scientific reports

Check for updates

OPEN

Genome-wide association study identifies candidate genes for agronomic traits under heat- and combined recurrent drought-and heat-stress in wheat

Amare KebedeB[1]✉, Wassu Mohammed[1], Gemechu Keneni[2], Jones Christopher Stephen[3], Alok Kumar[1], Khaled Al-Sham'aa[4] & Wuletaw Tadesse[4]

Combined heat and drought stress, mainly during sensitive growth stages, considerably limits wheat yields. Herein, we report on an experiment conducted to analyze and predict the phenology, architecture, and kernel yield component (PAKyC) traits of a diverse bread wheat population grown under heat stress and fully irrigated conditions (HSFIC), and combined recurrent drought and heat stress (CRDHS). 17,711 high-quality single-nucleotide polymorphisms (SNPs) were used to perform a GWAS for 13 PAKyC traits on 187 genotypes, a subset of 234 genotypes, using the BLINK model. Nine constitutive SNPs, 24 marker-trait associations (MTAs), broad-effect pleiotropic SNPs, conditional pleiotropic SNPs, adaptive pleiotropic SNPs under CRDHS, and two robust SNPs were identified. Furthermore, 35 candidate genes were associated with cellular components, biological processes, and molecular functions under CRDHS and HSFIC. Among these genes, 12 were highly expressed (> 0.5 TPM) in response to drought, heat, and combined stresses. Nine genes were identified to be transcription factors regulating drought and heat responses, while the functions of the other three genes remain unclear. Overall, HSFIC and CRDHS environments enabled the dissection of heritable and strongly correlated traits, MTAs, and pleiotropy types: constitutive, adaptive, and conditional. They also helped identify robust SNPs and candidate genes that are important for bread wheat breeding.

Keywords Blink, Marker-trait associations, SNPs, *Triticum aestivum*

Bread wheat (*Triticum aestivum* L., AABBDD genomes, 2n = 6x = 42) is a major global food crop that provides starch, protein, minerals, and vitamins[1]. The demand for wheat is increasing as the world population grows. However, wheat yield growth is low, mainly due to drought and heat stresses. Global warming in a changing climate amplifies the impacts of these abiotic stresses, contributing to food insecurity; for example, up to 8.8% of the World's population faced hunger in 2024, compromising the goal of hunger eradication by 2030[2].

Erratic rainfall patterns, rising atmospheric temperatures and CO_2 concentrations, and hot, dry winds are the major causes of drought stress, which impacts crop architecture, phenology, and biochemical activities, leading to reduced yields[3,4]. Compared with other natural hazards, drought is the primary factor affecting wheat yields, leading to an average kernel yield reduction of 27%[5–7], and it may reach up to 58–92%, depending on the plant's growth stage (flowering and grain-filling stages), duration, and severity of drought[8]. The optimal and maximum temperatures for wheat growth during root and shoot formation, anthesis, and leaf initiation range from 16–26 to 20–26 °C, respectively[9]. Plant architecture, photosynthesis, reproductive efficiency, and carbohydrate partitioning to the kernel determine kernel yield. Temperatures exceeding 28 °C during the anthesis and post-anthesis phases of wheat development can substantially reduce kernel weight[10]. For example, high temperatures can adversely affect wheat pollen, causing collapse, desiccation, deeply pitted structures, rough exine walls, and loss of columellae heads, leading to a reduced number of viable pollen grains. It may also impact the pistil, leading to desiccation of the style, stigma, and ovary, as well as a reduced number of viable pollen grains. These

[1]Genetics and Plant Breeding, School of Plant Sciences, Haramaya University, Haramaya, Ethiopia. [2]EIAR (Ethiopian Institute of Agricultural Research), Addis Ababa, Ethiopia. [3]ILRI (International Livestock Research Institute), Nairobi, Kenya. [4]ICARDA (International Center for Agricultural Research in the Dry Areas), Rabat, Morocco. ✉email: amarekebede@gmail.com

changes ultimately affect floret fertility and kernel yield[11–13]. High temperatures shorten the duration of grain-filling and impact the starch content, accounting for nearly three-quarters of the grain dry weight[14]. Overall, wheat yields decrease by 6% for every 1 °C increase in temperature[15].

Genotype-by-environment interactions present a substantial challenge in crop cultivar development[16]. Drought and heat stress frequently occur together, exacerbating the negative effect on crop yields[17–19]. Furthermore, plants respond differently to the individual and joint effects of drought and heat stresses[20]. It is, therefore, essential to understand the environmental, morphological, physiological, and genetic foundations of complex traits to improve bread wheat yield. The other challenge is that varieties that perform well during drought at one phenological phase may not be effective when drought occurs at a different stage. For example, the jointing stage of bread wheat is more sensitive to drought than the tillering stage[21–23]. Furthermore, drought stress does not occur in every season. Thus, the examination of target traits of plants exclusively under stressed conditions may hinder the full utilization of genetic resources that perform in less stressful environments.

To date, numerous studies have identified candidate genes in bread wheat associated with drought and heat stress using various experimental approaches[24], including drought stress during the early growth stage[25]; before heading[26]; flowering stage[27,28]; and the seedling and reproductive stages[29], under contrasting water availability[30], and in combined heat and drought stress environments[31,32]. However, comprehensive studies, including the separate and joint effects of heat and recurrent drought, diverse and advanced populations that best fit target environments, high-quality dense SNP markers, drought-heat stress-sensitive growth and development stages, and the integration of conventional and advanced tools to understand plant-environment complexities, are limited.

Hence, in this study, we integrated a wide-range of resources supporting experimental precision, incorporating: 1) six heat and combined recurrent drought- and heat-stress environments; 2) drought in synchrony with three sensitive bread wheat growth stages; 3) a diverse stress panel derived from different ICARDA breeding programs along with commercial cultivars; that have never been used for HSFIC and CRDHS in synchrony with sensitive growth stages; 4) 17,711 high-quality dense SNPs, 5) 14 PAKyC traits, and 6) stringent procedures and computationally efficient models. Using these resources we aimed to (1) identify the heritability and associations of PAKyC traits; 2) dissect constitutive SNPs, and MTAs; and (3) identify the candidate genes of constitutive SNPs and their expressions under CRDHS and HSFIC.

Results

Mixed-model ANOVA for individual traits under HSFIC and CRDHS conditions

Highly significant ($P<0.001$) differences were observed between dry seasons for all PAKyC traits, except for KPS and GFP of plants grown under HSFIC and CRDHS, respectively. Moreover, PAKyC trait values were lower (0.7–38%) under CRDHS compared with HSFIC. The interaction between genotype and season was significant for all traits except TKW and KY under CRDHS, and GFP and KPS under HSFIC. The interaction between genotypes and seasons was significant for all studied traits, except for TKW and KY under CRDHS, and GFP and KPS under HSFIC. Significant variation ($P<0.001$) was also observed among genotypes for the PAKyC traits in both environments, facilitating selection for improvement in HSFIC and CRDHS environments. Furthermore, the first and third quartiles of the traits indicate that 25% and 75% of the data, respectively, were represented by specified values of a trait. The broad-sense heritability under CRDHS was high (>0.6) for DH (0.91), DPM (0.81), KY (0.78), and KPS (0.72); medium (0.3–0.6) for SL (0.60), OL (0.53), PL (0.53), PH (0.48), TKW (0.44), and GFP (0.32); and low for CGC (0.23), and FLA (0.27). Under HSFIC, all traits had medium to high heritability, with the lowest for SPS (0.37) and the highest for DH (0.95). SPS (0.10) also had low heritability under CRDHS (Table 1).

Correlations among PAKyC traits

Multivariate analysis conducted under both HSFIC and CRDHS conditions demonstrated strong correlations ($r>0.5$) between PH and SL, and between DH and DPM. Furthermore, DH and DPM had moderate associations ($r=0.3$–0.5) with CGC, PH, and SL. Whereas, TKW was negatively correlated with both DH and KPS (Fig. 1a and b).

Under CRDHS, a strong correlation was identified between PH and PL (r = 0.55). Moderate associations were also observed between KY and TKW (r = 0.41), SPS and KPS (r = 0.48), FLA and SL (r = 0.31), and CGC with SL (r = 0.33) and PH (r = 0.43). Whereas KY was negatively correlated with DH (r = − 0.41) and DPM (r = − 0.42). Similarly, TKW was negatively correlated with CGC (r = − 0.30) and DPM (r = − 0.55). GFP was also negatively correlated with DH (r = − 0.37) (Fig. 1a).

Under HSFIC, strong positive correlations were observed between CGC and PH ($r=0.55$), CGC and SL ($r=0.51$), DPM and GFP ($r=0.57$), and SPS and KPS ($r=0.51$). Moderate associations were also found between PL and OL ($r=0.32$), PL and TKW ($r=0.33$), CGC and SPS ($r=0.35$), SL and SPS ($r=0.49$), and DH and SPS ($r=0.31$). In contrast, SL and OL ($r=-0.30$), and TKY and SPS ($r=-0.32$) were negatively correlated (Fig. 1b).

Marker distribution on A, B, and D sub-genomes and LD decay

Effective high-quality SNPs were mapped to the three sub-genomes of bread wheat. The B sub-genome contained the highest number of markers, constituting 42.1% (n = 7454), followed by the A sub-genome at 36.6% (n = 6487), and the D sub-genome at 21.3% (n = 3770) (Fig. 2a). Similar distributions of markers have been reported previously, with 42.5% in the B sub-genome, 35.5% in the A sub-genome, and 22.3% in the D sub-genome[24] and generally the highest proportion in the B, A, and then D sub-genomes[25,26]. Among the 21 chromosomes, the highest number of markers were mapped to chromosomes 2B (7.7%, n = 1364), 5B (7.1%, n = 1262), 3B (6.7%, n = 1183), 5A (6.5%, n = 1152), and 1B (6.5%, n = 1148). In contrast, chromosome 4D had the lowest number of markers (0.9%, n = 161) (Fig. 2b).

CRDHS													
	PAKyC trait												
Variation source	CGC (%)	DH	DPM	GFP	FLA (cm^2)	PH (cm)	PL (cm)	SL (cm)	OL (cm)	SPS	KPS	TKW (g)	KY (tha^{-1})
Season (S)	4462.3***	14,928.0***	1887.01***	6200.1***	613.40***	35,274***	1405.57***	433.96***	298.55***	489.67***	0.386 ns	15,714.9***	53.796***
Replication	1894.7***	53.6***	40.62***	187.6***	1.45 ns	51*	219.33***	1.77 ns	3.17***	35.89**	288.27**	3.2 ns	0.007 ns
Genotype (G)	34.8***	39.3***	35.10***	12.0***	18.57***	99***	19.25***	2.34***	1.21***	5.40***	75.769***	57.8***	1.149***
G X S	25.5***	3.9***	6.63***	8.1***	27.45***	54***	9.32***	0.97***	0.57***	5.23***	17.16 ns	36.1***	0.112 ns
Mean	49.22	50.96	86.76	35.8	14.21	58	21.91	8.57	4.63	14.13	42.06	30.87	3.02
Minimum	40.22	38.2	75.42	28.45	9.7	38.45	12.9	5.34	2.61	10.68	25.62	15.25	0.82
Maximum	59.63	63.87	100.78	42.02	21.32	75.71	30.86	12.04	7.17	18.39	65.61	46.56	5.91
Q1	46.63	46.28	83.69	34.16	12.51	52.51	19.54	7.69	4.03	13.23	35.94	26.54	2.25
Q3	51.73	54.84	90.82	37.39	15.73	63.67	24.06	9.34	5.28	15.02	47.12	34.43	3.68
CV (%)	9.5	10.2	4.1	9.6	24.8	15.0	13.7	13.3	18.9	13.0	12.9	21.1	22.3
H^2	0.23	0.91	0.81	0.32	0.27	0.48	0.53	0.60	0.53	0.10	0.72	0.44	0.78
HSFIC													
Season (S)	767.05***	880.84***	924.04***	0.471 ns	11,978.6***	202,380***	13,938.5***	334.39***	132.103***	1415.38***	6051.7***	30,722.7***	83.283***
Replication	109.32**	201.74***	2.04 ns	169.82***	2.3 ns	128**	39.5**	2.74**	1.684*	9.55*	5.1 ns	0.1 ns	0.761 ns
Genotype (G)	31.34***	32.07***	43.87***	9.111***	44.9***	104***	20.7***	3.29***	1.664***	5.80***	125.8***	49.6***	2.295***
G X S	15.37***	1.74**	2.21**	1.827 ns	26.2***	63***	9.9***	0.72***	0.607***	3.87***	1.9 ns	28.6***	1.442***
Mean	54.96	51.56	90.61	36.05	20.55	66.51	25.51	9.073	4.916	16.19	47.96	36.19	4.872
Minimum	43.86	41.77	75.73	26.58	7.56	48.65	12.99	4.89	2.754	12.17	20.57	21.77	1.557
Maximum	67.54	68.57	105.55	44	36.6	87.12	35.22	13.64	7.838	22.09	75.98	52.92	8.599
Q1	52.31	51.03	86	34.17	16.78	60.82	22.7	7.97	4.091	14.9	41.41	33.31	3.885
Q3	57.85	59.19	95.77	38.19	24.07	72.11	28.3	10.13	5.699	17.3	55.62	39.05	5.899
CV (%)	6.6	5.9	4.1	5.2	27.1	24.3	18.8	13.0	17.4	12.4	13.4	20.0	21.7
H^2	0.54	0.95	0.92	0.79	0.58	0.58	0.57	0.75	0.65	0.37	0.89	0.44	0.52
Significance of CRDHS over HSFIC													
%	10.4	1.2	4.2	0.7	30.9	12.8	14.1	5.5	5.8	12.7	12.3	14.7	38.0

