scientific reports

Check for updates

OPEN

Identification of quantitative trait nucleotides for grain quality in bread wheat under heat stress

Latief Bashir[1], Neeraj Budhlakoti[2], Anjan Kumar Pradhan[1,6], Divya Sharma[1], Antil Jain[1], Saman Saim Rehman[1], Vishal Kondal[1], Sherry R. Jacob[1], Rakesh Bhardwaj[1], Kiran Gaikwad[3], Dwijesh Chandra Mishra[2], Abhishek Pandey[4], Satinder Kaur[4], Pradeep Kumar Bhati[5], Rakesh Singh[1], Gyanendra Pratap Singh[1] & Sundeep Kumar[1]✉

Heat stress is a critical factor affecting global wheat production and productivity. In this study, out of 500 studied germplasm lines, a diverse panel of 126 wheat genotypes grown under twelve distinct environmental conditions was analyzed. Using 35 K single-nucleotide polymorphism (SNP) genotyping assays and trait data on five biochemical parameters, including grain protein content (GPC), grain amylose content (GAC), grain total soluble sugars (TSS), grain iron (Fe), and zinc (Zn) content, six multi-locus GWAS (ML-GWAS) models were employed for association analysis. This revealed 67 stable quantitative trait nucleotides (QTNs) linked to grain quality parameters, explaining phenotypic variations ranging from 3 to 44.5% under heat stress conditions. By considering the results in consensus to at least three GWAS models and three locations, the final QTNs were reduced to 16, with 12 being novel findings. Notably, two novel markers, AX-94461119 (chromosome 2A) and AX-95220192 (chromosome 7D), associated with grain Fe and Zn, respectively, were validated through Kompetitive Allele Specific Polymerase Chain Reaction (KASP) approach. Candidate genes, including the P-loop-containing nucleoside triphosphate hydrolases (NTPases), Bowman-Birk type proteinase inhibitors (BBI), and the NPSN13 protein, were identified within associated genomic regions. These genes could serve as potential targets for enhancing quality traits and heat tolerance in future wheat improvement programs.

Keywords Heat stress tolerance, Wheat, Quality traits, QTNs, ML-GWAS, KASP validation

Wheat is an important crop in over 40 countries of the world, serving as a primary source of calories (85%) and protein (82%) for a significant portion of the global population[1–3]. Biochemical traits in wheat like GPC, GAC, TSS, grain Fe and Zn content etc. play a vital role in determining its nutritional value, end-use quality, health benefits, breeding potential, disease resistance, environmental impact, and industrial applications[4]. Researchers and breeders continue to focus on studying and improving these traits to meet the growing demands of a global population and ensure a sustainable and nutritious food supply. However, the growth of wheat is greatly impacted by both biotic and abiotic stresses, leading to reduced grain quality and yield. Among these stresses, heat and drought pose a major challenge, affecting wheat production and productivity on a global scale. Climate models predict a potential rise in mean ambient temperature beyond 1.5 °C by the end of the twenty-first century[5]. Wheat thrives within an optimal temperature range of 12 to 22 °C during anthesis and grain-filling stages, and exposure to temperatures beyond this range significantly diminishes grain quality and yield[6–9]. Terminal heat stress occurs when the mean temperature during grain filling exceeds 31 °C, causing a reduction in yield by approximately 3–4% for every 1 °C increase above 28 °C[10].

High temperatures during the grain-filling stage lead to changes in various physiological, biological, and biochemical processes in wheat. Crucial determinants of wheat grain quality include grain protein content and its size[11]. Heat stress during grain filling affects grain protein content by reducing starch deposition, resulting in increased protein concentration due to higher nitrogen content per unit of starch[12–14]. However, the functionality of the increased protein diminishes significantly under heat stress, thereby reducing the overall end-use quality

[1]ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110012, India. [2]ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India. [3]ICAR-Indian Agricultural Research Institute, New Delhi, India. [4]Punjab Agricultural University, Ludhiana, India. [5]Borlaug Institute for South Asia (BISA) & CIMMYT-India, BISA Farm Ladhowal, Ludhiana, Punjab 141008, India. [6]Present address: Louisiana State University, Baton Rouge, LA, USA. ✉email: Sundeep.Kumar@icar.gov.in

of wheat[15]. Furthermore, heat stress tends to reduce glutenin synthesis while maintaining or increasing gliadin synthesis[16]. Additionally, the heat stress positively correlates with zinc and iron contents and total soluble sugars, affecting these quality parameters in wheat[17].

To ensure food security for millions of people worldwide, it is essential to explore untapped genetic diversity and develop climate-resilient cultivars. Developing heat stress (HS) tolerant varieties could be an effective and economical way to mitigate the negative impact of heat stress on crop quality. Wild relatives and landraces serve as valuable sources of heat stress-tolerant genes[18,19]. Moreover, heat-stress tolerance is known to be polygenic and quantitative in nature (Blum, 1988). In recent years, efforts have been made to improve grain quality in wheat and understand the mechanisms involved in heat tolerance through breeding and quantitative trait loci (QTL) mapping studies[20–22]. With advancements in sequencing techniques, there is a vast amount of omics information available for various crops, which can be utilized to identify genes or genomic regions associated with useful traits from germplasm collections in Gene banks.

Conventionally, QTL mapping using bi-parental mapping populations was used to identify genetic variations associated with specific traits. However, this method has limitations due to low allelic diversity and limited recombination events between the crossing parents, resulting in poor mapping resolution[23]. To overcome these limitations, Association Mapping (AM) or Linkage Disequilibrium (LD) mapping has been employed. AM identifies genetic correlations among unrelated individuals by leveraging genotypic-phenotypic interactions based on LD between functional loci and molecular markers. This approach allows the identification of potential candidate genes from associated genomic regions[24–28]. In recent years, the availability of advanced high-throughput genotyping assays like the 90 K SNP array, 35 K SNP array, and various genotyping by sequencing (GBS) methods has enabled researchers to conduct genome-wide association studies (GWAS) in wheat. Some studies, namely[21,22,29,30], have successfully identified several QTLs associated with heat tolerance using the GWAS approach. Furthermore, the application of ML-GWAS models has significantly improved the detection of significant marker-trait associations and the understanding of the genetic architecture of complex traits, surpassing the constraints of traditional single-locus GWAS (SL-GWAS) methods. The various multi-locus GWAS models in the current scenario includes multi-locus random-SNP-effect mixed linear model (mrMLM)[31], FASTmrMLM[32], fast multi-locus random-SNP-effect EMMA (FASTmrEMMA)[33], pKWmEB[34], pLARmEB[35], and Iterative modified-Sure Independence Screening EM-Bayesian LASSO (ISIS EMBLASSO)[36]. These models were implemented in the current study using the R package mrMLM v3.1[37]. mrMLM and FASTmrMLM improve speed and accuracy; FASTmrEMMA optimizes mixed models; pKWmEB and pLARmEB enhance power using kinship and effect estimates; ISIS EM-BLASSO refines variable selection for complex traits using Bayesian LASSO. Recently[38,39], identified the associated genomic reasons using GWAS for grain protein content and starch content in wheat, respectively. Despite these discoveries, there remains a necessity to identify new genomic regions associated with terminal heat tolerance using the LD mapping approach, particularly focusing on naturally occurring variations in landraces and locally adapted wheat cultivars. Due to the highly complex nature of heat tolerance, QTNs identified for specific locations or environments, sometime may not be universally applicable across different environments[30].

Considering the aforementioned points, this study aims to map and validate genomic regions or QTNs linked to grain quality traits in wheat under timely and late sowing conditions, which may confer tolerance to heat stress. This research utilizes a diverse panel of wheat genotypes, primarily from Indian wheat germplasm lines. Identifying these genomic regions is significant, as they could serve as potential targets for wheat breeders working to develop heat-tolerant varieties. The rising frequency and severity of heat stress events due to climate change make the development of such resilient wheat varieties crucial for sustaining production and ensuring food security.

Results

Phenotypic data analysis

Extensive phenotypic variation was observed among 126 bread wheat genotypes while evaluating the GPC, GAC, TSS content, grain Fe, and Zn content. These genotypes were cultivated under both normal and late sown conditions over two growing seasons "2020–2021 & 2021–2022", at three different locations of India thus covering twelve different environments for evaluation of phenotypic data (Tables S2-a,b, S3-a,b and S4-a,b). To further understand the distribution of the data in different environments and stress conditions in relation to various studied traits, same is presented through box plot (Figs. 1a–e). Under normal sown conditions the best linear unbiased estimate (BLUE) values of genotypes for GPC, GAC, TSS, grain Fe and Zn content was 12% (±0.57), 20% (±1.0), 3% (±0.52), 44 ppm (±6.57) and 31 ppm (±2.76) respectively. Contrarily, the BLUE values for the same attributes under late sown conditions were 13% (±0.69), 18% (±1.26), 4% (±0.75), 48 ppm (±6.21) and 37 ppm (±3.08) for GPC, GAC, TSS, grain Fe and Zn content respectively (Table S5). Moreover, the BLUE values of all wheat genotypes for quality attributes was also calculated under different environments (Tables S6-a,b and Fig. S1). The overall impact of HS on the qualitative attributes of wheat under study is summarized in the Table 1.

The analysis of variance (ANOVA), conducted using individual evaluation data from twelve different environments, revealed significant differences ($P<0.0001$) among the genotypes for all the studied traits (Table 2). Furthermore, the broad-sense heritability values for GPC, GAC, TSS, grain Fe, and zinc content were 0.85, 0.83, 0.92, 0.95, and 0.90, respectively, suggesting that the majority of the phenotypic variation was due to genetic factors. The heritability has been calculated using line-mean basis by following the approach of Glimour et al.[40]. In addition to that, correlation analysis among quality traits was conducted for both normal and late sown conditions. It was observed that overall correlations under normal sown conditions were weak. Results suggests that for normal sown condition, there is a positive correlation between grain Fe and Zn content (0.17, $P<0.1$), while GPC exhibits a negative association with GAC (− 0.24, $P<0.01$) and TSS (− 0.11, $P<0.1$) respectively

