7, tmax_9, tmin_4, tmin_11, tmin_12 and prec_5) between the two subspecies of *Ae. tauschii* (Fig.1E). Analysis of variance showed significant differences between the accessions from these subspecies (Table 2). Results suggest that L1E lineage is adapted to warmer and drier conditions of Eastern Iran, Afghanistan, Turkmenistan, Uzbekistan, Tajikistan and Kyrgyzstan (Table S1), indicating that these *Ae. tauschii* accessions could be a good source of drought and heat stress tolerance. The lowest precipitation of driest month characterized by high maximum temperature from May up to September was one of the major differentiating ecogeographic factors for L1E. In contrast, L1W is represented by accessions mostly from Eastern Turkey and Northwestern Iran where a high precipitation is recorded in May and significantly lower maximum temperature from May to September. The

393 factors that contribute most to the genetic differentiation of this sublineage are tmin_4, tmax_5 and
394 prec_5.

395 **Table 1** The geographic, climatic and bioclimatic variables that contributed most to SNP variation
396 in *Ae. tauschii*. The R^2.adj are cumulative values.

Variable	R^2.adj*	Df	AIC	F	Pr(>F)	
Temperature annual range (bio7)	0.27664	1	1236.8	43.4509	1.00E-04	***
Min temperature in November (tmin_11)	0.35655	1	1224.7	14.6606	1.00E-04	***
Max temperature in July (tmax_7)	0.39818	1	1218.2	8.5394	0.0008	***
Max temperature in June (tmax_6)	0.43887	1	1211.3	8.832	0.0011	**
Min temperature in December (tmin_12)	0.49921	1	1199.5	13.8934	0.0002	***
Longitude (lon)	0.52234	1	1195.1	6.1326	0.0043	**
Max temperature in September (tmax_9)	0.53963	1	1191.9	4.9424	0.0093	**
Precipitation of driest month (bio14)	0.55291	1	1189.6	4.0891	0.0148	*
Min temperature in April (tmin_4)	0.57148	1	1185.7	5.4646	0.004	**
Max temperature in May (tmax_5)	0.59765	1	1179.6	7.6331	0.0017	**
Precipitation in May (prec_5)	0.60705	1	1177.8	3.4162	0.025	*

397

398 Generally, *Ae. tauschii* ssp. *strangulata* lineages are found around the Caspian Sea with
399 some accessions found in Syria and Turkey. L2E is adapted to a relatively uniform precipitation
400 and mild temperature which are characteristic of the Southern Caspian Sea in Northern Iran. The
401 L2W accessions are mostly found near Western Caspian Sea in Azerbaijan and parts of
402 Northwestern Iran. The region is characterized by moderate to high variation in climatic and
403 bioclimatic factors. Amongst the main variables differentiating L2W from other sublineages are
404 tmin_4 and prec_5 (Table 2).

Table 2 Comparison of geographic and climatic factors between the main sub-lineages of *Ae. tauschii*.

Variable	Lineage 1 East (L1E)	Lineage 1 West (L1W)	Lineage 2 East (L2E)	Lineage 2 West (L2W)
lon	62.83^{a}	45.98^{c}	52.69^{b}	49.09^{bc}
bio7	40.08^{a}	37.01^{b}	30.11^{c}	34.88^{b}
bio14	3.13^{c}	7.10^{b}	19.90^{a}	10.16^{b}
tmax_5	269.44^{a}	195.90^{c}	244.45^{b}	221.77^{b}
tmax_6	324.80^{a}	251.92^{b}	283.60^{b}	273.94^{bc}
tmax_7	345.08^{a}	295.97^{b}	300.45^{b}	305.41^{b}
tmax_9	292.64^{a}	253.33^{c}	272.45^{b}	262.71^{bc}
tmin_4	80.82^{a}	32.77^{c}	84.80^{a}	53.00^{b}
tmin_11	18.26^{b}	0.03^{c}	78.25^{a}	28.65^{b}
tmin_12	-17.46^{b}	-49.28^{c}	42.35^{a}	-23.00^{b}
prec_5	26.72^{c}	58.36^{a}	30.60^{c}	46.41^{b}

Means with the same superscript letters are not significantly different. The following factors were considered: longitude (lon), temperature annual range (bio7), precipitation of driest month (bio14), maximum temperature in May, June, July and September (tmax_5, tmax_6, tmax_7 and tmax_9), minimum temperature in April, November and December (tmin_4, tmin_11 and tmin_12) and precipitation in May (prec_5). Statistical significance is based on Tukey's honestly significant difference test at 95% confidence level.

Mapping adaptive SNPs in *Ae. tauschii*

The results of both redundancy analysis (RDA) and genome-wide association mapping (GWAS) were used combined to identify SNPs that are significantly associated with variation in geographic, climatic and bioclimatic variables across the *Ae. tauschii* sampling locations. The first three RDA loadings for each SNP were extracted from the RDA model, transformed to Z-scores, and climate associated alleles (CAA) were defined as outlier SNPs with three standard deviations from the mean Z-score. After removing duplicate SNPs, a total of 10,149 D-genome SNPs showed a positive correlation (mean = 0.58, range 0.14-0.82) with 51 out of 58 variables analyzed in our study (Table S2, Table S3).

Temperature annual range (bio7) was ranked as the most significant variable accounting for 27.66% of the SNP variation in *Ae. tauschii* population based on the adjusted R^2 values ($P < 0.001$). It had the highest number of correlated SNPs (747) amongst other most significant variables. Chromosomes 1D and 7D had the highest number of SNPs showing significant correlation with the geographic, climatic and bioclimatic variables (Fig. 2A). Genome-wide association mapping is another approach that was previously used for studying genome-by-environment interactions (Wallace et al. 2016). By using a compressed mixed linear model, we identified a total of 10,569 D-genome SNPs significantly associated with 42 out of 58 variables (Table S2, Fig. 2B). Most of these variants were located on chromosomes 1D and 5D (Fig. 2A and Table S4). Unlike in RDA analysis, where each SNP was assigned to a single highly correlated variable, in GWAS, many SNPs were associated with more than one variable at FDR ≤ 0.05. Combined, RDA and GWAS identified 18,096 SNPs with significant association to geographic, climatic and bioclimatic variables (Table S5). Among the SNPs identified, a set of 2,622 SNPs were detected using both methods (Fig. 2C, Table S6). The functional annotation of these SNPs using SnpEff (Cingolani et al., 2012) showed that only 29 of them were stop codon gain, missense, synonymous, intronic or splice region variants (Table S6). The majority of SNPs (2,316) were intergenic variants, and 277 SNPs were located 5 kb upstream or downstream of gene models.