Table 1. Statistical description of the phenotypic data of the ICARDA diverse stress panel of 234 bread wheat genotypes grown under HSFIC and CRDHS conditions. PAKyC traits, phenology, architecture, and kernel yield component traits; CV, coefficient of variation; G x S, genotype by season interaction; Q1 and Q3, first and third quartiles; H^2, broad sense heritability; CGC, crop ground cover; DH, days-to-heading; DPM, days-to-physiological maturity; GFP, grain filling period; FLA, flag leaf area; PH, plant height; PL, peduncle length; SL, spike length; OL, awn length; SPS, spikelet number per spike; KPS, kernel number per spike; TKW, thousand-kernel weight; KY, kernel yield; CRDHS, combined recurrent drought and heat stress; HSFIC, fully irrigated conditions with heat stress; ns, not significant, * Significant at the 0.05 probability level. ** Significant at the 0.01 probability level. *** Significant at the 0.001 probability level. Blocks within replication and residuals were random effects.

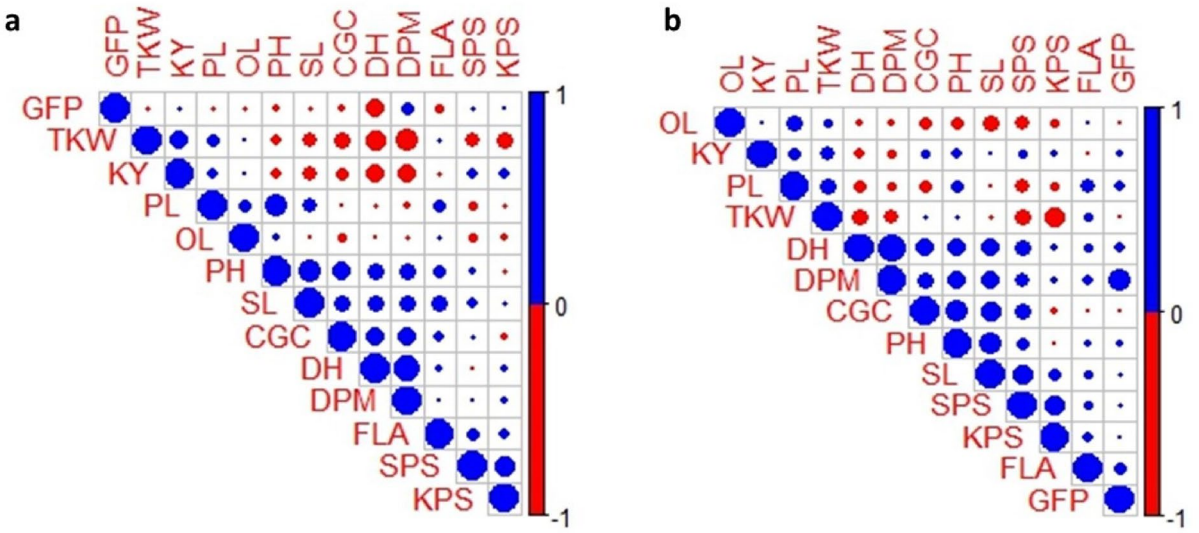

Fig. 1. Correlation among 13 PAKyC traits of plants grown under CRDHS (**a**) and HSFIC (**b**). Positive correlations are shown in blue, whereas negative correlations are shown in red. The colour and the size of the circles correspond to the correlation coefficients. The legends, on the right side of the correlograms, display the colours that correspond to the correlation coefficients. CGC, crop ground cover; DH, days-to-heading; DPM, days-to-physiological maturity; GFP, grain filling period; FLA, flag leaf area; PH, plant height; PL, peduncle length; SL, spike length; OL, awn length; SPS, spikelet number per spike; KPS, kernel number per spike; TKW, thousand-kernel weight; KY, kernel yield.

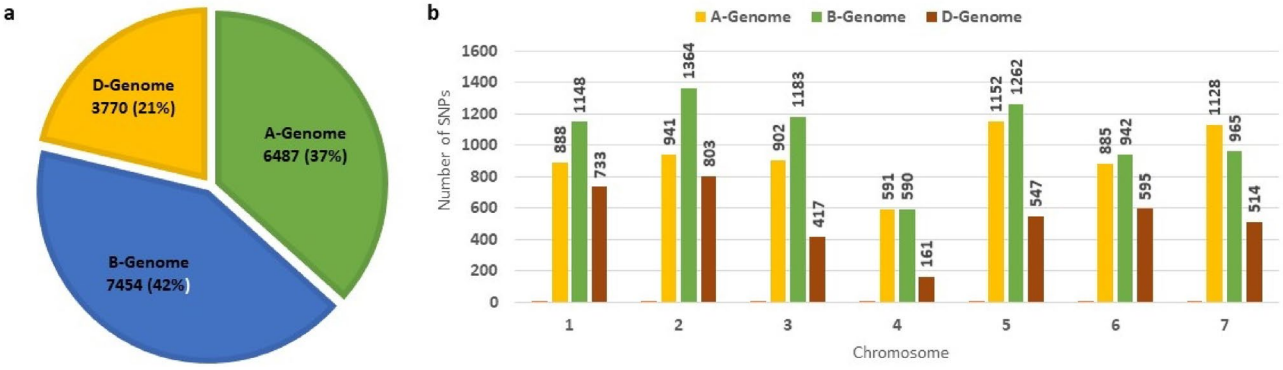

Fig. 2. Distribution of high-quality SNPs in the genomes (**a**) and chromosomes (**b**) resulting from 187 bread wheat genotypes used for GWAS

Linkage disequilibrium (LD) analysis enhances the efficiency of gene identification. LD values range from zero, indicating the independent inheritance of a pair of alleles, to one, which indicates a pair of alleles always inherited together. Hence, the strength of the association increases as the r^2 approaches 1. We found 1202748 marker pairs with average LD values of $R^2 = 0.2$. Twenty percent (n = 260,868) of the total pairs were significantly linked ($P \leq 0.01$). The B sub-genome contains the highest percentage (26%) of significantly linked marker pairs, followed by the A sub-genome (18%) and the D sub-genome (16%). The lowest LD value was observed among marker pairs located on chromosome 5D ($r^2 = 0.00$), while the highest was recorded for chromosome 1D ($r^2 = 0.25$) (Table 2). Collectively, the LD analysis within the stress panel indicates a strong predictive value for drought and heat stress tolerance.

The LD was estimated from the squared correlation coefficient (r^2) for all 17,711 high-quality markers. The genome-wide LD decayed with genetic distance, halving at 4.52 Mb for the whole genome, 3.11 Mb for the A sub-genome, 6.41 Mb for the B sub-genome, and 2.52 Mb for the D sub-genome (Fig. 3).

Population structure and kinship analysis

The K-value was used to determine the most likely number of subpopulations (Fig. 4). Genetic relationships among individual genotypes were evaluated using a kinship matrix derived from the SNPs. Population structure analysis revealed a maximum K-value (K = 2), thereby identifying two major subpopulations (Fig. 4a): subgroup I comprised 64 genotypes, while subgroup II consisted of 123 genotypes (Fig. 4b).

Chromosome	NMP	NSMP ($P<0.01$)	NSMP (%)	MDMP (Mbp)	Mean r^2
1A	393,828	71,157	18.1	122.53	0.19
2A	442,270	92,196	20.8	230.91	0.20
3A	406,351	79,959	19.7	220.98	0.17
4A	174,345	32,502	18.6	204.70	0.20
5A	662,976	102,763	15.5	119.96	0.19
6A	391,170	83,124	21.3	184.21	0.18
7A	635,628	111,196	17.5	199.35	0.18
1B	658,378	210,405	32.0	194.47	0.19
2B	929,566	249,094	26.8	262.86	0.16
3B	699,153	153,407	21.9	172.58	0.19
4B	173,755	80,306	46.2	240.94	0.14
5B	795,691	128,052	16.1	144.09	0.18
6B	443,211	122,488	27.6	201.81	0.24
7B	465,130	129,418	27.8	224.74	0.15
1D	268,278	40,945	15.3	154.46	0.25
2D	322,003	67,089	20.8	184.61	0.23
3D	86,736	11,171	12.9	160.99	0.22
4D	12,880	2295	17.8	167.85	0.20
5D	149,331	18,090	12.1	91.65	0.00
6D	176,715	25,268	14.3	145.03	0.19
7D	131,841	15,154	11.5	203.01	0.16
A sub-genome	443,795.4	81,842.4	18.4	183.23	0.19
B sub-genome	594,983.4	153,310.0	25.8	205.93	0.18
D sub-genome	163,969.1	25,716	15.7	158.23	0.18
Mean	1,202,748.0	260,868.4	20.0	182.50	0.20

Table 2. Summary of linkage disequilibrium between significantly ($P \leq 0.01$) linked pairs in each chromosome and genome-wide. NMP, number of marker pairs; NSMP, number of significant marker pairs; MDMP, mean distance between marker pairs; r^2, equivalent to the correlation coefficient between a pair of alleles.

Marker-trait associations

The efficiency of genome-wide association studies (GWAS) in detecting significant MTAs depends on the marker distribution within the genotypic dataset. An even spread of markers across the genome allows detection of MTAs with greater precision. A substantial reduction of rare alleles, inclusion of the population structure, and kinship also improve the reliability of GWAS[27,28]. In this study, data of 13 PAKyC traits, population structure, kinship information, and high-quality SNPs were used for GWAS.

SNP markers and best linear unbiased prediction (BLUP) data for the PAKyC traits were used for GWAS across six environments. The BLUP values were recorded from two seasons of CRDHS (E1 and E2) and one pooled dataset from E1 and E2 (E3), similarly from two seasons of HSFIC (E4 and E5) and one pooled dataset from E4 and E5 (E6).

The highest number of significant MTAs was identified in the A sub-genome, accounting for 64.7% (n = 44), followed by the B and D sub-genomes, with an equal proportion (17.6%, n = 12). SNP markers linked to the traits were found on all chromosomes except 4D, 6B, and 7B. The highest number of MTAs were identified on 5A chromosome (52.9%, n = 36) (Supplementary Tables S2, S3, and Table 3).

In this study, 68 MTAs (P values < Bonferroni threshold) were identified from 40 SNPs and 12 PAKyC traits (CGC, DH, DPM, GFP, PH, PL, SL, OL, FLA, SPS, TKW, and KY) using GWAS across six environments (Fig. 6). The MTAs with a frequency ranging from one to five across six environments were identified (Fig. 5). Among these markers, nine SNPs were detected in two or more environments, referred as constitutive o**r** stable SNPs, indicating the sites of most significance in bread wheat breeding for heat and recurrent drought stress. These include: 1) RAC875_c9984_1003; 2) AX-158538950; 3) TG0019; 4) IAAV1650; 5) wsnp_Ex_c23383_32628864; 6) wsnp_Ex_c31799_40545376; 7) AX-94430 584; 8) BS00039218_51; and 9) RAC875_c30566_230. The description of these SNPs is presented in Supplementary Table S2.