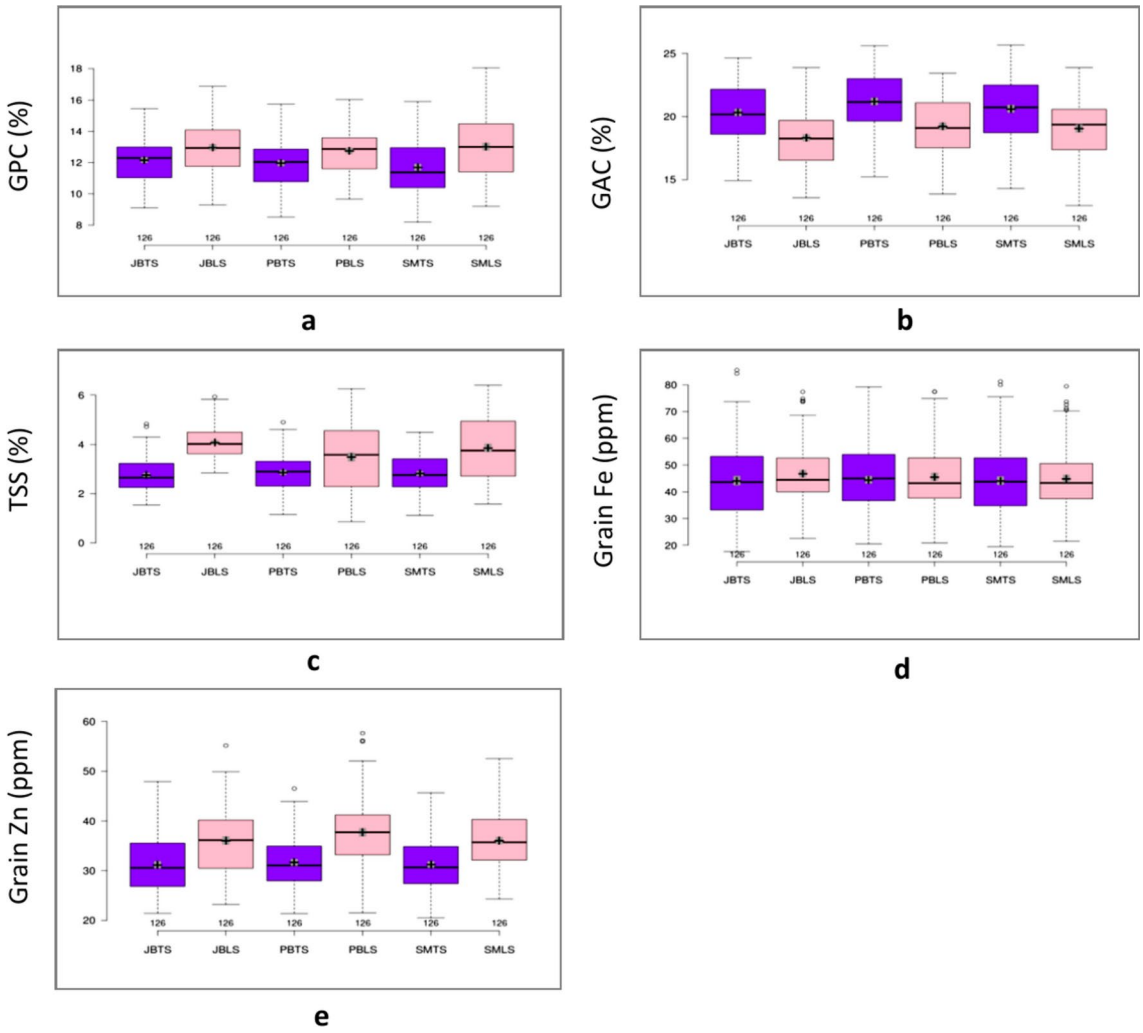

Fig. 1. Box plot of five biochemical traits under timely and late sown conditions: (**a**) GPC, (**b**) GAC, (**c**) TSS, (**d**) Grain Fe content and (**e**) grain Zn content. Each box in the plot represents the average performance under particular environmental condition. *JBTS* Jabalpur timely sown, *JBLS* Jabalpur late sown, *PBTS* Punjab timely sown, *PBLS* Punjab late sown, *SMTS* Samastipur timely sown, *SMLS* Samastipur late sown.

Quality parameter	Effect of heat stress
Grain protein content	+ 1.26%
Grain amylose content	− 2.36%
Total soluble sugars	+ 0.74%
Grain iron content	+ 1.35%
Grain zinc content	+ 6.39%

Table 1. Summary of effect of heat stress on quality parameters of wheat with respect to normal sown conditions; + (increases), − (decreases).

(Fig. S2). Under HS circumstances, there was a significant negative correlation between GPC and GAC (− 0.56, $P < 0.001$), while under the same conditions, there is a positive correlation between grain Fe and Zn content (0.43, $P < 0.001$). Additionally, under HS circumstances, the negative correlation between GPC and GAC (− 0.56, $P < 0.001$) has significantly risen compared to the normal conditions of sowing. Similarly, the positive correlation between grain Fe and Zn content (0.43, $P < 0.001$) has increased significantly under the same conditions (Fig. S3).

Source	DF*	GPC	GAC	TSS	Grain Fe	Grain Zn
Genotype	125	12.491***	28.839***	6.594***	884.501***	193.244***
Location	2	7.107NS	102.173***	7.766***	109.413NS	183.321NS
Year	1	80.275***	3125.755***	34.683***	2121.743***	847.964***
Condition	1	358.128***	1282.079***	375.614***	836.001NS	10,440.459***
Replication	1	0.70NS	2.03NS	0.32NS	1.99NS	1.65NS
Error	1382	3.126	5.633	0.741	102.788	28.024

Table 2 . ANOVA of seed biochemical (quality) traits evaluated at three locations under timely and late sown conditions. **DF* Degree of freedom, *NS* Non-significant, ***Indicates < 0.0001 level of significance. *GPC* grain protein content, *GAC* grain amylose content, *TSS* Total soluble sugar, *Fe* Iron, *Zn* Zinc.

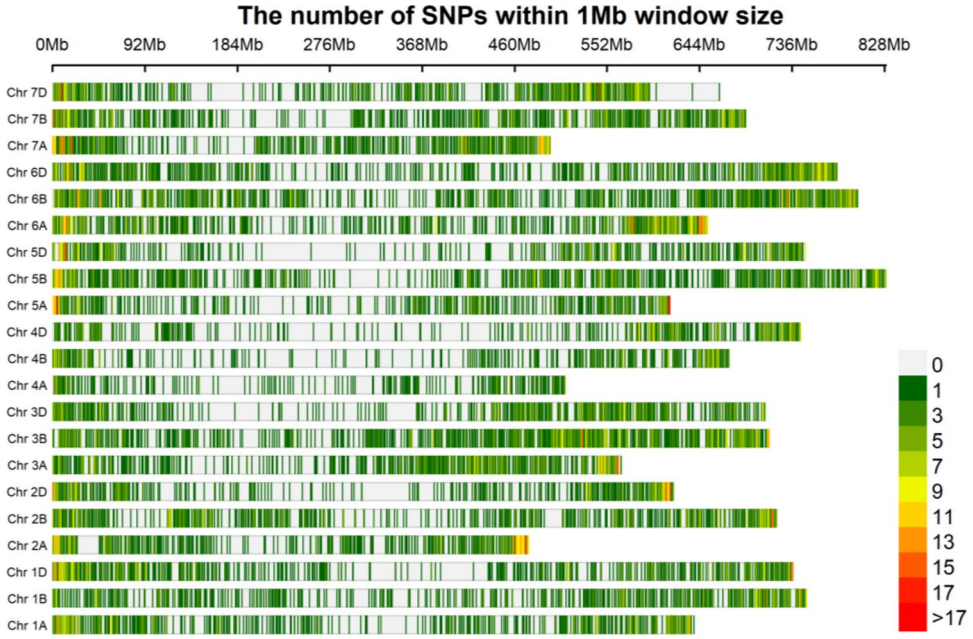

Fig. 2. Marker density map showing distribution of SNP markers within 1 Mb window size across twenty one chromosomes of *T. aestivum*.

Marker distribution and population structure analysis

A 35 K Axiom SNP array was used to genotype 126 wheat genotypes for the AM panel. The genotyping resulted in a total of 35,143 SNPs, which were further filtered for various quality parameters. Minor allele frequency (MAF < 10%), missing genotype and monomorphic SNPs were further filtered as a quality control step, and as a final number 15805 polymorphic SNPs (44.97%) were used for AM (Table S7). For better visualization, the distribution of markers in the current studied panel is presented through marker density map (Fig. 2). The average number of polymorphic SNP markers per chromosome was 752.6, and varied from 287 (Chr4D) to 1140 (Chr2B). Maximum SNPs on sub-genome A were mapped on 2A (940), followed by 1A (823). Similarly, the sub-genome B has the most SNPs on Chr2B (1140), followed by Chr5B (1107) and Sub-genome D's Chr2D had the most SNPs (917), followed by Chr7D (694) (Table S8-a–c).

To determine the population structure of AM panel, plot was generated to represent the relationship between ΔK value and the varying number of hypothetical subgroups K. The highest ΔK value was recorded at K = 2, indicating the maximum membership probability (Fig. 3a). Consequently, the panel consisting of 126 wheat genotypes was classified into two distinct clusters i.e. C1 and C2 (Fig. 3b). Among the 126 genotypes, 61 clustered to cluster 1 (C1- red), and 65 genotypes to cluster 2 (C2-green). In addition, 83% of the genotypes were pure line (similarity > 80%), of which 16% and 84% were EC (exotic collection—genotypes from foreign countries) and IC (indigenous collection—genotypes within India) respectively (Table S1), while remaining 17% of the genotypes were categorized as admixtures in C1. Furthermore, 75% of the genotypes from C2 were pure, with 5% of EC and 95% of IC collections, while remaining 25% were categorized as admixtures. Presence of two clusters among the genotypic panel was also confirmed from the phylogeny analysis (Fig. S4).

Linkage disequilibrium analysis

A total of 15,805 polymorphic SNPs were used to assess the level of LD within the AM panel. The extent of LD between marker pairs was estimated on a chromosome-by-chromosome basis. The sub-genome B comprised the

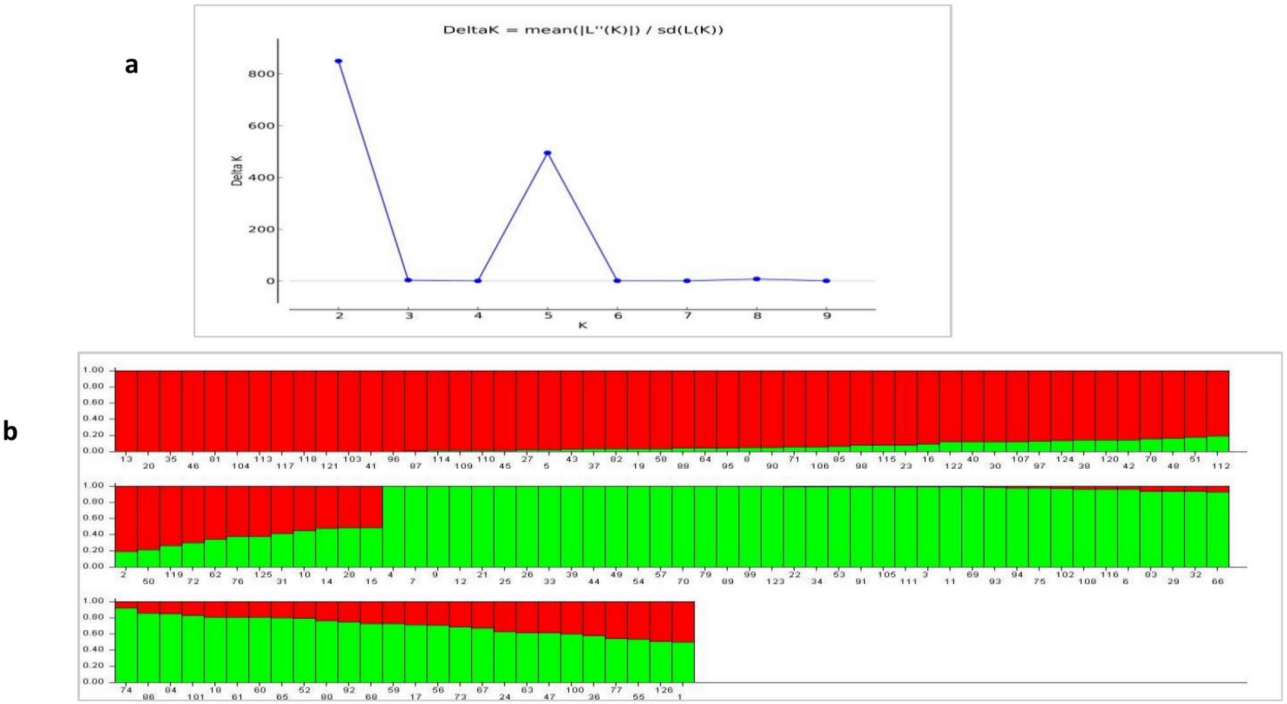

Fig. 3. Structure and kinship analysis. (**a**) Estimated ΔK values for a given K, the ΔK showed steep declined with a clear peak value at K = 2 (**b**) Bar plot of the AM panel showing the details of two clusters (C1 and C2) in different colors, viz. C1 (red, 61 genotypes), C2 (green, 65 genotypes).

most number of polymorphic SNPs (6113, 38.67%), followed by the sub-genome A (5096, 32.26%), and the sub-genome D (4596, 29.07%) (Table S8-a–c). Further, to examine the LD decay, the LOESS (locally weighted linear regression) curve was fitted for each sub-genomes and entire genome. Additionally, a threshold was determined using the confluence of the LD decay curve at $r^2 = 0.17$ based on the fitted model. In sub-genome A, LD decayed the fastest, followed by sub-genomes D and B. LD declined at 3.16 Mb in sub-genome A, as compared to 3.18 Mb in sub-genome D and 4.81 Mb in sub-genome B (Fig. 4a–d). Furthermore, LD decayed for the entire genome at 3.75 Mb.