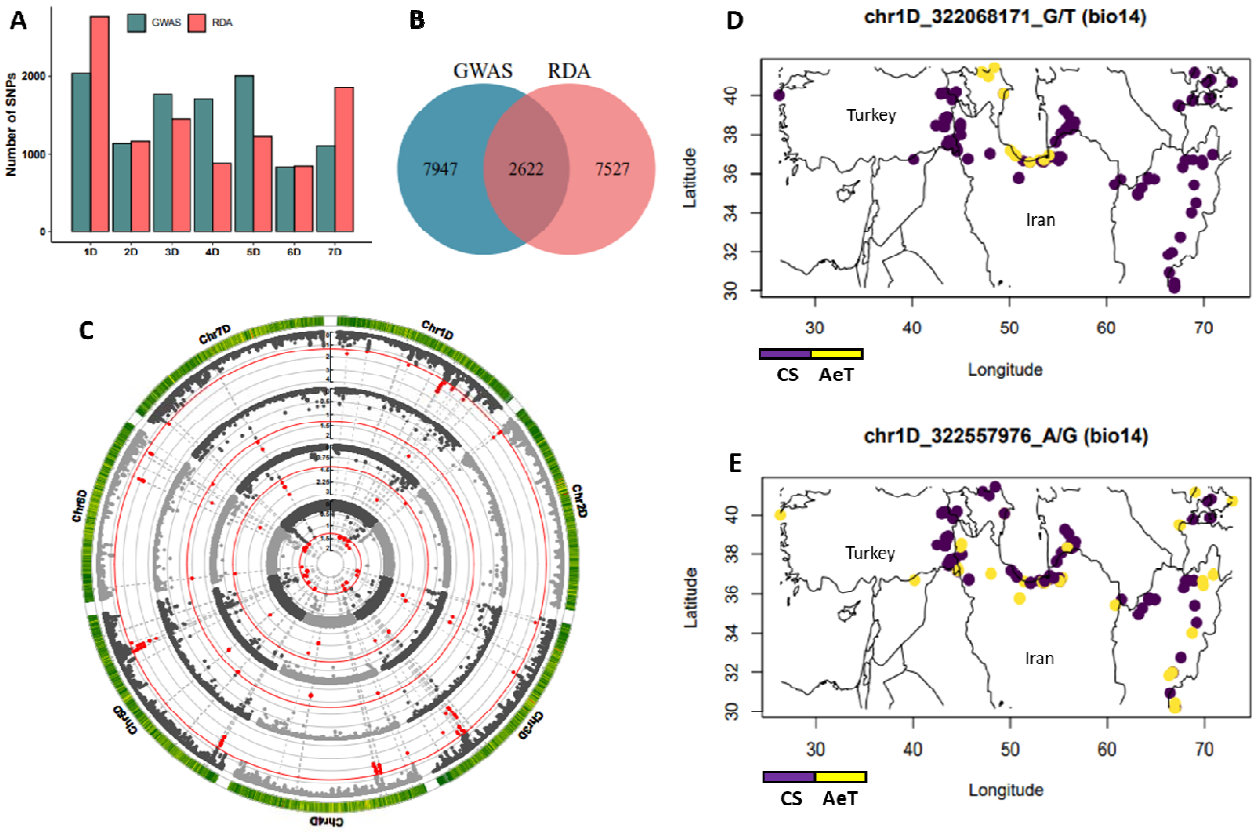


Fig. 2 Number of climate associated SNPs per chromosome identified by the redundancy analysis (RDA) and genome-wide association analysis (GWAS), and the geographical distribution of SNP alleles associated with precipitation of driest month (bio14). A) Chromosome distribution of SNPs identified through RDA and GWAS that were associated with different geographic, climatic and bioclimatic variables. B) Venn diagram showing the total number of climate associated SNPs identified by RDA and GWAS. C) Circular Manhattan plot showing GWAS for four of the most significant variables. Starting from the innermost circle outward are minimum temperature in April, maximum temperature in May and June (tmax_5 and tmax_6) and bio14. The red lines show an FDR threshold of 0.05 and the red dots are the significant SNPs on each chromosome. D) *Ae. tauschii* (AeT) specific allele (yellow) on chromosome 1D showing adaptation to areas with high precipitation of driest month near the Caspian Sea. E) *Ae. tauschii* specific allele (yellow) on chromosome 1D showing adaptation to areas with a wide range of reduced precipitation in the driest month. The purple color shows the reference allele similar to Chinese Spring (CS).

Consistent with prior studies, the geographic extent of CAAs could primarily be explained by the distribution of climatic factors (Hancock et al., 2011). For example, among SNPs significantly associated with bio14, there is an *Ae. tauschii* allele (chr1D_322068171) that was found only in accessions from the region near the Caspian Sea (Fig. 2D), which shows a high precipitation of driest month (Table 1). Another bio14-associated SNP (chr1D_322557976) had an allele identified in accessions from a broad geographic region characterized by low precipitation of driest month experiencing extreme drought stress due to high temperature from May up to September (Fig. 2E, Table 1). These results suggest that *Ae. tauschii* could be the source of adaptive alleles to a broad range of climatic factors useful for addressing the impact of climate change on wheat productivity.

Evaluation of the adaptive potential of *Ae. tauschii* CAAs in winter wheat

To evaluate the ability of CAAs from *Ae. tauschii* to improve the adaptive potential of bread wheat, we developed a wild relative introgression population using a set of 21 diverse accessions that were selected to capture the ecogeographical and allelic diversity of species (Nyine et al., 2020; Nyine et al., 2021). To facilitate comparison with parental lines, ILs in the populations were selected to match development and phenology of hexaploid wheat parents (Nyine et al., 2020). Out of 18,096 climate adaptive SNPs identified by RDA and GWAS, 31.4% (5,675 CAA SNPs) were present in the introgression population. It is likely that loss of some of the CAAs in introgression population could be caused by their linkage with deleterious alleles selected against during population development (Nyine et al., 2020) (Fig. 3). Among the introgressed CAAs, a total of 1,089 SNPs were detected using both environmental association scan methods.

Introgression of beneficial alleles occurs in both high and low recombining regions of the genome. While introgressions found in high recombining regions become shorter after a few generations of recombination, those in the low recombining regions tend to persist as large linkage blocks. The large introgression blocks in the pericentromeric regions of the chromosome could have unintended consequences on non-targeted traits due to linkage with deleterious alleles linked to adaptive SNPs and epistatic interactions with adapted genetic background. When selection is applied, the frequency of adaptive alleles in the high recombining regions usually increases

whereas the frequency of adaptive alleles in low recombining regions may (i) increase if SNP effect on an adaptive trait is stronger than the combined negative effects of linked alleles or (ii) reduce if the combined negative effects of linked alleles are stronger than the effects of adaptive SNPs. In a breeding population, shifts in allele frequency are best observed in the tails of distributions for phenotypes targeted by selection. We compared the frequency spectra of CAAs derived from *Ae. tauschii* in the tails of phenotype distributions in the introgression population. Our expectation was that if CAAs affect wheat performance, the tails of distribution for yield and yield component, CT and DDTH traits should show distinct frequency spectra. For this purpose, we ranked the ILs based on trait values and compared the CAA frequency spectrum in the whole population (WP) with CCA frequency spectra in the lower and upper 5th percentile tails of the phenotype distribution. The frequency spectra were generated for the 5,675 CAAs in the introgression population by counting *Ae. tauschii* alleles in the 9 allele frequency bins. Lines in the tails of the trait distribution were filtered to retain only those that ranked in the same percentile group for at least two traits.

A significant shift from the mean CAA frequency (0.226) in the WP was observed in the tails of phenotype distribution for various traits (Fig 3). The shift in the CAA frequency spectrum in the tails of yield distribution was significantly different from WP mean (Kolmogorov-Smirnov test: lower tail $P < 2.2e\text{-}16$; upper tail $P < 2.2e\text{-}16$). In the lower tail of yield distribution, the frequency of CAAs was high suggesting that lines with large introgression segments that had many CAAs that are likely in LD with deleterious alleles contributed to yield penalty. The top yielding lines showed a bimodal distribution of CAA frequency, suggesting the occurrence of both negative and positive selection at different CAA loci in these lines. For SNS however, a decrease in the *Ae. tauschii* allele frequency was linked with low spikelet number per spike whereas a combination of both low and high frequency *Ae. tauschii* alleles were associated with with higher SNS (Kolmogorov-Smirnov test: lower tail $P < 2.2e\text{-}16$; upper tail $P < 2.2e\text{-}16$).