The constitutive SNPs located on chromosomes 1D, 3B, 5B, and 6D are presented in Supplementary Tables S2 and S3. Manhattan and quantile–quantile (QQ) plots were generated for significant and constitutive SNPs to identify key associations (Fig. 6).

In silico analysis

The genetic basis of the PAKyC traits was associated with nine constitutive SNPs located on chromosomes 5A, 1D, 3B, and 6D. The SNPs were linked to 35 candidate genes associated with cellular components, biological processes, and molecular functions, and expressed under CRDHS and HSFIC (Supplement Table S4).

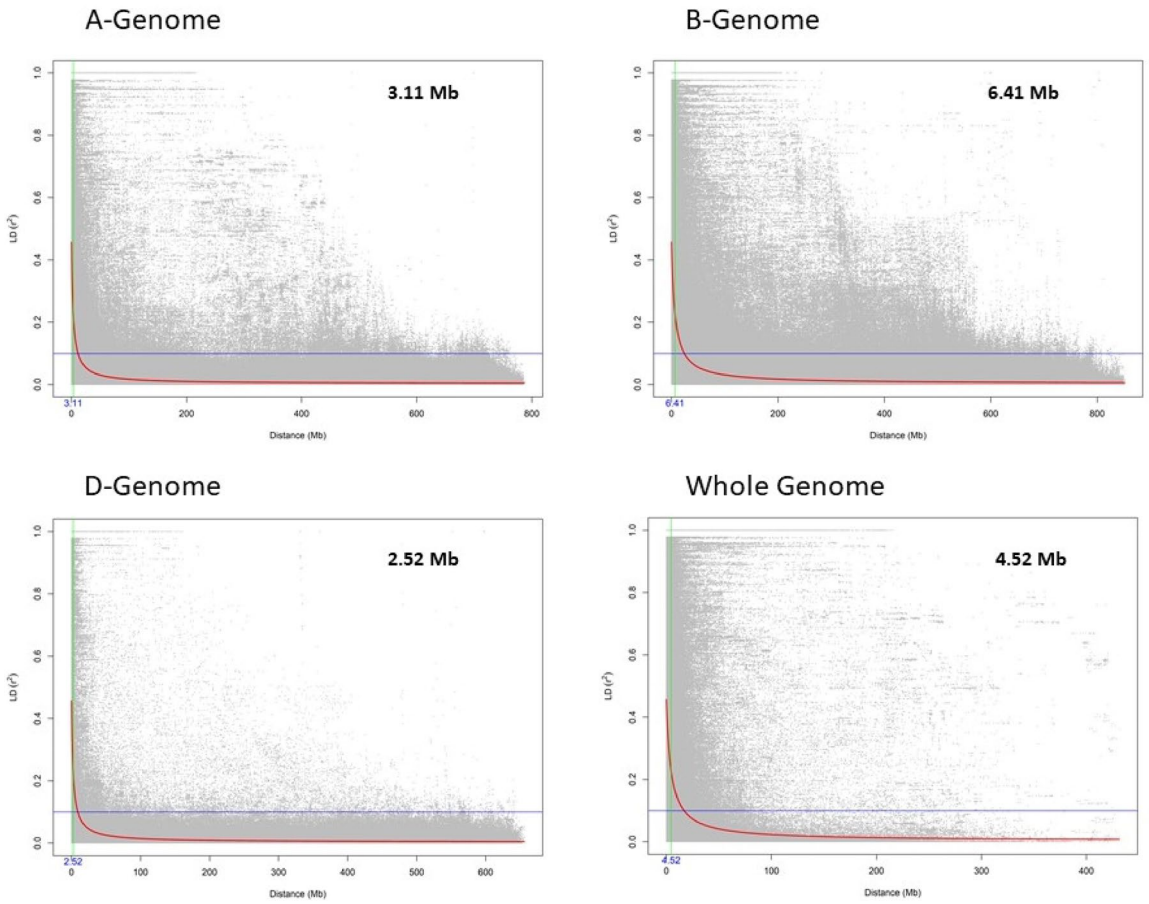

Fig. 3. Scatter plots showing linkage disequilibrium (LD) decay estimated by plotting (r^2) against physical distance (Mb) in 187 diverse genotypes. The green line indicates the threshold point where the LD dropped to 50% of its maximum value. The LD decay value at the cutoff point is depicted on the x-axis in blue font and at the top right in a larger font size.

Protein-coding candidate genes with high expression levels were further studied to uncover the molecular mechanism of the most significant associations. This study identified constitutive SNPs associated with 35 protein-coding transcripts. Among these 12 transcripts were highly expressed (>0.5 TPM) mainly in response to separate and combined drought and heat stress, underscoring their relevance for drought- and heat-tolerance breeding and wider adaptability. The expression of transcripts in response to combined drought and heat was less compared with the control, which may have resulted from the study being limited to the seedling growth stage and hours of stress period (Fig. 7).

The encoded proteins and functions of the candidate genes are presented in Table 4. The genetic components of bread wheat contain candidate genes associated with drought, heat, and their combined stresses[29].

In addition to these nine candidate genes, TraesCS1A03G0469000, TraesCS4B03G0670000, and TraesCS5A03G0938500 were highly expressed under drought and heat stress (Fig. 7), but, to the best of our knowledge, their specific role in bread wheat remains unclear. The chromosomal locations, biological processes, cellular components, and molecular functions of the candidate genes are presented in Supplementary Table S4.

Discussion

Heat and drought are the major abiotic stresses that globally limit wheat production, and plants mount a multigene response. The simultaneous occurrence of heat and drought worsens adverse effects on wheat phenology, plant architecture, and kernel yield. Wheat's varying sensitivity to these stresses across growth stages further complicates the situation. Therefore, extensive research is essential to understand the genetic basis of bread wheat response and to develop cultivars that are tolerant to heat, drought, and their combined effects. The success of breeding for heat and drought tolerance depends on testing environments and the genetic diversity of the materials used. To this end, we used a large, diverse stress panel sourced from various breeding programs at ICARDA, along with natural heat stress and controlled irrigation to simulate drought at different growth stages.

High variability ($P<0.001$) across genotypes for all traits, seasons, and genotype-by-season interactions for all traits, except for TKW and KY under CRDHS, and for GFP and KPS under HSFIC, was observed. This reveals the presence of important genes in the diverse stress panel that could be used to enable breeding for heat and recurrent drought tolerance. Compared with HSFIC, a greater negative impact (up to 38%) of CRDHS

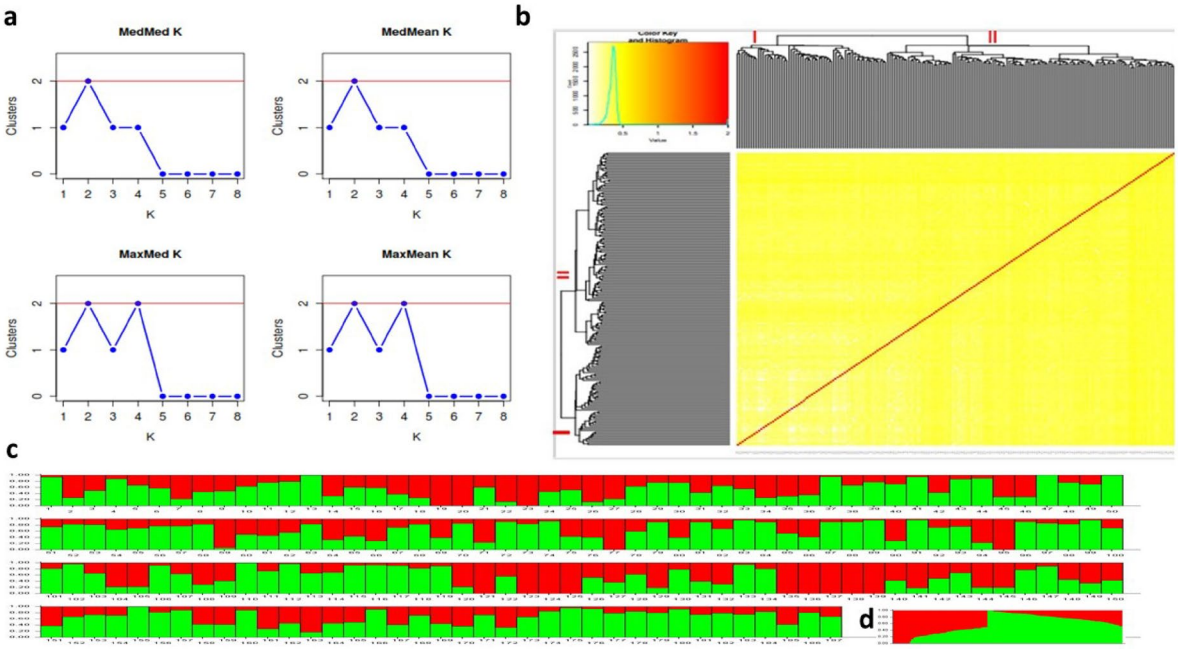

Fig. 4. Population structure and kinship matrix of the two major subpopulations of 187 ICARDA bread wheat genotypes: K value using the Puechmaille method (**a**), Heatmap of the kinship matrix (**b**), and STRUCTURE plots inferring the population structure (**c**) and (**d**): vertical bars display individual genotypes. The segment colour length in each vertical bar indicates the proportion contributed by each of the two populations in the model for that individual. The genotypes coded (pedigree) corresponding to the bar plot are presented in the supplementary file (Supplementary Table S1).

conditions on overall genotypes' performance, expressed in PAKyC traits, was observed. The pronounced combined effects of heat and drought on agronomic traits have been discussed elsewhere[48]. The combined stress induces oxidative stress through excessive production of reactive oxygen species (ROS) (O2 −, H2O2, OH −) and increased malondialdehyde levels, leading to reductions in photosynthetic attributes, nutrient uptake, and kernel yield.

Broad-sense heritability estimates varied across test environments and PAKyC traits, with lower values under CRDHS than HSFIC. The heritability estimates for DH, DPM, and KPS were high (>0.6) and medium (0.3–0.6) for PH, SL, PL, and TKW under both CRDHS and HSFIC conditions, indicating the predominant genetic control of these traits in a wide range of heat and drought environments. High heritability for KY and medium heritability for GFP were estimated under CRDHS, while high heritability for GFP and OL, and medium heritability for CGC, FLA, and SPS were estimated under HSFIC, indicating the higher genetic influence on these traits under CRDHS and HSFIC, respectively.

Strong (r > 0.5) correlations between PH and SL, and DH and DPM under both HSFIC and CRDHS environments were observed. PH and PL ($r = 0.55$) under CRDHS, and CGC and PH ($r = 0.55$); SL ($r = 51$), DPM and GFP ($r = 0.57$), and SPS and KPS ($r = 0.51$) under HSFIC were found. Furthermore, PH and SL are among the important traits for KY improvement. On the other hand, the negatively correlated traits, mainly KY and TKW, with DH and DPM, facilitate the selection for desirable traits early HD and DPM coupled with a higher KY and TKW under HSFIC and CRDHS. The correlation among traits primarily arises from pleiotropy or the close linkage of genes that influence breeding strategies (Sandhu, 2024). Pleiotropy is a complex phenomenon in which a single gene influences multiple quantitative traits through additive and dominance effects, acting separately or jointly. True pleiotropy involves a shared genetic architecture that influences multiple traits[49–51]. Hence, highly correlated traits may be used in multivariate analysis to predict genetic values. Overall, the strongly correlated traits facilitate simultaneous breeding for the targeted traits.