Association mapping

A significant marker-trait association analysis using six different ML-GWAS models (FASTmrMLM, mrMLM, FASTmrEMMA, ISIS EM-BLASSO, pLARmEB, and pKWmEB) revealed a total of 67 QTNs for five quality traits, with a LOD score of ≥ 3. These QTNs were identified based on the consensus results obtained from at least two ML-GWAS models (Table S9). Among these, 37 QTNs were found to be consistent across three or more models. Out of these, 21 QTNs belongs to the multi-model (≥ 3) and single environment category (Table S10), while 16 QTNs (Fig. 5) were classified under the stable category, representing the multi-model (≥ 3) and multi-environments (≥ 3) (Table 3). The range of r^2 values, spanning from 3 to 44.5%, indicates that the wheat quality traits are influenced by multiple loci with varying magnitudes of impact, ranging from small to moderate effects. Furthermore, QTNs were recorded from all twelve environments under study (Tables S11-a–d, S12-a–d, and S133-a–d).

In addition, the different ML-GWAS models identified a varying number of significant QTNs. Furthermore, FASTmrEMMA, mrMLM, FASTmrMLM, ISIS EM-BLASSO, and pLARmEB models detected 13, 4, 30, 15, and 2 significant QTNs, respectively (Fig. S5). However, the pKWmEB model did not identify any QTN above the threshold. Among the quality traits, grain Fe content exhibited the highest number of reliable QTNs (15) across three or more models, followed by TSS (10), Zn (9), GAC (2), and GPC (2). The identified QTNs were distributed throughout the genome, except for chromosomes 7A and 7B. The highest number of QTNs was mapped to Chromosome 2B (10), followed by 2A (6), 5D (5), 1B (4), 3A (4), 3B (4), 3D (4), 1D (3), 2D (3), 6A (3), 6B (3), 6D (3), 7D (3), 1A (2), 4A (2), 4D (2), 4B (1), 5A (1), and 5B (1).

Stable QTNs for GPC

A total of 10 reliable QTNs associated with GPC were identified within the AM panel on 8 different chromosomal loci viz. 1B, 1D, 2A, 2B, 3A, 3D, 5D and 6D. These QTNs were detected by at least two GWAS models and accounted for 3–28% of the total phenotypic variation (Table S9). Among these QTNs, *Qql.iari-2B.2_pro*, linked to the SNP *AX-95110999* on chromosome 2B, emerged as the most stable QTN for GPC (Fig. 6a). It exhibited the highest LOD score ranging from 3.47 to 6.09 and was consistently identified across four GWAS models (FastMrEMMA, FASTMrMLM, pLARmEB, and ISIS EM-BLASSO) (Table 3). Additionally, *Qql.iari-2B.2_pro* showed consistency in four different HS environments.

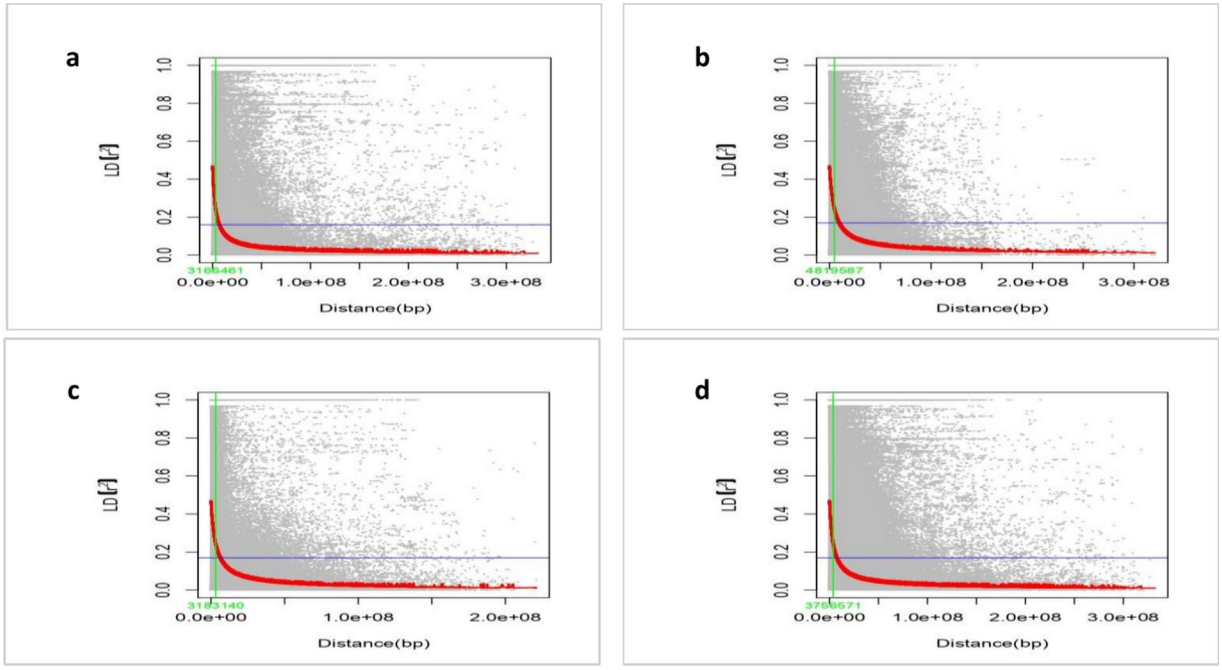

Fig. 4. LD decay plot with LOESS curve in red at $r^2 = 0.17$. (**a**) LD decay plot of sub-genome A, (**b**) LD plot of sub-genome B, (**c**) LD decay plot of sub-genome D and (**d**) LD decay plot of whole genome.

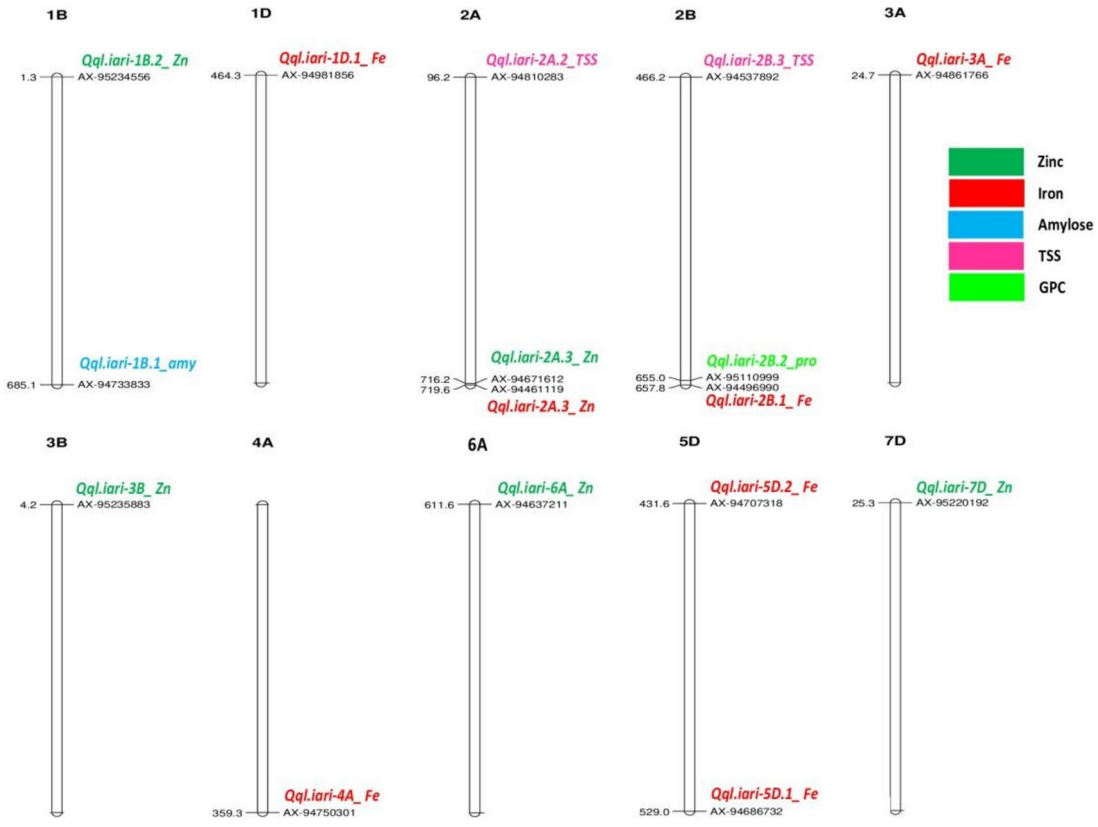

Fig. 5. Physical map for 16 highly stable QTNs. Different colors indicate different quality attributes, i.e., green: grain Zn content, red: grain Fe content, blue: GAC, pink: TSS and fast green: GPC.

QTN	Trait name	Marker*	Chr	Marker position (bp)	LOD score	'– log10(P)'	r^2 (%)	Detection method**	Environments
Qql.iari-1B.1_amy	GAC	AX-94733833	1B	6.85E + 08	3.52–10.57	4.25–4.86	3.87–32.00	1,2,3,4,5	5
Qql.iari-1D.1_ Fe	Fe	AX-94981856	1D	4.64E + 08	3.48–3.52	4.21–4.25	2.83–39.05	2,3,4,5	7
Qql.iari-2A.1_Fe	Fe	AX-94461119	2A	7.20E + 08	3.69–5.91	4.43–6.74	6.12–4.46	2,3,4,5	4
Qql.iari-2B.1_ Fe	Fe	AX-94496990	2B	6.58E + 08	3.08–3.67	3.78–4.41	5.49–11.99	2,3,4,5	5
Qql.iari-3A_ Fe	Fe	AX-94861766	3A	24,748,506	4.01–6.17	4.77–7.01	22.77 -30.24	1,2,3,4,5	5
Qql.iari-4A_ Fe	Fe	AX-94750301	4A	3.59E + 08	4.66–5.80	5.44–6.62	18.64–20.89	2,3,4,5	4
Qql.iari-5D.1_ Fe	Fe	AX-94686732	5D	5.29E + 08	3.10–6.61	3.80–7.46	3.05 -9.88	2,3,4,5	6
Qql.iari-5D.2_ Fe	Fe	AX-94707318	5D	4.32E + 08	3.06–4.30	3.76–5.06	6.85–13.21	2,3,4,5	3
Qql.iari-2B.2_pro	GPC	AX-95110999	2B	6.55E + 08	3.47–6.09	4.20–6.92	5.50–28.93	1,2,4,5	4
Qql.iari-2A.2_TSS	TSS	AX-94810283	2A	96,248,656	3.55–6.55	4.28–7.40	8.09–30.62	2,3,4,5	5
Qql.iari-2B.3_TSS	TSS	AX-94537892	2B	4.66E + 08	3.37–5.80	3.09–6.63	3.62–11.49	2,3,4,5	4
Qql.iari-1B.2_ Zn	Zn	AX-95234556	1B	1,266,231	3.16–4.35	3.86–5.12	5.55–25.06	2,3,4,5	5
Qql.iari-2A.3_ Zn	Zn	AX-94671612	2A	7.16E + 08	3.36–4.47	4.07–5.24	3.36–10.78	2,3,4,5	6
Qql.iari-3B_ Zn	Zn	AX-95235883	3B	4,214,485	3.24–5.61	3.95–6.42	7.71–14.62	2,3,4,5	3
Qql.iari-6A_ Zn	Zn	AX-94637211	6A	6.12E + 08	3.96–7.49	4.71–8.33	21.82–32.56	1,2,3,4,5	4
Qql.iari-7D_ Zn	Zn	AX-95220192	7D	25,273,226	3.04–4.87	3.74–5.65	3.98–12.75	2,3,4,5	6

Table 3. Information of 16 stable QTNs detected for various quality parameters. *Detected by ≥ 3 models and at ≥ 3 environments. **ML-GWAS methods (MrMLM-1, FASTMrMLM-2, FASTMrEMMA-3, pLARmEB-4, ISIS EM-BLASSO-5).

Stable QTNs for GAC

In the AM panel, a total of 7 reliable QTNs were identified, which were linked to GAC. These QTNs were detected by at least two GWAS models, explaining a range of 3.9–44.5% of the overall phenotypic variation (Table S9 and Table 3). The most stable QTN identified for GAC was *Qql.iari.1B.1_amy*, associated with the SNP *AX-94733833* on chromosome 1B (Fig. 6b). This QTN exhibited a highest LOD score ranging from 3.52 to 10.57 for different models and consistently appeared in five GWAS models (MrMLM, FastMrEMMA, FASTMrMLM, pLARmEB, and ISIS EM-BLASSO) (Table 3). Moreover, *Qql.iari-1B.1_amy* consistently displayed its association with GAC across the AM panel tested in five distinct HS environments.