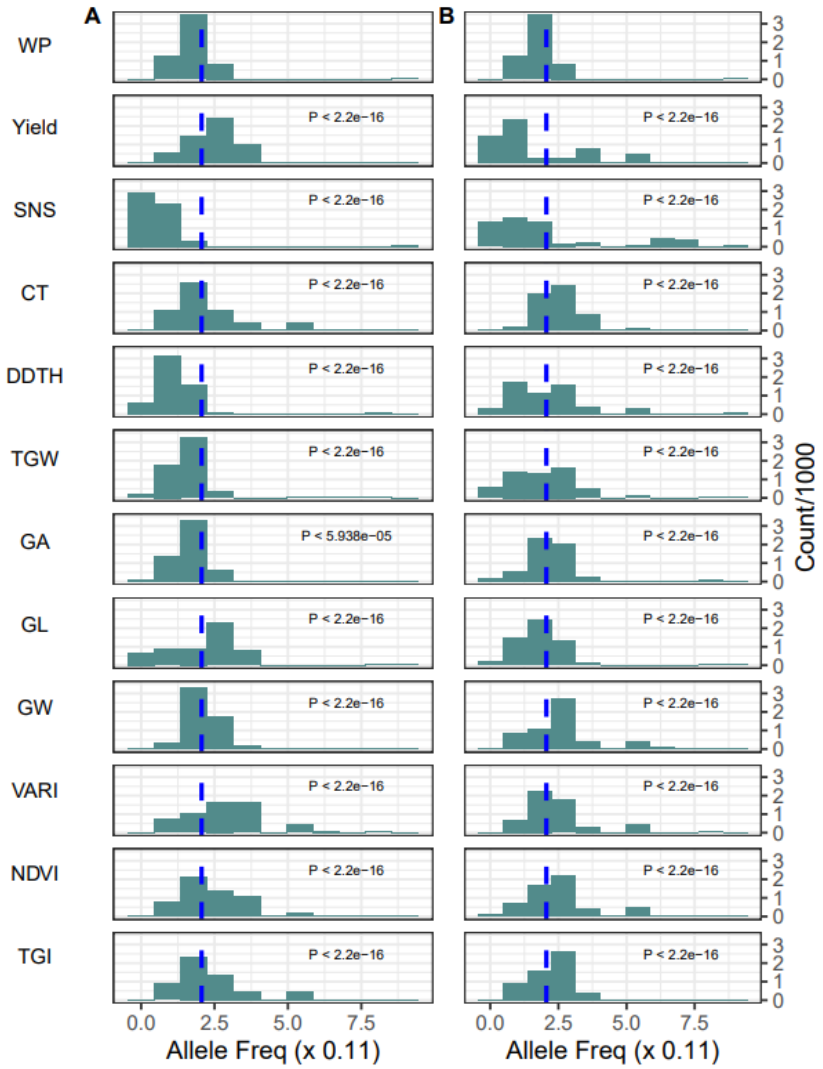
Previous studies have shown a positive relationship between yield and SNS, especially if all spikelets are fertile and produce seeds (Rawson 1970; Zhang et al., 2018, Kuzay et al., 2019). However, yield is a complex trait modulated by changes heading date and yield component traits, such as grain area, width and length. The shift in the CAA frequency spectrum for the deviation in days to heading (DDTH) distribution from the WP mean followed the same pattern as that

observed for SNS (Kolmogorov-Smirnov test: lower tail P < 2.2e-16; upper tail P < 2.2e-16) suggestive of the shared biological pathways between these two traits. These results are as expected because longer development period and hence delayed heading have been associated with increase in SNS (Rawson 1970). Guo et al. (2018) attributed the increase in spikelet number to delayed spikelet initiation and transition from double-ridge phase to terminal spikelet which coincided with delayed heading date.

Canopy temperature (CT) is one of the critical physiological traits that reflects the adaptive potential of plants in local environments (Kumar et al., 2017) (Still *et al.* 2021) and could be used to identify drought and heat stress tolerant plant genotypes. Previous studies demonstrated that CT in wheat is a complex quantitative trait mostly linked to QTLs that control root architecture necessary for improved water use efficiency and maintenance of transpiration rate (Pinto and Reynolds 2015). While in the low CT tail, most CAAs had lower than average allele frequency, we detected some CAAs that significantly increased in frequency compared to population mean suggestive of their contribution to regulation of CT. CT negatively correlated with yield under drought stress (r = -0.45, P = 0.0) which agreed with the previous studies that showed the importance of CT depression for increasing yield in wheat (Pinter et al. 1990; Amani et al. 1996). Introgression from *Ae. tauschii* into spring wheat was associated with low CT and improved yield under heat stress (Molero et al., 2023). Previously, we showed that the difference in mean yield of some *Ae. tauchii* introgression lines in our population reached 57% when compared to the checks under drought stress conditions (Nyine et al., 2021).

Besides CT, both visible atmospherically resistant index (VARI) and normalized difference vegetation index (NDVI) are correlated vegetation indices used to monitor plant health and biomass accumulation. Lines in both lower and upper tails showed a significant shift towards high frequency CAAs. The latter shows that some CAAs could be associated with the positive impact on vegetation indices and physiological status of the plants under stress. The finding of CAAs showing strong shift in the lower tails of both traits suggest that some CAAs or linked introgression variants could be associated with the negative impact on these traits. The triangular greenness index (TGI) is an indicator of total chlorophyll content in the leaves (Hunt et al., 2013) which is useful for estimating the stay green characteristics in wheat (Lopes and Reynolds 2012).

544 In this population, lines that matured late were those with the highest number of *Ae. tauschii*
545 alleles thus the TGI values were also high during the growing season.



546

547 **Fig. 3** Frequency spectra of climate associated alleles (CAA) in the tails of phenotype distribution
548 relative to the allele frequency spectra in the whole population (WP). The A and B panels show the
549 5th and 95th percentiles of the yield, spikelet number per spike (SNS), canopy temperature (CT)
550 deviation in days to heading (DDTH) from the control mean DTH, thousand grain weight (TGW),
551 grain area (GA), grain length (GL), grain width (GW), visible atmospherically resistant index
552 (VARI), normalized difference vegetation index (NDVI) and triangular greenness index (TGI)
553 traits.

554

Both natural and artificial selection results in genetic differentiation at target loci between the selected and non-selected populations. Since the efficiency of selection at target loci is higher in the high recombining regions, for CAAs linked with trait variation, we expect to observe higher frequency differentiation between the lines in phenotypic tails in the high-recombining terminal regions of chromosomes rather than in the low-recombining pericentromeric regions. By plotting CAAs identified for eleven most significant climatic and bioclimatic variables along the chromosomes, we show that differentiated CAAs are enriched in the high-recombing regions of chromosomes (Figs. 4C and 4D). These results suggest that 1) selection of lines in the phenotypic extremes of agronomic and physiological traits prioritizes those that carry CAAs located within the high-recombining regions of the genome likely due to the reduced linkage to deleterious alleles, and 2) introgressed CAAs are associated with variation in phenotypic traits linked with wheat performance in both irrigated and water-limiting conditions.

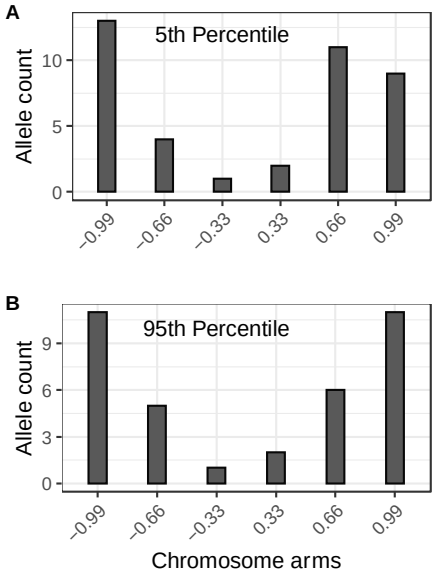


Fig. 4 Chromosome distribution of CAAs differentiated between the tails of phenotypic extremes in introgression population. The CAAs detected for most significant climatic and bioclimatic variables were included into the analyses. Each chromosome arm was split into three regions with each region representing 33.3% of arm length. The counts of CAAs in each region across all chromosomes in the wheat genome was combined. The A and B panels show chromosome distributions for differentiated CAAs in the 5^{th} and 95^{th} percentile tails of phenotype distribution, respectively.