Linkage disequilibrium (LD) measures the association between alleles. LD decay determines the marker density required to perform a GWAS. In this study, the B sub-genome had a greater degree of LD (6.41 Mb) compared to the A (3.11 Mb) and D (2.52 Mb) sub-genomes, indicating that genetic variants were more frequently inherited together in the B sub-genome, followed by the A and D sub-genomes. Factors such as population size, genetic drift, admixture, selection, mutation, recombination rate, reproductive method, and recombination frequency influence the size of LD blocks. In this study, the highest number of high-quality SNP markers was mapped to the B genome (n = 7454), followed by the A (n = 6487) and D (n = 3770) genomes. The number of SNPs mapped to the B sub-genome was roughly twice that of the D sub-genome (Fig. 2a). This is due to the initial hybridization (0.5–0.8 million years ago) between *Triticum urartu* (A genome) and *Aegilops speltoides* (B genome), which led to the formation of the tetraploid *Triticum turgidum* subsp. *dicoccoides* (2n = 4x = 28, AABB genome). After thousands of years, a second hybridization occurred between *Triticum turgidum* subsp.

E	Traits	SNP	Overlapping gene (protein coding, (IWGSC RefSeqv2.1))	Genomic location
E2, E3	PH	AX-94430584	TraesCS5D03G0629800	380864541–380874450(+)
			TraesCS5D03G0459500	300573238–300604383(−)
			TraesCS5A03G0066100	24128681–24138749(+)
			TraesCS2A03G0262100	80290658–80306954(−)
			TraesCS1D03G0222600	83841197–83850830(+)
			TraesCS1A03G0469000	311816693–311940249(+)
E1, E3, E4, E6	DPM, CGC, PH, SL, SPS, OL	AX-158538950	TraesCS6B03G0883400	560042918–560065255(+)
			TraesCS5B03G0959300	565319968–565342449(−)
E2, E3, E4, E5, E6	DH	RAC875_c9984_1003	TraesCS3A03G0275100	101346995–101347525(−)
			TraesCS1D03G0760200	416868897–416908969(+)
E1, E2, E3, E4, E5, E6	DPM, CGC, KY, OL, PL, SL, SPS	TG0019	TraesCS7D03G1033200	557181706–557191723(−)
			TraesCS7D03G1016800	551208747–551222140(−)
			TraesCS6D03G0912000	490947198–490979839(−)
E5, E6	SL, SPS	wsnp_Ex_c31799_40545376	TraesCS5B03G0233000	122529446–122544646(+)
			TraesCS5A03G0939300	591176877–591180614(+)
			TraesCS1D03G0134800	44895130–44912421(−)
			TraesCS1D03G0760200	416868897–416908969(+)
E3, E6	DH, TKW	wsnp_Ex_c23383_32628864	TraesCS7D03G0368300	115185078–115210035(+)
			TraesCS6D03G0407400	200388173–200398638(−)
			TraesCS1B03G0539500	335076300–335077166(+)
E4, E6	PH	BS00039218_51	TraesCS4D03G0093000	27578914–27584346(−)
			TraesCS3B03G1026600	666460077–666486390(+)
			TraesCS3B03G0104500	29312429–29316925(+)
			TraesCS1A03G0228400	91115673–91132505(−)
E3, E5	FLA, PH,SL	IAAV1650	TraesCS6B03G0294900	119099145–119126229(+)
			TraesCS4B03G0670000	511775354–511809898(−)
			TraesCS4A03G0793000	605662051–605680761(+)
			TraesCS3B03G0811900	524703146–524720091(+)
			TraesCS3A03G0903500	632675187–632700362(+)
			TraesCS3A03G0408300	198937370–198976506(+)
			TraesCS1B03G1247400	685341162–685360093(+)
			TraesCS1B03G0796800	503890262–503913076(+)
			TraesCS1A03G0151500	44904478–44929749(+)
E4, E6	DPM	RAC875_c30566_230	TraesCSU03G0136300	108551952–108565014(+)
			TraesCS7A03G0322100	93566847–93585025(−)

Table 3. Environmentally stable SNPs linked to phenology, architecture, and kernel yield component traits of bread wheat and 35 candidate genes identified under CRDHS and HSFIC**.** E, environment; E1, season one under CRDHS; E2, season two under CRDHS; E3, pooled environments (E1 and E2); E4, season one under HSFIC; E5. season two under HSFIC; E6, pooled environments (E4 and E5); CGC, crop ground cover; DH, days-to-heading; DPM, days-to-physiological maturity; FLA, flag leaf area; PH, plant height; PL peduncle length; SL, spike length; OL, awn length; SPS, spikelet number per spike; TKW, thousand-kernel weight; KY, kernel yield.

dicoccoides and *Aegilops Tauschii* (2n = 2x = 14, DD genome)[52,53]. Consequently, gene flow over an extended period has contributed to greater diversity in the AABB genome compared to the DD genome[54,55].

Chromosome-wide, the largest (n = 1364) and smallest (n = 161) numbers of markers were found on the 2B and 4D chromosomes, respectively. This was more than twice the amount reported previously at the sub-genome and chromosome levels, ranging from 2B (n = 599) to 4D (n = 23)[45]. The lower number of SNPs at the sub-genome and chromosome levels may have resulted from a comparatively small number of genotypes (n = 92) and SNP markers (n = 6,349). Therefore, the observed SNP distribution suggested the opportunity for fine mapping on 2B, 5B, 3B, 5A, and 1B with a greater proportion of MTAs on 5A for drought- and heat-associated traits. Others have also reported that chromosome 5A harbors significant MTAs for drought and heat tolerance in wheat[56–58].

Constitutive SNPs were associated with DH, DPM, PH, and SL in both CRDHS and HSFIC environments. These SNPs were also associated with TKW and FLA under HSFIC, and OL, PL, and KY under CRDHS. SPS showed low heritability under CRDHS and FICHS. However, it was associated with TG0019, AX-158538950, and wsnp_Ex_c31799_40545376. Similarly, CGC showed low heritability under CRDHS, while linked with

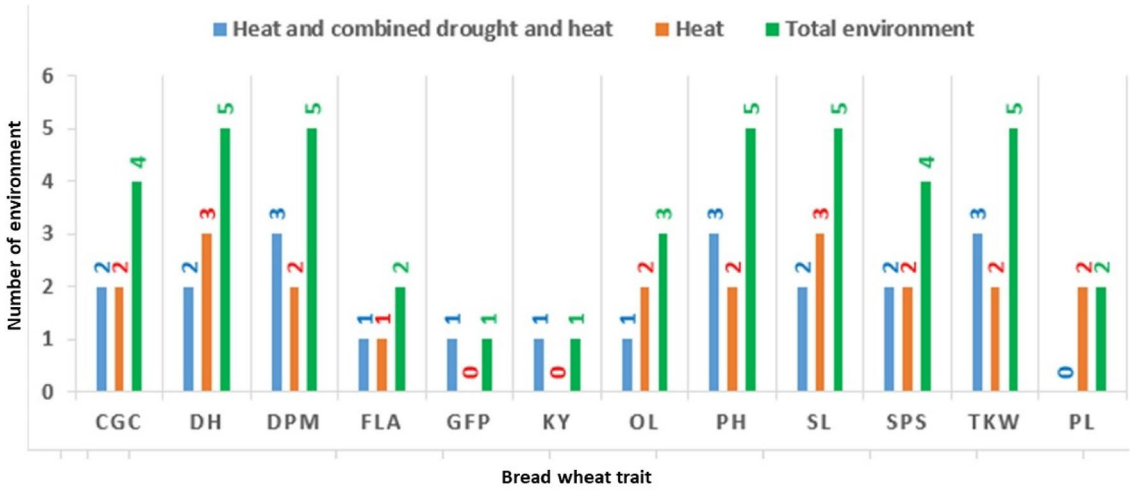

Fig. 5. Number of detected marker trait associations (MTAs) of 187 diverse genotypes across six environments in the GWAS using the BLINK model. CGC, crop ground cover; DH, days-to-heading; DPM, days-to-physiological maturity; GFP, grain filling period; FLA, flag leaf area; PH, plant height; PL peduncle length; SL, spike length; OL, awn length; SPS, spikelet number per spike; TKW, thousand-kernel weight; KY, kernel yield.

TG0019 and AX-158538950. These low heritabilities and associations with constitutive SNPs revealed a pattern of inheritance in which multiple genes influenced a specific trait (polygenic inheritance). Overall, the associations between these traits and stable SNPs could facilitate breeding for drought and heat stress in bread wheat.

Additionally, nine constitutive SNPs are associated with putatively pleiotropic loci. Pleiotropic quantitative trait nucleotides (QTNs) control the expression of more than one PAKyC trait, and the genes that control these traits can be located at the same genomic position, allowing simultaneous breeding for multiple traits[59–61]. This study revealed three types of pleiotropic SNPs: 1) Adaptive pleiotropic (SNPs that influence two or more traits under CRDHS (E1 to E3) or HSFIC (E4 to E6); 2) Broad-effect pleiotropic (SNPs that influence one trait under both CRDHS and HSFIC; and 3) conditional pleiotropic (SNPs that influence a trait under CRDHS and another trait under HSFIC). Furthermore, trait-specific SNPs were identified. In this context, RAC875_c9984_1003, AX-158538950, and TG0019 were broad-effect pleiotropic SNPs associated with DH, PH, and both CGC and OL, respectively, under both CRDHS and HSFIC environments, revealing wide adaptation. In addition, two SNPs exhibited conditional pleiotropy, influencing different traits. AX-158538950 was associated with SL and SPS in the HSFIC environment, and DPM and CGC in the CRDHS environment. Similarly, TG0019 was associated with PL and SL under HSFIC, and KY, DPM, and SPS under CRDHS conditions. IIAAV1650 was another conditional pleiotropic SNP that was associated with FLA under HSFIC, and with PH and SL under CRDHS. Overall, AX-158538950 and TG0019 demonstrated robust and stable genetic associations in both HSFIC and CRDHS environments.

IAAV1650, wsnp_Ex_c23383_32628864, and wsnp_Ex_c31799_40545376 were adaptive pleiotropic SNPs identified under CRDHS. Whereas AX-94430584 (PH), BS00039218_51 (PH), RAC875_c30566_230 (DPM), and RAC875_c9984_1003 (DH) were trait-specific SNPs; associated traits are presented alongside in parentheses. Uniquely, both AX-94430584 and BS00039218_51 were associated with PH under the CRDHS environment, while wsnp_Ex_c31799_40545376 was an adaptive, pleiotropic SNP specific to HSFIC that was associated with both SL and SPS. Interestingly, among the studied traits, GFP had low heritability under CRDHS, but still showed a significant association with a SNP marker (AX-94531491). However, it was not linked to any constitutive SNPs. Constitutive SNPs and pleiotropic architectures under stress environments have been reported on 1B, 2A, 2B, 2D, 3A, 4A, 4B, 5A, 6A, 6D, and 7A, which is consistent with our findings. Furthermore, constitutive and pleiotropic SNPs have also been reported on 4D, 6B, and 7B, with pleiotropic associations with PH, TKW, HD[62–64].

Overall, the larger LD decay values in the B sub-genome and A sub-genome could result in the detection of a greater proportion of MTAs, facilitating marker-assisted selection. However, more detailed analysis, e.g., effective fine mapping that uses linkage disequilibrium decay patterns between constitutive SNPs and causal variants, is required to distinguish the true pleiotropic SNPs from tightly linked intragenic SNPs for each quantitative trait nucleotide (QTN).

Drought and heat stress trigger the production of various proteins, including enzymes, transporters, and transcription factors. Transcription factors have diverse DNA-binding domains and are classified into gene families which regulate gene expression through either activating or repressing the transcription of downstream genes[65,66]. Transcription factors are multifunctional proteins that influence multiple genes and can be used to guide breeding strategies. Thus, further study on the complete regulatory mechanism of the individual and combined effects of these upstream and downstream regulators is indispensable[66,67].