Stable QTNs for grain TSS content

The association analysis revealed the discovery of a total of 16 reliable QTNs (located on 11 chromosomes, namely 1A, 2A, 2B, 2D, 3B, 3D, 5A, 5B, 6B, 6D and 7D) associated with grain TSS content. These QTNs were identified by at least two GWAS models and accounted for a considerable range of phenotypic variation, ranging from 3 to 27% (Table S9). Notably, two stable QTNs of these, namely *Qql.iari-2B.3_TSS* and *Qql.iari-2A.2_TSS*, were linked to grain TSS content (Table 3 and Fig. 6c). These QTNs were further associated with specific SNPs (*AX-94537892* and *AX-94810283*) located on chromosomes 2B and 2A, respectively. Furthermore, both *Qql.iari-2B.3_TSS* and *Qql.iari-2A.2_TSS* QTNs demonstrated the highest LOD score, ranging from 3.55 to 6.55, and consistently appeared in four out of six GWAS models (FastMrEMMA, FASTMrMLM, pLARmEB, and ISIS EM-BLASSO) (Table 3). Additionally, *Qql.iari-2B.3_TSS* and *Qql.iari-2A.2_TSS* consistently showed their association with grain TSS content across the tested association panel in 5 and 4different HS environments respectively.

Stable QTNs for grain Fe content

Among the total identified QTNs, a set of 20 reliable QTNs consistently appeared in at least two GWAS models for grain Fe content (Table S9). These QTNs were distributed across twelve different chromosomal regions, namely 1B, 1D, 2A, 2B, 3A, 3D, 4A, 4B, 4D, 5D, 6A, and 6B. They accounted for a phenotypic variation ranging from 3 to 39% (Table S9). Notably, most of the identified QTNs were located on chromosome 5D and explained 3.05–27.84% of the phenotypic variance. Eight QTNs (*Qql.iari-2A.1_Fe, Qql.iari-2B.1_Fe, Qql.iari-1D.1_Fe, Qql.iari-5D.1_Fe, Qql.iari-5D.2_Fe, Qql.iari-4A_Fe, Qql.iari-3A_Fe,* and *Qql.iari.1D.2_Fe*) exhibited high consistency across four GWAS models (FastMrEMMA, FASTMrMLM, pLARmEB, and ISIS EM-BLASSO). These QTNs were associated with specific SNPs (*AX-94461119, AX-94496990, AX-94981856, AX-94686732, AX-94707318, AX-94750301, AX-94861766,* and *AX-94981856*) located on chromosomes 2A, 2B, 1D, 5D, 5D, 4A, 3A, and 1D, respectively (Table 3). Moreover, all eight QTNs consistently demonstrated their association with grain Fe content across the tested AM panel in at least three different HS environments. Notably, *Qql.iari-5D.1_Fe* and *Qql.iari-3A_Fe* displayed the highest LOD scores, ranging from 3.10 to 6.61 and 4.01–6.17, respectively (Fig. 6d).

Stable QTNs for grain zinc content

A total of 14 QTNs showed reliable associations with grain Zn content, as detected by at least two models. These QTNs were located on various chromosomal regions, including 1B, 2A, 2D, 3B, 4A, 4D, 5D, 6A, 6B, and 7D. They explained phenotypic variation ranging from 1.24 to 31.51% (Table S9). Among these QTNs, five (*Qql.iari-6A_Zn, Qql.iari-2A.3_Zn, Qql.iari-7D_Zn, Qql.iari-1B.2_Zn,* and *Qql.iari-3B_Zn*) consistently appeared in four or five GWAS models. These QTNs were associated with specific SNPs (*AX-94637211, AX94671612, AX-*

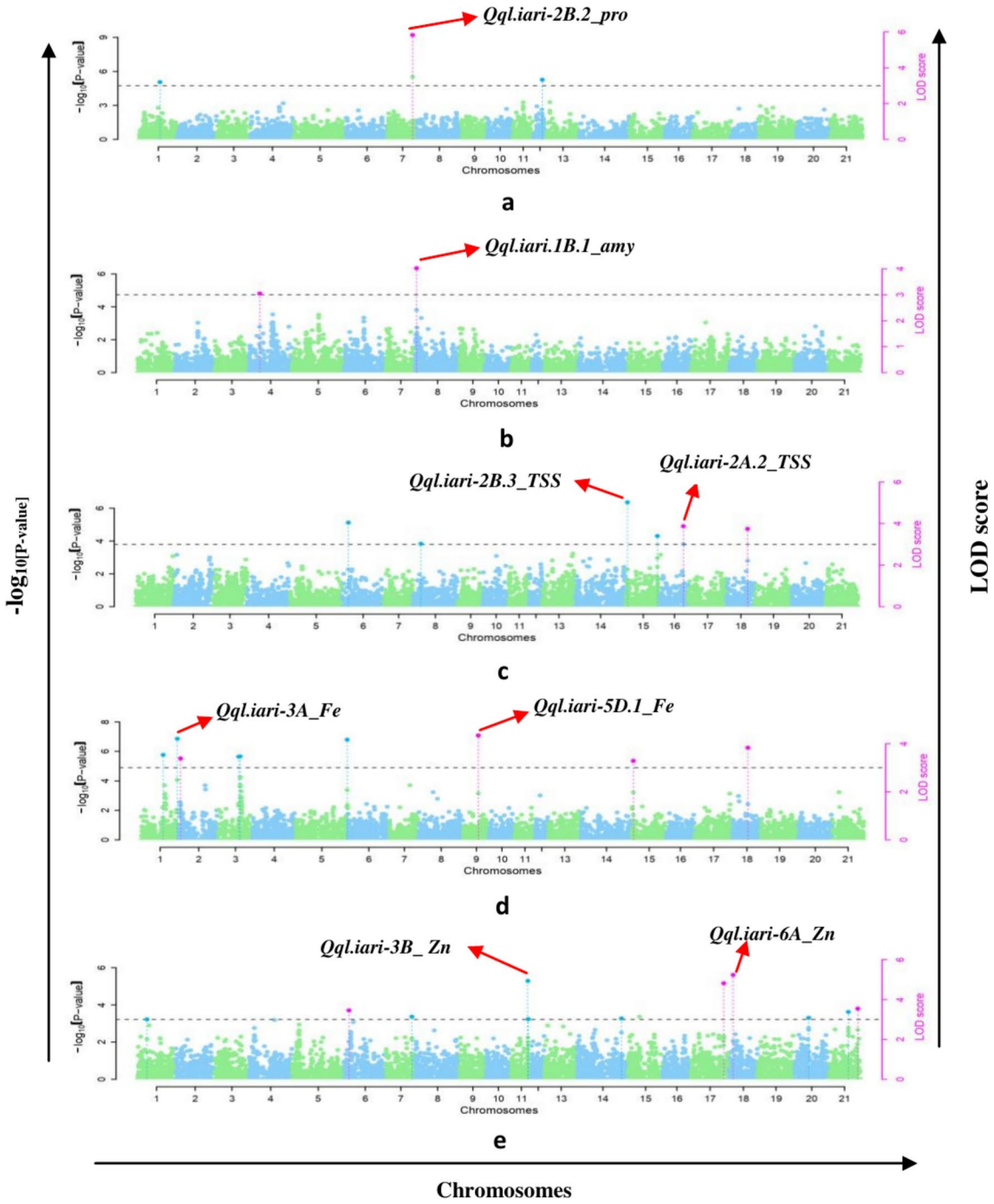

Fig. 6. Manhattan plot of stable QTNs for seed quality (a) GPC (b) GAC, (c) grain TSS, (d) grain Fe content and (e) grain Zn content.

95220192, AX-95234556, and *AX-95235883*) located on chromosomes 6A, 2A, 7D, 1B, and 3D, respectively (Table 3). Moreover, all five QTNs consistently demonstrated their association with grain Zn content in at least three different heat stress environments. Notably, the QTN *Qql.iari-6A_Zn* was detected by five GWAS models (MrMLM, FastMrMLM, FastMrEMMA, pLARmEB, and ISIS EM-BLASSO), indicating it as a highly reliable locus for grain Zn content under HS. It exhibited the highest LOD score ranging from 3.96 to 7.49, across the different GWAS models (Fig. 6e).

Annotation of putative candidate genes linked to QTNs

QTNs that demonstrated a high level of significance and consistency, were annotated due to their perceived reliability and ability to perform effectively in various environments. To identify potential candidate genes, SNPs (probe sequences) significantly linked to quality traits were cross-referenced with the *Triticum aestivum*

genome assembly IWGSC-refseq version 1.0 using the online web resource Ensemble plants (https://plants.ensembl.org/Triticumaestivum/Tools/Blast). To further validate and enrich our results, we conducted an additional Gene Ontology (GO) annotation analysis using the DAVID online tool. This cross-validation approach has helped to reinforce our findings and provide a more comprehensive understanding of the candidate genes' functions (https://david.ncifcrf.gov/summary.jsp) (Table S14). Sixteen stable QTNs (detected by ≥ 3 models at ≥ 3 locations) were subjected for annotation for their potential role/function, of which twelve markers were successfully annotated (Table 4).

KASP validation of identified QTNs

KASP primers (Table S15) have been designed from 51 QTNs of which 33 found to be polymorphic on panel of 101 lines consisting of resistant and susceptible lines (Tables S16-a,b). The panel has been procured from BISA, Ludhiana. The two pivotal KASP markers, denoted as AX-94461119 and AX-95220192, for grain Fe and Zn content respectively, were found to be associated with respective QTNs in this panel (Fig. 7). These markers, situated on distinct chromosomal locations, namely 2A and 7D, respectively, manifested remarkable statistical significance in the context of identified QTN concerning HS. Notably, the alternative alleles of these markers exerted discernible and substantial impacts on the phenomenon of HS, signifying tolerance.

In the subsequent validation phase, the study meticulously identified specific allelic variants correlated with the attributes of heat resistance or tolerance. Allelic variant G of marker AX-95220192, exhibiting a p-value < 0.001, as well as allelic variant C of marker AX-94461119, demonstrating a p-value < 0.001, emerged as prominently associated (gain Fe and zinc) to HS tolerance (Fig. 8).

Discussion

Ensuring global food security for millions of people requires urgent attention towards exploring untapped genetic diversity to develop climate-resistant crop varieties. Developing heat stress-tolerant cultivars can be an effective and cost-efficient approach to counter the negative impact of high temperatures on crop quality. Valuable genes and QTNs associated with heat stress tolerance can be found in wild relatives and landraces[18,19]. In recent years, the advancement in sequencing techniques has provided a wealth of omics information for various crops. This data can be leveraged to identify useful genes or genomic regions linked to beneficial traits from the vast germplasm collections preserved in Gene banks.

We observed a significant positive association between grain Fe concentration and grain Zn content under normal sown conditions which in agreement with the results of earlier study[17]. The uptake of Fe and Zn in wheat is closely linked, and the transport of one micronutrient can affect the uptake of the other. This may be due to the fact that both Fe and Zn share common transporters in the plant, which can result in their co-uptake and co-transport[17]. Therefore, under heat stress, the up-regulation of genes responsible for micronutrients transportation can result in a positive association between Fe and Zn in wheat. However, there was a significant negative correlation of GPC with grain amylose content and TSS. Studies have shown that the significant negative correlation between GPC, GAC, and TSS is under high nitrogen availability and/or genotypes with a high potential for protein accumulation. Therefore, managing nitrogen availability during grain development can be a key strategy to balance the allocation of resources between different metabolic pathways and maintain desirable grain quality traits, including GAC and TSS. These results are consistent with those of earlier studies[12–14]**.** All wheat grain quality parameters respond the same manner under late-sown conditions as they do under normal-sown ones. In contrast to normal-sown situations, the degree of association between the quality parameters in late-sown habitats has increased viz. negative correlation between GPC and GAC has dramatically increased under HS conditions. On the other hand, the positive association between grain Fe and Zn concentration has significantly enhanced under such circumstances. All these above results were in agreement with the earlier studies[12–14,17].