CAAs are linked with variation in adaptive traits in the introgression population.

Association analyses were performed in the introgression population using a set of 5,675 climate-adaptive SNPs to determine variants that contribute to improved performance of introgression lines under the water-limiting and irrigated conditions. The CT, heading date, yield and yield component traits were used as plant performance metrics. Significant CAA-trait associations with heading date, spikelet number per spike, grain width and length were observed in the introgression population (Supplementary File S1). Considering the importance of CT for assessing the physiological response of plants to drought stress (Pinto and Reynolds 2015; Kumar et al., 2017), we focused on the results of association analyses between CAAs and CT.

The CT data were collected using unmanned aerial system (UAS)-based thermal imaging from both irrigated and non-irrigated field trials at multiple time points during the growing season in 2018 and 2019. Variation in CT was influenced by both genotype and environment (Table 3). In the 2018 growing season, narrow sense heritability (h^2) for CT varied between 0.54 and 0.85 whereas in 2019 it ranged from 0.24 to 0.78. In the 2019 growing season, residual variance in CT was much higher than due to genotype effect. This could be linked to the fact that in 2019, Colby experienced high precipitation and low temperature conditions during the growing season. Best linear unbiased predictors for spatially corrected CT varied from -0.57 to 1.8 suggesting that some introgression lines were able to lower CT compared to others that had higher CT (Figure 5A). A comparison of yield and CT showed a strong negative relationship with CT accounting for 30% of yield variation in the *Ae. tauschii* introgression population (Fig. 5B). This result was confirmed by performing phenomic predictions using the random forest model with CT and yield component traits as predictors of yield (Fig. S2). These analyses showed that CT is the most significant factor for predicting yield followed by grain length and thousand grain weight in this population, consistent with previous observations (Wardlaw et al. 1989).

To further understand the impact of introgressed alleles from *Ae. tauschii* into hexaploid wheat background on yield, we identified ILs in the 5th and 95th percentiles of CT distribution and compared them to recurrent parents. The average yield for ILs in the 5th percentile of CT was 50 bpa, which was higher but not significantly different from the recurrent parents (48 bpa, P = 0.825). The lack of significant difference could be attributed to high variation in yield in ILs

605 showing low CT, suggesting that other genetic factors could contribute to final yield. The
606 introgression lines in the 95th percentile however, showed significant yield reduction (39 bpa)
607 relative to the recurrent parents (Tukey HSD, $P < 0.008$), confirming the importance of CT trait for
608 predicting grain yield in wheat.

609 **Table 3:** Effect of genotype and environment on canopy temperature variation and its heritability
610 in the *Ae. tauschii* introgression population.

Year	Flight date	Experiment	varG	varE	h^2	outliers	r^2	cv
2018	20180511	Irrigated	0.129	0.027	0.72	0	0.996	0.59
		Rainfed	0.342	0.101	0.69	1	0.977	1.1
	20180525	Irrigated	2.025	0.113	0.85	0	0.997	0.44
		Rainfed	2.552	0.608	0.73	1	0.966	1.17
	20180531	Irrigated	0.451	0.132	0.71	1	0.971	0.76
		Rainfed	0.442	0.289	0.54	2	0.895	1.26
2019	20190513	Irrigated	0.298	0.036	0.75	1	0.995	0.34
		Rainfed	0.061	0.083	0.37	2	0.967	1.04
	20190522	Irrigated	0.059	0.051	0.38	3	0.998	1.54
		Rainfed	0.084	0.051	0.53	2	0.996	1.06
	20190601	Irrigated	0.572	0.051	0.78	1	0.996	0.38
		Rainfed	0.077	0.221	0.24	5	0.837	1.78

20190607	Irrigated	0.612	0.082	0.76	1	0.99	0.51
	Rainfed*	0.002	0.166	0.01	4	0.793	1.49

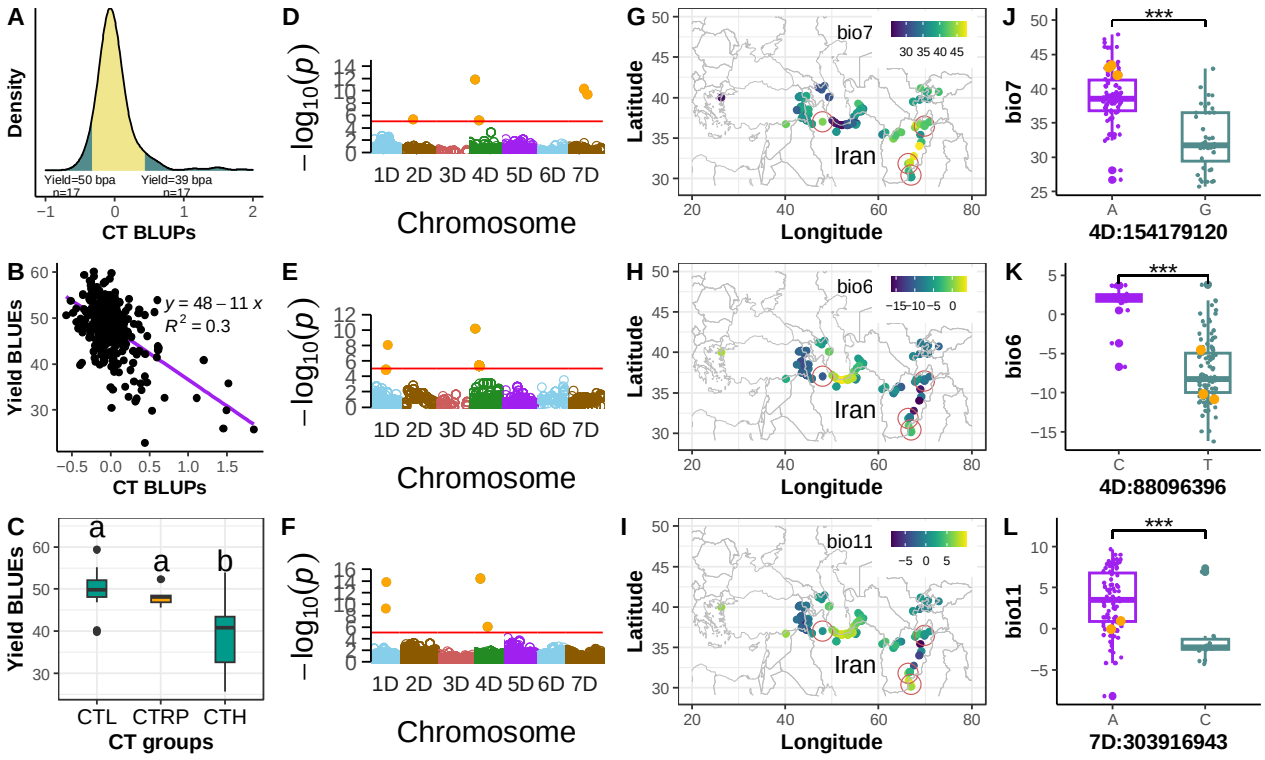
*Excluded from downstream analysis

To identify the genomic loci associated with CT depression, we performed GWAS using genotypes at CAAs (Segura et al., 2012; Wang and Zhang 2021). The multiple-locus mixed linear model (MLMM) and Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) revealed significant SNP-trait associations on chromosomes 1D, 2D, 4D and 7D (Fig. 5D, Table S7). Some of the most significant SNPs associated with CT on chromosomes 1D and 4D (chr1D_265243957 and chr4D_154179120) showed the highest correlation with temperature annual range (bio7, r = 0.59). Based on the RDA analyses, bio7 was identified as the most significant variable shaping SNP variation in *Ae. tauschii*. Other SNPs significantly associated with CT in introgression population correlated with the minimum temperature in the coldest month (bio6, r = 0.67), mean temperature in the coldest quarter (bio11, r = 0.6), longitude and precipitation in the driest months (August to October) (Table S7).