We identified nine key transcripts that were drought-and heat-responsive; thus, a need to leverage these resources to develop heat- and drought-resilient bread wheat varieties. Furthermore, a detailed examination

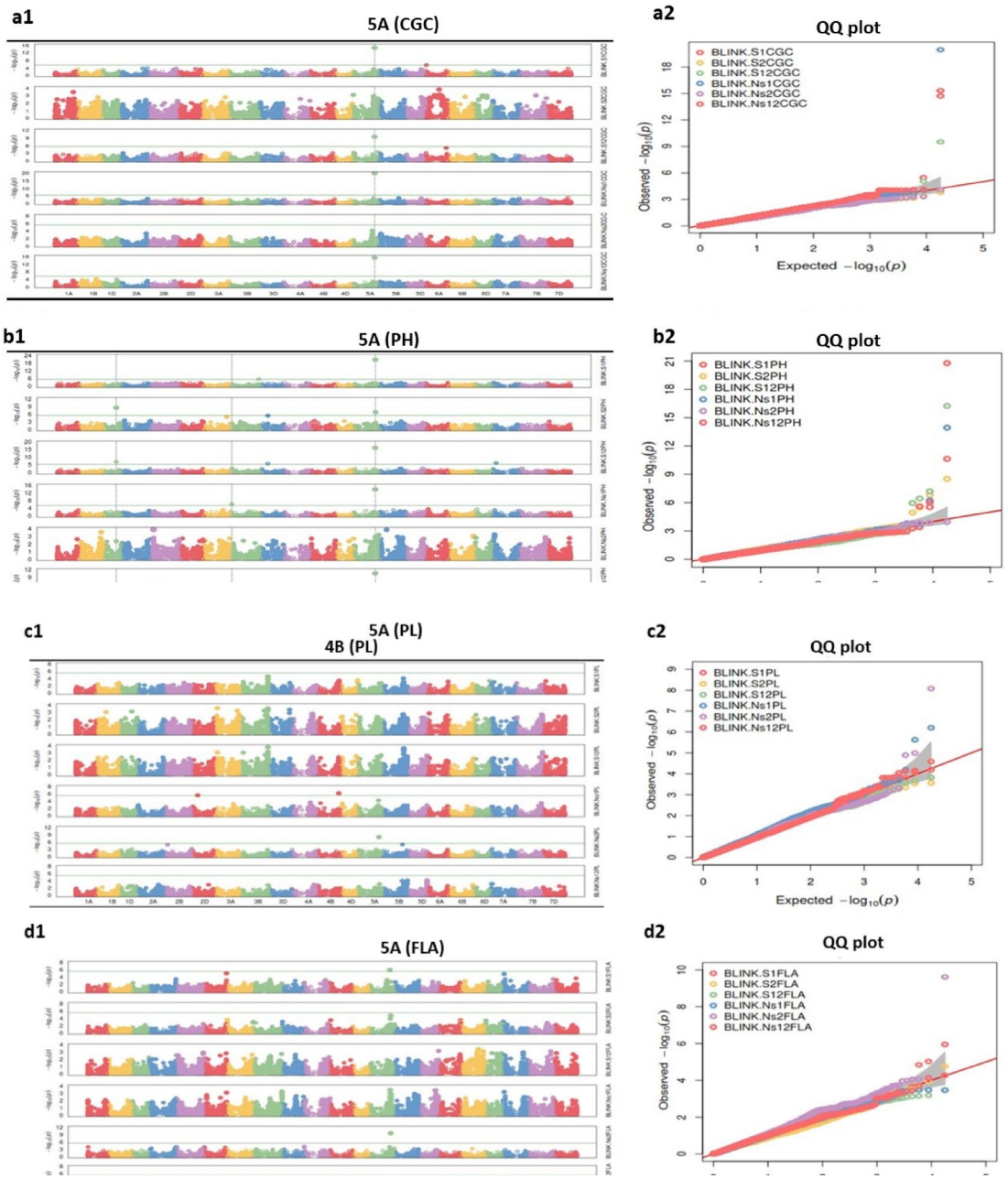

Fig. 6. Manhattan plots (**a1**–**j1**) and corresponding Q–Q plots (**a2**–**j2**) showing environmentally-stable MTAs of 187 genotypes grown under HSFIC and CRDHS. Significant associations were detected using the Bonferroni correction ($a = 0.05$). Legends in the Q–Q plots: S1, season one under CRDHS (**E1**); S2, season two under CRDHS (**E2**); S12, pooled S1 and S2 (**E3**); Ns1, season one under HSFIC (**E4**); Ns2, season two under HSFIC (**E5**); Ns12, pooled Ns1 and Ns2 (**E6**); CGC, crop ground cover; DH, days-to-heading; DPM, days-to-physiological maturity; FLA, flag leaf area; PH, plant height; PL peduncle length; SL, spike length; OL, awn length; SPS, spikelet number per spike; TKW, thousand-kernel weight.

of the role of the three unidentified transcripts (TraesCS1A03G0469000, TraesCS4B03G0670000, and TraesCS5A03G0938500 is required to exploit their potential in drought and heat environments.

In this study, genes encoding drought-and heat-related transcription factors and proteins were identified (Table 4). For instance, 1) NRT1.1 is the first nitrate transport protein cloned. NRT1s and PTRs are in the same family and are nitrate (NRT) and peptide (PTR) transporters. NRT1.1-mediated NO_3^- transport into guard cells may enhance plant resistance to drought stress; 2) subtilisin-like proteases (SBTs) are a large family of serine peptidases that are unique to plants and are associated with plant development and response to the environment[32,68]; 3) MYBS3 is a single DNA-binding repeat, a MYB transcription factor that is involved in stress responses[69]; and 4) drought stress triggers ABA production, leading to stomatal closure and the expression

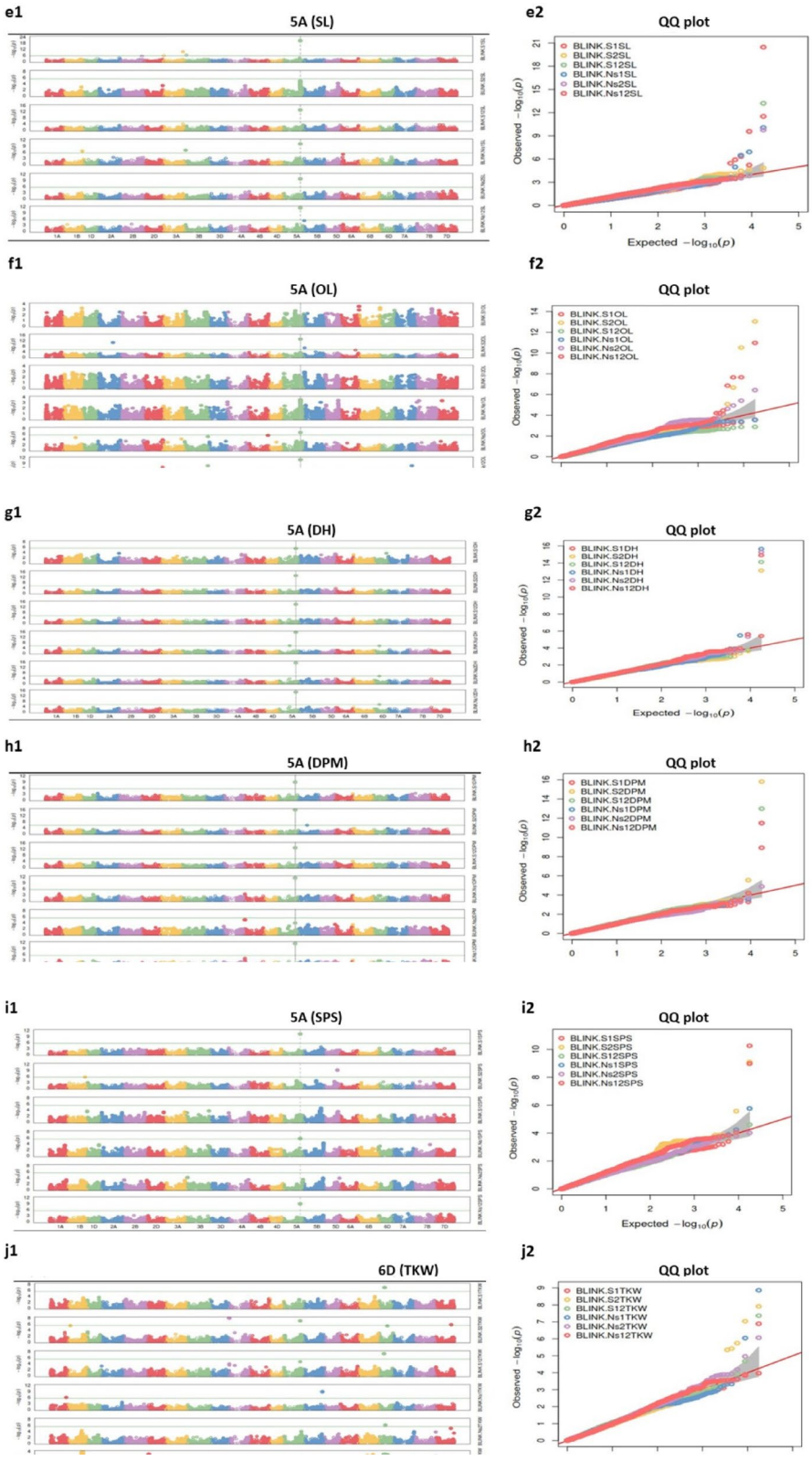

Fig. 6. (continued)

of drought stress-related genes[70,71]. Furthermore, the functional validation of candidate genes is required to confirm their influence, facilitate fine mapping, and advance bread wheat improvement.

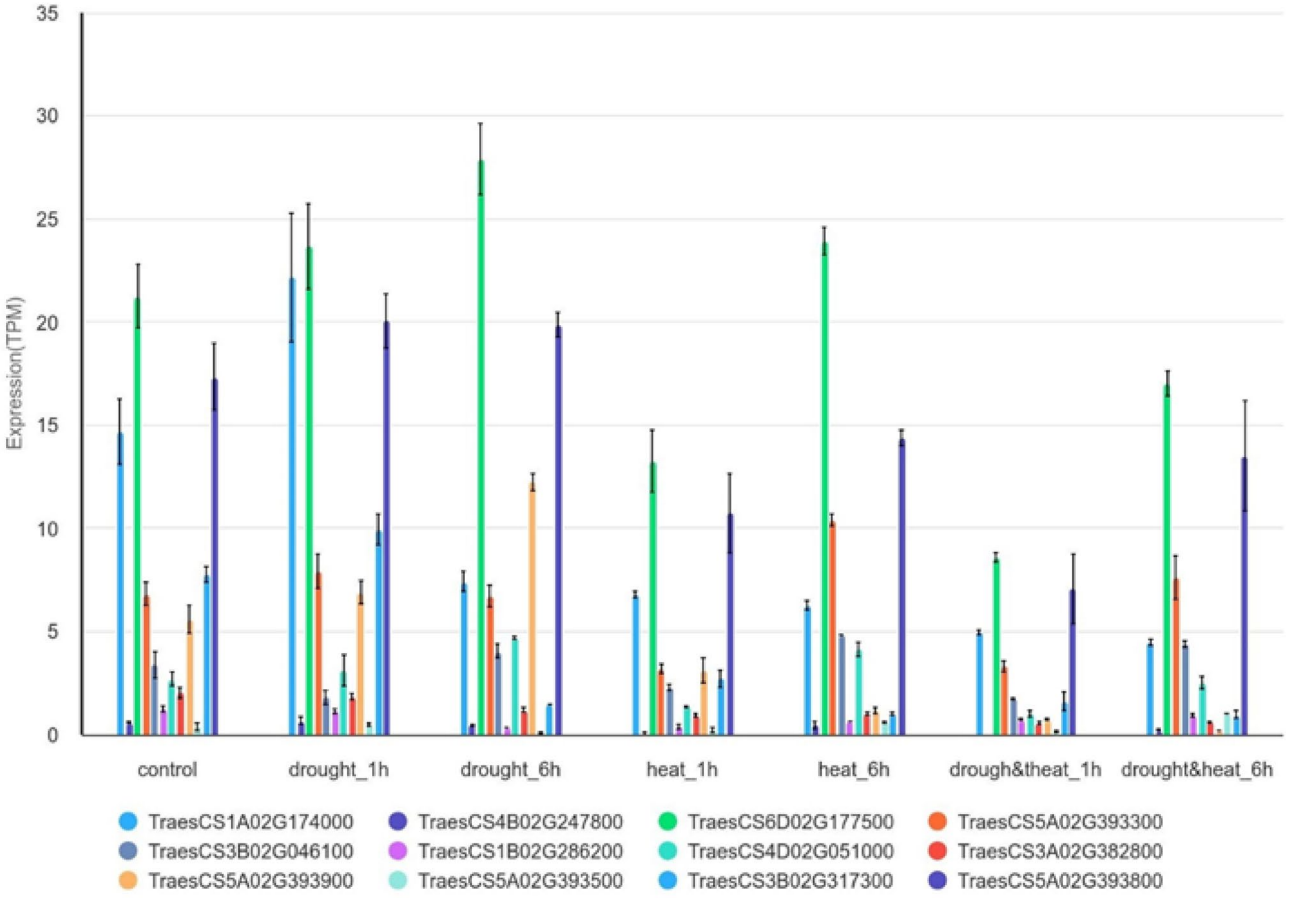

Fig. 7. Twelve genes with high expression (> 0.5 TPM) levels under drought, heat, and combined heat and drought stress conditions in bread wheat genotypes, representing the key transcripts acting in responses and adaptation to stress.