Trait	Marker	Chr	Gene	Start (pos)	End (pos)	Description/annotation
Fe	AX-94981856	1D	TraesCS1D02G395900	464,318,467	464,321,648	Protein of unknown function DUF2921
Fe**	AX-94461119	2A	TraesCS2A02G484500	719,565,499	719,566,884	Phylloplanin-like
Fe	AX-94496990	2B	TraesCS2B02G463800	657,790,256	657,793,959	Basic helix-loop-helix transcription factor, Regulator of tape*t al* programmed cell death, Male reproductive development
Fe	AX-94861766	3A	TraesCS3A02G046300	24,748,236	24,748,973	Bowman-Birk type proteinase inhibitor
Fe	AX-94750301	3D	TraesCS3D02G257000	359,318,802	359,324,276	Suppressor of white apricot
Fe	AX-94686732	5D	TraesCS5D02G500300	528,973,078	528,978,054	PTAC2
Fe	AX-94707318	5D	TraesCS5D02G343500	431,603,239	431,608,266	Tetratricopeptide-like helical domain superfamily
Zn	AX-95235883	3B	TraesCS3B02G008300	4,214,230	4,214,757	Invertase/pectin methylesterase inhibitor domain superfamily
Zn	AX-94637211	6A	TraesCS6A02G406100	611,575,953	611,577,930	Protein of unknown function DUF2921
Zn**	**AX-95220192**	7D	TraesCS7D02G049000	25,270,045	25,276,786	Oligopeptide transporter, OPT superfamily
GPC	AX-95110999	2B	TraesCS2B02G461500	655,014,129	655,017,199	P-loop containing nucleoside triphosphate hydrolase
TSS	AX-94810283	2A	TraesCS2A02G149700	96,248,330	96,251,720	NPSN13

Table.4.. Annotation of putative candidate genes linked to quantitative trait nucleotides (QTNs) of grain quality parameters. **Validated through KASP approach locations).

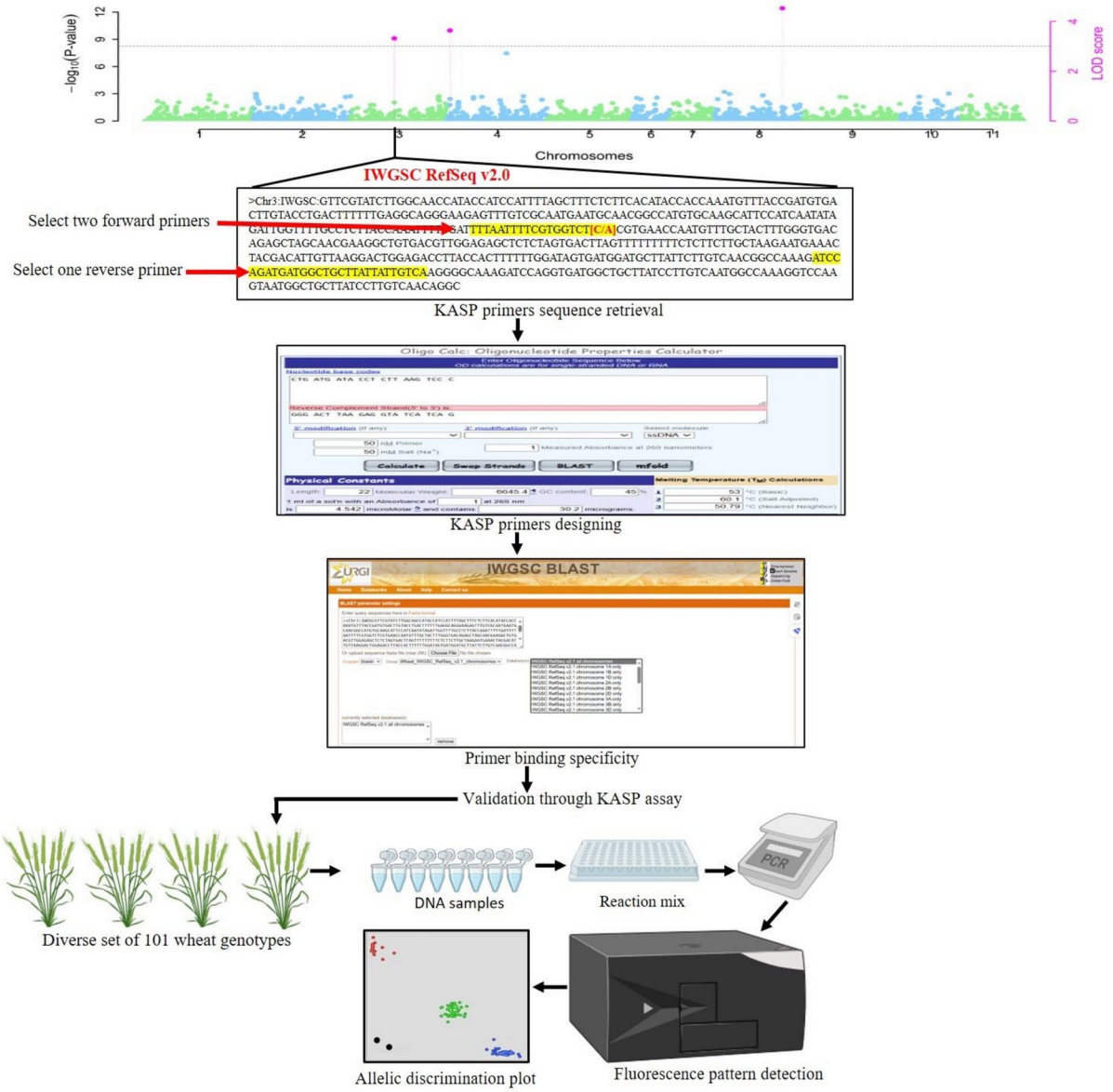

Fig. 7. Systematic representation of the workflow used to validate identified QTNs associated with the grain quality traits under heat stress through KASP assay.

After performing all the quality control to the raw SNP data obtained through 35 K SNP Axiom array, we finally proceed with a total 15,805 polymorphic SNPs for association mapping analysis. A total 15,805 polymorphic SNPs used as a final no. in the analysis, of which most SNPs (6113) are located in sub-genome B, followed by sub-genome A (5096 SNPs) and sub-genome D (4596 SNPs). To proceed for association mapping analysis, genotype panel was assessed for presence of population structure which is the one of the crucial factor as it minimizes the spurious associations. Population structure analysis revealed the presence of two clusters (k = 2) for the current genotypic panel i.e. C1 (61 lines) and C2 (65 lines), where among the clusters, genotypes can be further categorized as pure or admixture i.e. C1 (82% pure and remaining 18% admixtures) & C2 (75% pure and remaining 25% admixtures). Further, relation among the genotypes was also examined through phylogenetic analysis.

LD is another important factor that always needs to be taken into account while AM, particularly when identifying the range of highly associated SNPs. A high LD value indicates that fewer markers are required to cover the genome than a low LD value, which is useful to determine the number of loci necessary for the whole-genome scan[26,41]. In the present study, population structure in designated wheat germplasm increased the LD values because of admixtures in both clusters, resulting in numerous distinct loci displaying significant LD. LD decayed the fastest in sub-genome A, followed by sub-genomes D and B. The faster decay of LD in the A sub-genome of wheat is likely due to a combination of factors related to recombination rates, population history, and gene flow, which was in agreements of results from previous investigations[42,43].

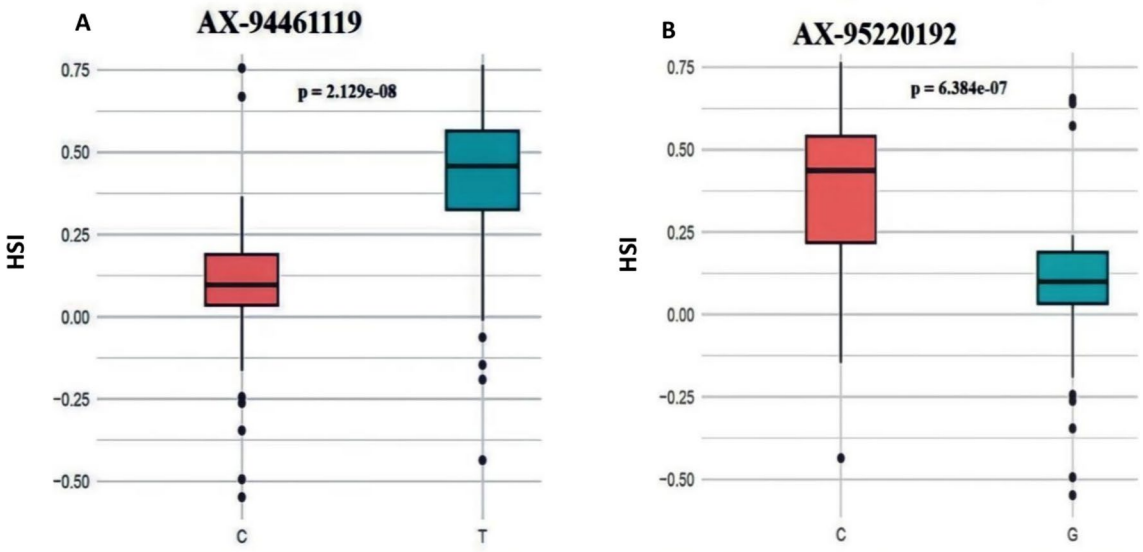

Fig. 8. Kruskal–Wallis test confirms significant Heat Susceptibility Index (HSI) differences between alleles of two KASP validated markers (**A**) AX-94461119, for grain iron content and (**B**) AX-95220192, for grain zinc content.

In current study, six different GWAS models (mrMLM, FASTmrMLM, FASTmrEMMA, pLARmEB, ISIS EM-BLASSO, and pKWmEB) were used to examine the marker-trait association to find out markers associated to the target trait[31,32,34,35,42]. Given that all-marker effects are concurrently assessed in ML-GWAS models, these are thought to be superior to SL-GWAS methods for mapping of complex traits[42,44,45] and hence were utilized for the current GWAS study. A total of 67 reliable QTNs for five quality parameters with a LOD score ≥ 3 were found in significant marker-trait association. The R^2 values of the QTNs ranged from 3 to 44.5%, demonstrating the large range of phenotypic variation seen for these parameters (Table S9). The finding suggests that quality traits are influenced by numerous loci with relatively modest effects, highlighting the intricate genetic regulation of these traits in the early stages of crop development. Additionally, among the six ML-GWAS models used in the analysis, the FASTmrMLM model demonstrated the highest number of associations, whereas the pLARmEB model revealed fewer associations above the threshold. To best of our knowledge there is hardly any study which reported the QTNs for quality parameters especially under HS regime. In the current study, we attempted to identify the novel QTNs for the quality traits under both in normal and late sown conditions. Fortunately, we were successful in identifying 12 annotated and stable QTNs and their respective candidate genes (Table 4).

Current study has identified a novel QTN for GPC (*Qql.iari-2B.2_pro*), which was further localized within the gene *TraesCS2B02G461500,* codes for *P-loop containing nucleoside triphosphate* hydrolase (NTPases). In wheat, P-loop containing NTPases are involved in many biological functions such as protein synthesis, plant growth, and stress response. These enzymes hydrolyze nucleoside triphosphates (NTPs) to nucleoside diphosphates (NDPs) and inorganic phosphate (Pi), releasing energy in the process. This energy can be used to drive cellular processes such as DNA replication or protein synthesis. P-loop containing NTPases in wheat also play a crucial role in plant growth and development by regulating cell division and differentiation[46,47]. For TSS SNP markers *AX-94537892* and *AX-94810283*, associated with QTNs *Qql.iari-2A.2_TSS* and *Qql.iari-2B.3_TSS* respectively, were found to be more reliable. Among these two, *AX-94537892* was located within the gene, *TraesCS2A02G149700*, which encodes the *NPSN13 protein*, a type of SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptors) protein that regulates vesicle trafficking. This protein is abundant in higher plant species, suggesting that its role in growth and development of wheat during heat stress[48]. Additionally, this gene is also responsible for vesicle-mediated resistance to stripe rust and powdery mildew in wheat[49]. Apart from this, NPSN13 has also been reported in *Arabidopsis* and rice (*Oryza sativa*) genomes[50–52].