Besides using only SNP sites with CAAs for GWAS, 5.3 million SNPs from *Ae. tauschii* introgression population were pruned based on LD resulting in 99,529 SNPs with $r^2 \leq 0.5$. When GWAS was performed using this set of SNPs, the MLMM model revealed three QTLs that were associated with CT, including one on chromosome 1D and two on 4D. The significant SNPs on 1D were chr1D_254646871 and chr1D_265399733 (Fig 5F, Table S7). Although these SNPs were not part of the SNP set detected in the environmental scans of *Ae. tauschii* accessions, they were located within the same genomic intervals identified by genome-wide association mapping in the introgression population using the climate-associated SNPs. Within the interval 254 – 266 Mb, there were 191 CAA SNPs (Table S3) that highly correlated with mean diurnal range (bio2), temperature annual range (bio7), mean temperature of driest quarter (bio9), minimum temperature in March (tmin_3), precipitation in September (prec_9) and longitude (lon). Similarly, the first QTL on 4D contains SNP chr4D_92689640 (Table S7), and in the interval 90 – 94 Mb on 4D, four CAA SNPs identified in *Ae. tauschii* accessions were strongly correlated with bio7. The second QTL contains SNP chr4D_229986737 (Table S7), and a search for CAA SNPs 5 Mb to the left and

637 right of the significant QTN identified 31 SNPs that were correlated with prec_9, tmin3 and lon
638 variables.

639 The source of alleles lowering CT in the introgression lines were mostly from *Ae. tauschii*
640 ssp. *tauschii* accessions (TA2388, TA2536, TA2521 and TA10177), collected from areas such as
641 Afghanistan, Iran and Pakistan known for high bio7 (43), on average with nearly no precipitation
642 in the driest quarter of the year (Figs. 5G, 5J). These results suggest that, *Ae. tauschii* growing in
643 high temperature and low precipitation conditions could improve wheat adaptation to water-
644 limiting conditions when introgressed into adapted wheat background.



645

646

647 **Fig. 5.** Relationship between **c**anopy temperature (CT) and yield performance of the introgression
648 lines, the genomic loci associated with CT and origin of *Ae. tauschii* providing CT lowering alleles
649 in hexaploid wheat background. (A) CT distribution for the introgression population at Colby in
650 2018 and 2019 growing seasons. The shaded tails represent the 5th and 95th percentiles. (B)
651 Regression of yield on CT, (C) Yield of ILs showing low CT (CTL) and high CT (CTH) relative to
652 recurrent parents CT (CTRP). Different letters on top of the box plots indicate significant

differences at 95% confidence level, (D-F) Manhattan plots showing the quantitative trait nucleotides (QTN) associated with CT. Where D and E are based CAA SNPs with MLMM and BLINK models, respectively and F is based LD pruned SNPs with MLMM. The red line in the Manhattan plot indicates the P-value corresponding to a threshold FDR 0.05. (F-H) Heatpoint maps showing temperature annual range (bio7), min temperature in the coldest month (bio6) and mean temperature in coldest quarter (bio11) in the geographic origin of *Ae. tauschii* accessions. Accessions in red circles are the parents for introgression lines with low CT. (I-K) Distribution of bioclimatic variables (bio7, bio6 and bio11) in geographical origin of 137 *Ae. tauschii* accessions. Orange dots represent *Ae. tauschii* accessions used to generate introgression lines that rank in the 5th percentile for CT distribution.

Discussion

The lineages of *Ae. tauschii* are spread over a large geographic area with a wide range of variation in climatic factors, including some of the locations with extremely dry and hot environments (Dvorak *et al.* 1998; Wang *et al.* 2013; Gaurav *et al.* 2022). The existence of strong SNP-climate correlations reported here provides effective means for detecting climate adaptive variants in diploid *Ae. tauschii* using environmental genome scans. The range of geographic distribution for adaptive variants varied broadly with some alleles showing narrow geographic distribution, and other alleles showing broad distribution across large geographic areas. Consistent with prior studies, these spatial patterns of allele distribution can primarily be explained by the distribution of climatic factors (Hancock et al., 2011) (Lasky *et al.* 2012, 2015). In our analyses the environmental and bio-climatic factors alone accounted for a substantial proportion (57.8%) of spatial genetic variation in *Ae. tauschii* with relatively small contribution from geographic dispersal (5.6%). These results suggest consistent environmental gradients across the *Ae. tauschii* distribution range likely shaped the spatial structure of genomic variation in this wild ancestor of wheat and contributed to genetic differentiation between its main lineages.

Correlations between genomic diversity and climatic factors indicate that the temperature and precipitation gradients during the growth periods coinciding with flowering, grain filling and maturation contributed to genetic differentiation among the four main lineages of *Ae. tauschii* and

its two subspecies, *strangulata* and *tauschii*. The lowest precipitation in driest month was one of the major differentiating ecogeographic factors for L1E, whereas temperature and precipitation gradients in April and May contributed most to the genetic differentiation of L1W lineage. The variation in precipitation and temperature in April and May were among the main ecogeographic factors explaining differentiation between the L2W and L2E lineages of *Ae. tauschii* ssp. *strangulata*. The lineage L2E of *Ae. tauschii* ssp. *strangulata*, which contributed 10,000 years ago to the origin of bread wheat (Wang et al., 2013; Luo et al., 2017), grows in a narrow geographic region south of the Caspian Sea with limited variation in climatic factors characteric of the humid mild subtropical environments. As a result, adaptive diversity captured by the D genome of bread wheat is primarily restricted to those alleles that are represented in this region. The limited levels of gene flow detected between wheat and *Ae. tauschii* ssp. *strangulata* did not have dramatic impact on the genetic diversity of the D genome (Wang *et al.* 2013; He *et al.* 2019; Zhou *et al.* 2020; Gaurav *et al.* 2022). Thus, the polyploidization bottleneck associated with wheat origin resulted in not only the overall loss of genetic diversity in the wheat D genome (He et al., 2019; Gaurav et al., 2022) but also in the massive loss of adaptive alleles represented in all four sublineages of *Ae. tauschii*. While the consequences of the loss of these alleles in wheat are hard to predict, we might expect that it had a negative impact on the adaptive potential of hexaploid wheat and offset progress with development of drought-resilient wheat varieties.

Introgression from *Ae. tauschii* into hexaploid wheat had positive effects on traits playing an important role in increasing crop productivity and improving adaptation to drought. In our previous study, we showed that 3.2% of introgression lines carrying *Ae. tauschii* haplotypes outperformed parental lines in drought trials (Nyine et al., 2021). Consistent with these results, several high-yielding drought tolerant cultivars have been derived from synthetic wheat lines created using *Ae. tauschii* as one of the parents (Rosyara et al. 2019; Molero et al., 2023; Pinto and Reynolds 2015). Our analyses suggest that improved performance of wheat introgression lines could be largely attributed to introduction of climate-adaptive alleles that show association with environmental variation in *Ae. tauschii*. Statistically significant shifts in allele frequency in the extreme tails of phenotypic trait distributions and significant associations detected for climate-adaptive alleles in GWAS for canopy temperature and productivity traits support this conclusion.