Trait	SNP	Traes ID (IWGSC RefSeqv2.1)	Candidate genes	Function
FLA, PH, SL	IAAV1650	TraesCS1B03G0796800	MYB proto-oncogene, transcription factor (MYBS3)	heat stress tolerance in Arabidopsis[30]; Drought and salt stress tolerance in rice[31]
		TraesCS3A03G0903500	Nitrate Transporter 1 (NRT1/ PTR FAMILY 5.10 (NPF5.10))	Drought stress tolerance[32]
		TraesCS3B03G0811900	Subtilisin-like proteases (SBTs)	Heat stress tolerance[33]; Drought stress tolerance[34]
			Dehydration-responsive element-binding protein 1G (DREB1G)	Drought stress tolerance in wheat[35] and Drought stress tolerance in cereals[36]
PH	BS00039218_51	TraesCS3B03G0104500	Ethylene-responsive transcription factor 5 (ERF5)	Drought stress tolerance[30,37]; Drought and salt tolerance[38]
		TraesCS4D03G0093000	Ethylene-responsive transcription factor 1 (EREBP1)	Drought stress tolerance in maize[39]
			Ethylene-responsive transcription factor 5 (ERF5)	Drought stress in *Arabidopsis thaliana*[40]
DPM, CGC, OL, PL, SL, SPS	TG0019	TraesCS5A03G0938200	Cullin Associated and Neddylation Dissociated 2 (CAND2)	Osmotic[41] and Heat stress[42]
		TraesCS5A03G0938800	Dehydration-responsive element-binding protein 1G (DREB1G)	Heat, drought, and salt stress tolerance in rice[34,43]
		TraesCS5A03G0939000	Abscisic acid response element binding factor 1 (TRAB1)	Drought stress[44] and abiotic stress tolerance in rice[45]:
SL, SPS	wsnp_Ex_c23383_32628864	TraesCS6D03G0407400	Ethylene-responsive transcription factor (ERF034)	Heat stress tolerance in wheat[46]; Drought and salt stress tolerance in durum wheat[47]

Table 4. The expression of candidate genes (> 0.5 TPM) under CRDHS and HSFIC. CGC, crop ground cover; DPM, days-to-physiological maturity; FLA, flag leaf area; PH, plant height; PL, peduncle length; SL, spike length; OL, awn length; SPS, spikelet number per spike.

Conclusions

Compared with HSFIC, the synergistic effects of stressors (CRDHS) considerably impacted the expression of PAKyC traits and their heritability. PAKyC traits had variations in broad-sense heritability, with the highest heritability for DH (0.91), DPM (0.81), KY (0.78), and KPS (0.72) under CRDHS.

Strong associations ($r > 0.5$) were observed between PH and SL, and DH and DPM under CRDHS and HSFIC. The strong association between PH and PL ($r = 0.55$) was also observed under CRDHS, while between CGC and PH ($r = 0.55$), CGC and SL ($r = 0.51$), DPM and GFP ($r = 0.57$), and SPS and KPS ($r = 0.51$) were identified under HSFIC.

The largest distribution of MTAs was identified in the A sub-genome (64.7%), mainly on the 5A chromosome (52.9%), and equally in the B and D sub-genomes (17.6%). Nine constitutive SNPs were involved in 24 MTAs. Associations with yield-influential traits DH, DPM, PH, and SL were identified in both CRDHS and HSFIC environments. Constitutive pleiotropic SNPs were detected under CRDHS and HSFIC, with some unique to one or the other environment. Three SNP markers were classified as broad-effect pleiotropic, and three as conditional pleiotropic. Additionally, constitutive pleiotropic SNPs, four unique to CRDHS and five unique to HSFIC, were discovered. Conversely, three constitutive SNPs that exhibited no pleiotropic effects were also found.

Furthermore, 35 candidate genes were identified. Among the candidate genes, 12 were highly expressed (> 0.5 TPM) under drought, heat, and combined stress environments. Among the highly expressed genes, nine were transcription factors involved in drought- and heat-regulation, while the functions of three genes remain unclear and require further studies.

This study provides insights into the expression of PAKyC traits and their molecular mechanisms under HSFIC and CRDHS environments. Overall, a comprehensive study combining heat stress, CRDHS, sensitive wheat growth stages, a large-size stress panel, high-quality dense SNPs, and technically sound experimental and computationally efficient procedures facilitated the dissection of the genetic basis of several quantitative traits. The identified candidate genes provide precise targets, advancing bread wheat breeding for heat- and CRDHS-resilience.

Methods

Research materials and phenotyping

A bread wheat stress panel, composed of 234 diverse genotypes from the breeding program of the International Centre for Agricultural Research in the Dry Areas (ICARDA) and commercial cultivars, was used in this study. These genotypes were selected for their potential for drought and heat tolerance. The genotypes pedigrees are presented in Supplementary Table S1.

This study was conducted at the Tony Farm Research Station of Haramaya University in Ethiopia, located in the semiarid tropical zone at a longitude of 9°36′48"N and a latitude of 41°50′24"E, and at an elevation of 1152 m above sea level. The soil is a sandy clay loam (73% sand, 15% silt, and 12% clay) with a neutral pH (6.8) and low organic matter content (1.105%). The climate is warm and dry, with a mean annual rainfall of 604 mm and an average temperature varying from 20.5 to 31.4 °C. The study was executed during two dry seasons, April to July 2021 and October to January 2021/2022. In the first season, rainfall (0–29.7 mm), rainy days (0 to 2), high temperatures (32.8–35.8 °C), low temperatures (21.1–23.8 °C), and sunshine hours (187 to 329), and in the second season, rainfall (0.0–13.0 mm), rainy days (0–1), high temperatures (29.1–33.6 °C), low temperatures (18.6–21.0 °C), and sunshine hours (198.5–312.0) were recorded. The high temperatures exceeded the maximum heat recommendation (25 °C) for wheat.

The stress panel was grown in a 13×18 alpha lattice design, consisting of 13 blocks each with 18 plots, with two replications under HSFIC and CRDHS. The alpha lattice design, which effectively partitions soil gradients into smaller and more uniform blocks, lowers the error mean square, thereby improving precision over a wide range of environments[72]. The two replicated trials were conducted side by side over the two dry seasons. Consequently, four separate environments were tested. Details of the HSFIC and CRDHS are explained below. Seeds were sown at 312 seeds per 1 m^2 plot, following the recommended sowing rate of 125 kg of seeds per hectare. A 40 g thousand-kernel weight was used to compute the sowing rate. They were planted in paired rows of 2.5 m length, with row-to-row, plot-to-plot, and block-to-block spacing of 0.2 m, 0.5 m, and 1 m, respectively. The soil was fertilized with NPS and urea, each at 100 kg ha^{-1}. NPS was applied at sowing, while urea was applied during the seedling and booting growth stages.

Tensiometers were used to measure soil water potential (SWP) which has a strong correlation with the leaf water potential[73–75]. Five tensiometers were installed diagonally across the CRDHS experiment at a depth of 50 cm to monitor the SWP. A hole was made in the soil with a diameter slightly larger than the tensiometer tip. Before the tensiometer was inserted into the hole, a soil–water slurry was added to ensure contact between the soil and the tensiometer tip. Low soil moisture was induced by suspending irrigation, and drought inductions were considered when three or more installed tensiometers read above 45 kPa. Water was supplied through furrow irrigation until the plot was ponded to a depth of 2.5 cm. A 3 m buffer zone separated the HSFIC and CRDHS experiments to prevent lateral water movement. Adequate soil moisture was maintained throughout all growth stages as needed for HSFIC. A decimal code for growth stages was used[76]. In the CRDHS experiment, recurrent drought (low-moisture stress at successive growth stages) was induced during 1) tiller formation, from main shoot (three fully expanded leaves) to the formation of the first tiller (GS21), 2) stem elongation, from the first node detection (GS31) to the appearance of the flag leaf (GS37), and 3) terminal drought, from 50% heading (GS59) to the hard dough stage (GS87). Irrigation was suspended before the crop reached the target growth stage to induce severe drought to impact plant growth and development, so that the plants could recover following drought relief.

Data on 14 phenology, architecture, and kernel-yield component (PAKyC) traits in HSFIC and CRDHS environments were recorded following the wheat growth stages[76]. The PAKyC traits recorded were crop ground cover (CGC), days-to-heading (DH) when 25% of spike had emerged in 50% of the plants (DGS 53), days-to-physiological maturity (DPM), grain filling period (GFP), flag leaf area (FLA) calculated using the length and width of leaves from five randomly selected plants, and an empirical coefficient (0.75). In addition, five plants from each plot were used to measure plant height (PH) from the soil surface to the tip of the spike, excluding the awns. Peduncle length (PL), spike length (SL), and awn length (OL) were also measured. Upon physiological maturity, spikelet number per spike (SPS), kernel number per spike (KPS), thousand-kernel weight (TKW), kernel yield (KY), and kernel moisture content at harvest (MC) were recorded. The MC was used to standardize the TKW and KY to 12 percent.

Phenotypic data analysis

Statistical analyses of the HSFIC and CRDHS experiments were conducted separately in R (version 4.4.1, R Foundation for Statistical Computing, Platform: x86_64-w64-mingw32/×64). The PAKyC traits were analyzed using a linear mixed-effects model (LMM). The best linear unbiased estimate (BLUE) mean for each season (environment) was estimated with genotype as a fixed effect. The variance components were estimated using a customized script and the lmer function in the lme4 package[77]. The statistical significance of fixed predictors was assessed using Satterthwaite's degrees of freedom, and of random effects using likelihood ratio tests using the package lmerTest[78]. The BLUP values and the Pearson correlation coefficient (r) were computed in the Meta-R software[79]. The correlation coefficient (r) was visualized with the Corrplot package at $P<0.05$[80]. BLUP values were used to estimate broad-sense heritability (H^2), characterized as low (0–0.3), moderate (0.3–0.6), and high (0.6 and above)[81].

Genotyping and quality control

196 genotypes of diverse origin, excluding commercial cultivars, were genotyped. TraitGenetics Section, at SGS Institut Fresenius GmbH, Gatersleben, Germany (http://www.traitgenetics.com/en/), performed DNA extraction and SNP genotyping. Genomic DNA was extracted from leaf tissue samples pooled from ten young seedlings of each genotype. DNA was extracted using an in-house CTAB-based protocol with chloroform extraction and isopropanol precipitation, followed by ethanol washing and resuspending the dried pellet in distilled deionized water. An optimized wheat 25 K Infinium iSelect array was utilized to generate the SNPs. The array contains 24,145 SNPs, combining 17,229 markers from the Illumina 20 K array, 6916 new markers from the 135 K Axiom array, and additional trait and gene-specific markers. SNP discovery was performed via the GenomeStudio Project software package (Illumina, San Diego, CA, USA). The latest reference genome, the Chinese Spring genome assembly from the International Wheat Genome Sequencing Consortium (IWGSCv2p1), was used for SNP chromosome alignment[82]. Quality control of SNPs involved removing those with more than 20% missing data, heterozygosity above 15%, and minor allele frequency (MAF) below 5% to prevent confounding effects from spurious alleles, using Trait Analysis by Association, Evolution, and Linkage (TASSEL v5.2.96) software[83]. LD kNN, k-nearest neighbours, imputation in the TASSEL was used to impute missing markers. Ultimately, 17,711 high-quality SNP markers, spread across 187 genotypes, were retained for further analysis.