Grain Fe and Zn are important micronutrients for wheat growth and development[53]. We have identified 7 novel QTNs in relation to grain Fe content in wheat (Table 4). The most potent SNP marker (*AX-94861766*) located on chromosome 3A, associated with the QTN *Qql.iari-3A_ Fe*, and with the gene *TraesCS3A02G046300*, which is responsible for coding the *Bowman-Birk type proteinase inhibitor (BBI)*. In wheat, BBI has a crucial function in regulating endogenous proteases during grain development and is induced in response to biotic and abiotic stresses, suggesting a role in plant defense mechanisms[54]. The other proteins that has been encoded by rest of the genes for grain iron content include, *phylloplanin-like, basic helix-loop-helix transcription factor* (regulator of tapetal programmed cell death, male reproductive development), *PTAC2* and *tetratricopeptide-like helical domain superfamily*. In addition, we have also pinpointed four QTNs related to grain Zn content. The QTN with the highest potency (*Qql.iari-6A_ Zn*) was associated with the marker *AX-94637211* as this particular marker was detected by five out of six models and located on gene *TraesCS6A02G406100*. This gene encodes a protein with an unknown function, *DUF2921*, while the other genes encode proteins belonging to the *invertase/ pectin methylesterase inhibitor domain superfamily* and *Oligopeptide transporter, OPT superfamily*.

The present investigation has successfully validated two potential SNP markers, namely AX-94461119 and AX-95220192, using the KASP approach. Notably, this study represents the first instance of the AX-94461119 marker being linked to grain Fe content in bread wheat which is located on chromosome 2A and is correlated with the gene TraesCS2A02G484500, responsible for encoding Phylloplanin-like proteins. Similarly, our research marks the initial documentation of the KASP marker AX-95220192 being associated with grain Zn content in bread wheat. This marker is positioned on chromosome 7D and is linked with the gene TraesCS7D02G049000, which codes for an Oligopeptide transporter within the OPT superfamily. Recent studies have indicated the significant involvement of peptide transporters (OPT family) in crucial roles during both biotic and abiotic stresses, such as heat stress[55]. This suggests their importance in maintaining wheat quality under adverse environmental conditions. The identified KASP marker holds promise for successful utilization in genetic introgression strategies to enhance or preserve grain zinc content in bread wheat under challenging environmental conditions. It can be employed in the development of heat-tolerant (HT) cultivars through pre-breeding or genetic enhancement, as well as in creating HT bio-fortified varieties by introgressing QTNs/genes from heat-tolerant genotypes to cultivated ones.

Our research identified a total of 67 reliable QTNs for various grain quality parameters in bread wheat under HS. These QTNs were distributed across most of the wheat chromosomes, with the exception of a few chromosomes like 7A and 7B, for the five different quality traits. Prior studies have also pinpointed several quality genes/QTNs scattered throughout almost all wheat chromosomes[56,57]. In our study, we compared the locations of the identified QTNs with those from previous research. However, some comparisons proved challenging due to differences in marker platforms, mapping populations, and the absence of a consensus map for position comparison. Below, we discuss the novel associations for grain quality traits found in this study, alongside those previously reported.

For GAC, ML-GWAS detected 7 reliable QTNs. Notably, *Qql.iari-1B.1_amy* (685.09 Mb) and *Qql.iari-1D_amy* (419.73 Mb), located on chromosomes 1B and 1D, respectively, are in proximity to the markers, *QGAsC1B*[57] and *QGAsC1D*[57]. Regarding GPC, we found four QTNs, *Qql.iari-1B_pro* (5.58 Mb), *Qql.iari-2A_pro* (715.41 Mb), *Qql.iari-3D_pro* (305.89 Mb), and *Qql.iari-3D.2_pro* (426.48 Mb) on chromosomes 1B, 2A, and 3D, respectively, which were close to QTLs, *WMC419*[58], *QGPC1B*, *QGPC3A*, and *QGPC3D*[57]. Furthermore, the QTN *Qql.iari-6D_pro* (363.26 Mb) for GPC reported on chromosome 6D was found to be closely associated with the QTN *wsnp_Ex_c1988_3742022*[38]. Additionally, three QTNs related to grain Fe content, *Qql.iari-3B_Fe* (0.219 Mb) located on chromosome 3B, *Qql.iari-4D_Fe* (359.11 Mb) on chromosome 4D, and *Qql.iari-6A_Fe* (560.20 Mb) on chromosome 6A were found to be in close proximity to identified QTLs from previous studies, specifically *QFe.caas-3BL*, *QFe.caas-4DS*, and *QFe.caas-6AS*[59]. Furthermore, five QTNs associated with grain Zn content, *Qql.iari-2A_Zn* (572.81 Mb) on chromosome 2A, *Qql.iari-3B.1_Zn* (21.38 Mb), and *Qql.iari-3B.2_Zn* (4.21 Mb) both on chromosome 3B, *Qql.iari-4D_Zn* (375.50 Mb) on chromosome 4D, and *Qql.iari-6A_Zn* (611.57 Mb) on chromosome 6A were found to be in vicinity of QTLs, *QZn.caas-2AS*, *QZn.caas-3BS*, *QZn.caas-4DS*, and *QZn.caas-6AS*, respectively[59]. However, when it comes to grain TSS, there is hardly any identified QTL or QTN in wheat associated with HS tolerance.

The ultimate goal of our current study was to identify and validate QTNs in selected wheat germplasm for quality traits under HS conditions. The findings of our study can be utilized variously in crop improvement programs. Some of the potential benefits for future prospectus of crop improvement programs are worthy to mention here: Utilization of validated QTNs in marker assisted breeding, value addition to national gene bank by conserving identified trait-specific wheat germplasm genotypes, development of HT cultivars through pre-breeding or genetic enhancement and development of HT bio-fortified varieties by introgression of Fe & Zn QTLs from HT genotypes to cultivated ones.

Materials and methods

Plant material, experimental design and trait evaluation

A total of 500 wheat germplasm lines were accessed for phenotypic evaluation out of which 126 diverse genotypes were used for quality assessment. The diverse lines (126) were selected from the main panel (500) on the basis of their superior biochemical profile. The material mainly consists of landraces, and few exotic genotypes etc. as detailed in Table S1. The experimental material was grown at three distinct locations in India, operating under the Borlaug Institute for South Asia (BISA). These locations include the BISA farm at Pusa, Bihar (25.9599°N, 85.6696°E), falling under the North East Plain Zones (NEPZ); BISA Jabalpur (23.2241°N, 80.0709°E), situated in the Central Zone (CZ); and BISA Ludhiana (31.267°N, 75.7686°E), located in the North West Plain Zones (NWPZ).

The experimental trials were conducted for two consecutive years, "2020–2021 & 2021–2022" *Rabi* season and each season includes two sowing regimes. The first regime was the normal sowing, carried out in mid-November under non-stress conditions. The second regime involved late sowing, conducted in mid-December, which subjected the plants to HS. Overall, the evaluation of the experimental material was carried out under twelve different environments, encompassing all possible combinations of growing conditions, locations, and years. The evaluation of genotypes for heat tolerance was conducted based on the results from the 1st year trial under all environments. The experimental design utilized in this study was an alpha lattice, with each plot measuring 3.024 square meters (2.8 meters x 1.08 meters). The plots were organized into six rows with a row-to-row spacing of 18 cm. Throughout the experiment, five checks were used, namely RAJ3765 (CK1), DBW107 (CK2), DBW71 (CK3), HD2932 (CK4), and PBW771 (CK5). Each genotype from all twelve environments (combinations of growing conditions, locations, and years) was replicated twice.

Seed was broadcasted, for normal sown, sowing was conducted in the 2nd week of November and in the 2nd week of December for the late-sown conditions. The maximum and minimum temperatures were recorded at different growth stages of wheat under both growing regimes. Standard cultural practices were followed

throughout the crop season, and two sprays of systemic fungicide, Tilt (Propiconazole 25% EC), were applied at 40–45 days after sowing and 70–75 days after sowing for disease protection. Harvesting took place in the first week of April for the normal sown crop and the first week of June for the late-sown crop. Seeds were separated from the plants using mechanical threshers. The same set of practices was maintained for both crop seasons. After threshing and drying, a total of 3024 samples were collected for further quality evaluation. These samples comprised 126 genotypes, each replicated twice, under two conditions (normal and late sown), at three different locations, and across two years (2020–2021 and 2021–2022).

Phenotyping for wheat quality traits

In the present study, five different parameters related to wheat quality were investigated. These parameters included GPC, GAC, TSS, and the grain Fe and Zn content. To measure these traits, standard biochemical procedures were employed. The experiment involving analysis of GPC, GAC, and TSS was carried out in two main steps. Firstly, near-infrared spectroscopy (NIRS) was used to acquire spectra from 126 genotypes across multiple years and locations. Based on the analysis of the NIR spectral data, a diverse set of samples (95) with varying biochemical properties was selected using Statistical Package for Social Sciences (SPSS) software. The second step involved evaluating various biochemical traits for the selected wheat samples.

For the NIRS spectra acquisition, grains were scanned using a monochromator-based FOSS NIRS 6500 instrument, and the reflectance spectra (Log 1/R) were recorded in the range of 400 to 2500 nm. The spectra were then analyzed using the Global program in WINISI software version 3.10 (Infrasoft International, Port Matilda PA, USA). Subsequently, the selected wheat samples were converted into flour and homogenized using a cyclotech mill. Reflectance spectra data were recorded for the homogenized samples, which were later utilized in developing calibration equations and prediction models for detecting GPC, GAC and TSS in wheat. The homogenized flour was used for analyzing total GPC, TSS, minerals (Fe and Zn), and GAC content. Calibration equations for NIRS were developed by correlating the spectral data with the results obtained from actual chemical analysis. To validate the spectral data, laboratory values were compared with the predicted ones.

Additionally, to confirm the accuracy of the predicted values obtained from the spectral data, standard laboratory tests were conducted for each parameter. The GPC was estimated using the Dumas method (ISO/TS 16634-2:2009) for protein estimation. The TSS was extracted and estimated using the anthrone-sulfuric acid reagent method (Du Bois et al. 1956), employing the principle of spectrophotometry. For estimating the GAC, the potassium-iodine (K2-I2) method[60] was employed. The estimation and extraction of minerals (grain Fe and Zn) were carried out using the atomic absorption spectrometry (AAS) method[61].

Statistical analysis of the phenotypic traits

To assess the significance of the data across all environments, an ANOVA was conducted using SAS v9.4. Additionally, the interrelationships among the biochemical traits were examined through correlation analysis i.e. Pearson correlation coefficient. In order to gain a better understanding of the distribution of phenotypic values for the traits, BLUE values were calculated for each trait. These statistical analyses were carried out using R (R Core Team, 2020)[62]. Additionally, the broad-sense heritability of the quality traits was calculated using the restricted maximum likelihood method by following the approach of Glimour et al.[40].

DNA extraction and SNP genotyping

DNA extraction was performed using the CTAB method. The genotyping process utilized a 35 K SNP Axiom array, which was developed through the GBS approach. To ensure the quality of DNA, NanodropTM 2000 (Thermo Fisher Scientific, Wilmington, DE, USA) was used, and only samples with good quality DNA were chosen for genotyping. For the genotyping step, the Axiom® Wheat Breeders' Array (Thermo Fisher Scientific, Wilmington, DE, USA) was employed, following the procedure outlined by Affymetrix (Axiom® 2.0 Assay for 384 samples P/N 703,154 Rev. 2). To ensure reliable data, SNP markers with more than 10% missing data, less than 10% MAF, and those with monomorphic alleles were excluded from the analysis. After filtering, a total of 15,805 polymorphic SNPs remained, which were then utilized for AM and other downstream analyses.

Population structure, kinship and linkage disequilibrium analysis

Presence of sub-groups among the current wheat AM panel was assessed using Bayesian inference program STRUCTURE v2.3.4[63]. The STRUCTURE program is a model based clustering method which identifies the sub-groups (i.e. K value) among genotypic panel using the information from the genotypic data, for this purpose an admixture model was used. Initially, 10,000 burn-in length and 20,000 Monte Carlo Markov chain replicates were provided as initial parameters, and 10 runs for each value of K between 1 and 10 were performed. Finally, using the methodology developed by[64], which is implemented through web utility named STRUCTURE HARVESTER[65], the number of clusters (K value) was calculated. Further, kinship matrix was also estimated using Tassel v5.0[66]. The genotypes were then categorized into distinct subgroups based on the highest membership probability within each subgroup. All this information like kinship and population structure was further used as auxiliary information in the association mapping analysis.