The signatures of adaptation detected by environmental scans in the genome could be driven by complex historic gradients of environments and associated with diverse adaptive mechanisms (Lasky *et al.* 2012; Anderson and Song 2020). As a result, it is difficult to establish relationships between adaptive alleles from environmental scans and specific phenotypic traits measured for introgression populations in field trials. The temperature annual range (bio7) was among the main bio-climatic factors that contributed most to shaping the spatial genomic variation in *Ae. tauschii*. The SNP locus located on chromosome 4D showing strongest association with bio7 in environmental scans also showed strongest association with variation in canopy temperature in introgression lines. This result indicates that among the targets of selection imposed by variation in bio7 are variants associated with pathways controlling physiological processes responsible for maintaining canopy temperature under drought stress (Jackson *et al.* 1981). In the field trails, lines carrying *Ae. tauschii* alleles at loci associated with reduced canopy temperature were among the top yielding introgression lines suggesting that these *Ae. tauschii* alleles improve adaptation to drought stress and likely act as the main drivers of increased yield. Likewise, detection of *Ae.tauschii* introgression into chromosome 6D associated with reduction in canopy temperature and increased yield under dry conditions, confirms the importance of this adaptive mechanism for drought tolerance (Molero et al., 2023) (Still *et al.* 2021). These results indicate that environmental scans focusing on the relevant bio-climatic variables are an effective means for uncovering variants in wild relatives to improve wheat adaptation to water-limiting conditions and increase its yield potential.

Our study shows that whole genome sequencing of diverse collections of wild relatives integrated with environmental scans could provide an effective strategy for prioritizing wild relatives from germplasm banks for introgression into wheat. Continued reduction in the cost of genome sequencing and availability of reference genomes for the increasing number of wild relative species makes this strategy an attractive option for even large genebank collections including tens of thousands of lines (Mascher *et al.* 2019; Bohra *et al.* 2022). These approaches could quickly help to detect accessions enriched for alleles providing adaptation to target environments (Brunazzi *et al.* 2018; Anderson and Song 2020). As it was shown in our study, genome-wide introgression of prioritized diversity into adapted germplasm followed by fast high-throughput phenotyping using UAS-based imaging platforms could help to quickly identify promising germplasm for improving the adaptive potential of wheat. Expansion of these efforts

742 from direct ancestors of bread wheat to include more distant *Aegilops* and *Triticum* species have
743 potential to further broaden adaptive diversity accessible to wheat breeders for climate-proofing
744 food production systems.

745

746 **Acknowledgements**

747 This research was supported by the Agriculture and Food Research Initiative Competitive Grants
748 2022-68013-36439 (WheatCAP) and 2020-68013-30905 (USDA NIFA WWBI Hub) from the
749 USDA National Institute of Food and Agriculture.

750

751 **Author contribution**

752 MN - data collection, analysis and interpretation, manuscript writing; DD - field data collection,
753 UAS based phenotyping; EAd - population development, data analysis and field data collection;
754 MC - population development and field data collection; HW - UAS imaging data analysis; AA -
755 next-generation sequencing; AF - population development, data collection, experimental design;
756 EAk - data analysis and interpretation, conceptualization, conceiving idea, manuscript writing.

757

758 **References**

759 Amani, I., Fischer, R.A. and Reynolds, M.P., 1996. Canopy temperature depression association
760 with yield of irrigated spring wheat cultivars in a hot climate. Journal of Agronomy and
761 Crop Science, 176(2), pp.119-129.

762 Balfourier, F., Bouchet, S., Robert, S., De Oliveira, R., Rimbert, H., Kitt, J., Choulet, F.,
763 International Wheat Genome Sequencing Consortium, BreedWheat Consortium and Paux,
764 E., 2019. Worldwide phylogeography and history of wheat genetic diversity. Science
765 advances, 5(5), p.eaav0536.

766 Blanchet, F.G., Legendre, P. and Borcard, D., 2008. Forward selection of explanatory variables.
767 Ecology, 89(9), pp.2623-2632.

768 Börner, A., Ogbonnaya, F.C., Röder, M.S., Rasheed, A., Periyannan, S., Lagudah, E.S. 2015.
769 Aegilops tauschii Introgressions in Wheat. In: Molnár-Láng, M., Ceoloni, C., Doležel, J.
770 (eds) Alien Introgression in Wheat. Springer, Cham. https://doi.org/10.1007/978-3-319-
771 23494-6_10.

772 Caye, K., Deist, T.M., Martins, H., Michel, O. and François, O., 2016. TESS3: fast inference of
773 spatial population structure and genome scans for selection. Molecular Ecology Resources,
774 16(2), pp.540-548.

775 Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and
776 Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide
777 polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-
778 2; iso-3. fly, 6(2), pp.80-92.

779 Chang, C.W., Fridman, E., Mascher, M., Himmelbach, A. and Schmid, K., 2022. Physical
780 geography, isolation by distance and environmental variables shape genomic variation of
781 wild barley (*Hordeum vulgare* L. ssp. *spontaneum*) in the Southern Levant. Heredity,
782 128(2), pp.107-119.

783 Frachon, L., Bartoli, C., Carrère, S., Bouchez, O., Chaubet, A., Gautier, M., Roby, D. and Roux,
784 F., 2018. A genomic map of climate adaptation in Arabidopsis thaliana at a micro-
785 geographic scale. Frontiers in plant science, 9, p.967.

786 Frankel, O.H., 1984. Genetic perspectives of germplasm conservation. In: W. K. Arber, K.
787 Llimensee, W. J. Peacock, and P. Starlinger, Eds., Genetic manipulation: impact on man
788 and society, Cambridge University Press, Cambridge, 161, p.170.

789 Gaurav, K., Arora, S., Silva, P., Sánchez-Martín, J., Horsnell, R., Gao, L., Brar, G.S., Widrig, V.,
790 John Raupp, W., Singh, N. and Wu, S., 2022. Population genomic analysis of Aegilops
791 tauschii identifies targets for bread wheat improvement. Nature biotechnology, 40(3),
792 pp.422-431.

Guo, Z., Chen, D., Röder, M.S., Ganal, M.W. and Schnurbusch, T., 2018. Genetic dissection of pre□anthesis sub□phase durations during the reproductive spike development of wheat. The Plant Journal, 95(5), pp.909-918.

Hancock, A.M., Brachi, B., Faure, N., Horton, M.W., Jarymowycz, L.B., Sperone, F.G., Toomajian, C., Roux, F. and Bergelson, J., 2011. Adaptation to climate across the Arabidopsis thaliana genome. Science, 334(6052), pp.83-86.

Hao, M., Zhang, L., Ning, S., Huang, L., Yuan, Z., Wu, B., Yan, Z., Dai, S., Jiang, B., Zheng, Y. and Liu, D., 2020. The resurgence of introgression breeding, as exemplified in wheat improvement. Frontiers in plant science, 11, p.252.

He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P., Forrest, K., Fritz, A., Hucl, P., Wiebe, K. and Knox, R., 2019. Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. Nature Genetics, 51(5), pp.896-904.

Hunt Jr, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S., Perry, E.M. and Akhmedov, B., 2013. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. International journal of applied earth observation and Geoinformation, 21, pp.103-112.

International Wheat Genome Sequencing Consortium, 2018. Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. Science, 361.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature, 484(7392), pp.55-61.

Kumar, M., Govindasamy, V., Rane, J., Singh, A.K., Choudhary, R.L., Raina, S.K., George, P., Aher, L.K. and Singh, N.P., 2017. Canopy temperature depression (CTD) and canopy greenness associated with variation in seed yield of soybean genotypes grown in semi-arid environment. South African Journal of Botany, 113, pp.230-238.

820 Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z., Fraser, M., Anderson, J.A., Brown-
821 Guedira, G., DeWitt, N. and Peters Haugrud, A., 2019. Identification of a candidate gene for
822 a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution
823 genetic mapping. Theoretical and Applied Genetics, 132, pp.2689-2705.