Linkage disequilibrium (LD) and population structure

Pairwise LD values (r^2) between the markers across chromosomes were estimated in TASSEL version 5.2.96. A higher r^2 value indicates a higher level of LD. The $r^2=0$ represents complete linkage equilibrium between the two variants, whereas $r^2=1$ indicates complete linkage disequilibrium. The LD block size across the whole genome and three sub-genomes of bread wheat genotypes was determined as the distance at which LD values fell to half their maximum[84].

The genetic structure in the association panel was determined using a model-based Bayesian cluster analysis in STRUCTURE v2.3.4 software, which is appropriate for polyploid genetic data and provides unbiased inference in situations of weak differentiation[85]. The best-fitting number of assumed subpopulations (K values), ranging from one to eight, was evaluated by performing three independent Markov chain Monte Carlo (MCMC) runs of 100,000 iterations following a burn-in period of 10,000 steps for each value of K.

The optimum number of subpopulations was determined using four estimators: median of means (MedMeaK), maximum of means (MaxMeaK), median of medians (MedMedK), and maximum of medians (MaxMedK) employed in the Puechmaille method[86], which fits the dataset of this study, which consists of an uneven population size drawn from six distinct populations of different parental origins, and cultivars. This method was replicated using increasing thresholds (0.5, 0.6, 0.7, and 0.8) to evaluate the performance of the estimators. The best alignment of the results across the range of K values was determined using CLUMPAK through the StructureSelector web interface[87]. Then, STRUCTURE v2.3.4 (Pritchard et al., 2000) was run with the identified K value.

Genome-wide association study (GWAS)

Combinations of 1) phenotypic data: the best linear unbiased prediction (BLUPs) values of CGC, DH, DPM, GFP, PH, PL, FLA, SL, OL, SPS, KPS, TKW and KY in six environments: CRDHS in season one, S1 (environment one, E1), CRDHS in season two, S2 (environment two, E2) and pooled mean CRDHS S1 and S2, S12 (environment one, E3), and HSFIC in season one, Ns1 (environment four, E4), HSFIC in season two, Ns2 (environment five, E5) and pooled mean HSFIC Ns1 and Ns2, Ns12 (environment six, E6), and of the 187 diverse genotypes 2) genotypic data: 17,711 high-quality SNPs 3) kinship matrix data and 4) population structure (covariate) were used for the GWAS.

Significant marker-trait associations (MTAs) were identified using the genome association and prediction integrated tool (GAPIT) version 3.03[88] using the Bayesian information and linkage-disequilibrium Iteratively-nested keyway (BLINK) model[89], which has superior computational efficiency and enhanced statistical power compared to the multilocus mixed linear model (MLMM) and fixed and random model circulating probability unification (FarmCPU). Basically, this is the improved version of FarmCPU; it removes the assumption that causal genes are evenly distributed across the genome, which may lead to the inclusion of non-causal genes and the exclusion of causal genes present in the same bin.

Additionally, BLINK implements a computationally efficient fixed effect model using Bayesian information criteria, simplifying the computational procedures. It iteratively incorporates associated markers as covariates for testing markers and eliminates their connection to the cryptic relationship among genotypes as one of the strategies to reduce false positives. The associated markers are selected according to linkage disequilibrium, optimized for Bayesian information content, and reexamined across multiple tests to reduce false negatives[88,89]. A Bonferroni threshold (a/N, where $a = 0.05$, and $N =$ the number of total markers) of high-level stringency was used to determine the statistical significance of each SNP marker. R^2 was used to describe the variation explained (PVE) by significant MTAs. The Bonferroni correction is a widely used method for setting the statistical significance threshold (*P*-value) in bread wheat GWAS. It is a conservative technique that controls the family-wise error rate by dividing the significance level by the number of comparisons, yielding an adjusted threshold. This approach is assumed to reduce statistical power and increase the risk of type II errors (false negatives); however, the loss of power is often overstated, and it still offers a level of confidence that exceeds that of less stringent methods, which require additional validation. In wheat GWAS, the Bonferroni correction effectively manages both false positives and false negatives within the BLINK model, thereby enhancing replicability[90,91].

In silico analysis

The environmentally stable SNPs associated with the studied PAKyC traits under CRDHS and HSFIC were used to identify candidate genes. Bread wheat has a large and complex genome, approximately 16 gigabases, composed of repetitive sequences (> 85%), and many homoeologous regions between the sub-genomes[92,93], and greater than 11 Mb LD decay[94]. These, consequently, hinder the efficiency of GWAS in identifying candidate genes. However, GWAS with high-resolution and dense markers effectively identify candidate genes influencing quantitative traits[94–96]. MTAs and pleiotropic loci are also effectively identified using fixed and narrow genomic intervals, including 0.1 Mb, 0.2 Mb, and 0.5 Mb, in bread wheat[97–99]. In this study, high LD measures (r^2) and slow depreciation rates in the dense SNP panel, coupled with a high expectation of recombination frequency in the stress panel led to narrowing the window size to 0.2 Mb to retrieve high-confidence candidate genes. Furthermore, a similar window size is used in the wheat reference genome, representing the average length of conserved regions[100]. Consequently, a 0.2 Mb sequence centered on each stable SNP was extracted from the Ensembl Plants database (http://plants.ensembl.org/index.html) of the bread wheat genome for in silico analysis. The sequences were annotated using the basic local alignment search tool (BLAST) and aligned to the Wheat Chinese Spring IWGSC genome assembly[101]. Protein-coding transcripts were further examined for expression, partitioning under drought, heat, and drought-heat stress conditions. The expression of transcripts was measured on a transcripts per million (TPM) basis at the root, stem, leaf, spike, and grain levels using GeneExpression on the WheatOmics platform[102]. Candidate genes with high expression levels (> 0.5 TPM) were selected for further analysis[29]. Then, molecular functions, biological processes, and cellular components (Gene Ontology) related to these genes were retrieved from Triticeae-GeneTribe, TGT[103]. The roles of the identified candidate genes in regulating phenotypic traits were determined based on previous reports.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 8 December 2025; Accepted: 19 August 2026
Published online: 28 August 2026

References

1. Shewry, P. R. & Hey, S. J. The contribution of wheat to human diet and health. *Food Energy Secur.* **4**, 178–202 (2015).
2. SDG. The Sustainable Development Goals Report 2025. (2025).
3. Hochman, Z. Effect of water stress with phasic development on yield of wheat grown in a semi-arid environment. *Field Crops Res.* **5**, 55–67 (1981).
4. Safár, J. Development of chromosome-specific BAC resources for genomics of bread wheat. *Cytogenet. Genome Res.* **129**, 211–223 (2010).
5. Khadka, K., Earl, H. J., Raizada, M. N. & Navabi, A. A physio-morphological trait-based approach for breeding drought tolerant wheat. *Front Plant Sci.* **11**, 715 (2020).
6. Wang, L. Drought in agriculture and climate-smart mitigation strategies. *Cell Rep. Sustain.* **2**, 100386 (2025).
7. Zhang, J. et al. Effect of drought on agronomic traits of rice and wheat: A meta-analysis. *Int. J. Environ. Res. Public Health* **15**, 839 (2018).
8. Ma, J. et al. Transcriptomics analyses reveal wheat responses to drought stress during reproductive stages under field conditions. *Front. Plant Sci.* **8**, 592 (2017).
9. Khan, A. Rising atmospheric temperature impact on wheat and thermotolerance strategies. *Plants* **10**, 43 (2020).
10. Kino, R. I. High post-anthesis temperature effects on bread wheat (*Triticum aestivum* L.) grain transcriptome during early grain-filling. *BMC Plant Biol.* **20**, 1–7 (2020).
11. Ferris, R. Effect of high temperature stress at anthesis on grain yield and biomass of field-grown crops of wheat. *Ann. Bot.* **82**, 631–639 (1998).

12. Prasad, P. V. V. Response of floret fertility and individual grain weight of wheat to high temperature stress: Sensitive stages and thresholds for temperature and duration. *Funct. Plant Biol.* **41**, 1261–1269 (2014).
13. Saini, H. S. Abnormal sporogenesis in wheat (*Triticum aestivum* L.) induced by short periods of high temperature. *Ann. Bot.* **49**, 835–846 (1982).
14. Zhang, W. et al. Low-temperature at booting reduces starch content and yield of wheat by affecting dry matter transportation and starch synthesis. *Front Plant Sci.* **14**, 1207518 (2023).
15. Asseng, S. et al. Rising temperatures reduce global wheat production. *Nat. Clim. Change* **5**, 143–147 (2014).
16. Kebede, B. A. & Getahun, A. Adaptability and stability analysis of groundnut genotypes using AMMI model and GGE-biplot. *J. Crop Sci. Biotechnol.* **20**, 343–349 (2017).
17. Lama, S., Leiva, F., Vallenback, P., Chawade, A. & Kuktaite, R. Impacts of heat, drought, and combined heat–drought stress on yield, phenotypic traits, and gluten protein traits: Capturing stability of spring wheat in excessive environments. *Front. Plant Sci.* **14**, 1179701 (2023).
18. Lesk, C. et al. Compound heat and moisture extreme impacts on global crop yields under climate change. *Nat. Rev. Earth Env.* **3**, 872–889 (2022).
19. Tricker, P. The physiological and genetic basis of combined drought and heat tolerance in wheat. *J. Exp. Bot.* **69**, 3195–3210 (2018).
20. Awasthi, R. et al. Exploring the synergistic effects of drought and heat stress on chickpea seed development: Insights into nutritional quality and seed yield. *Plant Stress* **14**, 100635 (2024).
21. Bapela, T. Genetic improvement of wheat for drought tolerance: Progress, challenges and opportunities. *Plants* **11**, 1331 (2022).
22. Daryanto, S. Global synthesis of drought effects on cereal, legume, tuber and root crops production: A review. *Agric. Water Manag.* **179**, 18–33 (2017).
23. Senapati, N. Drought tolerance during reproductive development is important for increasing wheat yield potential under climate change in Europe. *J. Exp. Bot.* **70**, 2549–2560 (2019).
24. Mekonnen, T. et al. Genome-wide association study reveals novel genetic loci for quantitative resistance to septoria tritici blotch in wheat (*Triticum aestivum* L.). *Front Plant Sci* **12**, 671323 (2021).
25. Govta, N. Genome-wide association study in bread wheat identifies genomic regions associated with grain yield and quality under contrasting water availability. *Int. J. Mol. Sci.* **23**, 10575 (2022).
26. Soumya, P. R. Population structure and genome-wide association studies in bread wheat for phosphorus efficiency traits using 35 K Wheat Breeder's Affymetrix array. *Sci. Rep.* **11**, 7601 (2021).
27. Sidhu, N. S. et al. Genome-wide association studies for next-generation crop breeding in the postgenomic era. *Plant Breed.* **145**, 54–84 (2026).
28. Wu, Y. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome. Biol.* **18**, 86 (2017).
29. Liu, Z. et al. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* **15**, 152 (2015).
30. Wang, R. et al. A novel transcription factor PtMYB77 from *Pinellia ternata* enhances heat tolerance in *Arabidopsis* by inducing early expression of heat shock factor genes. *Ind. Crops Prod.* **227**, 120791 (2025).
31. Yang, A. A R2R3-type MYB gene, OsMYB2, is involved in salt, cold, and dehydration tolerance in rice. *J. Exp. Bot.* **63**, 2541 (2012).
32. Fang, X. Z. NRT1.1 dual-affinity nitrate transport/signalling and its roles in plant abiotic stress resistance. *Front. Plant. Sci.* **12**, 715694 (2021).
33. Singh, P. Genome-wide identification and characterization of subtilisin-like serine protease encoding genes in *Vigna radiata* L. Wilczek. *Sci. Rep.* **15**, 1–16 (2025).
34. Dai, M. Genome-wide analysis of the SBT gene family involved in drought tolerance in cotton. *Front. Plant Sci.* **13**, 1097732 (2023).
35. Peremarti, A. et al. Transcriptomic and proteomic analyses of a pale-green durum wheat mutant shows variations in photosystem components and metabolic deficiencies under drought stress. *BMC Genom.* **15**, 125 (2014).
36. Todingan, A. K., Aprilian Pailang, A. & Hala, Y. DREB1 gene in regulating drought tolerance in cereal plants. *Bioeksperimen: Jurnal Penelitian Biologi* **11**, 16–28 (2025).
37. Nazir, F. et al. Plant hormone ethylene: A leading edge in conferring drought stress tolerance. *Physiol. Plant.* **176**, e14151 (2024).
38. Pan, Y. et al. An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. *Plant Cell Rep.* **31**, 349–360 (2012).
39. Maghraby, A. Genome-wide identification and evolutionary analysis of the AP2/EREBP, COX and LTP genes in *Zea mays* L. under drought stress. *Sci. Rep.* **14**, 7610 (2024).
40. Tang, J. & Bassham, D. C. Autophagy in crop plants: What's new beyond *Arabidopsis* ?. *Open Biol.* **8**, 180162 (2018).
41. Wang, L. F. et al. CAND2/PMTR1 is required for melatonin-conferred osmotic stress tolerance in *Arabidopsis*. *Int. J. Mol. Sci.* **22**, 4014 (2021).
42. Buttar, Z. A. et al. Melatonin suppressed the heat stress-induced damage in wheat seedlings by modulating the antioxidant machinery. *Plants* **9**, 809 (2020).
43. Agarwal, P. Overexpression of PgDREB2A transcription factor enhances abiotic stress tolerance and activates downstream stress-responsive genes. *Mol. Biol. Rep.* https://doi.org/10.1007/s11033-009-9885-8 (2010).
44. Taghipour, A. *et al.* (2016) Expression analysis of five critical transcription factors (TFs) OsbHLH148, OsbZIP72, OsMYB2, OsNAC6 and TRAB1 in response to drought stress in contrasting Iranian rice genotypes **9**, 327–333.
45. Paul, S. Transcript analysis of abscisic acid-inducible genes in response to different abiotic disturbances in two *indica* rice varieties. *Theor. Exp. Plant Physiol.* **31**, 249–272 (2019).
46. Magar, M. M. Genome-wide analysis of AP2/ERF superfamily genes in contrasting wheat genotypes reveals heat stress-related candidate genes. *Front. Plant Sci.* **13**, 853086 (2022).
47. Faraji, S. et al. The AP2/ERF gene family in *Triticum durum*: Genome-wide identification and expression analysis under drought and salinity stresses. *Genes* **11**, 1464 (2020).
48. Hussain, H. A. Interactive effects of drought and heat stresses on morpho-physiological attributes, yield, nutrient uptake and oxidative status in maize hybrids. *Sci. Rep.* **9**, 1–12 (2019).
49. Mackay, T. F. C. Pleiotropy, epistasis and the genetic architecture of quantitative traits. *Nat. Rev. Genet.* **25**, 639–657 (2024).
50. Solovieff, N. Pleiotropy in complex traits: Challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495 (2013).
51. Stearns, F. W. One hundred years of pleiotropy: A retrospective. *Genetics* **186**, 767–773 (2010).
52. Wang, J. et al. *Aegilops tauschii* single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.* **198**, 925–937 (2013).
53. Zhao, X. Population genomics unravels the Holocene history of bread wheat and its relatives. *Nat. Plants* **9**, 403–419 (2023).
54. Ahmed, H. I. Einkorn genomics sheds light on history of the oldest domesticated wheat. *Nature* **620**, 830–838 (2023).
55. Cavalet-Giorsa, E. Origin and evolution of the bread wheat D genome. *Nature* **633**, 848–855 (2024).
56. Abou-Elwafa, S. F. & Shehzad, T. Genetic diversity, GWAS and prediction for drought and terminal heat stress tolerance in bread wheat (Triticum aestivum L.). *Genet. Resour. Crop Evol.* **68**, 711–728 (2021).