LD is the non-random co-segregation of alleles at two or more loci. It can be measured as the correlation (r^2) frequency among allele pairs across a pair of markers. LD was investigated using TASSEL v5.0[66]. To assess the significance of pair-wise LD, 1000 permutations were conducted. Furthermore, LD decay distance across the sub-genomes and whole genome was estimated by plotting the scatterplot of LD r^2 values between marker pairs and the physical distance.

Genome-wide association analyses

In many GWAS studies, the single locus model is commonly used, relying on single marker associations at a time throughout genome-wide scans. However, these models have certain drawbacks, such as repeated test adjustments for critical values and not considering the overall impact of multiple loci[44]. This is due to their one-dimensional genome scan nature. To address this issue, ML-GWAS models have been proposed, which simultaneously take input from multiple loci. Utilizing such multi-locus techniques has become prevalent and practical, especially for detecting minor effect loci in complex polygenic traits like heat tolerance[31–36]. Additionally, it is recommended to adopt a multi-model approach, wherein multiple models are used, and only loci consistent across different models are reported. This approach leverages the algorithmic strengths of various models and reinforces the findings by corroborating them with multiple approaches[42,67–72].

In the current GWAS study, six different multi-locus models were utilized, including multi-locus random-SNP-effect mixed linear model (mrMLM)[31], FASTmrMLM[32], fast multi-locus random-SNP-effect EMMA (FASTmrEMMA)[33], pKWmEB[34], pLARmEB[35], and Iterative modified-Sure Independence Screening EM-Bayesian LASSO (ISIS EMBLASSO)[36]. These models were implemented in the current study using the R package mrMLM v3.1[37]. A LOD score of ≥ 3.00 was used as the threshold to declare stable QTNs. The statistical power of ML-GWAS has been found to outperform SL-GWAS methods, as reported in research by[45].

Favourable QTN analysis

In the GWAS analysis, the significant QTNs detected by at least three GWAS models were considered as reliable associations. To understand the potential of underlying genes associated with the SNPs, a physical map of wheat chromosomes was created using Map Chart 2.3 (https://www.wur.nl/en/show/Mapchart.htm), showing the SNPs associated with their respective genomic locations. QTNs detected by at least three models and present in a single environment were considered as significant QTNs. On the other hand, QTNs detected by at least three models and present in at least three different environments were classified as stable QTNs.

To identify potential candidate genes, only stable QTNs were annotated. The significantly associated SNPs with grain quality traits were searched against the *Triticum aestivum* genome assembly IWGSC-refseq version 1.0 through the online web resource Ensemble Plants (https://plants.ensembl.org/Triticumaestivum/Tools/Blast). The genes located in the overlapping region, within 1 Mb upstream and downstream of the matched regions were chosen as candidate genes. Furthermore, for cross validation of results from Ensemble Plants, DAVID online tool (https://david.ncifcrf.gov/summary.jsp) was utilized, which provides functional annotation and GO information for the detected genes. Additionally, the GWAS-related analyses were performed using the R package mrMLM v3.1. The stable QTNs were given more importance, as they were considered more reliable and expected to perform well across different environments. This approach aimed to identify robust associations between the genomic regions and the grain quality traits in wheat.

Validation of identified QTNs

The QTNs were validated through the creation KASP markers derived from significant SNPs (Fig. 8). To accomplish this, a 100 bp sequence flanking each side of the significant SNPs was extracted from the Chinese Spring reference genome RefSeqv1.0 based on their physical positions. Subsequently, PCR-based markers were developed utilizing bio-polyploidtools v0.9.7.[73]. KASP primers have been designed from 51 QTN of which 33 found to be polymorphic on panel of 101 lines consisting of heat stress tolerant and susceptible lines (Table S16-a,b). The panel has been procured from BISA, Ludhiana and the validation set was partitioned into two categories (Tolerant and Susceptible). Among the 101 genotypes, 66 were classified as Tolerant (0–49%), while 35 genotypes were identified as Susceptible (29–76%). To design markers targeting the putative SNPs, 100 bp sequence surrounding the candidate SNP on either side was extracted from the wheat reference genome IWGSC RefSeq version 2.1. KASP primers were designed from the putative SNPs using Polymarker (https://doi.org/10.1093/bioinformatics/btv069). Two locus-specific forward primers were labeled with FAM (5′-GAAGGTGACCAAGTTCATGCT-3′) and HEX (5′-GAAGGTCGGAGTCAACGGATT-3′), respectively, while the common reverse primer remained unlabeled. The PCR assay was conducted in a 4 µl reaction volume, which included 2 µl of genomic DNA (30 ng/µl), 1.96 µl of 2X KASP reaction mix, and 0.04 µl of KASP assay mix (containing the three primers). This setup was performed in a 384-well plate using a thermocycler (Applied Biosystems Veriti). The thermal cycling protocol started with an initial denaturation at 95 °C for 15 min, followed by 10 touchdown cycles at 95 °C for 10 s and 61–57 °C for 1 min, with a 0.6 °C reduction per cycle. This was followed by 25 cycles at 95 °C for 20 s and 57 °C for 1 min for annealing. These three primers for a single KASP assay were ordered from Barcode Biosciences (www.barcodebiosciences.com). The genotypic data from homozygous alleles of the KASP markers were analyzed using the Kruskal test to assess statistically significant differences between alternative alleles[74].

Conclusion

In conclusion, our finding has shown that heat stress under late sown conditions has more drastic effects on the quality parameters of wheat than normal sown conditions. However, iron and zinc were the most stable parameters under both the regimes. Our study has identified 12 annotated and stable QTNs, for five different quality parameters with their putative candidate genes. Additionally, two KASP markers for grain iron and zinc content has been validated on chromosome 2A and 7D respectively, which will be useful in SNP assisted pyramiding of multiple QTNs to attain a high level of tolerance against heat stress and subsequently in development of heat tolerant bio-fortified wheat varieties. Our research highlights the large variability found in Indian wheat germplasm for quality traits, and the novel genomic regions regulating them, which makes this germplasm a valuable source for improving wheat quality and yield. By utilizing these germplasm, we can develop wheat varieties that are better adapted to local conditions and can withstand environmental stresses, ultimately leading

to improved food security and sustainability. Overall, this study provides a foundation for further research on quality traits of wheat under different environmental conditions and highlights the importance of conserving and utilizing germplasm for crop improvement.

Data availability

The datasets generated and analysed during the current study are available in the DRYAD repository, https://doi.org/10.5061/dryad.3n5tb2rs1.

Received: 4 May 2024; Accepted: 18 February 2025
Published online: 24 February 2025

References

1. Poudel, P., Poudel, M. R. & Puri, R. R. Evaluation of heat stress tolerance in spring wheat (*Triticum aestivum* L.) genotypes using stress tolerance indices in western region of Nepal. *J. Agric. Food Res.* **5**, 100179. https://doi.org/10.1016/j.jafr.2021.100179 (2021).
2. Pal, N., Saini, D. K. & Kumar, S. Breaking yield ceiling in wheat: progress and future prospects. *Int. Open* https://doi.org/10.5772/intechopen.102919 (2022).
3. Nahid, N. et al. Introductory chapter: Current trends in wheat research. *Curr. Trends Wheat Res.* https://doi.org/10.5772/intechopen.103763 (2022).
4. Sharma, D., Singh, R., Tiwari, R., Kumar, R. & Gupta, V. K. Wheat responses and tolerance to terminal heat stress: A review. In *Wheat Production in Changing Environments* (eds Hasanuzzaman, M. et al.) (Springer, 2019). https://doi.org/10.1007/978-981-13-6883-7_7.
5. Brouns, F. J. P. H., van Buul, V. J. & Shewry, P. R. Does wheat make us fat and sick. *J. Cereal Sci.* **58**, 209–215. https://doi.org/10.1016/j.jcs.2013.06.002 (2013).
6. Pörtner, H.-O. et al. *IPCC: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2022). https://doi.org/10.1017/9781009325844.
7. McDonald, G. K., Sutton, B. G. & Ellsion, F. W. The effect of time of sowing on the grain yield of irrigated wheat in Namoi Valley, New South Wales. *Aust. J. Agric. Res.* **34**, 224–229. https://doi.org/10.1071/AR9830229 (1983).
8. Macas, B., Gomes, C. & Dias, A. S. Efeito das temperaturas elevadas durante o enchimento do grao em trigo mole e rijo no Sul de Portugal. *Melhoramento* **36**, 27–45 (1999).
9. Macas, C., Gomes, M. C., Dias, A. S. & Coutinho, J. The tolerance of durum wheat to high temperatures during grain filling. In *Durum Wheat Improvement in the Mediterranean Region: New Challenges*, 257–261. (2000). http://om.ciheam.org/om/pdf/a40/00600039.pdf
10. Reynolds, M. P., Balota, M., Delgado, M. I. B., Amani, I. & Fischer, R. A. Physiological and morphological traits associated with spring wheat yield under hot, irrigated conditions. *Aust. J. Plant Physiol.* **21**, 717–730. https://doi.org/10.1071/PP9940717 (1994).
11. Coles, G. D. et al. Environmentally-induced variation in starch and non-starch polysaccharide content in wheat. *J. Cereal Sci.* **26**, 47–57. https://doi.org/10.1006/jcrs.1996.0102 (1997).
12. Wardlaw, I. F., Blumenthal, C., Larroque, O. & Wrigley, C. W. Contrasting effects of chronic heat stress and heat shock on kernel weight and flour quality in wheat. *Funct. Plant Biol.* **29**, 25–34. https://doi.org/10.1071/PP00147 (2002).
13. Gooding, M. J., Ellis, R. H., Shewry, P. R. & Schofield, J. D. Effects of restricted water availability and increased temperature on the grain filling, drying and quality of winter wheat. *J. Cereal Sci.* **37**, 295–309. https://doi.org/10.1006/jcrs.2002.0501 (2003).
14. Stone, P. J. & Nicolas, M. E. Comparison of sudden heat stress with gradual exposure to high temperature during grain-filling in two wheat varieties difference in heat tolerance. II. Fractional protein accumulation. *Aust. J. Plant Physiol.* **25**, 1–11. https://doi.org/10.1071/PP96103 (1998).
15. Corbellini, M. et al. Effect of the duration and intensity of heat shock during grain filling on dry matter and protein accumulation, technological quality and protein composition in bread wheat and durum wheat. *Aust. J. Plant Physiol.* **24**, 245–250. https://doi.org/10.1071/PP96067 (1997).
16. Majoul, T., Bancel, E., Triboi, E., Ben Hamida, J. & Branlard, G. Proteomic analysis of the effect of heat stress on hexaploid wheat grain: Characterization of heat-responsive proteins from total endosperm. *Proteomics* **3**, 175–183. https://doi.org/10.1002/pmic.200390026 (2003).
17. Narendra, M. C., Roy, C., Kumar, S., Virk, P. & De, N. Effect of terminal heat stress on physiological traits, grain zinc and iron content in wheat (*Triticum aestivum* L.). *Czech J. Genet. Plant Breed.* **57**, 43–50. https://doi.org/10.17221/63/2020-CJGPB (2021).
18. Hossain, A. et al. *Wild Relatives of Plants as Sources for the Development of Abiotic Stress Tolerance in Plants. In Plant Perspectives to Global Climate Changes* 471–518 (Academic Press, 2022).
19. Sharma, D. et al. Population structure, marker-trait association and identification of candidate genes for terminal heat stress relevant traits in bread wheat (*Triticum aestivum* L. em Thell). *Plant Genet. Res.* **18**(3), 168–178 (2020).
20. Hays, D. B., Do, J. H., Mason, R. E., Morgan, G. & Finlayson, S. A. Heat stress induced ethylene production in developing wheat grains induces kernel abortion and increased maturation in a susceptible cultivar. *Plant Sci.* **172**, 1113–1123. https://doi.org/10.1016/j.plantsci.2007.03.004 (2007).
21. Qaseem, M. F. Genome-wide association mapping in bread wheat subjected to independent and combined high temperature and drought stress. *PLoS ONE* **13**, 199–221. https://doi.org/10.1371/journal.pone.0199121 (2018).
22. Qaseem, M. F., Qureshi, R., Shaheen, H. & Shafqat, N. Genome-wide association analyses for yield and yield-related traits in bread wheat (*Triticum aestivum* L.) under pre-anthesis combined heat and drought stress in field conditions. *PLoS ONE* **14**, 213–407. https://doi.org/10.1371/journal.pone.0213407 (2019).
23. Hall, D., Tegström, C. & Ingvarsson, P. K. Using association mapping to dissect the genetic basis of complex traits in plants. *Briefings Funct. Genom.* **9**(2), 157–165 (2010).
24. Yu, J. & Buckler, E. S. Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* **17**, 155–160. https://doi.org/10.1016/j.copbio.2006.02.003 (2006).
25. Yu, J. et al. Unified mixed model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**, 203–208. https://doi.org/10.1038/ng1702 (2006).
26. Semagn, K., Bjørnstad, A. & Xu, Y. The genetic dissection of quantitative traits in crops, Electron. *J. Biotechnol.* **13**, 16–17. https://doi.org/10.2225/vol13-issue5-fulltext-21 (2010).
27. Gupta, P. K., Kulwal, P. L. & Jaiswal, V. Association mapping in crop plants: Opportunities and challenges. *Adv. Genet.* **85**, 109–147. https://doi.org/10.1016/B978-0-12-800271-1.00002-0 (2014).
28. Gupta, P. K., Kulwal, P. L. & Jaiswal, V. Association mapping in plants in the post-GWAS genomics era. *Adv. Genet.* **104**, 75–154. https://doi.org/10.1016/bs.adgen.2018.12.001 (2019).
29. Maulana, F., Huang, W., Anderson, J. D. & Ma, X. F. Genome-wide association mapping of seedling drought tolerance in winter wheat. *Front. Plant Sci.* **11**, 573786 (2020).