824 Lasky, J.R., Upadhyaya, H.D., Ramu, P., Deshpande, S., Hash, C.T., Bonnette, J., Juenger, T.E.,
825 Hyma, K., Acharya, C., Mitchell, S.E. and Buckler, E.S., 2015. Genome-environment
826 associations in sorghum landraces predict adaptive traits. Science advances, 1(6),
827 p.e1400218.

828 Li, Y., Cao, K., Li, N., Zhu, G., Fang, W., Chen, C., Wang, X., Guo, J., Wang, Q., Ding, T. and
829 Wang, J., 2021. Genomic analyses provide insights into peach local adaptation and
830 responses to climate change. Genome research, 31(4), pp.592-606.

831 Lipka, A.E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P.J., Gore, M.A., Buckler, E.S. and
832 Zhang, Z., 2012. GAPIT: genome association and prediction integrated tool. Bioinformatics,
833 28(18), pp.2397-2399.

834 Lopes, M.S. and Reynolds, M.P., 2012. Stay-green in spring wheat can be determined by spectral
835 reflectance measurements (normalized difference vegetation index) independently from
836 phenology. Journal of experimental botany, 63(10), pp.3789-3798.

837 Luo, M.C., Gu, Y.Q., Puiu, D., Wang, H., Twardziok, S.O., Deal, K.R., Huo, N., Zhu, T., Wang,
838 L., Wang, Y. and McGuire, P.E., 2017. Genome sequence of the progenitor of the wheat D
839 genome Aegilops tauschii. Nature, 551(7681), pp.498-502.

840 Martins, K., Gugger, P.F., Llanderal□Mendoza, J., González□Rodríguez, A., Fitz□Gibbon, S.T.,
841 Zhao, J.L., Rodríguez□Correa, H., Oyama, K. and Sork, V.L., 2018. Landscape genomics
842 provides evidence of climate□associated genetic variation in Mexican populations of
843 Quercus rugosa. Evolutionary Applications, 11(10), pp.1842-1858.

844 McArdle, B.H. and Anderson, M.J., 2001. Fitting multivariate models to community data: a
845 comment on distance□based redundancy analysis. Ecology, 82(1), pp.290-297.

Molero, G., Coombes, B., Joynson, R., Pinto, F., Piñera-Chávez, F.J., Rivera-Amado, C., Hall, A. and Reynolds, M.P., 2023. Exotic alleles contribute to heat tolerance in wheat under field conditions. Communications Biology, 6(1), p.21.

Nyine, M., Adhikari, E., Clinesmith, M., Jordan, K.W., Fritz, A.K. and Akhunov, E., 2020. Genomic patterns of introgression in interspecific populations created by crossing wheat with its wild relative. G3: Genes, Genomes, Genetics, 10(10), pp.3651-3661.

Nyine, M., Adhikari, E., Clinesmith, M., Aiken, R., Betzen, B., Wang, W., Davidson, D., Yu, Z., Guo, Y., He, F. and Akhunova, A., 2021. The haplotype-based analysis of Aegilops tauschii introgression into hard red winter wheat and its impact on productivity traits. Frontiers in Plant Science, 12, p.716955.

Ortiz-Bobea, A., Wang, H., Carrillo, C.M. and Ault, T.R., 2019. Unpacking the climatic drivers of US agricultural yields. Environmental Research Letters, 14(6), p.064003.

Pestsova, E.G., Börner, A. and Röder, M.S., 2001. Development of a set of Triticum aestivum□Aegilops tauschii introgression lines. Hereditas, 135(2□3), pp.139-143.

Pinter Jr, P.J., Zipoli, G., Reginato, R.J., Jackson, R.D., Idso, S.B. and Hohman, J.P., 1990. Canopy temperature as an indicator of differential water use and yield performance among wheat cultivars. Agricultural Water Management, 18(1), pp.35-48.

Pinto, R.S. and Reynolds, M.P., 2015. Common genetic basis for canopy temperature depression under heat and drought stress associated with optimized root distribution in bread wheat. Theoretical and Applied Genetics, 128, pp.575-585.

Rawson, H.M., 1970. Spikelet number, its control and relation to yield per ear in wheat. Australian Journal of Biological Sciences, 23(1), pp.1-16.

Rosyara, U., Kishii, M., Payne, T., Sansaloni, C.P., Singh, R.P. and Braun, H.J., 2019. Genetic contribution of synthetic hexaploid wheat to CIMMYT's spring bread wheat breeding germplasm. Sci Rep. 2019.

871 Schlenker, W. and Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to
872 US crop yields under climate change. Proceedings of the National Academy of sciences,
873 106(37), pp.15594-15598.

874 Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q. and Nordborg, M., 2012.
875 An efficient multi-locus mixed-model approach for genome-wide association studies in
876 structured populations. Nature genetics, 44(7), pp.825-830.

877 Singh, N., Wu, S., Tiwari, V., Sehgal, S., Raupp, J., Wilson, D., Abbasov, M., Gill, B. and
878 Poland, J., 2019. Genomic analysis confirms population structure and identifies inter-
879 lineage hybrids in Aegilops tauschii. Frontiers in Plant Science, 10, p.9.

880 Smith, O., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., Gaffney, V. and Allaby,
881 R.G., 2015. Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000
882 years ago. Science, 347(6225), pp.998-1001.

883 Shewry, P.R., 2009. Wheat. Journal of experimental botany, 60(6), pp.1537-1553.

884 Tack, J., Barkley, A. and Nalley, L.L., 2015. Effect of warming temperatures on US wheat yields.
885 Proceedings of the National Academy of Sciences, 112(22), pp.6931-6936.

886 Turner, T.L., Bourne, E.C., Von Wettberg, E.J., Hu, T.T. and Nuzhdin, S.V., 2010. Population
887 resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. Nature
888 genetics, 42(3), pp.260-263.

889 Van Den Wollenberg, A.L., 1977. Redundancy analysis an alternative for canonical correlation
890 analysis. Psychometrika, 42(2), pp.207-219.

891 Wallace, J.G., Zhang, X., Beyene, Y., Semagn, K., Olsen, M., Prasanna, B.M. and Buckler, E.S.,
892 2016. Genome□wide association for plant height and flowering time across 15 tropical
893 maize populations under managed drought stress and well□watered conditions in
894 Sub□Saharan Africa. Crop Science, 56(5), pp.2365-2378.

895 Wang, J., Luo, M.C., Chen, Z., You, F.M., Wei, Y., Zheng, Y. and Dvorak, J., 2013. Aegilops
896 tauschii single nucleotide polymorphisms shed light on the origins of wheat D□genome

genetic diversity and pinpoint the geographic origin of hexaploid wheat. New phytologist, 198(3), pp.925-937.

Wang, J. and Zhang, Z., 2021. GAPIT version 3: boosting power and accuracy for genomic association and prediction. Genomics, proteomics & bioinformatics, 19(4), pp.629-640.

Wardlaw, I.F., Dawson, I.A., Munibi, P. and Fewster, R., 1989. The tolerance of wheat to high temperatures during reproductive growth. I. Survey procedures and general response patterns. Australian Journal of Agricultural Research, 40(1), pp.1-13.

Zabel, F., Müller, C., Elliott, J., Minoli, S., Jägermeyr, J., Schneider, J.M., Franke, J.A., Moyer, E., Dury, M., Francois, L. and Folberth, C., 2021. Large potential for crop production adaptation depends on available future varieties. Global Change Biology, 27(16), pp.3870-3882.

Zhang, J., Gizaw, S.A., Bossolini, E., Hegarty, J., Howell, T., Carter, A.H., Akhunov, E. and Dubcovsky, J., 2018. Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. Theoretical and Applied Genetics, 131, pp.1741-1759.