57. Sukumaran, S., Reynolds, M. P. & Sansaloni, C. Genome-Wide Association Analyses Identify QTL Hotspots for Yield and Component Traits in Durum Wheat Grown under Yield Potential, Drought, and Heat Stress Environments. *Front. Plant Sci.* **9**, 81 (2018).
58. Khan, A. et al. Identifying the physiological traits associated with DNA marker using genome wide association in wheat under heat stress. *Sci. Rep.* **14**, 20134 (2024).
59. Afzal, F. Genome-wide analyses reveal footprints of divergent selection and drought adaptive traits in synthetic-derived wheats. *G3 Genes|Genomes|Genetics* **9**, 1957–1973 (2019).
60. Ahmed, H. G. Genome-wide association mapping of bread wheat genotypes using yield and grain morphology-related traits under different environments. *Front Genet* **13**, 1008024 (2023).
61. Devate, N. B. Genome-wide association mapping for component traits of drought and heat tolerance in wheat. *Front Plant Sci* **13**, (2022).
62. Guan, P. Global QTL Analysis Identifies Genomic Regions on Chromosomes 4A and 4B Harboring Stable Loci for Yield-Related Traits Across Different Environments in Wheat (Triticum aestivum L.). *Front Plant Sci* **9**, (2018).
63. Qaseem, M. F. Genome-wide association mapping in bread wheat subjected to independent and combined high temperature and drought stress. *PLoS ONE* **13**, 1–22 (2018).
64. Schulthess, A. W. The roles of pleiotropy and close linkage as revealed by association mapping of yield and correlated traits of wheat (*Triticum aestivum* L.). *J. Exp. Bot.* **68**, 4089–4101 (2017).
65. Shao, H., Wang, H. & Tang, X. NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Front Plant Sci* **6**, 902 (2015).
66. Wang, H., Wang, H., Shao, H. & Tang, X. Recent Advances in Utilizing Transcription Factors to Improve Plant Abiotic Stress Tolerance by Transgenic Technology. *Front Plant Sci* **7**, 67 (2016).
67. Hrmova, M. & Hussain, S. S. Plant transcription factors involved in drought and associated stresses. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms22115662 (2021).
68. Jia, J. et al. Comparative transcriptomic analysis reveals the roles of overlapping heat-/drought-responsive genes in poplars exposed to high temperature and drought. *Sci. Rep.* **7**, 43215 (2017).
69. Su, C. F. et al. A novel MYBS3-dependent pathway confers cold tolerance in rice. *Plant Physiol* **153**, 145–158 (2010).
70. Jogawat, A. et al. Crosstalk between phytohormones and secondary metabolites in the drought stress tolerance of crop plants: A review. *Physiol. Plant.* **172**, 1106–1132 (2021).
71. Mittler, R. & Blumwald, E. The roles of ROS and ABA in systemic acquired acclimation. *Plant Cell* **27**, 64–70 (2015).
72. Redhu, M. Assessing alpha lattice design for heat stress indices and yield stability in wheat genotypes. *Sci. Rep.* **15**, (2025).
73. Olczyk, T., Li, Y. & Munoz-Carpena, R. Using Tensiometers for Vegetable Irrigation Scheduling in Miami-Dade County. *EDIS* **2002**, (1969).
74. Sundaram, P. K., Kumar, S., Shivani, S., Kumar, U. & Mondal, S. Tensiometer based irrigation scheduling in wheat (*Triticum aestivum*) in middle Indo-Gangetic plains. *Indian J. Agric. Sci.* **92**, 231–235 (2022).
75. Yadvinder-Singh, Kukal, S. S., Jat, M. L. & Sidhu, H. S. Improving Water Productivity of Wheat-Based Cropping Systems in South Asia for Sustained Productivity. in *Advances in Agronomy* vol. 127 157–258 (Elsevier, 2014).
76. Zadoks, J. C., Chang, T. T. & Konzak, C. F. A decimal code for the growth stages of cereals. *Weed Res.* **14**, 415–421 (1974).
77. Bates, D. "Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* **67**, 1–48 (2015).
78. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. LmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
79. Alvarado, G. *et al.* "META-R (Multi Environment Trail Analysis with R for Windows). CIMMYT Research Data & Software Repository Network, V23 (2016).
80. Wei, T. et al. Package 'corrplot'. *Statistician* **56**, e24 (2017).
81. Robinson, H. F., Comstock, R. E. & Harvey, P. H. Genotypic and Phenotypic Correlations in Corn and Their Implications in Selection[1]. *Agron. J.* **43**, 282–287 (1951).
82. IWGSC. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, (2018).
83. Bradbury, P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
84. Remington, D. L. et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc Natl Acad Sci U A* **98**, 11479–11484 (2001).
85. Stift, M., Kolar, F. & Meirmans, P. G. STRUCTURE is more robust than other clustering methods in simulated mixed-ploidy populations. *Heredity* **123**, 429–441 (2019).
86. Puechmaille, S. J. The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Mol. Ecol. Resour.* **16**, 608–627 (2016).
87. Li, Y. L. & Liu, J. X. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* **18**, 176–177 (2018).
88. Wang, J. GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics Proteomics Bioinforma.* **19**, 629–640 (2021).
89. Huang, M. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience.* **8**, 154 (2019).
90. Johnson, R. C. et al. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics* **11**, 724 (2010).
91. Sandhu, K. S. Comparing performances of different statistical models and multiple threshold methods in a nested association mapping population of wheat. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2024.1460353 (2024).
92. Aury, J. M. Long-read and chromosome-scale assembly of the hexaploid wheat genome achieves high resolution for research and breeding. *Gigascience* https://doi.org/10.1093/gigascience/giac034 (2022).
93. Zimin, A. V. et al. The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *Gigascience* **6**, gix097 (2017).
94. Pang, Y. et al. High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Mol. Plant* **13**, 1311–1327 (2020).
95. DeWitt, N. et al. Sequence-based mapping identifies a candidate transcription repressor underlying awn suppression at the *B1* locus in wheat. *New Phytol.* **225**, 326–339 (2020).
96. Walkowiak, S. et al. Multiple wheat genomes reveal global variation in modern breeding. *Nature* **588**, 277–283 (2020).
97. IWGSC et al. Loci and candidate genes controlling root traits in wheat seedlings—A wheat root GWAS. *Funct. Integr. Genomics* **19**, 91–107 (2019).
98. Muhammad, A. Uncovering genomic regions controlling plant architectural traits in hexaploid wheat using different GWAS models. *Sci. Rep.* https://doi.org/10.1038/s41598-021-86127-z (2021).
99. Rathan, N. Genome-wide association study identifies loci and candidate genes for grain micronutrients and quality traits in wheat (*Triticum aestivum* L.). *Sci. Rep.* https://doi.org/10.1038/s41598-022-10618-w (2022).
100. Vasudevan, A. Lifting of the 1,000 wheat exome project SNPs from *Triticum aestivum* cv. Chinese spring assembly RefSeq v1.0 to RefSeq v2.1. *BMC Res. Notes* https://doi.org/10.1186/s13104-023-06496-8 (2023).

101. Zhang, Z. et al. Integrating high-throughput phenotyping and genome-wide association studies for enhanced drought resistance and yield prediction in wheat. *New Phytol.* **243**, 1758–1775 (2024).
102. Ma, S. et al. WheatOmics: A platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol. Plant* **14**, 1965–1968 (2021).
103. Chen, Y. et al. A collinearity-incorporating homology inference strategy for connecting emerging assemblies in the Triticeae tribe as a pilot practice in the plant pangenomic era. *Mol. Plant* **13**, 1694–1708 (2020).

Acknowledgements
We thank Haramaya University and ICARDA for providing the research facilities for the field experiments and plant materials, respectively.

Author contributions
Conceptualization, investigation, data curation, writing the original draft by A.KB.; funding acquisition by A.KB., W. T., W.M.; supervision by W. T., W.M., G.K., J.C.S., A.K.; data analysis by A.KB., K.A.; editing by J.C.S., W. T., A.K., G.K., W.M., and research materials by W.T.

Funding
The costs for the genotyping and field experiments were funded by ICARDA and Haramaya University, respectively.

Declarations

Competing interests
The authors declare that they have no conflict of interest.

Ethical approval
The genotypes utilized in this study were obtained from the Kulumsa Agricultural Research Centre in Ethiopia, a partner of the ICARDA for bread wheat research. We obtained permissions and rights to use these genotypes for research purposes, and any released varieties were approved by the National Variety Release Committee.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-68092-7.

Correspondence and requests for materials should be addressed to A.K.