30. Kumar, S. et al. Genome-wide association study reveals genomic regions associated with ten agronomical traits in wheat under late-sown conditions. *Front. Plant Sci.* **10**, 33889. https://doi.org/10.3389/fpls.2020.549743 (2020).
31. Wang, S. B. et al. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* **6**, 1–10 (2016).
32. Tamba, C. L. & Zhang, Y. M. A fast mrMLM algorithm for multi-locus genome-wide association studies. *Biorxiv* https://doi.org/10.1101/341784 (2018).
33. Wen, Y. J. et al. Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Brief Bioinform.* **19**(4), 700–712. https://doi.org/10.1093/bib/bbw145 (2018).
34. Ren, W. L. et al. pKWmEB: Integration of Kruskal–Wallis test with empirical Bayes under polygenic background control for multi-locus genome-wide association study. *Heredity* **120**, 208–218. https://doi.org/10.1038/s41437-017-0007-4 (2018).
35. Zhang, J. et al. pLARmEB: Integration of least angle regression with empirical Bayes for multilocus genome-wide association studies. *Heredity* **118**, 517–524. https://doi.org/10.1038/hdy.2017.8 (2017).
36. Tamba, C. L., Ni, Y. L. & Zhang, Y. M. Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies. *PLOS Comput. Biol.* **13**(1), 1005357. https://doi.org/10.1371/journal.pcbi.1005357 (2017).
37. Yang, Y., Chai, Y., Chen, L. & Hu, Y. G. Multi-locus GWAS of quality traits in bread wheat: Mining more candidate genes and possible regulatory network. *Front. Plant Sci.* **11**, 548041 (2020).
38. Alemu, S. et al. Multi-model genome-wide association and genomic prediction analysis of 16 agronomic, physiological and quality related traits in ICARDA spring wheat. *Euphytica* **217**(11), 1–22. https://doi.org/10.1007/s10681-021-02933-6 (2021).
39. Muqaddasi, Q. H. et al. Prospects of GWAS and predictive breeding for European winter wheat's grain protein content, grain starch content, and grain hardness. *Sci. Rep.* **10**(1), 12541 (2020).
40. Gilmour, A. R., Thompson, R. & Cullis, B. R. Average Information REML: An efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51**(4), 1440. https://doi.org/10.2307/2533274 (1995).
41. Myles, S. et al. Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell* **21**(8), 2194–2202. https://doi.org/10.1105/tpc.109.068437 (2009).
42. Vikas, V. K. et al. Multi-locus genome-wide association studies (ML-GWAS) reveal novel genomic regions associated with seedling and adult plant stage leaf rust resistance in bread wheat (*Triticum aestivum* L.). *Heredity* **128**, 434–449. https://doi.org/10.1038/s41437-022-00525-1 (2022).
43. Danakumara, T. et al. Genetic dissection of seedling root system architectural traits in a diverse panel of hexaploid wheat through multi-locus genome-wide association mapping for improving drought tolerance. *Int. J. Mol. Sci.* **22**, 7188. https://doi.org/10.3390/ijms22137188 (2021).
44. Mir, Z. A. et al. Recent advances on genome-wide association studies (GWAS) and genomic selection (GS); prospects for Fusarium head blight research in Durum wheat. *Mol. Biol. Rep.* **50**, 3885–3901. https://doi.org/10.1007/s11033-023-08309-4 (2023).
45. Zhang, Y., Jia, Z. & Dunwell, J. M. Editorial: The applications of new multi-locus GWAS methodologies in the genetic dissection of complex traits. *Front. Plant Sci.* **10**, 1–6. https://doi.org/10.3389/fpls.2019.00100 (2019).
46. Chauhan, H., Khurana, N., Nijhavan, A., Khurana, J. P. & Khurana, P. The wheat chloroplastic small heat shock protein (sHSP6) is involved in seed maturation and germination and imparts tolerance to heat stress. *Plant Cell Environ.* **35**(11), 1912–1931. https://doi.org/10.1111/j.1365-3040.2012.02525.x (2012).
47. Jadon, V. et al. Molecular mapping of biofortification traits in bread wheat (*Triticum aestivum* L.) using a high-density SNP based linkage map. *Genes* **14**(1), 221 (2023).
48. Gaggar, P., Kumar, M. & Mukhopadhyay, K. Genome-scale identification, in silico characterization and interaction study between wheat SNARE and NPSN gene families involved in vesicular transport. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **18**(6), 2492–2501. https://doi.org/10.1109/TCBB.2020.2981896 (2021).
49. Wang, X. et al. Wheat TaNPSN SNARE homologues are involved in vesicle-mediated resistance to stripe rust (*Puccinia striiformis* f. sp. tritici). *J. Exp. Bot.* **65**(17), 4807–4820. https://doi.org/10.1093/jxb/eru241 (2014).
50. Zheng, H. et al. NPSN11 is a cell plate-associated SNARE protein that interacts with the syntaxin KNOLLE. *Plant Physiol.* **129**, 530–539. https://doi.org/10.1104/pp.003970 (2002).
51. Uemura, T. et al. Systematic analysis of SNARE molecules in Arabidopsis: Dissection of the post-Golgi network in plant cells. *Cell Struct. Funct.* **29**(2), 49–65. https://doi.org/10.1247/csf.29.49 (2004).
52. Bao, Y. M., Wang, J. F., Huang, J. & Zhang, H. S. Molecular cloning and characterization of a novel SNAP25-type protein gene OsSNAP32 in rice (*Oryza sativa* L.). *Mol. Biol. Rep.* **35**, 145–152. https://doi.org/10.1007/s11033-007-9064-8 (2008).
53. Wani, S. H. et al. Improving zinc and iron biofortification in wheat through genomics approaches. *Mol. Biol. Rep.* **49**, 8007–8023. https://doi.org/10.1007/s11033-022-07326-z (2022).
54. Xie, Y., Ravet, K. & Pearce, S. Extensive structural variation in the Bowman-Birk inhibitor family in common wheat (*Triticum aestivum* L.). *BMC Genomics* **22**(1), 218. https://doi.org/10.1186/s12864-021-07475-8 (2021).
55. Kurt, F. & Filiz, E. *Functional Divergence in Oligopeptide Transporters in Plants* 167–183 (Elesiver, 2022). https://doi.org/10.1016/B978-0-323-85790-1.00018-X.
56. Yang, L. et al. QTL mapping for grain yield-related traits in bread wheat via SNP-based selective genotyping. *Theor. Appl. Genet.* **133**, 857–872. https://doi.org/10.1007/s00122-0199-03511-0 (2020).
57. Li, Q. et al. Quantitative trait locus (QTLs) mapping for quality traits of wheat based on high density genetic map combined with bulked segregant analysis RNA-seq (BSR-Seq) indicates that the *Basic 7S Globulin* gene is related to falling number. *Front. Plant. Sci.* **11**, 600788. https://doi.org/10.3389/fpls.2020.600788 (2020).
58. Guo, Y. et al. Genetic dissection of protein and starch during wheat grain development using QTL mapping and GWAS. *Front. Plant Sci.* **14**, 1189887. https://doi.org/10.3389/fpls.2023.1189887 (2023).
59. Shahzad, M. et al. Identification of QTLs on chromosome 1B for grain quality traits in bread wheat (*Triticum aestivum* L.). *Cytol. Genet.* **50**, 89–95. https://doi.org/10.3103/S0095452716020110 (2016).
60. Gilbert, G. A. & Spragg, S. P. *Methods in Carbohydrate Chemistry* Vol. 4, 168 (Academic Press, 1964). https://doi.org/10.1002/ange.19650770520.
61. Shar, G. Q., Kazi, T. G., Jakhrani, M. A. & Sahito, S. R. Determination of iron, zinc and manganese in nine varieties of wheat (*Triticum aestivum* L.) And wheat flour by using atomic absorption spectrophotometer. *Asian J. Plant Sci.* **1**, 208–209. https://doi.org/10.3923/ajps.2002.208.209 (2002).
62. Venables, W. N. & Smith, D. M. The R development core team, An Introduction to R, Version, 1(0) (2003).
63. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multi-locus genotype data. *Genetics* **155**, 945–959. https://doi.org/10.1093/genetics/155.2.945 (2000).
64. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**(8), 2611–2620 (2005).
65. Earl, D. A. & VonholdT, B. M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7 (2012).
66. Bradbury, P. J. et al. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308 (2007).
67. Sachdeva, S. et al. Multi-model genome-wide association studies for appearance quality in rice. *Front. Plant Sci.* **14**, 1304388 (2024).

68. Zhong, H. et al. Uncovering the genetic mechanisms regulating panicle architecture in rice with GPWAS and GWAS. *BMC Genomics* **22**, 86. https://doi.org/10.1186/s12864-021-07391-x (2021).
69. Gahlaut, V., Jaiswal, V., Balyan, H. S., Joshi, A. K. & Gupta, P. K. Multi-locus GWAS for grain weight-related traits under rain-fed conditions in common wheat (*Triticum aestivum* L.). *Front. Plant Sci.* https://doi.org/10.3389/fpls.2021.758631 (2021).
70. Malik, P. et al. Single-trait, multi-locus and multi-trait GWAS using four different models for yield traits in bread wheat. *Mol. Breed.* **41**, 1–21 (2021).
71. Kumar, M. et al. GWAS and genomic prediction for pre-harvest sprouting tolerance involving sprouting score and two other related traits in spring wheat. *Mol. Breed.* **43**(3), 14 (2023).
72. Pradhan, A. K. et al. Identification of novel QTLs/defense genes in spring wheat germplasm panel for seedling and adult plant resistance to stem rust and their validation through KASP marker assays. *Plant Dis.* **107**(6), 1847–1860 (2023).
73. Ramirez-Gonzalez, R. H., Uauy, C. & Caccamo, M. PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics* **31**, 2038–2039. https://doi.org/10.1093/bioinformatics/btv069 (2015).
74. Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1**, 583–621. https://doi.org/10.2307/2280779 (1952).

Acknowledgements
We are thankful to director ICAR-NBPGR, New Delhi, for providing funding and technical assistance to our study. We are also thankful to Dr. Uttam Kumar, Dr. Pradeep K Bhati and Mr Manish for his support in filed trails at different locations.

Author contributions
Conceptualization and methodology: LB, SK, AP, NB; Data analysis: LB, AK, NB, DS; Writing–original draft preparation: LB, NB; Writing–reviewing and editing: all authors; Funding acquisition: SK, GP; Resources: SK, RB, DC, SK; Supervision: SK, SR, RS, KG, NB.

Funding
This was students research work and was funded by the ICAR-NBPGR, IARI, New Delhi.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91199-2.

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

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.