Supporting Information

Fig. S1 The effect of geographic, climatic and bioclimatic variables on the SNP diversity in *Ae. tauschii* population.

Fig. S2 Importance of canopy temperature and yield component traits for yield prediction based on random forest model in *Ae. tauschii*-Wheat introgression population.

Table S1 Average historical climatic and bioclimatic data from the geographic origin of the diverse *Ae. tauschii*.

Table S2 Number of climate associated SNPs for geographic climatic and bioclimatic variables discovered through redundancy analysis and genome-wide association analysis in *Ae. tauschii* population.

Table S3 SNPs showing the highest correlation with geographic, climatic and bioclimatic variables in *Ae. tauschii* population.

Table S4 Effect of SNPs that are significantly associated with geographic, climatic and bioclimatic variables in *Ae. tauschii* population based on genome-wide association analysis.

Table S5 SNPs identified by redundancy analysis (RDA) and/or genome-wide association (GWAS) to be involved in *Ae. tauschii* ecogeographic adaptation.

Table S6 High confidence climate associated SNPs identified by both redundancy analysis (RDA) and genome-wide assocition (GWAS) analysis in *Ae. tauschii* population.

Table S7 SNPs significantly associated with canopy temperature in Ae. tauschii introgression population.

File S1 Genome-wide association mapping of CAAs for adaptive traits in *Ae. tauschii*-wheat introgression population.

Additional References.

Anderson, J. T., and B. H. Song, 2020 Plant adaptation to climate change—Where are we? J. Syst. Evol. 58: 533–545.

Araus, J. L., J. P. Ferrio, R. Buxó, and J. Voltas, 2007 The historical perspective of dryland agriculture: lessons learned from 10,000 years of wheat cultivation. J. Exp. Bot. 58: 131–45.

Avni, R., M. Nave, O. Barad, K. Baruch, S. O. Twardziok *et al.*, 2017 Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science (80-.). 357:.

Balfourier, F., S. Bouchet, S. Robert, R. DeOliveira, H. Rimbert *et al.*, 2019 Worldwide phylogeography and history of wheat genetic diversity. Sci. Adv. 5: eaav0536.

Bari, A., K. Street, M. Mackay, D. T. F. Endresen, E. de Pauw *et al.*, 2012 Focused identification

947 of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental
948 variables. Genet. Resour. Crop Evol. 59: 1465–1481.

949 Bohra, A., B. Kilian, S. Sivasankar, M. Caccamo, C. Mba *et al.*, 2022 Reap the crop wild relatives
950 for breeding future crops. Trends Biotechnol. 40: 412–431.

951 Brunazzi, A., D. Scaglione, R. F. Talini, M. Miculan, F. Magni *et al.*, 2018 Molecular diversity
952 and landscape genomics of the crop wild relative Triticum urartu across the Fertile Crescent.
953 Plant J. 94: 670–684.

954 Dvorak, J., M. C. Luo, Z. L. Yang, and H. B. Zhang, 1998 The structure of the Aegilops tauschii
955 genepool and the evolution of hexaploid wheat. Theor. Appl. Genet. 97: 657–670.

956 Exposito-Alonso, M., H. A. Burbano, O. Bossdorf, R. Nielsen, and D. Weigel, 2019 Natural
957 selection on the Arabidopsis thaliana genome in present and future climates. Nature.

958 Gaurav, K., S. Arora, P. Silva, J. Sánchez-Martín, R. Horsnell *et al.*, 2022 Population genomic
959 analysis of Aegilops tauschii identifies targets for bread wheat improvement. Nat. Biotechnol.
960 40: 422–431.

961 Gill, B. S., B. Friebe, W. J. Raupp, D. L. Wilson, T. S. Cox *et al.*, 2006 Wheat Genetics Resource
962 Center: The First 25 Years. Adv. Agron. 89: 73–136.

963 He, F., R. Pasam, F. Shi, S. Kant, G. Keeble-Gagnere *et al.*, 2019 Exome sequencing highlights the
964 role of wild relative introgression in shaping the adaptive landscape of the wheat genome.
965 Nat. Genet. 51: 896–904.

966 Jackson, R. D., S. B. Idso, R. J. Reginato, and P. J. Pinter, 1981 Canopy temperature as a crop
967 water stress indicator. Water Resour. Res. 17: 1133–1138.

968 Kihara, H., 1944 Discovery of the DD-analyser, one of the ancestors of Triticum vulgare.
969 Agricuture Hortic. 19: 889–890.

970 Kishii, M., 2019 An Update of Recent Use of Aegilops Species in Wheat Breeding. Front. Plant
971 Sci. 10:.

Lasky, J. R., D. L. Des Marais, J. K. McKay, J. H. Richards, T. E. Juenger *et al.*, 2012 Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. Mol. Ecol. 21: 5512–29.

Lasky, J. R., H. D. Upadhyaya, P. Ramu, S. Deshpande, C. T. Hash *et al.*, 2015 Genome-environment associations in sorghum landraces predict adaptive traits. Sci. Adv. 1: e1400218–e1400218.

Luo, M.-C., Z.-L. Yang, F. M. You, T. Kawahara, J. G. Waines *et al.*, 2007 The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor. Appl. Genet. 114: 947–59.

Mascher, M., M. Schreiber, U. Scholz, A. Graner, J. C. Reif *et al.*, 2019 Genebank genomics bridges the gap between the conservation of crop diversity and plant breeding. Nat. Genet. 51: 1076–1081.

Nesbitt, M., and D. Samuel, 1996 From staple crop to extinction? The archaeology and history of hulled wheats., pp. 41–100 in *International Workshop on Hulled Wheats*, edited by S. Padulosi, K. Hammer, and J. Heller. Italy International Plant Genetic Resources Institute, Rome.

Nyine, M., E. Adhikari, M. Clinesmith, K. W. Jordan, A. K. Fritz *et al.*, 2020 Genomic patterns of introgression in interspecific populations created by crossing wheat with its wild relative. G3 Genes, Genomes, Genet. 10: 3651–3661.

Ozkan, H., G. Willcox, A. Graner, F. Salamini, and B. Kilian, 2011 Geographic distribution and domestication of wild emmer wheat (Triticum dicoccoides). Genet. Resour. Crop Evol. 58: 11–53.

Sharma, S., A. W. Schulthess, F. M. Bassi, E. D. Badaeva, K. Neumann *et al.*, 2021 *Introducing beneficial alleles from plant genetic resources into the wheat germplasm*.

Sohail, Q., T. Inoue, H. Tanaka, A. E. Eltayeb, Y. Matsuoka *et al.*, 2011 Applicability of Aegilops tauschii drought tolerance traits to breeding of hexaploid wheat. Breed. Sci. 61: 347–57.

998 Still, C. J., B. Rastogi, G. F. M. Page, D. M. Griffith, A. Sibley *et al.*, 2021 Imaging canopy
999 temperature: shedding (thermal) light on ecosystem processes. New Phytol. 230: 1746–1753.

1000 Tanno, K.-I., and G. Willcox, 2006 How fast was wild wheat domesticated? Science 311: 1886.

1001 Wang, J., M.-C. Luo, Z. Chen, F. M. You, Y. Wei *et al.*, 2013 Aegilops tauschii single nucleotide
1002 polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint
1003 the geographic origin of hexaploid wheat. New Phytol. 198: 925–937.

1004 Zhao, X., Y. Guo, L. Kang, C. Yin, A. Bi *et al.*, 2023 Population genomics unravels the Holocene
1005 history of bread wheat and its relatives. Nat. Plants 9: 403–419.

1006 Zhou, Y., X. Zhao, Y. Li, J. Xu, A. Bi *et al.*, 2020 Triticum population sequencing provides
1007 insights into wheat adaptation. Nat. Genet. 52: 1412–1422.

1008